Typological Studies in Language 124

Nominalization in Languages of the Americas

edited by Roberto Zariquiey, Masayoshi Shibatani and David W. Fleck

John Benjamins Publishing Company

Nominalization in Languages of the Americas

Typological Studies in Language (TSL)
ISSN 0167-7373

A companion series to the journal Studies in Language. Volumes in this series are functionally and typologically oriented, covering specific topics in language by collecting together data from a wide variety of languages and language typologies.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/tsl

Volume 124

Nominalization in Languages of the Americas
Edited by Roberto Zariquiey, Masayoshi Shibatani and David W. Fleck

Nominalization in Languages of the Americas

Edited by

Roberto Zariquiey
Pontificia Universidad Católica del Perú

Masayoshi Shibatani
Rice University

David W. Fleck
American Museum of Natural History

John Benjamins Publishing Company
Amsterdam / Philadelphia

DOI 10.1075/tsl.124

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2018058728 (PRINT) / 2019010765 (E-BOOK)

ISBN 978 90 272 0244 4 (HB)
ISBN 978 90 272 6273 8 (E-BOOK)

John Benjamins Publishing Company · https://benjamins.com

Table of contents

Part III. Case studies on nominalization in individual languages

CHAPTER 1

Nominalization in languages of the Americas

An introduction

Roberto Zariquiey, Masayoshi Shibatani and David W. Fleck
Pontificia Universidad Católica del Perú / Rice University / American Museum of Natural History

Recent scholarship on relativization and subordination in the languages of the Americas has confirmed the crucial role that nominalization plays in the formation of complex constructions in these languages (Comrie & Estrada-Fernández 2012; van Gijn, Haude & Muysken 2011). The interaction between nominalization and relativization is pervasive in Amerindian languages (and elsewhere), as demonstrated by nine out of the ten papers devoted to specific languages (or language families) in Comrie and Estrada-Fernández (2012). The papers in this collection show that nominalization plays a vital role in relativization in Tuscarora (Iroquoian), Northern Paiute (Uto-Aztecan), Pima Bajo (Uto-Aztecan), Yaqui (Uto-Aztecan), Seri (isolate/Hokan), Yucatec Maya (Mayan), Hup (Nadahup), and Toba (Guaycuruan). A similar panorama is revealed by the papers in van Gijn, Haude & Muysken (2011), which indicate that, together with serialization/clause-chaining and switch-reference, nominalization is among the most important and widespread structures underlying subordinate constructions in South American languages.

Although nominalizations are among the most salient and widespread features of Amerindian languages (and elsewhere), they have not been considered in their own right; but rather have been habitually looked at through the glass of their manifestations in specific constructions such as relative and complement clause constructions and other types of subordination. This distorted perspective has resulted from the imbalance in the past scholarship that has tended to focus on grammatical constructions such as relative clauses and subordinate clauses at the expense of nominalization structures themselves. No introductory linguistics textbooks fail to mention relative clauses, but how many of them discuss nominalizations, even lexical nominalizations? Wikipedia has lengthy and detailed discussions on relative clauses, but dedicates only a small space to the discussion of nominalization. This is surely at odds with the insight of the philosopher

https://doi.org/10.1075/tsl.124.01zar

of language Zeno Vendler, who more than fifty years ago pointed out that "the grammar of nominalizations is a centrally important part of linguistic theory" (Vendler 1967: 125). In this context, the foremost goal of the present volume is to correct the bias in the field, in particular the prevailing lopsided perspective on the relationships between nominalizations and their uses in specific grammatical constructions.

Most of the papers included in this collection treat grammatical nominalizations as such, by looking at their multiple uses and functions, their internal and external syntax, and their diachronic development in a representative sample of Amerindian languages, with detailed discussions focused mainly on those in South America. The list of all the languages from the Americas that are the objects of detailed case studies and those that have been analyzed for theoretical and areal studies in this volume is presented in Table 1. Map 1 provides the approximate location of all the extant languages in Table 1.

Table 1. List of Amerindian languages examined in this volume

Language	ISO 639-3	Language family
Aguaruna	agr	Jivaroan
Apurinã	apu	Arawakan
Apyãwa	taf	Tupí-Guaraní
Awa Pit	kwi	Barbacoan
Southern Aymara	ayc	Aymaran
Ayoreo	ayo	Zamucoan
Barasano	pok	Tucanoan
Baure	brg	Arawakan
Bora	boa	Witotoan
Borôro	bor	Bororoan
Cahita	yaq	Uto-Aztecan
Kapanawa	kaq	Panoan
Cavineña	cav	Tacanan
Central Alaskan Yup'ik	esu	Eskimo-Aleut
Central Aymara	ayr	Aymaran
Cherokee	chr	Iroquoian
Chipaya	cap	Uru-Chipaya
Creek	mus	Muskogean
Crow	cro	Siouan-Catawban
Cubeo	cub	Tucanoan
Ecuadorian Siona	sey	Tucanoan

Table 1. (*continued*)

Language	ISO 639-3	Language family
Embera	emp	Chocoan
Emerillon	eme	Tupí-Guaraní
Gavião	gvo	Tupian
Halkomelem	hur	Salish
Harakmbut	amr	isolate
Hixkaryana	hix	Cariban
Hup	jup	Nadahup
Iskonawa	isc	Panoan
Jamul Tiipay	dih	Cochimí-Yuman
Jarawara	jaa	Arawan
K'iche'	quc	Mayan
Kakataibo	cbr	Panoan
Kamaiurá	kay	Tupian
Karitiana	ktn	Tupí-Arikem
Kiowa	kio	Kiowa-Tanoan
Kipeá	kzw	Isolate
Kwazá	xwa	Isolate
Lakhota	lkt	Siouan-Catawban
Leko	lec	isolate
Lule	–	Lule-Vilela (?) – extinct
Mapudungun	arn	isolate
Matses	mcf	Panoan
Mẽbengokre	txu	Jêan
Mekens	skf	Tupian
Millcayac	–	Huarpean – extinct
Mojave	mov	Cochimí-Yuman
Mosetén	cas	Mosetenan
Movima	mzp	isolate
Musqueam	hur	Salish
Navajo	nav	Eyak-Athabaskan
Nheengatú	yrl	Tupí-Guaraní
Nieves Mixtec	mxv	Otomanguean
Nivaclé	cag	Mataguayan
Northern Paiute	pao	Uto-Aztecan
Páez	pbb	isolate
Parkatêjê	gvp	Jêan

(*continued*)

Table 1. (*continued*)

Language	ISO 639-3	Language family
Piapoco	pio	Arawakan
Pilagá	pig	Guaycuruan
Pima Bajo	pia	Uto-Aztecan
Pirahã	myp	Muran
Quechua (Ayacucho)	quy	Quechuan
Quechua (Bolivian)	quh	Quechuan
Quechua (Cuzco)	quz	Quechuan
Quechua (Huallaga)	qub	Quechuan
Rainy River Ojibwa	ojs	Algonquian
River Thompson Salish	str	Salish
Shawi	cbt	Kawapanan
Shoshone	shh	Uto-Aztecan
Slave	den	Eyak-Athabaskan
St'át'imcets	lil	Salish
Tapiete	tpj	Tupí-Guaraní
Tariana	tae	Arawakan
Timbira	xri	Jêan
Toba or Qom	tob	Guaykuruan
Trió	tri	Cariban
Trumai	tpy	isolate
Tupinambá	tpk	Tupí-Guaraní
Tuyuca	tue	Tucanoan
Uchumataqu	–	Uru-Chipaya
Vilela	vil	Lule-Vilela (?)
Waiwai	waw	Cariban
Wampis	hub	Jibaroan
Wari'	pav	Chapacuran
Wichí	mzh	Matacoan
Xavante	xav	Jêan
Yagua	yad	isolate
Yaminawa	yaa	Panoan
Yaqui	yaq	Uto-Aztecan
Yucatec Maya	yua	Mayan
Yurakaré	yuz	isolate

Map 1. Approximate location of the Amerindian languages examined in this volume

Several papers in the present collection openly state that the structures that accomplish the relativization and complementation functions are not independent clauses apart from grammatical nominalizations. Thus, the novelty of the present volume, and perhaps one of its main strengths, is its innovative approach, in which nominalizations are taken as the starting point, rather than interpreted from the perspective of their usage-based manifestations such as relative clauses and complement or adverbial clauses. Despite the variety of approaches followed by the papers that compose this volume, all of them share this basic understanding of nominalization. For example, multiple papers in this volume discuss the relativizing or complementation function of grammatical nominalizations, rather than

describing them as "nominalization strategies" or complement or relative clauses as if these constructions exist apart from nominalizations. The data in some contributions not strictly aligning with this orientation can easily be reinterpreted in the newer light, where functional unity, rather than formal manifestations, is considered to be a starting point for the comparison and understanding of the inevitable crosslinguistic variations in form.

This approach largely follows the ideas that one of the editors of this volume, Masayoshi Shibatani, has been exploring and expounding during the last decade in multiple conferences and seminars, like the First International Workshop on Nominalization in Indigenous languages of the Americas, hosted by the Pontificia Universidad Católica del Perú in 2014, which resulted in the collection presented in this volume. Shibatani's approach to nominalization is presented in this volume for the first time as a comprehensive and systematic monograph. Shibatani's chapter focuses mainly on grammatical nominations, which have been understudied compared to lexical nominalizations. Assuming a strictly functional stance, indeed a perspective more radical than ones taken by many functionalists, Shibatani defines nominalizations as denotative expressions like ordinary nouns, but metonymic in nature. His paper presents a convincing argumentation on how these functional properties help understand the prototypical grammatical behavior of nominalizations in the languages of the world. Shibatani's ideas on nominalization open fundamental questions not only about the nature of nominalization itself, but also about the notion of clause, sentence, subordination, clause-chaining, finiteness, predication, denotation and reference, all of which have been ill-defined despite the central role they play in grammatical descriptions and debates. These questions will surely be the focus of consequential debates as intended by this volume and are likely to be relentlessly pursued in our discipline. While Shibatani's contribution in this volume has far-reaching theoretical implications, its focus is mainly on empirical and conceptual issues, including discussions of different types of nominalization, the crosslinguistic usage patterns of different types of nominalization structures, the form-function correlations, and the crosslinguistic and diachronic patterns of nominalization markers.

The other theory-oriented paper in this volume is Cristofaro's contribution, which focuses on the available diachronic evidence for the possible origins of nominalization. This evidence provides a perspective useful for understanding the synchronic properties that nominalizations exhibit cross-linguistically and poses important challenges for a number of traditional assumptions about the nature of nominalization. Like Shibatani's discussion on the rise and spread of nominalization markers, Cristofaro's diachronic approach to the typology of nominalization is in direct relation with various papers included in this volume that focus on the historical development of nominalizations in a diverse range of languages and language families (see, for instance, Álvarez, Bruil, Cruz and Praça, and Gipper and Yap).

Two of the papers in this volume focus on areal linguistics. Van Gijn explores the interaction between subordination and case marking in the languages of South America by offering a systematic account of case-marked adverbial structures. His study reveals that there is a well supported diachronic path from grammatical nominalization to adverbial-like constituents (as predicted by Shibatani's analysis of the adverbial use of nominalizations). Furthermore, the methods used by van Gijn prove that the presence and distribution of case markers on adverbial structures exhibit systematic patterns that suggest transparent areal effects and interesting genealogical retentions.

Golluscio, Hasler and de Reuse offer a comparison of the formal differences in complementation structures (Shibatani's NP-use of event nominalizations) in languages of the Chaco with those in Andean and Amazonian languages, finding several important differences among these. For example, whereas Andean and Amazonian languages typically exhibit dedicated morphosyntactic marking for lexical and grammatical nominalizations, Chacoan languages generally utilize more general dependency markers construable as markers of grammatical normalizations. Chacoan languages are also distinguished from Amazonian and Andean languages in that the subordinators that mark grammatical nominalizations do not mark lexical nominalizations. They conclude that these and other properties associated with nominalizations and their complementation-use are promissory criteria for establishing the Chaco as a linguistic area.

The remaining 12 papers included in the present volume focus on the synchronic and/or diachronic properties of nominalization in specific Amerindian languages or language families. Detailed discussions of nominalization in 14 languages, belonging to 10 different genetic units (8 language families and 2 isolates) are included in this volume (see Table 2). Summaries of the topics and principal findings of each of these contributions are provided in the following paragraphs.

In line with Shibatani's ideas on nominalization, Yuki-Shige Tamura studies nominalization in Central Alaskan Yup'ik. The crucial point of his paper is that a single process (i.e., nominalization) is sufficient to account for multiple morphosyntactic phenomena that have been treated separately in previous descriptions; specifically, various types of lexical nominalizations, relativization and complementation. Tamura shows the importance that apposition has played in the development of the multiple functions of nominalizations in Central Alaskan Yup'ik.

Albert Álvarez presents a diachronic account of non-subject nominalizers in Cahita, a dialect continuum in the Uto-Aztecan family. Based on what he calls "relativization/nominalization syncretism", Álvarez argues against the relative-clause analysis of relativizing constructions in the Cahita dialects: the restricting function of relative clauses is accomplished simply by means of grammatical nominalizations. The evidence seems irrefutable, both from a synchronic and a

Table 2. List of Amerindian languages that are the topics of chapters in this volume

Author	Language(s)	ISO 639-3	Family
Tamura	Central Alaskan Yup'ik	esu	Eskimo-Aleut
Álvarez	Cahita (Tehueco, Yaqui, Mayo)	yaq	Uto-Aztecan
Cahlon	Cuzco Quechua	quz	Quechuan
Gipper and Yap	Yurakaré	yuz	Isolate
Bruil	Ecuadorian Siona	sey	Tucanoan
Peña	Wampis	hub	Jibaroan
Van linden	Harakmbut	amr	Isolate
Rojas-Berscia	Shawi	cbt	Kawapanan
Valle and Zariquiey	Kakataibo	cbr	Panoan
Zariquiey	Iskonawa	isc	Panoan
Fleck	Matses	mcf	Panoan
Estevam	Xavante	xav	Jê
Cruz and Praça	Tupinambá	tpk	Tupí-Guaraní
Cruz and Praça	Apyãwa	taf	Tupí-Guaraní
Cruz and Praça	Nheengatú	yrl	Tupí-Guaraní

diachronic perspective. The author shows that this relativizing function is the consequence of a more general function of noun modification that is typologically common for nominal constituents. The paper is also devoted to the study of the diachronic development of each of the markers that accomplish the function of non-subject nominalization. Interesting paths that revolve around possessive classifiers and nominal adpositions are discussed and formalized, supporting the view that these are likely sources for nominalizers (Shibatani this volume).

Also taking a diachronic approach in line with grammaticalization theory, Rammie Cahlon presents a discussion of habitual periphrasis of Cuzco Quechua. It becomes clear from his data (taken from published Cuzco Quechua sources) that habitual periphrasis has come from the reanalysis of lexical nominalizations functioning as noun phrases in a copula construction as a periphrastic verbal construction with an auxiliary. One of the most interesting findings of his paper is that this grammaticalization process is still ongoing and is enriching the TAM system of the language. Five different stages in association with this process are carefully distinguished and described in Cahlon's chapter.

Sonja Gipper and Fong Ha Yap explore the functions and development of the grammatical nominalizer *=ti* in Yurakaré, a language isolate spoken in the eastern foothills of the Bolivian Andes. The functions of this marker include relativization,

complementation, and other adverbial functions. Gipper and Yap's paper, like several others in this volume, is particularly important for establishing the functional relation between nominalization and so-called adverbial clauses, as discussed by Shibatani (this volume). Based on a corpus of spoken Yurakaré, the authors propose a grammaticalization path for *=ti* and argue that the most plausible source for this marker is a demonstrative, very much in line with the findings that Shibatani and Cristofaro present in this volume. One final issue discussed by Gipper and Yap is the process of the insubordination of *=ti*, which, according to the authors, is developing into a stance marker for 'intersubjective commitment'.

Taking as the starting point the strikingly ample homophony in the verbal morphology of Ecuadorian Siona (particularly in association with the morpheme *-kɨ*, which is used in reportative, dependent and nominalized constructions), Martine Bruil offers fascinating insights into the grammaticalization paths that give Siona's verbal morphology its atypical profile. The data on Ecuadorian Siona not only show the important role that nominalization played in the development of the complex system of clause types that the language exhibits, but also presents important support for the intrinsic relationship between nominalization and (numeral) classifiers explored by Shibatani in this volume. It turns out that the synchronic homophony that is characteristic of Siona verbal morphology is a consequence of the origin of various synchronic constructions of the language (reportatives, interrogatives and so on) from grammatical nominalizations.

Jaime Peña discusses grammatical nominalizations in Wampis, a Jibaroan (Chicham) language spoken at the western edge of the Peruvian Amazon. As is the case in other languages described in this volume, nominalizations in Wampis perform functions such as complementation and relativization. What is particularly interesting in this language is that nominalizations are developing new functions, some of which are equivalent to predication. In this function, Wampis nominalizations are able to predicate on their own and receive finite verb morphology (see the contributions by Valle and Zariquiey and Estevam for similar situations in Kakataibo and Xavante, respectively).

An Van linden presents a discussion of verb-based nominalization in the Amarakaeri variety of the Harakmbut language of Amazonian Peru, with some comments on noun-based nominalizations, a topic pursued in earnest in Shibatani's contribution. The author shows that in Harakmbut different nominalizers exhibit different (and clear-cut) semantic properties (some are dedicated participant/argument nominalizers and others are dedicated event nominalizers), when they are combined with different additional suffixes. The various morphological combinations exhibit different functions in discourse; for example, some nominalizations have a modification function, others present a complementation function, and another set has adverbial functions. This study also reveals a

recurrent grammatical patterning underlying this widely diverse set of nominalizations: all nominalizations exhibit NP-like external syntax and – when consisting of more than one word – clause-like internal syntax, similarly to other languages described in this volume.

Luis Miguel Rojas-Berscia offers an overview of nominalization in Shawi, a Kawapanan language spoken in Amazonian Peru. In his analysis, Rojas-Berscia proposes a distinction between two main types of nominalization in this language: strong and weak. The distinction between strong and weak nominalization is directly associated with the number of verbal features that are retained by the nominalized construction and the degree of recategorization that it exhibits (i.e. the number of nominal properties that the nominalized construction acquires). Since strong nominalizations in Shawi are highly lexicalized and weak nominalizations are not, the author argues that this distinction largely corresponds to that between lexical and grammatical nominalization in Shibatani's theoretical frame presented in this volume. It is interesting that in Shawi these different types of nominalization processes are achieved by means of different morphological markers. This is the case for European languages like English or Spanish, but not for most of the languages described in this volume (see, for instance, Álvarez for Cahita, and Valle and Zariquiey for Kakataibo).

Daniel Valle and Roberto Zariquiey discuss grammatical nominalizations in two dialects of Kakataibo, a Panoan language spoken in the Peruvian Amazon. Grammatical nominalizations are so frequent in Kakataibo discourse that the syntax of the language can hardly be understood without an adequate analysis of these structures. Valle and Zariquiey describe the multiple functions of nominalization in Kakataibo, ranging from relativization and complementation to stand-alone (insubordination) constructions that have a function equivalent to that of independent sentences in the language. The paper also focuses on the historical development of switch-reference markers, many of which come from the reanalysis of nominalizers and case makers.

Roberto Zariquiey explores the form and functions of grammatical nominalizations in Iskonawa, an obsolete Panoan language. The paper focuses on the extreme structural similarity between clauses and grammatical nominalizations, and shows that various switch-reference categories, which exhibit dedicated markers in other Panoan languages (see Valle and Zariquiey, this volume) are expressed by nominalizations in Iskonawa. After demonstrating that this simplified pattern is not the result of obsolescence, the paper poses interesting questions about the development of the complex switch-reference systems found in the Panoan family.

David Fleck describes a third language in the Panoan family, Matses, which is spoken in Peru and Brazil. After providing an overview of the structure and functions of nominalizations in Matses, Fleck's chapter focuses on comparing the

internal and external syntax of Matses grammatical nominalizations in comparison to lexicalized nominalizations. The notable finding of this study is that grammatical nominalizations can become lexicalized without losing their clause-like internal syntax.

Spoken in the Mato Grosso state of Brazil by approximately 15,000 speakers, Xavante is a member of the central group of languages of the Jê family. Adriana Estevam investigates the pervasiveness of nominalization in Xavante discourse from both diachronic and synchronic perspectives. After discussing the various functions of nominalizations in Xavante, Estevam focuses on the use of nominalizations as finite constructions of different types, arguing that the predication function of nominalizations in Xavante do have functional motivations.

Finally, Aline da Cruz and Walkíria Neiva Praça explore nominalizations in three Brazilian Tupí-Guaraní languages: Tupinambá, Apyãwa and Nheengatú. Thanks to the conservative character of the first two languages, the authors are able to offer a careful and convincing scenario for the diachronic development of several nominalizers in Tupí-Guaraní, finding important differences between the conservative Tupinambá and Apyãwa, and the innovative Nheengatú. It has been proposed that Tupí-Guaraní languages are descendants of an omnipredictive language; that is, a language in which all lexical entries were predicative. Nheengatú is also innovative with respect to this domain, in the sense that it lacks omnipredicative patterns. The authors argue that the loss of omnipredicativity has led to a radical restructuring of the form and the functions of nominalizers in this innovative language.

We are extremely thankful to all the authors for their valuable contributions and their patience throughout this complex and long editorial process. Thanks to the quality of their work and their expertise in the languages, language areas and language families discussed in their papers, we have been able to prepare the present volume, which will enrich the debate on grammatical nominalizations and will surely set the basis for new approaches to this fascinating phenomenon of paramount importance for the adequate understanding of Amerindian languages, especially their similarities to and differences from the languages of the world.

References

Comrie, Bernard & Zarina Estrada-Fernández (eds). 2012. *Relative Clauses in Languages of the Americas: A Typological Overview* [Typological Studies in Language 102]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.102

van Gijn, Rik, Haude, Katharina & Muysken, Pieter (eds). 2011. *Subordination in Native South American Languages* [Typological Studies in Language 97]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.97

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca NY: Cornell University Press.

PART I

Nominalization theory, definitions and typology

CHAPTER 2

What is nominalization? Towards the theoretical foundations of nominalization

Masayoshi Shibatani
Rice University

過而不改、是謂過矣
(孔子)

This paper discusses foundational issues in nominalization, focusing on empirical, conceptual, and theoretical problems that have plagued the field for years. Current definitions of nominalization are based on narrow observations on verbal-based nominalizations, while languages across the globe display nominal-based ones, many of which share morphology with the former. Nominalization morphology in many languages also applies to units larger than words, yielding grammatical nominalizations besides lexical nominalizations. The imbalance in the past studies, which have focused on so-called relative clauses at the expense of grammatical nominalizations, has resulted in a lopsided view on the relationship between the two. This, in turn, has led to the mishandling of different manifestations of nominalizations as if they are derivatives of relative clauses, as suggested by the widely used terms such as "headless relative clause" and "internally-headed relative clause". We demonstrate that these, including the ordinary restrictive relative clause, are not independent grammatical structures but are epiphenomena arising from the uses of grammatical nominalizations. A clear distinction between structures and their use is a theoretical prerequisite in arriving at a satisfactory understanding of the nature of grammatical nominalizations and their role in grammar.

1. Introduction

Studies on nominalization, in both Western and Eastern grammatical traditions, have largely concentrated on lexical nominalizations, neglecting grammatical nominalizations, despite their theoretical importance and far-reaching implications to the descriptive practice. This imbalance is due to the fact that while lexical nominalizations (e.g. English *sing-er*) typically involve distinct morphology and

https://doi.org/10.1075/tsl.124.02shi

their lexical status as nouns is relatively clear-cut, grammatical nominalizations (e.g. [I know] *that John recklessly shoots trespassers*; [I saw] *John shoot trespassers*; *John's recklessly shooting trespassers* [angered the entire community]; *To shoot trespassers* [is unacceptable]) vary considerably in form, some of which displaying internal structural properties similar to clauses, and their nominal status is less fully realized compared to lexical nominalizations (e.g. *a/the shooting* [of trespassers]; *those terrible shootings* [of trespassers], but not **a/*the shooting trespassers* [is unacceptable]).[1]

Our main concerns in this paper are these understudied types of nominalization, whose position in grammar, we claim, has not been fully assessed despite their far-reaching theoretical implications. Our emphasis is on the conceptual and empirical issues pertaining to grammatical nominalizations, because there are deeprooted and widespread misconceptions about them, hampering a proper perspective on and correct understandings of the relevant data. Because of this, we provide a multitude of data, not only from the Americas, but also elsewhere around the globe that bear on the topic, which to a great extent are redundant, but which are felt necessary to dispel the misunderstandings found in many, if not all, current definitions of nominalization. An abundance of data also helps correct some narrow views on naminalization found in several statements and analyses in professional publications, such as *"a somewhat more rare* function of nominalization [is its use] as a relative clause modifying a head noun" (Comrie & Thompson 2007: 378; emphasis added) and "*in certain languages* relativization is indistinct from nominalization" (Comrie & Thompson 2007: 379; emphasis added).

This paper is organized as follows. In the next section the current definitions of nominalization are placed under close scrutiny, with our findings that they are either incorrect or at best insufficient. We offer an alternative definition of nominalization as a metonymy-based phenomenon that applies not only to verbals but also to nominals and that yields structures that form nouns as well as phrasal units. Section 3 starts earnest discussions of verbal-based grammatical nominalizations. A major distinction is made between event nominalizations (Section 3.1) and argument nominalizations (Section 3.2) and their formal representations are attempted in Section 4. Section 5 starts discussions of the usage patterns of these grammatical nominalizations, dividing them into two major uses, the NP-use and the modification-use (Section 5.1). The latter includes an adverbial use of nominalizations, which are also used in so-called clause-chaining constructions. Section 5.2 demonstrates that so-called internally-headed relative clauses are not really relative clauses at all and that they are no more than an instance of NP-use

1. See Lees (1963) for an early, but still the most comprehensive treatment of English nominalizations.

of event nominalizations. Section 5.3 discusses another issue of whether so-called complement clauses are distinct from nominalizations in NP-use.

Similar arguments are advanced for argument nominalizations, first showing that so-called headless relative clauses are an NP-use of argument nominalizations and are not relative clauses at all, contrary to the widely subscribed labeling of them as such (Section 5.4). Then, Section 5.5 demonstrates that ordinary relative clauses are no more than a modification-use of argument nominalizations and that there are no structures such as relative clauses independently from argument nominalizations in modification-use. A new analysis of relative clauses is proposed that minimizes the role of so-called relative pronouns that play a significant role in the traditional as well as generative analyses of relative clauses.

In view of the lack of proper understandings of what clauses and sentences are in the field, Section 6.2 offers functional definitions of clauses and sentences and shows how they differ from nominalizations. Section 6.3 discusses the phenomenon of insubordination, by which event nominalizations may become used as sentences. We then go on to provide some evidence supporting our claim that nominalizations are neither clauses nor sentences (Sections 6.4 and 6.5).

Section 7 takes up nominal-based nominalizations showing that many languages of the world show a morphological connection between them and verbal-based nominalizations. The importance of nominal-based nominalizations is shown in Section 8, where it is demonstrated that the NP-use of N-based nominalization is the locus of innovations in the development of NP-use markers, which, by spreading to the modification-use of both N-based and V-based nominalizations, may eventually become nominalizers. Finally, Section 9 takes stock of our lengthy discussions on grammatical nominalizations and demonstrate their implications for both theoretical and desciptive issues inherent in the treatment of relative clauses by Keenan and Comrie (1977) and Comrie and Keenan (1979).

2. Defining nominalization

Popular definitions of nominalization found in dictionaries, encyclopedias, and introductory textbooks, as well as more technical scholarly papers such as the following are all problematic.

> nominalization refers to the process of forming a noun from some other word-class or, especially in classical transformational grammar, the derivation of a noun phrase from an underlying clause. (Crystal 1980: 328)

> …operations that allow a verb to function as a noun …are called **nominalizations**, and can be described with a simple formula: V→ N. (Payne 1997: 223)

> The term *nominalization* means 'turning something into a noun'.
> (Comrie & Thompson 1985/2007: 334)

> 'nominalization' actually conflate[s] two properties: deverbalization …and substantivization (acquisition of noun-properties). (Malchukov 2004: 6)

Below we focus on three issues that these definitions raise, namely (i) the overall characterization of the process, (ii) the inputs to the process, which are restricted to verbs or members of non-nominal categories (Crystal, Payne, Malchukov), and (iii) the outputs, defined as nouns (Payne, Comrie and Thompson, Crystal), or noun phrases (Crystal).

2.1 Nominalization as a metonymic process

The definitions of nominalization above are all too general in that they do not specify the relationships between the inputs and the outputs other than that the former are verbal and the latter nominal. It is, however, not the case that any type of nominalization is derivable from any verbal input. Imagine an extreme case of trying to derive forms using the English nominalization suffix *-er* such as *singer* and *driver* from verb roots like *walk* and *kill*. No language would allow drivations like *walk* > *singer* and *kill* > *driver*, while all the popular definitions of nominalization simply say that the process derives nouns or nominal expressions from verbs or members from non-nominal categories. Along a similar but more plausible line, consider deriving nominalizations denoting agents like *singer* and *driver* from stative or simple processes verbs such as *resemble* and *die* that denote non-activity relational properties, e.g. **resembler* and **dier*. Agentive nominalizations are derivable only from activity verbs that predicate over an agentive subject. Simple activity verbs, on the other hand, would not yield resultative nominalizations of the type such as (*a*) *painting*, (*a*) *building*, and (*his*) *writings*, which are associated with verbs of production and certain change-of-state verbs. In other words, nominalization is far more constrained than the overall characterizations of the conventional definitions have it.

In view of the relationships between the inputs and the outputs seen above, we define nominalization as a metonymic process along the line of Fillmore's frame semantics (Fillmore 1976, 1982).[2] Namely, verbal-based nominalization, for example, yields only those forms that denote things and thing-like entities (both

2. Allan (2001: 251) characterizes a semantic frame as consisting of "characteristic features, attributes, and functions of a denotatum, and *its characteristic interactions with things necessarily or typically associated with it*" (emphasis added). A similar effort is seen in Langacker's (1987) Cognitive Linguistics framework in terms of the notion of "profiling".

concrete and abstract) that are in close association with the scenes/scripts evoked by the use of specific verbs.[3] The resulting nominalization structures often have marking that more narrowly circumscribes the range of denotations they evoke, as in the English forms *employment*, *employer*, and *employee*. As these examples show, nominalization yields structures denoting substantive or entity concepts that are metonymically evoked by the nominalization structures themselves such as events, facts, and propositions (cf. *employment, employing*), as well as concrete objects such as event participants (agents and patients) and entities conventionally associated with specific events like instruments, resulting objects and locations (cf. *employer, employee, plier, a building*, *a landing*). As products, nominalizations are like nouns (hence the term "nominalization") by virtue of their association with an *entity-concept* denotation, a property that provides a basis for the referential function of a noun phrase headed by such nominalizations.[4] Verbs and verb phrases, on the other hand, are associated with *relational* concepts (time-stable or transient properties *pertaining to an entity or entities*) and play a predication function in a clause by ascribing a relational concept to the referent of a subject noun phrase (see Section 6.2). They differ crucially from nouns and nominalizations in not denoting things and thing-like entity-concepts and thereby in being unable to play a referential function.

Metonymy is a powerful cognitive process that allows a variety of form-concept connections increasing the expressive power of a language with limited resources. By taking advantage of our knowledge that many things in the world occur in close association, a metonymic construal allows us to conceptualize and denote entities in alternative ways beyond the conventional form-concept connections. The result of such a process is a richer (and often colorful) description of a denoted entity focusing on some associated features that the speaker finds more informative and relevant to the context. Specific metonymic expressions are not random but are manifestations of experientially-grounded general conceptual metonymic schemas of the type, THE PART FOR THE WHOLE (Get *your butt* over here!), PRODUCER FOR PRODUCT (He bought a *Ford*), THE PLACE FOR THE INSTITUTION (The *White House* isn't saying anything), etc. (Lakoff & Johnson 2003: 38). Schemas that play important roles in nominalization include THE EVENT FOR THE STATE/PROCESS/ACTIVITY, THE EVENT FOR THE FACT, THE EVENT FOR THE PROPOSITION (all for event nominalizations), THE EVENT FOR THE PROTAGONISTS

3. A good description of nominalization intimates metonymic form-meaning relationships involved in nominalizations; e.g., "A nominalization on an oblique (i.e., non-subject) argument … typically denotes an object or location *closely associated* with generic performances of the activity named by the verb from which it is derived." (Miller 2001: 120; emphasis added).

4. See Section 4 on the difference between denotation and reference.

(argument nominalizations), THE EVENT FOR THE RESULT (resultative nominalizations), POSSESSOR FOR POSSESSED, PRODUCER FOR PRODUCT (both for nominal-based nominalizations), and others.

A single metonymic expression may denote a variety of entity concepts that are closely associated with the concepts denoted by the original words or larger structures, and it is the speech context that determines and selects the denotation/reference most relevant to the context per Gricean maxims of conversation (or the Cooperative Principle), one of which (the Maxim of Relevance) requires an expression to be contextually relevant at the time of the utterance. For example, *the United States* may metonymically evoke and denote a variety of entities closely associated with the country by this name (e.g. different types of representatives of the country in question), but only a contextually relevant interpretation would be intended by the speaker and would be chosen by the hearer – e.g. the sitting US president in *the United States has decided to pull out from the Paris agreement*, or a US women's soccer team in *the United States defeated China 1–0 to advance to the semifinals of the 2015 FIFA Women's World Cup*. Likewise, the lexical nominalization *half-pounder*, based on the noun *half-pound* and is used in an expression like *Give me a half-pounder*, may denote a hamburger in a fast-food restaurant, a can of tobacco in a smoke shop, a bag of jelly beans in a candy shop, or a steelhead trout among fishing aficionados.

While many lexical nominalizations tend to have more uniform denotations, grammatical nominalizations do not have fixed or uniform denotations, and speech context plays an important role in determining and selecting the denotation/reference most consistent with the context. For example, the Spanish grammatical nominalization [el [que Ø es blanco]$_{\text{NMLZ}}$] (the [NMLZR Ø is white]) 'the one which is white' can refer to a range of objects classed as masculine matching its denotation of an entity that is white. In actual usage, the context and the Gricean Cooperative Principle determine the reference. So, *El que es blanco* would be understood to be referring to a white car when uttered in response to the question *¿Qué coche te gusta?* "Which car do you like?" and a white hat when it answers the question *¿Cuál sombrero usarás hoy?* "Which hat will you wear today?". Our claim is that there is nothing like a deletion of a head noun or a pronominal element involved here. The construction [el [que Ø es blanco] is a complete structure, a grammatical argument nominalization (see Section 3.2), whose reference in discourse is determined by the context, exactly like the determination of the actual reference of a metonymic expression such as *the United States* following questions like "Who has decided to pull out from the Paris agreement?" or "Who defeated China 1–0 to advance to the semifinals of the 2015 FIFA Women's World Cup?" or the like.

For the purposes of this paper, we offer the following definition of nominalization:

> Nominalization is a metonymy-based grammatical derivation process yielding constructions associated with a denotation comprised of entity (thing-like) concepts that are metonymically evoked by the nominalization structures, such as events, facts, propositions, resultant products and event participants. Nominalizations, as grammatical structures, are similar to nouns by virtue of their association with an entity-concept denotation; they both denote thing-like concepts, which provide a basis for the referential function of an NP headed by these nominals.

Notice at this juncture that nominalization is not a morphosyntactic notion. In particular, nominalization structures may or may not have associated morphology, as the comparison between the two relevant froms below indicates.

(2-1) a. *drive* > *driver*
b. *cook* > (*a*) *cook*

(2-2) a. Quechua (Cuzco dialect; Lefebvre and Muysken 1988)
[[*Xwan papa-ta mikh-sqa-n*]$_{\text{NMLZ}}$]$_{\text{NP}}$*-ta yacha-ni*.
Juan potato-ACC eat-NMLZR-3-ACC know-1SG
(lit.) 'I know Juan('s) eating potatoes.'
b. *I know* [[*Juan eats potatoes*]$_{\text{NMLZ}}$]$_{\text{NP}}$

The noun *cook* has no marking indicating that it is a nominalized from like the noun *driver* with a nominalization marker, yet *cook* and *driver* are functionally alike both semantically and syntactically. It is on the basis of this functional similarity that we treat the nouns *driver* and *cook* as nominalized forms despite the difference in morphology.[5] The same logic applies to the relevant structures in (2-2a) and (2-2b); semantically both denote a fact, and syntactically both function as an object argument of the verb meaning "know". Treating the relevant structure in (2-2a) as a nominalization, while not recongnizing a nominalization in (2-2b) on account of the presence and absence of nominalization morphology is like treating the Quechua form *wanu-či* (die-CAUS) 'kill' as causative but not the English verb *kill*. Just as causation is not a morphosyntactic notion, nominalization is not a morphosyntactically definable phenomenon.

2.2 Meaning range and the nature of lexical nominalizations

As seen above, metonymy allows various form-meaning connections beyond the conventional lexicon of basic, underived nouns. Nominalization as a grammatical process yields forms (nouns and larger structures or constructions), which are called "nominalizations" and are labeled as […]$_{\text{NMLZ}}$ in this paper, that denote

5. Labelling the derivation *cook* > (*a*) *cook* as "conversion" is besides the point.

things and thing-like entities evoked by the derived nominalization structures. Verbal-based nominalizations, as noted above, evoke concepts intimately related to what the verbal bases denote, namely states (*slowness, freezing*), processes (*flowing, freezing*), activities (*fight*, *skating*, *building, employment*), facts, propositions, attendant protagonists such as agents (the noun *cook, employer*) and patients (*fryer, keeper* "a fish that is of sufficient size to be caught and retained without violating the law", *employee*), resultant products (the nouns *freeze*, *building, painting, writing*), as well as instruments (*plier, screwdriver*) and locations (*landing, bus stop*) inherently or conventionally associated with particular events. While some lexical nominalizations involve morphology that delimits the range of meanings associated with the derived nominals, as in the case of the English suffix *-er* and others seen above, some others may form nominalizations with a greater range of meanings as with the so-called gerundive *-ing* form in English. Of the Japanese lexical nominalizations, stem nominalizations, involving *-i*/*-Ø* suffix, display a diverse array of meaning patterns on their own, but more typically in forming compounds with another nominal element.[6] However, as in the English cases discussed above, the form-meaning connections are not random and are metonymically bound such that some meaning patterns are more consistently observed while others are not. What follows summarizes major form-meaning patterns that Japanese stem nominalizations display.

(2-3) a. Process/Activity: *nagare* 'flowing', *suberi* 'sliding', *ugoki* 'movement', *oyogi* 'swimming', *ake-sime* 'opening and closing', *mawasi-yomi* (rounding-reading) 'reading by circulating reading materials in a group', *yama-nobori* 'mountain climbing', *hito-gorosi* (person-killing) 'manslaughter'

b. State/Characteristic person: *hare* 'fine weather condition', *yuu-yake* (evening-burning) 'evening glow', *zikan-gire* 'time expiration', *ame-agari* (rain-stopping) 'after the rain', *hanasi-zuki* (talk-liking) 'a talkative person', *Tookyoo-umare* 'a Tokyo-born person'

c. Agent/Natural force: *suri* 'pickpocket', *tasuke* 'helper,' *hito-gorosi* 'killer', *uso-tuki* (lie-telling) 'liar', *arasi* 'storm', *hubuki* 'snow storm'

d. Instrument/Chemical agent: *hasami* 'scissors', *hakari* 'scale', *nezimawasi* 'screwdriver', *tume-kiri* 'nail cutter', *ha-migaki* 'tooth paste', *simi-nuki* 'stain remover'

e. Patient: *yatoi* 'employee', *tukai* 'errand runner', *tumami* 'what is picked/hors d'oeuvre', *ture* 'one taken along/companion'

6. These nominalizations, often labeled 'infinitive', involve the suffix *-i* attaching to a consonant-final root/stem (e.g. *odor-* > *odor-i* 'dancing/dance'), and *-Ø* to a vowel-final root/stem (e.g. *kake-* > *kake* 'betting'). See Shibatani (2018b) for discussions of these lexical nominalizations and their theoretical implications, including critiques of earlier treatments of them.

f. Resultant product: *koori* 'ice', *yogore* 'stain', *age* 'thin, fried tofu', *kangae* 'thought', *kasi* 'loan', *sirase* 'message', *sasayaki* 'a whisper', *saezuri* 'a chirp'
g. Location: *hanare* 'detached room/house', *nagasi* 'sink', *watasi* 'landing pier', *mono-hosi* 'cloth-drying place'
h. Game name: *tako-age* 'kite flying', *karuta-tori* 'playing Japanese cards', *nawa-tobi* 'rope skipping'
i. Sports technique: *seoi-nage* (Jūdō), *oosoto-gari* (Jūdō), *uwate-nage* (Sumō), *osi-dasi* (Sumō)

What we observe in the remainder of this paper is that grammatical nominalizations to a large extent parallel these lexical nominalizations in the range of meaning extension, indicating that these two are closely connected phenomena, a fact that is also indicated by morphology in many languages (see Section 2.4). Indeed, there are historical connections between grammatical nominalizations and lexical nominalizations such that the latter often arise from the former (see Fleck, this volume, for an extensive discussion on this). The fact that the Japanese stem nominalization seen above displays such a diverse range of meaning, rather than more narrowly circumscribed patterns as in the case of *employment*, *employer* and e*mployee*, indicates that it was once a productive grammatical nominalization process (see Shibatani (2018a) for the synchronic data suggesting this and Section 5.1 below, where so-called clause-chains are discussed).

Despite these connections between lexical and grammatical nominalizations, there can be differences between the two. While the above exposition takes a formal difference as a criterion for distinguishing between lexical (those being *single words*, possibly with internal structural complexity as in the case of compound *words* and those containing vestiges of certain affixes) and grammatical nominalizations (those having phrasal structures larger in size than words), the distinction between these two types of nominalization can be quite subtle and difficult to pinpoint in many cases.[7] Since this paper is mainly concerned with grammatical nominalizations, we spend the rest of this subsection discussing the

7. Part of this problem comes from the difficulty of distinguishing between "words", especially compound words, and "phrases". Some compounds show clearer differences vis-à-vis phrasal counterparts as in the case of *blackboard* vs. *black board*, where the former has primary-tertiary stresss pattern (*blâckbòard*) and the latter secondary-primary stresss pattern (*blâck bóard*), the former has a semantically bleached adjective (*green blackboard*) and the latter an adjective with the literal meaning (?*green black board*), and while the former does not allow internal structures to intereact with external elements (**very blackboard*, meaning a blackboard which is very black), but the latter does (*very black board*). There are, however, many compounds that are not as straightforward as in the case of *blackboard* vs. *black board* (see Lees (1963, Appendix A) for a related discussion).

nature and issues pertaining to what may be called lexical nominalizations and the distinctions between them and grammatical nominalizations.

One difference observed across languages is a formal/morphological difference. Lexical nominalizations are often associated with specific morphology, as in *employ* > *employment*, but grammatical nominalizations may, in many languages, show no morphological marker at all, as in *I know* [*John employs Bill*]$_{\text{NMLZ}}$, where the nominalization denoting a fact has the same *internal structure* as the sentence *John employs Bill*. A formal characteristic of grammatical nominalizations like this has led many to view them as clauses or even sentences.[8] But, as noted above, even lexical nominalizations may not have any morphological indication, e.g. *cook* > *a cook, walk* > *a walk*, and thus the notion of nominalization is independent from morphological marking.

Another characteristic that distinguishes lexical nominalizations from grammatical nominalizations is that the former have irregular gaps in the meaning patterns. Returning to the Japanese stem nominalizations discussed above, many of them based on action verb roots allow both activity and agent/instrument readings (*suri* 'pickpocketing/pickpocket', *hito-gorosi* 'manslaughter/killer', *simi-nuki* 'stain removing/stain remover'), whereas many other similar forms have only one reading. Forms like *yama-nobori* 'mountain climbing', *uo-turi* 'fish catching', and *sumi-yaki* 'charcoal-making' only denote activities, whereas *uta-utai* (song-singing) 'singer', *e-kaki* (picture-drawing) 'painter', and *sumoo-tori* (sumo-taking) 'sumo-wrestler' name only agents and not activities such that while [*yama-nobori*]-*suru* 'do mountain-climbing' is possible, *[*uta-utai*]-*suru* 'do song-singing' is not.

Finally, grammatical nominalizations differ from lexical nominalizations in that their meanings tend to be compositional, while lexical nominalizations having a word status may undergo meaning specialization, as in the case of *e-kaki* (picture-drawing) 'a painter' denoting a professional artist, as opposed to the grammatical counterpart *asoko de e o kaku no* (there LOC picture ACC draw NPM) '(one) who draws a picture there', which is free from such a conventionalized restrictive meaning. The term "lexicalization" is sometimes used to refer to this kind of specialization in meaning. In the case of *e-kaki* 'a painter' above, this sense of lexicalization coincides with the status of the form as a noun. But since there are phrases whose meanings are specialized/noncompositional, namely idioms (e.g. *all hell breaks loose*, *kick the bucket*), meaning specialization itself does not define words. And there are many lexical nominalizations that do not have specialized

8. Bear in mind that structural resemblances do not guarantee that we are dealing with the same grammatical units. In particular, we reject a view that certain nominalizations are clauses/sentences because their internal structural properties are similar to those of clauses and sentences. (see Section 6.2 on the definitions of clauses, sentences, and nominalizations).

meanings. Indeed, there are cases where a nominalized form may convey both lexicalized meaning and literal, compositional meanings. Consider the following forms from Central Alaskan Yup'ik:

(2-4) Cental Alaskan Yup'ik (Tamura 2017 quoting Jacobson 1984: 450, 560)
- i. *pi-* 'do' > *pi-sta* 'doer, servant'
  pi-sta-i 'the one who did something to him', 'his servant'
- ii. *angu-* 'pull' > *angu-n* 'the one who is pulling', 'man'
- iii. *mikel-* 'small' > *mikel-nguk* 'the one who is small', 'child'
- iv. *kipus-* 'buy' > *kiups-vik* 'the place for buying', 'store'

Since meaning does not provide a definitive criterion for lexical categories (cf. *resemble* vs. *similar* in English, where the former is a verb and the latter an adjective), the morphosyntactic status of the form in question must be ascertained in order to determine its lexical (word-unit) status.[9] If the form has morphosyntactic properties of ordinary nouns, then it is a lexical nominalization. In the case of *singing* in English, for example, we must recognize two distinct forms. One of them occurs as a word by itself, as in *her beautiful* ***singing*** (*impressed us*). The other *singing* does not, e.g. **her beautifully* ***singing*** *(was quite a show).* This *singing* occurs only in a phrasal form, as in *her beautifully* ***singing the national anthem*** (*was quite a show*). Compare this with the former *singing*, which is modified by an adjective and which does not form a phrase with a noun phrase, as in *her beutiful* ***singing*** **(of) the national anthem*, where the preposition *of* is required. *Singing the national anthem* is not a word and therefore it is a grammatical nominalization, while *singing* as in *singing of the national anthem* is a word and displays many of the essential properties of ordinary nouns. However, the grammatical nominalization *singing the national anthem* also displays some major properties of ordinary nouns, such as being modified by a genitive determiner, as in ***her*** *singing the national anthem*, and it heads a subject and an object NP, as in [*singing the national anthem*] *would be quite appropriate*, yet it does not allow the marking by an article (**a/*the singing the national anthem*), unlike ordinary nouns or the lexical counterpart (*a/the singing of the national anthem*).

As the above discussion demonstrates, the question of whether a nominalization structure is a noun (lexical) or not boils down to the degree of similarity of the form to ordinary nouns. It was on observations like this that John Robert Ross (aka Haj Ross) proposed what he called the *nouniness squish*, capturing the cline of nouniness among various types of nominals:

9. There are also cases where English does not have corresponding nouns and descriptive grammars must resort to an analytic translation such as 'one who does X'.

> *that* clauses > *for to* clauses > embedded questions > Acc *ing* complements > poss *ing* complements > action nominals[10] > derived nominals > underived nominals
> (Ross 1973)

One may use the term "lexical nominalization" in reference to those forms that display the total or a majority of morphosyntactic properties of ordinary nouns and "grammatical nominalization" for those divergent from ordinary nouns yet showing varying degrees of nouniness in morphosyntax. These two types of nominalizations are best considered to form a continuum on formal grounds. What unifies them (and their subtypes) as nominalizations is their semantic function (denoting things and thing-like entities), with varying degrees of morphosyntactic repercussions of this function depending on their types.

In view of the infeasibility of applying nouniness tests to the data from a diverse array of languages below, the tactic adopted in this paper is to treat those forms discussed under "nominalization" in the literature as lexical nominalizations, while treating as grammatical nominalizations those that are discussed elsewhere in the grammar under such headings as "relative clauses" or "subordination", which tend to include verbal morphosyntactic properties (e.g. person and voice marking, the presence of an argument NP and an adverbial modifier) not associated with ordinary nouns (and "true" lexical nominalizations). It is interesting and relevant to examine if the two readings for each of the Yup'ik forms in (2-4) above correlate with differences in morphosyntactic properties.[11] Either way, we would not know whether a given form is a lexical or grammatical nominalization until we apply available morphosyntactic nouniness tests and determine the degree to which the given form resembles ordinary nouns in the language.[12]

---

10. The *singing* of *singing of the national anthem* appears to retain a certain degree of "verbiness" in that, while *Her* ***beautifully*** *singing of the national anthem impressed them* is totally ungrammatical, *Her singing of the national anthem* ***beautifully*** *impressed them* does not appear to be totally ungrammatical.

11. Fleck's contribution to this volume exactly addresses this important issue and shows that the formally identical forms display a different external property (whether or not the nominalized forms can be modified by a possessive form) depending on whether they are associated with a conventionalized meaning denoting a specific object (e.g. a stick or club) or with a schematic meaning denoting a variety of objects (e.g. any object that one may use to hit something/someone) (see also Tamura, this volume).

12. More can be said about this distinction from various theoretical perspectives such as the distinction between lexicon/morphology and syntax, as in the Generative Grammar framework. The framework that better accommodates the lexical-grammatical continuum is Langacker's Cognitive Grammar (Langacker 1987, 1991) that recognizes degrees of entrenchment and schematicity, both of which attempt to capture the degree of productivity of construction types and the associated difference in meaning (idiomaticity/composionality). We are unable to go into

2.3 Inputs to nominalization processes

Perhaps influenced by the term "nominalization", which suggests turning something in one state into something in another state, many researchers, as indicated by the quotes above, believe that nominalization is restricted to derivations that turn verbal inputs into nominal outputs, as in Payne's (1997) and Malchukov's (2004) definitions or to those that take members of non-nominal categories as their inputs, as in Crystal's (1980) definitions. While recognizing nominalizations based on nominal inputs, Comrie and Thompson (1997/2007) give short shrift to such cases by allocating only one page and a few additional lines in their 47-page discussions on lexical nominalizations. A similar bias toward verbal-based nominalization is also clearly seen in all the papers collected in Yap, Grunow-Hårsta & Wrona (2011).

It is unclear why these researchers have decided to focus more on verbal-based nominalizations, when even such a well-known language as English exhibits nominal-based nominalizations. The case in point involves the so-called agentive suffix *-er*, deriving verbal-based forms such as *play* > *play**er***, *sing* > *sing**er***, which, everyone would agree, is a clear case of lexical nominalization. But this process takes a wide range of nominal inputs, as demonstrated by *villag**er***, *New York**er***, *rear-end**er***, *right-wing**er***, *leftfield**er***, *knuckleball**er***, *tenn**er***, *49**ers***, *lif**er***, *spring break**er***, the aforementioned *half-pound**er***, *18-wheel**er***, etc. While many of these are not strictly agentive, they denote entities that are closely associated with the meaning of the base forms. Whether a derived form denotes an agent or non-agentive entity simply depends on the nature of the base form; verb-based nominalizations denote an entity most closely associated with activities, namely an agent (and possibly an instrument and others), whereas nominal-based ones denote other types of entities metonymically evoked in close association with the denotations of the base nouns, such as the people associated with specific locations one way or another, and those entities associated with a specific quantity, time, or manner.

The above pattern certainly is not limited to English. Parkatêjê, a Je language in northern Brazil, has the agentive suffix *-katê*, which nominalizes verb roots (e.g. *krere* 'sing' > *krere-**katê*** 'singer'; *jakre* 'write' > *jakre-**katê*** 'writer'). But this suffix productively applies to animal names as well, producing forms like *pryre* 'animal' > *pryre-**katê*** '(animal) hunter', *rop* 'jaguar' > *rop-**katê*** 'jaguar hunter', and *kukryt* 'tapir' > *kukryt-**katê*** 'tapir hunter' (author's field notes).

these alternative perspectives due to other empirical, and more pressing, issues that need to be addressed, but how the difference such as the degree of productivity and meaning compositionality correlates with the degree of formal nouniness discussed here would be an interesting and perhaps worthwhile topic to pursue.

Central Alaskan Yup'ik has several nominalizers that apply to verbs to derive nouns (lexical nominalizations) and gramatical nominalizations. But most, if not all, also take nouns as inputs.

(2-5) Locative/Time nominalizer *-vik* 'place toV; place or time of V-ing; **place for N**': (Tamura 2017 quoting Jacobson 1984: 586; Tamura this volume)
  a. Verb-based lexical nominalization
    i. *nere-* 'eat' > *nervik* 'dining hall, restaurant, table'
    ii. *qavar-* 'sleep' > *qavarvik* 'sleeping bag, bedroom'
  b. Noun-based lexical nominalization
    qimugta 'dog' > *qimugtevik* 'kennel, dog house'

(2-6) Instrumental nominalizer *-cuun* 'device for V-ing, **device associated with N**' (Tamura 2017 quoting Jacobson 1984: 450)
  a. Verb-based lexical nominalization
    i. *mingqe-* 'saw' > *mingqesuun* 'sawing machine'
    ii. *igar-* 'write' > *igarcuun* 'pencil, pen'
    iii. *nere-* 'eat' > *nerrsuun* 'fork, eating utensil'
  b. Noun-based lexical nominalization
    i. *anuqa* 'storm' > *anuqessuun* 'wind generator, storm lantern'
    ii. *arnassagaq* 'old woman' > *arnassagarcuun* 'old-age pension for a woman'

Yagua in northwestern Amazonia has *nominalizing* classifiers that apply not only to verbal roots but also to adjectival as well as nominal roots (e.g. *tiryǫ́ǫ́-**jay*** (sleep-CLF.PELT) 'sleeping mat', *jąąmu-**daisiy*** (big-CLF.THIN.POLE) 'big blowgun, pole', *nǫǫnoo-**jąą*** (light-CLF.LIQUID) 'kerosene') (Payne 1985). The Salish language Halkomelem has similar nominalizing classifiers that also apply to verbal, adjectival, and nominal roots (e.g. *ʔitət=**ə'wtx**^{w}* (sleep=CLF.HOUSE) 'hotel, bedroom', *q̓aq'iy =**e'wtx**^{w}* (sick=CLF.HOUSE), 'hospital', *tel=**e'wtx**^{w}* (money=CLF.HOUSE) 'bank') (Gerdts and Hinkson 2004).

Languages outside of the Americas also allow nominalization to apply to nouns. Gã, a Kwa language spoken in Ghana, has the "agentive" suffix *-lɔ*, which nominalizes verbs, as in *jù* 'steal' > *jù-**lɔ́*** 'thief' and *tsɔ̃́ɔ̃́* 'show, teach' > *tsɔ̃́ɔ̃́-**lɔ́*** 'teacher'. Just like the English and the Parkatêjê counterpart, this suffix also applies to nouns, as in *àníháó* 'laziness' > *àníháó-**lɔ́*** 'lazy person', *bé!í* 'quarrel(n.) > *béì-**lɔ́*** 'quarrelsome person', and *àmálé* 'lie (n.)' > *àmálé-**lɔ́*** 'liar'. Gã also has the "agentive" nominalizer ***-tsɛ̀***, which applies to both adjectives and nouns; *àgbò* 'big' > *àgbò-**tsɛ̀***

'big one', *bíbìóó* 'small' > *bíbìóó-**tsɛ̀*** 'small one'; *àfú* 'hump' > *àfǘ-**tsɛ̀*** 'hunchback', *shìká* 'money' > *shìká-**tsɛ̀*** 'rich person' (Campbell 2017).

Chinese has a number of agentive suffixes that derive nouns from verbs, which are similar in function to the English *-er* suffix seen above. They are also similar in that they apply to noun inputs as well, as shown in the Mandarin forms below:

(2-7)		V > N	N > N
	-zhě (者 'person');	jì-zhě to record-SUF' 'reporter'	yè-zhě writing.board-SUF 'dealer'
		zuò-zhě to make-SUF 'author'	dìguó zhǔyì-zhě imperialism-SUF 'imperialist'
		zhǎng-zhě to be elderly-SUF 'elderly man'	bǐ-zhě pen-SUF 'author'
	-shǒu (手 'hand');	zhù-shǒu to help-SUF 'assistant'	pào-shǒu canon-SUF 'gunner'
		hǎo-shǒu to be good-SUF 'skilled person'	qí-shǒu flag-SUF 'standard bearer'
		xiōng-shǒu to be ferocious-SUF 'murderer'	gē-shǒu song-SUF 'singer'
	-jiā (家 'house');	zuò-jiā 'writer' to make-SUF 'writer'	yìshù-jiā art-SUF 'artist'
		huà-jiā to paint-SUF 'painter'	yùndòng-jiā movement-SUF 'athlete'
		pīpíng-jiā to criticize-SUF 'critic'	kēxué-jiā science-SUF 'scientist'

In Section 7 we offer a novel analysis of the genitive/possessive form (e.g. *my, mine*) as an instance of nominal-based nominalizations. As a way of summarizing the discusions above, observe the following examples from the Athabaskan language

Slave and the Austronesian language Tagalog that shed light on two important issues that are pursued in this paper. Namely, (i) that nominalization applies to nominals as well (the point proven above), and (ii) that lexical nominalizations may share morphology with grammatical nominalizations indicating that the two constitute a unified phenomenon. The Slave nominalizing suffix *-i* and its phonological variants productively derive verbal-based grammatical nominalizations that denote an event protagonist (e.g. (2-8c)). It also derives lexical nominalizations (e.g. (2-8a)), and it applies to numerals as well, which are assumed to be nouns in this language, as in most other languages (e.g. (2-8b)).

(2-8) Salve (Rice 1989)[13]

a. Verbal-based lexical nominalization
*ʔehdzo.**i***
trap.NMLZR
'a trap'
cf. *ʔehdzo* 'S/he traps something.'

b. Nominal-based nominalization
i. *dı̨-**i*** *whehk'é* (Bearlake dialect) (258)
four-NMLZR 3.shot
'S/he shot four {ANIMALs}.'
ii. *lake'e-**e*** *ragokedéhwe* (Hare dialect) (258)
five-NMLZR 3.start back
'Five {PEOPLE} started back.'

c. Verbal-based grammatical nominalization
[*nı̨wą́ kedaw'í*] ***i*** *ke gogháyeyida* (Hare dialect) (258)
long 3PL.sat NMLZR PL 1SG.saw.3PL
'I met ones who stayed a long time.'

In a similar vein, the Tagalog locative nominalizer -(*h*)*an*, which productively derives verbal-based locative grammatical nominalizations, also attaches to nouns and yields new nouns (lexical nominalizations) denoting locations that the referents of the base nouns are conventionally associated with (Schachter & Otanes 1972/1983: 98ff).

13. The glosses in the examples from other sources are mostly original, except for what I consider to be nominalizing morphemes, which in the originals are glossed variously as NOM, N, REL, COMP, SBR (subordinator) etc. I take the liberty of glossing them uniformly as NMLZR. The nominalization structures, on the other hand, are marked as $[\ldots]_{\text{NMLZ}}$. It is amusing indeed to notice that Comrie (2006) relabels as REL(!) the Tibetan nominalizers so recognized and glossed as NOM by the Tibetan specialist Scott DeLancey (DeLancey 2002).

(2-9) Tagalog nominal-based lexical nominalization

i.	*aklat*	'book'	>	*aklat-**an***	'library'
ii.	*halaman*	'plant'	>	*halaman-**an***	'garden'
iii.	*tarangka*	'lock'	>	*tarangka-**han***	'gate'

(2-10) Tagalog verbal-based grammatical nominalization

Ang [*b < in > ili-**han** ng lalaki ng saging*] *ang tindahan ni Dodong.*
TOP buy<PFV>-LF GEN man GEN banana TOP store GEN Dodong
'Dodong's store is *where the man bought the bananas*.'

Having demonstrated that nominalization applies to nominals and having taken a quick view on the possibility that nominalization morphology produces units larger than words, we shall now more closely look at morphological connections between lexical and grammatical nominalizations as a way to motivate the recognition of grammatical nominalizations as such.

2.4 Outputs of nominalization processes

While Payne's (1997) and Comrie and Thompson's (1985/2007) definition of nominalization restricts the outputs to nouns, and the translations of the term "nominalization" in some grammatical traditions may literally mean noun-formation, as the term *meishi-ka* in Japanese linguistics does, the process actually creates units larger in size than words, as we have already seen above. The Slave form [[nįwą́ kedaw'í] i] 'one who stayed long' and the Tagalog form *ang* [b < in>ili-han ng=lalaki ng=saging] 'where the man bought bananas' are cases in point. Our arguments for treating these as (grammatical) nominalizations, rather than, say, verbal phrases, rest on two crucial facts pertaining to these structures. The foremost is the fact that these phrasal strctures denote substantive concepts just like ordinary nouns and lexical nominalizations. We shall dwell on this fundamental property of grammatical nominalizations throughout this paper, but especially in Section 6.2, where we discuss differences among nominalizations, clauses, and sentences.

The other fact that motivates analyzing the structures like the Salve and Tagalog grammatical nominalizations above as such is that they share morphological marking with uncontroversial lexical nominalizations. This kind of morphological connections between lexical and grammatical nominalizations are not limited to Slave and Tagalog, as we saw earlier in terms of the forms *her **singing** of the national anthem* and *her **singing the national anthem,*** although many other languages show different marking patterns for the two types of nominalization. While nominalization is not definable in terms of morphology, morphological connections between lexical nominalizations and grammatical nominalizations provide concrete and positive evidence that the latter represent a nominalization

phenomenon. Below we offer a sampling of languages from a diverse array of language families in the Americas and a few additional examples from outside this region to show that our treatment of grammatical nominalizations as such is also morphologically supported.

Let us start with languages in South America beginning with the Tapiete examples below.

(2-11) Tapiete
(Tupí-Guaraní; Bolivia, Paraguay, Argentina; González 2005; Ciccone 2008)
a. Lexical nominalizations
i. *hë'ë* 'be sweet' > *hé'ë-**wä*** 'sugar, honey'
ii. *arɨka'e* 'long ago' > *arɨká'e-**wa*** 'ancestors, history'
b. Grammatical nominalizations
i. [*mbiri-iwi ou-**wa***] *kö'ëin-wai ou*
far-DIR 3.come-NMLZR in the morning come
'(The one) who comes from far away has arrived this morning.'
ii. [*kä'ä tenta-pe hau-**wa***] *hayasi*
yesterday town-LOC 1.eat-NMLZR be.rotten
'What I ate yesterday in town was rotten.'

The first grammatical nominalization above is a subject argument nominalization that denotes an agentive entity (similar to the English agentive lexical nominalization *employer*) metonymically evoked by the nominalization structure marked by the nominalizer *-wa*, which also derives lexical nominalizations. The second form, on the other hand, is an object grammatical nominalization denoting a patientive entity (similar to the English patientive lexical nominalization *employee*).

Kakataibo has several different nominalizers for lexical and grammatical nominalization but the two processes do share the nominalizer *-kë*, similarly to the Tapiete nominalizer illustrated above. The first grammatical nominalization is a patient argument nominalization like the patient lexical nominalization *tua-**kë*** 'son of a woman', while the second is an event nominalization that metonymically denotes a fact pertaining to the event.

(2-12) Kakataibo (Panoan; Peru; Zariquiey 2011: 297, 632, 638)
a. Lexical nominalization
i. *mapun-* 'to cover' > *mapun-**kë*** 'house'
ii. *tua-* 'to give birth' > *tua-**kë*** 'son of a woman'
b. Grammatical nominalization
i. *ashi ka ‘ën ñuikaskë ‘iashín*
a=ishi ka [‘ë=n ñui-kas-**kë**] ‘i-a-x-ín
that=only NAR.3p 1sg=A tell-DES-NMLZR be-PERF-3p-prox
'Only that was what I wanted to tell.'

ii. *Juan hotelnu tëëkë (a-x) ka asabi ‘ikën*
[Juan hotel=nu tëë-**kë**] (a-x) ka asabi ‘ikën
Juan.ABS hotel=LOC work-NMLZR 3sg=S NAR.3p good be.3p
‘(The fact) that Juan works in the hotel is good.’

Rojas-Berscia’s contribution to this volume also shows that the Kawapanan language Shawi derives both lexical and grammatical nominalizations by the use of the nominalizer *-napi*, as seen below.

(2-13) Shawi (Kawapanan)
a. Lexical nominalization
i. *shawi-**napi*** ‘the gossiper’
ii. *ni’i-**napi*** ‘the watcher’
iii. *u’u-**napi*** ‘drinker’
b. Grammatical nominalization
*atari nusha a-ka’n-**napi***
chicken meat CAUS-eat-NMLZR
‘the one who makes someone eat chicken’

The next examples come from Bolivian Quechua.

(2-14) Bolivian Quechua (Bills et al. 1971: 101, 106)
a. Lexical nominalization
i. *llank’a-**q*** ‘worker’
ii. *yanapa-**q*** ‘helper’
iii. *ranti-**q**-kuna* ‘buyers’
iv. *wayk’u-**q*** ‘cook’
b. Grammatical nominalization
i. *haway runa* [*qulqí muna-**q***] (based on Bill’s et al. 1971: 274)
that man money want-NMLZR
‘That man is the one who wants the money.’
ii. *Munaku-ni* [*runa maqa-**q***]
love-1SG man hit-NMLZR
‘I love the one who hit the man.’

Turning to North America, we again observe widespread morphological connections between lexical and grammatical nominalizations demanding recognition of the latter as a type of nominalization.

(2-15) Northern Paiute (Numic, Uto-Aztecan; Thornes 2003: 118, 431)
a. Lexical nominalization
i. *tiničui-**dɨ***
teach-NMLZR
‘teacher’

ii. *hoawai-**dɨ***
hunt-NMLZR
'hunter'
iii. *nayakwi-**dɨ***
(play)handgame-NMLZR
'gambler'

b. Grammatical nominalization
i. *oʔo iwa-ʔyu su=uuni-ʔyu naʔa*
DEM many-NOM NOM=that.kind-NOM grow
su=[*na-tihona-**dɨ***]
NOM=MM-dig.roots-NMLZR
'There's a lot of that kind growing out there for the digging.' (i.e. that which is dug).
ii. *ta=sakwa umɨ-no* [*na-kwii-**dɨ***]
1DL.INCL=MOD 3PL-COM MM-smoke-NMLZR
'We should (go) with them to the smoke-pit.' (i.e. that which smokes)

Shoshone, which also belongs to the Numic branch of Uto-Aztecan family, like Northern Paiute, has several types of nominalizers used for lexical and grammatical nominalizations (cf. Dayley 1989).

Kiowa, forming the Kiowa-Tanoan family with the Tanoan pueblos of New Mexico and Arizona, has a nominalizer that derives nouns from indefinite and interrogative roots as well as from some locative roots and phrases, according to Watkins (1984). This suffix, as seen below, appears to have a much more productive use as a nominalizer deriving grammatical nominalizations, which are used as modifiers of nouns-in so-called relative clause constructions-as in other languages.

(2-16) Kiowa (Kiowa-Tanoan; Watkins 1984: 108, 109, 230)

a. Lexical nominalization
i. *hɔ́n-**dé*** 'someone', 'something' 'some kind of' (< *hɔ́n-* indefinite root)
ii. *hɔ̂n-**dé*** 'who?, what? 'what kind of?' (< *hɔ̂n* interrogative root)
iii. *tʰáy-**dè*** 'lid, cover' (< *tʰáy* 'atop')
iv. *tô-y-**dè*** 'household goods, furniture' (< *tô-y* 'in the house')
v. *gû-y-**tè*** 'other(s), different one(s), outsider(s)' (*gû-y* 'outside')

b. Grammatical nominalization
i. *óy-gɔ̀ k'yą́ˈhį̂ˈ* [*ø-déˈ-**dè***] *k'yátaʾy-k'ìˈ ø-dɔ́ˈ*
there-ADV man 3SG-stand-NMLZR chief-male 3SG-be
'The man (who is) standing there is a chief.'

ii. *k'íˑ* [*k'ɔ́dáˑl-ɔ̂ˑ ø-òl + sɔ́l-**dè***]
wood wagon-on 3SG-load+be-MMLZR
gyà-p'éttɔ̀
1SG/AGT:SGOBJ-take=down-IMPF
'I am unloading wood that was loaded in the wagon.'

The Siouan language Crow spoken in Montana has, among others, the agentive (*ak-*) and the locative/temporal/manner (*ala-*) nominalizers that derive lexical nominalizations. These are also used in grammatical nominalizations in a parallel manner, as below.

(2-17) Crow (Siouan; Graczyk 2007: 254, 255)
a. Lexical nominalization
i. ***ak**-disshí* 'dancer' < *disshí* 'dance'
ii. ***ak**-kummí* 'singer' < *kummí* 'sing'
iii. ***ala**-chiwakáa-u* 'church' < *chiwakíi* + PL ('where they pray')
iv. ***ala**-sáhta* 'fork in a river' < *sáhta* 'forked, pronged'
b. Grammatical nominalization
i. *hileen* [***ak**-isshí-ss-aa-lee-sh*] *awe-taa xémm-ak*
these NMLZR-top-GOAL-PORT-go-DET ground-PATH lie-SS
dupesaa-(a)-ahk-uu-k
pant-CONT-remain-PL-DECL
'The ones who had brought him to the top were lying on the ground panting.'
ii. *púaee* [*bale **ala**-satché*] *ko kukaá húu-ssaa-k*
smoke wood NMLZR-thick PRO SOURCE come-NEG-DECL
'The smoke isn't coming where the trees are thick [the forest].'

A close parallel to the Crow pattern is seen in the Yuman language Jamul Tiipay in southern California, in which lexical nominalizations and grammatical nominalizations share nominalization elements (prefixes, suffixes, length ablaut), as observed below.

(2-18) Jamul Tiipay (Yuman; Miller 2001: 117, 118, 122, 214)
a. Lexical nominalization
i. ***kwe-ch**-cheyaaw* 'singer' < *cheyaw* 'to sing' (Subject nominalization)
ii. ***ke**-meyaally* 'tortilla maker' < *meyally* 'to make tortillas' (Subject nominalization)
iii. *ya**'**-ma-**ch*** 'powwow' < *iima* 'to dance' (Oblique nominalization)
iv. ***a**'-sii-**ch*** 'liquor' < *si* 'to drink' (Oblique nominalization)

b. Grammatical nominalization
   i. [*met'aar-i* ***ke***-*pa*-***ch***]-*pu* *yaach Uuyaaw*
   outdoors-LOC NMLZR-be.PRS.SRS-NMLZR-DEM 1.SJ know
   (Subject nominalization)
   'I know the one who is out there.'
   ii. [*nya'ru me'a-y* *stu*-***ch***]-*pu* *nyaach uuyaaw*
   money where-LOC pick.up-NMLZR I.SUB know
   (Oblique nominalization)
   'I know where you got that money=I know the place where you got the money.'

Salish languages of the Pacific north coast all appear to contain the morpheme *s-* that produces both lexical and grammatical nominalizations, as in Musqueam below.

(2-19) Musqueam (Halkomelem, Salish; Suttles 2004: 77, 101 264, 265)
a. Lexical nominalization
   i. ***s***-*ʔə́ɬtən* 'food' < *ʔə́ɬtən* 'eat'
   ii. ***s***-*ʔítᶿəm* 'clothing' < *ʔítᶿəm* 'get dressed'
   iii. ***s***-*píw*' 'ice' < *píwət* 'freeze it'
   iv. ***s***-*kʷix* 'name' < *kʷíxət* 'name it'
   v. ***s***-*yá'ys* 'work' < *yá'ys* 'work'
b. Grammatical nominalizations
   i. *kʷθə* [*niʔ* *nə*-***s***-*ʔəx̌i*-*t*]
   ART AUX my-NMLZR-borrow-TR
   'what I rent him'
   ii. *ʔə́y* [*k'wə* [***s***-*q'en*-*əθət*-*ct*]]
   good ART NMLZR-return-self-our
   'We'd better return.' (lit. Our returning would be good.)

Finally, Eskimo languages display several nominalizers that yield different types of nominalizations such as agentive, instrumental, and locative nominalizations. The following illustrates the use of the agentive nominalizer -(*s*)*ta* in Central Alaskan Yup'ik marking both lexical and grammatical nominalizations.

(2-20) Central Alaskan Yup'ik (Miyaoka 2012: 532, 533, 535)
a. Lexical nominalization
   i. *cali*-***sta*** 'worker'
   ii. *kuvya*-***sta*** 'one who fishes by net'
   iii. *ikayur*-***ta*** 'helper'
b. Grammatical nominalization
   i. [*neqe-m* *nere*-***sti***-*i*]
   fish-REL.SG eat-NMLZR-ABS
   'one that eats the fish'
   cf. *qimugta ner'-uq neq-mek*
   'the dog is eating fish'

ii. carayi-i-m *tangvag-**ti**-i*
ghost-EV-REL.SG see-NMLZR-ABS.3SG.SG
'the one who saw the ghost'

Just to complete the picture, let us look at a few languages outside the Americas. Thai (Tai-Kadai) has the nominalizer *thîi*, which marks lexical nominalizations, many of which form noun compounds, as well as grammatical nominalizations of both those that denote event protagonists and those that denote state of affairs or facts.

(2-21) Thai (Tai-Kadai)
a. Lexical nominalization
i. ***thîi**-nâŋ*
NMLZR-sit
'seat'
ii. ***thîi**-tàt lép*
NMLZR-cut nail (Iwasaki & Ingkaphirom 2005: 45)
'nail clipper'
iii. ***thîi**-lăaw dinsɔ̌ɔ*
NMLZR-sharpen pencil
'pencil sharpener'
iv. ***thîi**-pɔ̀ɔt krapǒŋ*
NMLZR-open can
'can opener'
b. Grammatical nominalization
i. *chăn càʔ sày* [***thîi** khwěɛn nay tûu*]
I will wear NMLZR hang in closet
(courtesy of Kingkarn Thepkanjana)
'I will wear the one that hangs in the closet.'
ii. *tɛ̀-wâa dii ná* [***thîi** mây mii khay pen alay*]$_{\text{NMLZ}}$
but good PP NMLZR NEG have who CPP what
(Iwasaki & Ingkaphirom 2005: 255)
'But it was good that no one was hurt.'

Korean has a fair number of nouns derived from verbs by *-um* suffixation. This suffix is also involved in grammatical nominalizations, which are somewhat archaic sounding, according to the Korean speakers consulted.

(2-22) Korean (Courtesy of Sung-Yeo Chung)

a. Lexical nominalization

i. *cwuk-* 'die' > *cwuk-**um*** 'death', *cwuk-em* 'corpse'

ii. *mwut-* 'bury' > *mwut-**um*** 'burial', *mwut-em* 'grave'

iii. *kuli-* 'draw' > *kuli-**m*** 'drawing, picture'

b. Grammatical nominalization

[*mayil swul-ul masi-**m***]*-un* [*kenkang-ul ilh-**um***]*-ul*
every.day alcohol-ACC drink-NMLZR-TOP health-ACC lose-NMLZR-ACC
[*uymiha-m*]*-i-yo*.
mean-NMLZR-COP-ASSERTIVE
'To drink alcohol every day means to lose one's health.'

As final examples for morphological connections between lexical and grammatical nominalizations, let us examine the following data from the Austronesian language Malagasy, where, besides so-called focus morphology, which is a nominalizing morphology itself (see Section 3.2), there are several nominalizing prefixes that combine with different focus morphology. For example, in (2-23a) below, the nominalizer is a circumfix combining the *f-* prefix and the circumstantial focus suffix *-ina* (and its variants). The same circumfix is used in forming event grammatical nominalizations, as in (2-23b).

(2-23) Malagasy (Austronesian)[14]

a. Lexical nominalizations

i. ***f**-i-anar-**ana***
NMLZR-AF.MM-advice-CF
'school'

ii. ***f**-am-ono-**na** olona*
NMLZR-AF-kill-CF human
'human killing, murder'

b. Grammatical nominalizations

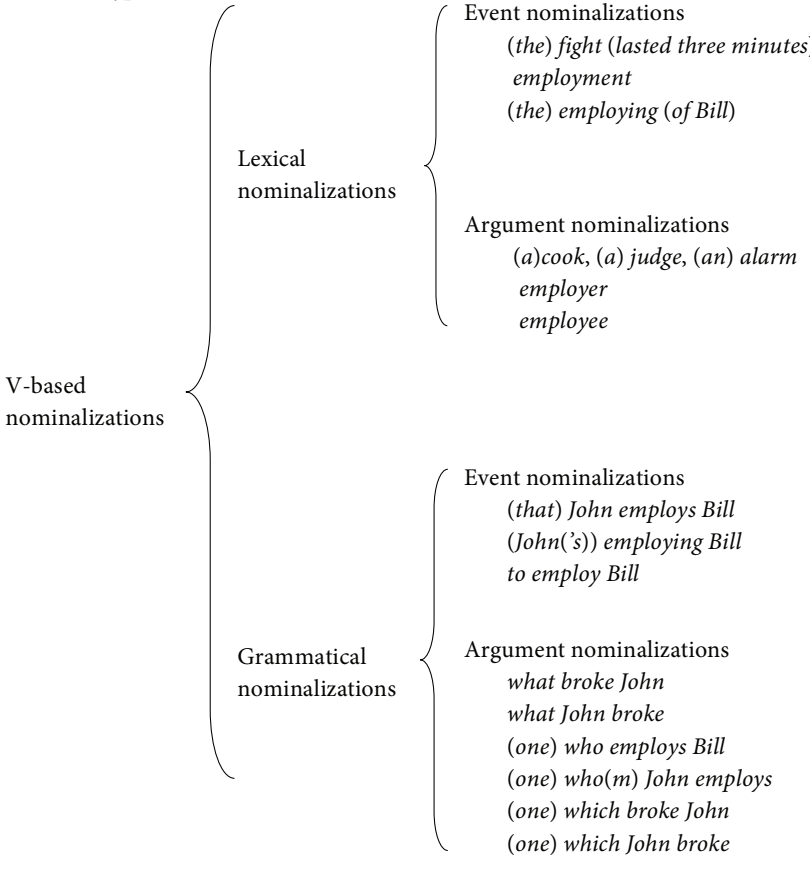
i. *mahafinaritra ny fandehadehanana miaraka amin'ny*
mahafinaritra ny [**f**-an-dehadeha-**nana** miaraka amin-ny
fun INDEF NMLZR-AF-walk(rdpl)-CF outside with-IND
ankizy
ankizy]
child
'Walking outside with children is fun.'

ii. *hitako ny fikapohana ilay alike*
hita-ko ny [**f**-i-kapoh-**ana** ilay alike]
see.PF-1P.SG INDEF NMLZR-AF.MM-hit-CF DEF dog
'The hitting of the dog was seen by me.'

14. The examples without mention of the sources are from the author's own research.

As seen above, languages across the globe mark both lexical and grammatical nominalization similarly highlighting the underlying unity of the two types of nominalization. Many other languages, however, opt for marking lexical and grammatical nominalizations differently. This option is also motivated in that such a marking pattern reflects the differences in the types of nominalization, namely lexical nominalization vis-à-vis grammatical nominalization. We shall see in Section 8 how languages respond to the two opposing functional demands, one a desire to mark an underlying unity in the face of functional diversities, and the other a drive to mark functional differences of the forms at the expense of their underlying unity. Table 1 below summarizes the types of verbal-based nominalizations illustrated by some representative English forms, where the function-based classification cuts across the distinctions between lexical and grammatical nominalizations as well as formal morphosyntactic differences.

Table 1. Types of verbal-based nominalizations



Our focus in this paper is on grammatical nominalizations, the treatment of which has far-reaching theoretical implications calling into question many descriptive practices and theoretical claims made in the field.

3. Verbal-based grammatical nominalization

Verbal-based grammatical nominalizations, like the ones given above, involve a verbal head possibly with nominal arguments and modifiers.[15] We distinguish between event nominalizations and argument nominalizations. The former denote events and such abstract concepts as a state of affairs, a fact, or a proposition associated with the denoted events. They also denote event protagonists as well as resultant products, similar to resultative lexical nominalizations (e.g. *a building*, *a painting*). The argument nominalizations, on the other hand, denote in a clearer manner concrete things and thing-like entities, such as an agentive event protagonist, a patientive protagonist, a beneficiary, an instrument, or a location in close association with the concepts denoted by the verbal stems. That these two types of construction represent a unified phenomenon, namely grammatical nominalization, is indicated by a shared morphology in many languages such as the sharing of the particle *no* in the following Japanese examples, where the particle, glossed NPM (NP-use marker), marks a use of grammatical nominalizations as the head of a noun phrase – the *NP-use* of grammatical nominalizations (see Section 5).

(3-1) Event nominalization
Masako wa [*otto ni sonna onna ga ita*] ***no*** *o sitta.*
Masako TOP husband LOC such woman NOM existed NPM ACC learned
'Masako learned that (her) husband had such a woman.'

(3-2) Argument nominalization
i. [*Asoko de Ø koi o tutte iru*] ***no*** *wa boku no otooto da.*
there LOC carp ACC fish be.PRS NPM TOP GEN y.brother COP
'(The one) who is fishing a carp there is my younger brother.'
ii. [*Asoko de otooto ga Ø tutte iru*] ***no*** *wa koi da.*
there LOC y.brother NOM fish be.PRS NPM TOP carp COP
'What my younger brother is fishing there is a carp.'

The entity-denoting function of grammatical nominalizations allow them to head an NP, the most telling syntactic property of nominals. In addition, they may function as a modifier in an NP, or they may play an adverbial function, where they denote such notions as simultaneously- or sequentially-occurring events, purpose,

15. These cover verb-based and adjective-based grammatical nominalizations in those languages in which adjectives also inflect as in Japanese.

reason, time, and location metonymically associated with events. These are all *uses* of grammatical nominalizations, not what grammatical nominalizations are per se, as we shall see below.

3.1 Event nominalizations

Event nominalizations have clause-like internal structures, often with a full array of NP arguments overtly expressed, as in (3-1) above. They have, however, external syntagmatic properties like nouns in that they head an NP, playing both syntactic and referential functions of arguments of a clause. Bear in mind that grammatical categories are determined on the basis of external properties, not by internal properties, meaning that even if a structure is clause-like internally, it does not follow that the structure in question is a clause (see Section 6). Event nominalizations denote the following kinds of concepts:

(3-3) a. Event/Activity
River Thompson Salish (Salish; Kroeber 1977)
cuk^{w} [*e s-pek*^{w}*-e-s*]$_{\text{NMLZ}}$
finish ART NMLZR-split-TRZ-3.TS
'S/he finished splitting them.'
English
[*John's falling off of the bed*]$_{\text{NMLZ}}$ *happened at 3: 00 AM.*
Cf. *The* ***event*** *of* [*John's falling off of the bed*]$_{\text{NMLZ}}$ *happened at 3: 00 AM.*

b. Fact
Yaqui (Uto-Aztecan; Guerrero 2012)
Nim achai [*jaibu enchi siika-m*]*-ta te'a-k.*
1SG.GEN father already 2SG.ACC go.SG.PFV-NMLZ-ACC find-PFV
'My father discovered that you already left.'
Quechua (Cuzco dialect; Lefebvre & Muysken 1988)
[*Xwan papa-ta mikh-sqa-n*]*-ta yacha-ni.*
Juan potato-ACC eat-NMLZ-3-ACC know-1SG
'I know that Juan eats potatoes.'

c. Proposition
Northern Paiute (Western Numic; Uto-Aztecan; Thornes 2012)
ni [*ɨ = čadua-na*] *naka-supidakwatu.*
1 2 = talk-NMLZ hear-understand
(lit.) 'I understood your talking.'
'I understood what you are saying.'
(Glosses modified and the literal translation supplied)
English
[*That John is honest*] *is absurd.*
Cf. *The* ***proposition*** [*that John is honest*]$_{\text{NMLZ}}$ *is absurd.*

d. Event protagonist/participant
Navajo (Athabaskan; Hale & Platero 1973)
[*Ashkii at'ééd yiyiiłtsą́-(n)ę́ę́*] *yidloh*.
boy girl saw-NMLZR laugh
i. 'The girl that the boy saw is laughing.'
ii. 'The boy that saw the girl is laughing.'
Cf. [*Ø ashkii yiyiiłtsą́-(n)ę́ę́*] ***at'ééd*** *yidloh*
boy saw-NMLZR girl laugh
'The girl that saw the boy is laughing.'

e. Resultant product ("resultative nominalizations")
Waiwai (Cariban; Derbyshire 1999: 57)
[*a-mok-rɨ*] *w-enta*
2-come-AC.NMLZR I-hear+IMM.PST
'I heard you/your coming.'
Bolivian Quechua
[*Maria laranjas-ta ch'irwa-sqa-n*]*-ta ujya-ni*
oranges-ACC squeeze-P.NMLZ-3SG-ACC drink-1SG
(lit.) 'I drink that Maria squeezing oranges.' 'I drank {juice resulting from} Maria's squeezing oranges.'

f. Location
Gavião of Rondônia (Mondé; Tupi; Moore 2012)
[*me-tá mát*] *ká téét méèy-ka paágáá*
2P-live NMLZ.CONCRETE in exact 2P-(AUX.IMPERAT.DEF)-go (3S)-open
kára-ále-á
yet-FUTURE-END
'Go open (it) where you live.'
Mosetén (Sakel 2004: 94, 95)
*chhiko'-ñi-ti-**dye'***
liquid-put-VD-NMLZR
'place where one washes oneself'
cf. *saeks-e-**dye'***
eat-VI-NMLZR
'food'
*fer-**dye'***
strong-NMLZR
'strength'

g. Time
Mosetén (Sakel 2004: 95)
*ïjts-ï-**dye**-dyedye-'-ra, añe-**dye'***
mature-VI-NMLZR-INC-F.S-IR rain-NMLZR
'at the beginning of the time of the ripening, the rainy season (the time of rain)'

h. Reason
Musqueam (Suttles 2004: 105, 267)
ʌ̓'íʔ [*kʷ'ə* [*nə-šxʷ-ʔəm'í*]]
important ART my-OBL.NMLZR-come
'I've come for an important reason.' (lit. 'The reason why I have come is important.')
Cf. *ʔéx̌əθ* 'lie down' > ***šxʷ**-ʔéx̌əθ* 'bed'
yá·əs 'be working' > ***šxʷ**-yá·əs* 'tools'

i. Manner
Mẽbengokre. (Je; Salanova 2011)
[*a-dju-jarẽnh*] *mex*
2-ANTIPASS-say.NMLZR good
'You spoke well.' (lit. 'Your saying was good.')

3.2 Argument nominalizations

When event nominalizations of the above type evoke event protagonists as denoted entities, potential ambiguity arises, as indicated by the Navajo example in (3-3d) above and the Slave example in (5-69b) below. Languages appear to cope with this problem in several ways.[16] The Siouan language Crow opts for marking the relevant argument by the indefinite specific determiner *-m* (Graczyk 2007: 262ff). In (3-4) below, the only possible reading is that it is "this one" who is being concerned as having children, not "his sister-in-law".

(3-4) [*hinne hawáta-**m** úake dútt-ak áxpee-sh*] *kalakoon*
this one-DET his.sister-in-law takes-SS marry-DET then
dáak-uu-wish-dak (Graczyk 2007: 266)
child-PL-exist-COND
'if this one, who took his sister-in-law and married her, then has children'.

By far the more widespread method of unambiguously indicating the argument to be evoked is leaving the relevant argument position lexically unfilled, as in (3-2) above. The gap then indicates the grammatical role that the denotation of argument nominalizations stand for, such that a nominalization with a gap in subject position denotes an entity playing the subject role. These, paralleling the Japanese examples in (3-2), are illustrated by Korean subject argument nominalization (3-5a)

16. Word order within a nominalization structure seems to be relevant in Diegueño (Gorbet 1974).

and object argument nominalization below.[17] They are grammatical counterparts of lexical argument nominalizations of the type, *employer* (agent nominalization) and *employee* (patient nominalization). The gap in these grammatical argument nominalizations functions as a variable that can refer to anyone matching the denotation, namely any individual fishing carp there for (3-5a) and anything that the speaker's younger brother is fishing for (3-5b). The Korean pattern is paralleled by a large number of languages, in which the only clue for the type of argument nominalized is the position of a gap (or a missing argument) in the nominalization structure, as also seen in the Chinese examples below.

(3-5) Korean

a. [*ceki-eyse* ***Ø*** *inge-lul nakk-ko iss-nun*] *key nay tongsayng-i-ya.*
there-at carp-ACC fish-GER be-NMLZR NPM my y.brother-COP-IND
'(The one) who is fishing carp there is my younger brother.'

b. [*ceki-eyse nay tongsayng-i* ***Ø*** *nakk-ko-iss-nun*] *key inge-i-ya.*
there-at my y.brother-NOM fish-GER-be-NMLZR NPM carp-COP-IND
'What my younger brother is fishing there is carp.'

(3-6) Mandarin Chinese

a. [***Ø*** *zài nàr diào lǐyú*]*=de shì Xiǎo Wáng.*
PROG there fish carp=NMLZR COP Little Wang
'(The one) who is fishing carp there is Little Wang.'

b. [*Xiǎo Wáng zài nàr diào* ***Ø***] *=de shì lǐyú.*
Little Wang PROG there fish =NMLZR COP carp
'What Little Wang is fishing there is carp.'

Other languages depart from these patterns of argument nominalizations in two ways. One is the pattern found in a fair number of languages in the mainland Southeast Asia, Semitic languages, Iranian, and some others, where the argument nominalized is marked by a pronoun rather than by a gap. These languages actually combine the gap strategy and the pronoun strategy, where subject nominalization is marked by a gap, object nominalization by either a gap or a pronoun, and oblique nominalizations by a pronoun.[18] Observe the following data from Thai,

17. Depending on the word order and the absence of case markers, argument nominalizations marked by a gap may still be ambiguous, as in Iskonawa discussed by Zariquiey (this volume). Zariquiey, however, points out an important difference between event nominalizations, which also evoke event protagonists, and argument nominalizations in that the former evoke only most promiment arguments (subject and object), while the latter may denote entities holding peripheral roles.

18. Kakataibo shows a pattern like this according to Roberto Zariquiey.

which may have a pronoun in subject position, and Modern Hebrew, where non-subject positions allow or require a pronoun.

(3-7) Thai (Yaowapat & Prasithrathsint 2009:7)
thəə mây khuan kin yaa [*thîi Ø/**man** mòtʔaayúʔ*][19]
2.SG not should eat medicine NMLZR Ø/3.SG expire
'You should not take the medicine which expired.'

(3-8) Modern Hebrew (Semitic)
a. *Zo she* = [*Ø boxa*] *xi xavera sheli*.
this.FEM NMLZR cries is friend mine
'The (one) who is crying is my friend.'
b. *Zo she* = [*Yoav raa Ø/ **ota** etmol*] *xi xavera sheli*
this.FEM NMLZR = Yoav saw her yesterday is friend my
'The (one) whom Yoav saw yesterday is a friend of mine.'
c. *Ze* [*sixakti ito etmol*] *haya shovav*.
this.MSC 1.played with.him yesterday was naughty
'The (one) with whom I played yesterday was naughty.'

3.2.1 *Role markers*

While in Japanese, Thai, Hebrew, and many other languages, a gap or a pronoun is the only clue in ascertaining the nature of an argument evoked, others in addition have morphological markers that indicate the grammatical role of the evoked argument. The pattern that appears more wide-spread than others distinguishes subject (or agentive) and non-subject (or patientive/result) argument nominalization possibly with some additional distinctions, as in some dialects of Quechua (*-q* vs. *-sqa*), Turkish (*-En* vs. *-dIk/-cEk* + personal suffix), Tibeto-Burman Qiang (*-m* (and others) vs. *-Ø* + GEN), and Yaqui, where, as seen below, the argument positions nominalized are represented by a gap for central grammatical relations and with person-marked relational particles, e.g. *a-mak* in (3-9d), for peripheral relations with additional morphology (*-me* for subject, *-ʼu* for non-subject, and *-ʼVpo* for locative) indicating the grammatical roles that are nominalized.

19. My Thai consultant finds the form with the pronoun in subject position less felicitous than the one with a gap. Riddle (1993: 60) gives White Hmong as another language that permits a "resumptive pronoun" in subject position of a relative clause. A Kwa language in Ghana, is a rare type *requiring* a pronoun in the subject position of argument nominalizations as well (Campbell 2017).

(3-9) Yaqui (Southern Uto-Aztecan; Albert Alvarez 2012)
Subject nominalizer: *-me.*
a. *Jabesa* [*Ø wa-me yabe-m tea-ka-**me***]
who DEM-PL key-PL find-PERF-S.NMLZR
'Who is the one that found the keys?'
Non-subject nominalizer: *-ʼu.*
b. *Jitasa* [*Joan-ta Ø tea-ka-ʼ**u***]
what John-GEN find-PERF-NS.NMLZR
'What is the one that John found?'
c. [*em rebo'osam ameu jinu-ka-ʼ**u***]*-m*
2SG.GEN mantilla 3PL.DAT buy-PERF-NS.NMLZR-PL
'ones from whom you bought mantilla'
d. [*nim a-mak yepsa-ka-ʼ**u***]
1SG.GEN 3SG.ACC-COM come-PERF-NS.NLZR
'one with whom I came'.
Locative nominalizer: *-ʼVpo*
e. [*jamuch-im Ø to'e-ʼ**epo***]
woman-PL sleep-L.NMLZR
'(place) where women are sleeping'.

Far more complex patterns are seen in South America. The Carib language Hixkaryana has markers of event nominalizations distinct from those marking the role types of argument nominalizations, both of which have variant forms marking tense as well.

(3-10) Hixkaryana (Carib; Derbyshire 1999: 48–49)
a. Event/Action nominalizer: *-(ni)ɾɨ*
a-wanota-niɾɨ
2-sing-AC.NMLZR
'your singing'
b. Event/Action nominalizer-Past tense: *-thɨɾɨ*
i-wanota-thɨɾɨ komo
3-sing-AC.NMLZR COLL
'their singing (in the past)'
c. Nominalizer of A (Agentive protagonist of transitive event): *-ɲe*
ɾo-hananɨh-ɲe
1-instruct-A.NMLZR
'one who instructs me'
d. Nominalizer of the S (Protagonist of intransitive event)/O (Patientive protagonist of transitive event)-Past tense: *-saho*
S: *ɨ-manho-saho uɾo/omoɾo/moki*
IMPERS-dance-S.NMLZR 1/2/3PRO
'I (am)/you (are)/he (is) the one who danced.'

O: *t-ono-saho* *koso*
IMPERS-eat-O.NMLZR deer
'The deer (was) what was eaten.'

e. Nominalizer of thing/time/place associated with State/Action: *-t(o)ho*
ɾ-omoh-toho
I-come-ASSOC.NMLZR
'the time of my coming'/'place to which I'll come'/'circumstance of my coming'
ɨ-hkoto-tho
3-cut-ASSOC.NMLZR
'saw, thing for cutting it'

f. Nominalizer of Negation: *-hɨnɨ/-hnɨ*
ɨ-to-hnɨ *uɾo*
IMPERS-go-NEG.NMLZR 1PRO
'I (am) one who does not/will not go.'

While in Hixkaryana O and S are treated alike showing an ergative pattern in nominalization (3-10d), the Tupian language Kamaiurá distinguishes not only between O and S, but also between O (patient) and P (theme).

(3-11) Kamaiurá (Tupí-Guaraní; Seki 2000 and p.c.)[20]
A: *-tat/-tar*

a. [*Ø ka'i-a* *juka-**tar**-er-a*]
monkey-NU kill-A.NMLZR-PST-NU
'(one) who killed a monkey'

S: *-ma'e*

b. *o-'ata-ma'e* (122)
3-walk-S.NMLZR
'that which/who walk'

c. *i-pitsun-ama'e* (179)
3-black-S.NMLZR
'that which is black'

O (patient): *-emi*

d. *je=r=emi-juka* (121)
1SG=RELATIONAL=O.NMZ-kill
'what I killed'

20. Apparently the functions of the cognates of these markers are somewhat different in other Tupí-Guaraní languages (see Jensen 1999: 160).

P (theme): *-ipyt*/*-pyr*

e. *i-mono-pyr-er-a* *Sau Paulo katy*
3-send-P.NMZR-PST-NU direction
'one sent to Saõ Paulo'

Location/Instrument: *-tap*/*-ap*/*-taw*/*-aw*

f. [*i-jo-taw-er-a*]
3-go-L.NMLZR-PST-N
'place where he went out from'

g. [*moĩ-a* *juka-taw-er-a*]
cobra-N kill-I.NMLZR-PST-N
'thing with which I killed the snake'

The original function of the so-called Austronesian focus-system appears similar to these role markers of argument nominalizations (Starosta, Pawley, and Reid 1982). The AF (actor-focus) form marks subject/agent argument nominalizations, PF (patient-focus) form object/patient argument nominalizations, LF (locative-focus) form locative argument nominalizations, and CF (circumstantial-focus) form beneficiary and instrumental argument nominalizations. Many Austronesian languages in Taiwan, Malaysia, and Indonesia have reduced the proto-Austronesian four-way focus pattern to a three- or a two-way contrast, but many others in Taiwan, the Philippines, as well as Malagasy largely have reflexes of the original four-way contrast, as seen in the following forms from Mayrinax Atayal in Taiwan.

(3-12) Mayrinax Atayal (Austronesian; based on Huang 1995)

a. Actor focus (Actor nominalization)
ßaq-un=mu *kuʔ* [***m***-*aquwas* Ø] *kaʔ haca ʔ*
know-PF=1SG.GEN NOM.REF AF-sing LIN that
'I know that singer/one who is singing.'

b. Patient focus (Patient nominalization)
ma-hnuq *kuʔ* [*ß*-***in***-*ainay* Ø *nukʔ* *naßakis*]
AF-cheap NOM.REF buy<PF.REALIS>buy GEN.REF old.man
'What the old man bought was cheap.'

c. Locative focus (Locative nominalization)
yayhapuyan kuʔ [*naniq*-***an*** Ø *cuʔ* *ßuŋaʔ nkuʔ* *ʔulaqiʔ*]
kitchen NOM.REF eat-LF ACC.NONREF yam GEN.REF child
'The kitchen is (the place) where the child eats yam.'

d. Circumstantial focus (Benefactive/instrumental nominalization)
ini=mu *sʔwaʔ=I kuʔ* [***si***=*ghahapuy* Ø *nkuʔ* *kanairil*]
NEG=1SG like=LF NOM.REF CF=cook GEN.REF woman
'I don't like the one for whom the woman cooks.'

Turning to European languages, German shows a very systematic pattern similar to the Austronesian focus system, but in terms of demonstrative-based nominalization markers. Observe:

(3-13) German

a. *Ich empfange den,* [***der*** [*Ø morgen kommt*]].
I receive ART SUB.NMLZR tomorrow comes
(SUB nominalization)
'I receive the one who comes tomorrow.'

b. *Ich empfange den,* [***den*** [*du mir Ø vorgestellt hast*]].
I receive ART DO.NMLZR you me introduce.P have
(DO nominalization)
'I receive the one whom you introduced to me.'

c. *Ich empfange den,* [***dem*** [*du Ø den Brief gegeben hast*]].
I receive ART IO.NMLZR you ART letter give.PP have
(IO nominalization)
'I recieve the one whom you gave the letter.'

d. *Ich empfange den*, [***dessen*** [*Ø Buch ich gelesen habe*]].
I receive ART GEN.NMLZR book I read.PP have
(GEN nominalization)
'I receive the one whose book I have read.'

Modern English has partially lost the marking distinction between subject and object nominalization, but the marker *whom* still uniquely marks an object nominalization. It also has distinct markers for adverbial nominalizations, denoting a place, a time, etc., as below, where we are reanalyzing so-called relative pronouns as nominalizers.

(3-14) English

a. *You should marry* [***who*** [*Ø loves you*]]. (SU nominalization)
b. *You should marry* [***who/whom*** [*you love Ø*]]. (OBJ nominalization)
c. *The man* [***whose*** [*Ø book has just been published*]] is in town.
(GEN nominalization)
d. *The book hits* [***where*** [*it hurts Ø most*]]. (Place nominalization)
e. *That was* [***when*** [*I ruled the world Ø*]]. (Time nominalization)[21]
f. [***Why*** [*he didn't come to the party Ø*]] *remains a mystery*.
(Reason nominalization)
g. *You might not like* [***how*** [*he eats his food Ø*]]. (Manner nominalization)

English also distinguishes forms for animate and inanimate denotations in terms of *who(m)* and *which/what*.

21. It is not clear if we have a gap in this and the structures in (f) and (g).

(3-15) English
a. *You may choose* [***who*(*m*)** [*you like Ø*]]. (Animate OBJ nominalization)
b. *You may choose* [***which*/*what*** [*you like Ø*]]. (Inanimate OBJ nominalization)

The classificatory function of nominalizers is observed in Newar and many other languages that have *nominalizing classifiers* (see Section 8).

(3-16) Newar (Tibeto-Burman; courtesy of Kazuyuki Kiryu)
a. [*ana Ø dan-ā cwã:=**mha***] *rām=yā macā kha:.*
there stand-CM exist.ND=NMLZR Ram=GEN child COP
(Animate SUB nominalization)
'The one standing there is Ram's child.'
b. [*ana Ø du=**gu***] *rām=yā gāri kha:.*
there exist.ND=NMLZR Ram=GEN car COP
(Inanimate SUB nominalization)
'The one that is there is Ram's car.'

These classifying nominalizations play a vital role in our understanding of nominalizations as nominal structures that denote substantive entities classifiable according to features such as gender, animacy, physical shape, and function. The overall importance of the role markers discussed above in the formation of relative clause constructions is discussed in Section 9.2.

4. Formal representations of verbal-based nominalizations

We assume that nouns have a denotation index that connects a noun with a set of substantive concepts it denotes, as in the following manner, where the arbitrary numbers are indices that connect nouns (forms) and the entity concepts that they denote (meanings).

(4-1) Noun CONCEPT

a. $[\text{dog}]_{N}^{25}$ 25 .....

b. $[\text{cat}]_{N}^{30}$ 30 .....

c. $[\text{flower}]_{N}^{45}$ 45 .....

Nominalization is a process that creates linguistic forms (words and complex structures) that have form-concept pairings similar to nouns, as below.

(4-2) Nominalization CONCEPT

a. $[[\text{cook}]_{\text{NMLZ}}]_{\text{N}}^{125}$

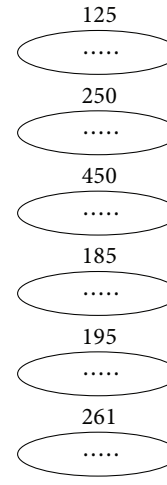
b. $[[\text{building}]_{\text{NMLZ}}]_{\text{N}}^{250}$

c. $[\text{that [John came yesterday]}]_{\text{NMLZ}}^{450}$

d. $[\text{who [Ø}^{185}\text{ loves you]}]_{\text{NMLZ}}^{185}$

e. $[\text{whom [you love Ø}^{195}\text{]}]_{\text{NMLZ}}^{195}$

f. [she [Yoav raa **ota**^{261} etml]]$_{\text{NMLZ}}^{261}$
NMLZR Yoav saw her yesterday
'(one) whom Yoav saw yesterday'



Lexical nominalizations (4-2a) and (b) are like ordinary nouns, except the former is derived via zero morphology, and are straightforward. While nouns have permanent form-concept relationships registered in the speaker's mind, grammatical nominalizations establish form-concept relationships for the nonce, whereby nominalization structures as a whole bear denotation indices. In event nominalization (4-2c), the entire structure may denote a kind of fact, namely the one associated with the event of John's coming yesterday. In the case of argument nominalizations, nominalization structures bind an argument position that is empty, as in the English examples or that may be occupied by a pronoun, as in the Hebrew example (4-2f). This binding means that what the nominalization structure denotes plays the grammatical role of co-indexed NP position. Subject nominalization (4-2d), for example, denotes a substantive entity that plays a subject role; e.g. one who loves the addressee. Object nominalizations (4-2e), (f) denote substantive entities that play an object role; e.g. one the addressee loves or one Yoav saw yesterday.

Unlike the usual analysis of relative clauses that assumes so-called relative pronouns to bind an empty argument position, our analysis makes the entire nominalization structures bear a denotation index, just like ordinary nouns, which directly binds an empty argument/adjunct position, or one filled by a pronoun, both of which function as a variable. In this analysis so-called relative pronouns play no pronominal role; they are treated as nominalizers whose functions are similar to nominalizing morphology or nominalizing particles in other languages. This is a desired result since many languages of the world, as we shall see presently, do

not involve any pronouns, like the so-called relative pronouns in English, in their formation of relative clause constructions.

As in the other cases of metonymy, argument nominalizations evoke a variety of denotations and context determines the appropriate denotations per the Gricean Cooperative Principle. For example, Spanish subject nominalization (4-3) below denotes a variety of masculine entities that are white, such as a white stallion, a white male dog, a white hat, and a white car. In the context of (4-4), the denotation of {WHITE HAT} will be chosen, while in the context of (4-5), the denotation of {WHITE CAR} will be selected.

(4-3) [*que* [*Ø es blanco*]]
NMLZR is white
'that which is white'

(4-4) a. *¿Qué tipo de sombrero te gusta?*
What kind of hat you like.3SG
'What kind of hat do you like?'
b. *Me gusta uno* [*que* [*Ø es blanco*]]
me like.3SG one NMLZR is white.
'I like one which is white.'

(4-5) a. *¿Qué tipo de coche te gusta?*
what kind of car you like.3SG
'What kind of car do you like?'
b. *Prefiero uno* [*que* [*Ø es blanco*]].
prefer.1SG one NMLZR is white
'I prefer one which is white.'

Having elaborated on the nature of grammatical nominalizations, we shall now turn to their use and function.

5. Structures and their use

One of the major problems we find in some current definitions of nominalization in dictionaries and encyclopedias, and even in professional writings is the confusion over a form and its use. For example, Crystal's definition discussed in the introduction says that nominalization may involve "the derivation of a noun phrase from an underlying clause" (Crystal 1980: 328). In a similar vein, Givón (2009: 66) defines nominalization as "the process via which a *finite verbal clause* – either in its entirety or only the subject-less verb phrase – is converted to a *noun phrase*."

Just as ordinary nouns are different from noun phrases, nominalizations are different from noun phrases, though both may be *used* as the head of a noun phrase. Nouns and nominalizations may also be *used* as a modifier that may restrict or identify the denotation of a head noun. These two major uses of nominals are illustrated below, first using an ordinary noun.

(5-1) NP-use/Referential function
[[*Cotton*]$_N$]$_{NP}$ *is a useful commodity*.

(5-2) Modification-use/Restrictive function
a. [[*cotton*]$_N$ [*mill*]$_N$]$_N$ (noun compound)
cf. *lumber mill*
b. [[[*cotton*]$_N$]$_{N'}$ [*shirt*]$_N$]$_{NP}$ (noun phrase)
cf. *silk shirt*

In (5-1) the noun [*cotton*]$_N$ heads an NP and has a referential function at the NP level, referring to a type of commodity in the real world. We distinguish between "denotation" and "reference". The former is the relation between words/structures and their associated mental concepts, as described in the preceding section, while the latter is the relationship between noun phrases headed by words/structures and entities in the real (or imaginary) world. Thus nouns and nominalizations themselves have denotations but they do not directly refer. When they head an NP, as in (5-1), they play a referential function as the main constituent of an NP. Reference is a type of speech act of pointing out a real world entity by the use of a nominal form as the head of nominal phrase of N′ or NP upon recognition of a real world entity in question as an instance of the denotation of the nominal form being used.[22] In the modification-use, as in (5-2), nouns do not refer. Instead, their denotations function to restrict the denotation of a head nominal. Notice that these modifying nouns remain nouns; in particular, they do not become adjective even when they play a modification function. This is shown by the fact that a modifying noun takes an adjectival modifier, not an adverbial modifier that adjectives would take. Observe:

(5-3) a. [[***Egyptian*** *cotton*]$_{N'}$ *shirt*]$_{NP}$
b. [[***new***/**newly car*]$_{N'}$ *smell*]$_{NP}$

Figure 1 shows the relationship between a structure (noun in our case) and its use and the functions associated with the use.

22. Cf. Searle (1969: 85): "[reference is] an act of identifying some entity that the speaker intends to talk about."

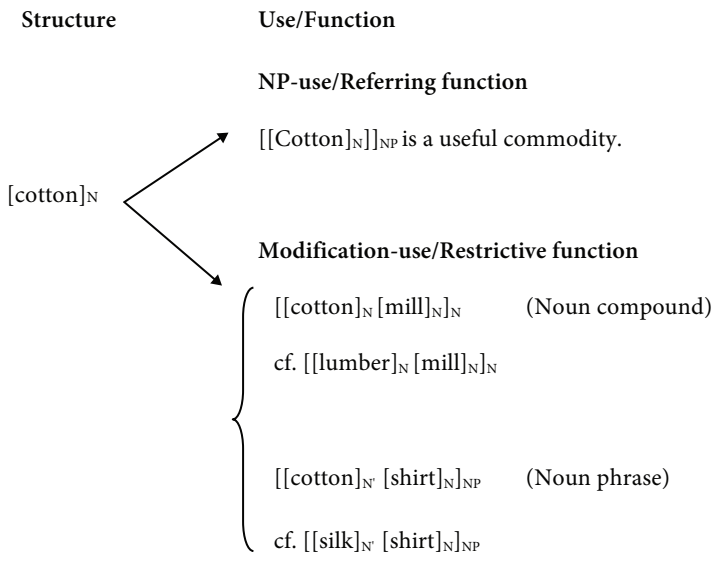


Figure 1. Two uses of noun

Grammatical nominalizations, qua quasi-nominals, behave like ordinary nouns in allowing both NP- and modification-use, as below:

(5-4) NP-use of event nominalization
a. [[*That John is honest*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *is absurd*.
b. *I know* [[*that John is honest*]$_{\text{NMLZ}}$]$_{\text{NP}}$

(5-5) Modification-use of event nominalization
a. [[*The proposition* [*that John is honest*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *is absurd*.
b. *I know* [*the fact* [*that John is honest*]$_{\text{NMLZ}}$]$_{\text{NP}}$

(5-6) NP-use of argument nominalization
a. [[*Who* [*Ø gets there first*]]$_{\text{NMLZ}}$]$_{\text{NP}}$ *defines the truth*. (David Baldacci)
b. *You should marry* [[*who* [*you love Ø*]]$_{\text{NMLZ}}$]$_{\text{NP}}$

(5-7) Modification-use of argument nominalization
a. [*The man* [*who* [*Ø gets there first*]]$_{\text{NMLZ}}$]$_{\text{NP}}$ *defines the truth*.
b. *You should marry* [*a man* [*who* [*you love Ø*]]$_{\text{NMLZ}}$]$_{\text{NP}}$

The two uses of an argument nominalization structure are diagrammatically shown in Figure 2 (next page).

5.1 NP-use and modification-use of grammatical nominalizations

As demonstrated by many of the contributions to this volume, these patterns of use of grammatical nominalizations are not unique to English or European languages.

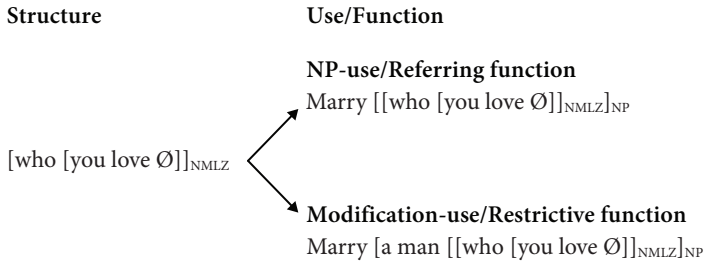


Figure 2. Two uses of argument nominalization

Just to drive this point home, we shall first provide random samples from genetically diverse languages of the Americas and elsewhere illustrating the NP- and modification-use of argument nominalizations paralleling the patterns of (5-6) and (5-7) above, whose far-reaching theoretical significance will be discussed in the subsequent sections. The examples labeled (a) represent the NP-use and those labeled (b) the modification-use of argument nominalizations.

(5-8) Toba (Guaicuruan; Messineo & Porta 2009 and Cristina Messineo p.c.)

a. [[*na* [*Ø-chigoqchigiña yi Espinillo*]$_{\text{NMLZ}}$]]$_{\text{NP}}$ *Ø-tayge da Salta ko'ollaGa*
DD 3-come from DD Espinillo 3-go DD Salta PST
'Those who came from the Espinillo went to Salta.'

b. *ajem si-kjen* [[*so* [*ſijaGawa* [*Ø-neta-ge da cako*]$_{\text{NMLZ}}$]$_{\text{N'}}$]$_{\text{NP}}$ (55)
I 1A-greet DET man 3-be-DIR DET Chaco
'I greeted the man who lives in Chaco.'

(5-9) Tapiete (Tupí-Guaraní; Ciccone 2008)

a. *á-ha-po a-hapi* [[*kwé(we) a-yasíya-wa*]$_{\text{NMLZ}}$]]$_{\text{NP}}$
1SG.AC-go-FUT 1SG.AC-light before 1SG.AC-cut-NMLZR
'I am going to light what I cut last time.'

b. *hau ye* [*waka ro'o* [*a-mbaku-wa*]$_{\text{NMLZ}}$]$_{\text{NP}}$ (27)
1:eat already cow meat 1.SG.AC-heat-NMLZR
'I already ate the meat that I heated.'

(5-10) Kipeá (Macro-Jê; Rodrigues 1999: 104, 195)

a. [[*di-te-ɾi*]$_{\text{NMLZ}}$]$_{\text{NP}}$
ABS-come-NMLZR
'the one who comes'

b. [*waɾe* [*du-di-ɾi udza*]$_{\text{NMLZ}}$]$_{\text{NP}}$
priest ERG-give-NMLR knife
'the priest who gave a knife'

(5-11) Kakataibo (Panoan; Zariquiey 2011: 634, 642)

a. *Marianën ʼakukë a kana pian*
[Maria-nën Ø ʼaku-kë]$_{\text{NMLZ}}$ a kana pi-a-n
Maria-ERG cook-NMLZR 3SG.O NAR.SG eat-PERF1/2P
'I ate what Maria cooked.'

b. *ain bënën ʼakë buë*
[ain bënë=n Ø ʼa-kë]$_{\text{NMLZ}}$ buë
3.GEN husband=ERG do-NMLZR fish.esp
'the fish that her husband fished'

(5-12) Bolivian Quechua (Quechuan)

a. [*Maria-q wayk'u-sqa-n*]*-ta mik"u-sayku*
Maria-GEN cook-O.NMLZR-3SG-ACC eat-PROG.1PL.EXCL
'We are eating what Maria cooked.'

b. [[*Maria-q wayk'u-sqa-n*] *wallpa*]*-ta mik"u-sayku*
Maria-GEN cook-O.NMLZR-3SG chicken-ACC eat-PROG.1PL.EXCL
'We are eating the chicken that Maria cooked.'

(5-13) Hixkaryana (Carib; Derbyshire 1999: 48, 57)

a. NP-use of O-nominalization (*-saho*)
[*t-ono-saho*]$_{\text{NMLZ}}$ *koso*
IMPERS-eat-O.NMLZR deer
'The deer (was) what was eaten.'

b. Modification-use of O-nominalization (*-nɨ*)
mɨrɨrɨ ɛrahma-phɨ-i-ya tɨ-ɾui nɨ-kupɨ-hpɨ
that see-PST-3-ERG 3REFL-brother O.NMLZR-do-PST
'He saw that, what his brother had done.'

(5-14) Tuyuca (Tucano; courtesy of Janet Barnes)

a. [*baco-á-ri-gɨ*]
to.have.been.bitten-recent-SG.NMLZR-CLF
(cylindrical.shape,long.and.solid)
'that which (=a cylindrically-shaped, long and solid thing) was bitten'

b. [*nɨḱá* [*baco-á-ri-gɨ*]]
leg to.have.been.bitten-recent-SG.NMLZR-CLF
'the leg that was bitten'

(5-15) Piapoco (Arawak; Klumpp & Burquest 1983: 390, 395)

a. *yà-a-wa* [*i-té-eyéi-ca yà-ana*]
3 m-go-aspect 3 m-carry-[+pl]-aspect 3 m-limb
'(the ones) who carry the animal's leg go'

b. *niái inanaía* [*ì-yamé-eyéi-cawa capìi ìricu*]
those women 3-stay-[+pl]-aspect house in
'those women who stayed in the house'

(5-16) K'ichee' (Mayan (Quichean); Larsen & Norman 1979: 357 and courtesy of Nora England and Telma Can Pixabaj)

a. *x-Ø-inw* [*lee* [*x-Ø-u-ch'ay* *lee*
ASP-3SG.ABS-1SG.ERG-see NMLZR ASP-3SG.ABS-3SG.ERG-hit the
achih]$_{NMLZ}$]$_{NP}$
man
'I saw the one whom the man hit.' (Or 'I saw the one who hit the man.')

b. [*lee ixoq* [*lee* [*x-Ø-u-ch'ay* *lee achih*]]$_{NMLZ}$]$_{NP}$
the woman NMLZR ASP-3SG.ABS-3SG.ERG-hit the man
'the woman whom the man hit' (Or 'the woman who hit the man')

(5-17) Yucatec (Mayan (Yucatecan); Gutièrrez-Bravo 2012: 262, 264)

a. *Yaan-Ø* [*k-u* *wéej taal bejla*]$_{NMLZ}$-*e'*
EX-ABS.3SG HAB-ERG.3SG still come today-CL
'There are those that still come today.'

b. *Tuláakal le gente* [*k-u* *taal*]$_{NMLZ}$-*o'*, *k-u*
all DM people HAB-ERG.3SG come-CL HAB-3.ERG.3SG
ts'a-ik-Ø *u jaal.*
give-IND-ABS.3SG ERG food
'All the people that came, he gave them their food.'

(5-18) Nieves Mixtec (Otomamguean; Caponigro, Torrence & Cisneros 2013: 70, 76)

a. [*yō*[23] *ni-kānī jēráldó*] *ni-kāni jwán*
NMLZR CMP-hit Geraldo CMP-hit Juan
'The one(s) who hit Geralido hit Juan too.'

b. *jwán kūt óó = ra* *ñáʔa* [*yō* *kūt óó* *jēráldó*]
Juan like.con = 3SG.M woman NMLZR like.con Geraldo
'Juan likes the woman who Geralido likes.'

(5-19) Yaqui (Southern Uto-Aztecan; Alvarez 2011)

a. [*Junu'u weyeka-me*] *nakapit*
DEM be.standing-SUB.NMLZR deaf
'This one who is standing is deaf.'

b. *Joan uka* *chu'u-ta* [*Maria-ta* *ke'e-ka-m*]*-ta* *me'a-k*
John DET.ACC dog-ACC Mary-ACC bite-PERF-SUB.NMLZR-ACC kill-PERF
'John killed the dog that bit Mary.'

23. The original gloss was "who", which, like English wh-nominalizations/relatives, is treated as a nominalizer in this paper. See Mithun (2012) for more examples illustrating the use of interrogative pronouns as nominalizers.

(5-20) Tümpisa (Panamint) Shoshone (Northern Uto-Aztecan; Dayley 1989: 476)

a. [*ke tamangkan-tü*]$_{\text{NMLZ}}$ *naamaa setü.* [*Nümmi appü*] *utü.*
not tooth.having-NMLZR was this our(EXCL) father that
'He is one who is missing a tooth. That is our father.'

b. *Tangumü* [*nü pusikwa-tü*]$_{\text{NMLZ}}$ *tühüyanna kuttihantü.*
man me know-NMLZR deer.O shot-STV
'The man who knows me is shooting the deer.'

(5-21) Navajo (Southern Athabaskan; Willie 1989: 415, 435)

a. [*'at'ééd yizts'ǫs-yéé*] *yałtí'*
girl 3sO: 3sS:kiss-NMLZR speaking
'The one who kissed the girl is speaking.'

b. [*'ashikii* [*'at'ééd yizts'ǫs-yéé*]]
boy girl 3sO: 3sS:kiss-NMLZR
'the boy who kissed the girl'

(5-22) Lakhota (Siouan; Van Valin 1977: 81)

a. [[*ŝu'kawakhâ wa ima'kicu ki he*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *wâyâ'ke*
horse a he.takes.it.from.me DET that he.sees.him
'He saw the one who took a horse from me.'

b. [*wicha'ŝə* [*ŝu'kawakhâ wa ima'kicu ki he*]$_{\text{NMLZ}}$]$_{\text{NP}}$
man horse a he.takes.it.from.me DET that
wâyâ'ke he?
he.sees.him Q
'Did he see the man who took a horse from me?'

(5-23) Cherokee (Iroquoian; Montgomery-Anderson 2008: 523, 560)

a. *júúskwakahli jituútóoʔa*
[ti-uu-ǩ skwakahli ji-tee-uu-at- óoʔa]
DST2-3B-striped NMLZR-DST-3B-be.called:PRC
'what is called "Striped" '

b. *askaya jijiiyaliì nohehtiiskv́ aàhnika*
a-skaya [ji-jii-ali-hnohehtiisk-v́v́ʔi] a-ahnika
3A-man NMLZR-1A.AN-MDL-talk.with:INC-EXP\SUB 3A-leave:IMM
'The man that I was talking to left.'

(5-24) Creek (Muskogee) (Muskogean; Martin 2011: 391, 394)

a. [*iłki-acól-i lêyk-a:t*] *ł-ółho:y-atí:s*
3PAT.father-old-1 sit.SG.FGR-REF DIR-reach.DU.LGR-PST5-IND
'They got to [where their elderly father lived].'

b. *asêy ifà* [*a:hôył-a:t*] *lopéyc-i:-t ô:-s*
that dog DIR.stand.FGR-REF nice-DUR-T be.FGR-IND
'That dog standing over there is friendly.'

(5-25) Rainy River Ojibwa (Algonquian; Johns 1982: 161, 162)

a. [*ka*[24]*:-nagamʌ-t*] *kino:zi*
NMLZR-sing-3c tall.3
'The one who is singing is tall.'

b. [*inini* [*ka:-nagamʌ-t*]] *kino:zi*
man NMLZR-sing-3c tall.3
'The man who is singing is tall.'

(5-26) St'át'imcets (Lillooet) (Salish; Davis 2010: 11)

a. ats'x-en=lhkácw=ha *i*=[[*t'íq*]=*a*]
see-TR=2SG.SU=YNQ PL.DET=[[arrive]=EXIS]
'Did you see [the ones who arrived]?'

b. *ats'x-en=lhkácw=ha i*= [[t'íq]=*a sqáyqeycw*]
see-TR=2SG.SU=YNQ PL.DET=[[arrive]=EXIS men]
'Did you see [the men who arrived]?'

(5-27) Central AlaskanYupik (Eskimo-Aleut; Miyaoka 2012: 533, 543)

a. [*Tau-na* [*neqe-m ii-ngan nere-sti-i*]$_{NMLZ}$
that-EE.ABS.SG fish-REL.SG eye-REL.3SG.SG eat-A.NMLZR-ABS.3SG.SG
kass'a-u-llini-uq
white.man-be-EVD-IND.3SG
'(I see now) that the one who is eating the fish eye is a white man.'

b. [*neqe-m nere-sti-i*] *qimugta*
fish-REL.SG eat-A.NMLZR-ABS.3SG.SG dog.ABS.SG
'the dog that eats the fish'

Just to round things out, let us observe similar uses of argument nominalizations elsewhere around the globe, starting in the East Asia.

(5-28) Korean

a. [*Yenghi-ka ilk-un*] *kes-un acwu elyep-ta*.
Yonghee-NOM read-NMLZR NPM-TOP very difficult-IND
'What Yonghee read is very difficult.'

b. [*Yenghi-ka ilk-un*] *chayk-un acwu elyep-ta*.
Yonghee-NOM read-NMLZR book-TOP very difficult-IND
'The book that Yonghee read is very difficult.'

(5-29) Mongolian (Chakhar, Kesigten subdialect); courtesy of Bayaerduleng and Benjamin Brosig)

a. [[t^{h}ənt tʃɔgsɔ-tʃ pæ:-ɢa:]$_{NMLZ}$=n^{j}]$_{NP}$ mɑn-æ: xʉ:xət
there stand-CVB be-PRS.NMLZR=3POSS we-GEN child
'The one standing there is our child.'

24. The original gloss for this was "WH-". See footnote 19.

b. [[tʰənt tʃɔgsɔ-tʃ pæː-gaː]$_{NMLZ}$ xɯːxət]$_{NP}$ pɔl man-æː
there stand-CVB be-PRS.NMLZR child TOP we-GEN
'The child who is standing there is ours.'

(5-30) Mandarin Chinese

a. [[*Wǒ zuótiān mǎi-de*]$_{NMLZ}$]$_{NP}$ *hěn guèi.*
I yesterday buy-NMLZR very expensive
'What I bought yesterday was very expensive.'

b. [[*Wǒ zuótiān mǎi-de*]$_{NMLZ}$ *shū*]$_{NP}$ *hěn guèi.*
I yesterday buy-NMLZR book very expensive
'The book that I bought yesterday was very expensive.'

(5-31) Thai (courtesy of Kingkarn Thepkanjana)

a. *chán chɔ̂ɔp* [[*thîi khwɛ̌ɛn nay tûu*]$_{NMLZ}$]$_{NP}$
I like NMLZR hang in closet
'I like the one that is hanging in the closet.'

b. *chán chɔ̂ɔp* [*kràproɔŋ* [*thîi khwɛ̌ɛn nay tûu*]$_{NMLZ}$]$_{NP}$
I like skirt NMLZR hang in closet
'I like the skirt that is hanging in the closet.'

(5-32) Mayrinax Atayal (Austronesian; based on Huang 1995)

a. *ßaq-un=mu kuʔ* [[*m-aquwas*]$_{NMLZ}$] *kaʔ hacaʔ*]$_{NP}$
know-PF=1SG.GEN NOM.REF AF-sing LIN that
'I know that one who is singing there.'/ 'I know that singer there.'

b. *ßaq-un=mu kuʔ* [*kanairil kaʔ* [*m-aquwas*]$_{NMLZ}$ *kaʔ hacaʔ*]$_{NP}$
know-PF=1SG.GEN NOM.REF woman LIN AF-sing LIN that
'I know that woman who is singing there.'

(5-33) Standard Indonesian (Austronesian)

a. *Aku makan* [[*yang diberikan ayah kepada-ku*]$_{NMLZR}$]$_{NP}$
I eat NMLZR give father to-me
'I ate what father gave to me.'

b. *Aku makan* [*ikan* [*yang diberikan ayah kepada-ku*]$_{NMLZR}$]$_{NP}$
I eat fish NMLZR give father to-me
'I ate the fish that father gave to me.'

(5-34) Sasak (Pancor ngeno=ngené dialect; Lombok Island, Indonesia; Austronesian)

a. *Beng oku* [[*si léq méje*]$_{NMLZ}$ *ino*]$_{NP}$
give I NMLZR on table the
'Give me the one that is on the table.'

b. *Beng oku* [*buku* [*si léq méje*]$_{NMLZ}$ *ino*]$_{NP}$
give I book NMLZR on table the
'Give me the book that is on the table.'

(5-35) Kalkatungu[25] (Pama-Nyungan; Australia; Blake 1979: 101)

a. *kaanta-n̪a pakaik-ka kalpuru-ṯiŋu* [[*ɲiṉ-ti ŋu-n̪a ṇaɲa*]$_{\text{NMLZ}}$]$_{\text{NP}}$
leave-PST that-Ø Boulia-ABL you-ERG NMLZR-ACC saw
'The one whom you saw left Boulia.' (Ø = a morpheme without a referential content)

b. *ŋai u̯antiji-n̪a* [*pa-u n̪aur-ku* [*ɲin-ti ŋu-n̪a l̪aji*]$_{\text{NMLZ}}$]$_{\text{NP}}$
I look after-PST that-DAT child-DAT you-ERG NMLZR-ACC hit
'I've been looking after that kid you belted.'

(5-36) Telugu (Dravidian, southern India; courtesy of K. V. Subbarao)

a. *neenu* [[*John icc-in-a*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *di*[26] *cadiveenu*
I give-PST-NMLZR NPM read
'I am reading what John gave me.'

b. *neenu* [[*John icc-in-a*]$_{\text{NMLZ}}$ *pustakam*]$_{\text{NP}}$ *cadiveenu*
I give-PST-NMLZR book read
'I am reading the book that John gave me.'

(5-37) Hindi (Indo-Aryan; courtesy of Miki Nishioka)

a. *main* [*us kii xariidii huii*] *paRh rahaa*
I s/he.OBL GEN.F.SG buy.PFV.F SG/PL be.PFV.F SG/PL read PROG.M.SG
huU
be.PRS.1SG
'I am reading what s/he bought.'

b. *main* [*us kii xariidii huii*] *kitaab*] *paRh*
I s/he.OBL GEN.F.SG buy.PFV.f.SG/PL be.PFV.F.SG/PL book.F.SG read
rahaa huU
PROG.M.SG be.PRS.1SG
'I am reading the book that s/he bought.'

(5-38) Abkhaz (North West Caucasian; courtesy of George Hewitt)

a. [[*'jy.b.taxy.w*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *d.ga*
whom.you(FEM).want.Non-Finite/STAT/PRS 3SG.take(IMP)
'Take whom you (Female) want!'

b. [[*'jy.b.taxy.w*]$_{\text{NMLZ}}$ *a-xàc'a*]$_{\text{NP}}$
whom.you(FEM).want.Non-Finite/STAT/PRS article-man
d-aa-wèit'
he-comes-PRS/Finite/Non-STAT
'Here comes the man whom you want.'

---

25. Kalkatungu being an extremely "flat" language, the constituency of relevant phrases below is not entirely certain. It is clear, though, that there is a nominalizer and that a nominalization can function as a subject and can modify a noun.

26. See the discussion on this particle labeled NPM (NP-use marker) in Section 6.2 below.

(5-39) Soqotri (Semitic; Yemen)

a. [[*lɛ jə-ʕɛːlʒɛn birhɛ*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *ɟəduːh*
NMLZR-PL PL.M.IMP-love children come.3PL.M.PER
'Ones (MASC) who love children came.'

b. *ʁivɟ* [[*lɛ jə-ʕɛːlʒɛn birhɛ*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *ɟəduːh*
men NMLZR.PL PL.M.IMP-love children come.3PL.M.PER
'Men who love children came.'

(5-40) German

a. *Ich treffe* [*den,* [*der morgen kommt*]$_{\text{NMLZ}}$]$_{\text{NP}}$
I meet ART SUB.NMLZR tomorrow comes
'I meet the one who comes tomorrow.'

b. *Ich treffe* [*den Mann,* [*der morgen kommt*]$_{\text{NMLZ}}$]$_{\text{NP}}$
I meet ART man SUB.NMLZR morning comes
'I meet the man who comes tomorrow.'

(5-41) Spanish

a. [*El* [*que está leyendo un libro*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *es mi padre.*
the NMLZR is reading a book is my father
'The one who is reading a book is my father.'

b. [*El hombre* [*que está leyendo un libro*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *es mi padre.*
the man NMLZR is reading a book is my father
'The man that is reading a book is my father.'

(5-42) Kanuri (Nilo-Sahalan; Hutchison 1981)

a. [*àwó*[27] [*nyíà gàlàngîn*] = *dɔ́*] $_{\text{NP}}$ *fànέ!*
THING to.you 1SG.advise=DET 2SG.listen.IMV
'Listen to what I am advising you.'

b. [*kâm* [*rúkɔ́ nà= dɔ́*]]$_{\text{NP}}$ *sáwànɔ́m*
person 1SG.saw=DET your.friend
'The person that I saw is your friend.'

(5-43) Akan (Kwa; Campbell 2013)

a. *Mè-ǹ-hú* [*dèὲ* [*ɔ̀-bɔ́-ɔ̀ màámé*] *nó*]
1SG-NEG-like NMLZR 3SG-PST-hit woman DEF
'I didn't like the one who hit the woman.'

b. *Pàpá nó* [*déὲ* [*mè-hyíâ nó ɛ́nŕ!á*] *no*]
man DEF NMLZR 1SG-meet him yesterday DEF
'the man that I met yesterday'

27. See Section 8 on this marker.

(5-44) Chichewa (Bantu; Mchombo 2004, p. c.)

a. [*zi-méné mú-kú-zí-fŭn-a*] *zi-li pa chulu.*
10-NMLZR 2.PL-PRS-10OM-want-fv 10-be 16.LOC 7.anthill
'What you want is on the anthill.'

b. *Mbuzí* [*zi-méné mú-kú-zí-fŭn-a*] *zi-li pa chulu.*
10.goats 10-NMLZR 2.PL-PRES-10OM-want-fv 10-be 16.LOC 7.anthill
'The goats that you want are on the anthill.'

Many more similar examples could easily be adduced, but the above is perhaps enough to dispel the widely-accepted analyses of so-called relative clauses as independent structures apart from nominalizations, analyses that give rise to limited observations such as "a somewhat more rare function of nominalization: as a relative clause modifying a head noun" (Comrie & Thompson 2007: 378) or claims like "in certain languages relativization is indistinct from nominalization" (Comrie & Thompson 2007: 379) that suggest that relativization and nominalization are in principle two distinct structures.[28]

It is worth noting that the NP-use of argument nominalization plays an important role in the formation of Wh-questions and so-called cleft (or focusing) constructions in a fair number of languages. Perhaps the most well-known of this is the case in Western Malayo-Polynesian, where we find the following patterns (see Bruil (this volume) for a pertinent discussion):

(5-45) Sasak (Pancor dialect, Western Malayo-Polynesian)
Argument nominalization in NP-use

a. [*Si Ø mbace buku=ne*] *batur=ku.*
NMLZR AF.read book=this friend=1SG
'The one who read this book is my friend.'

Argument nominalization in modification-use

b. *Dengan* [*si Ø mbace buku=ne*] *batur=ku.*
man NMLZR AF.read book=this friend=1SG
'The man who read this book is my friend.'

Wh-question

c. *Sai* [*si Ø mbace buku=ne*]
who NMLZR A.read book=this
'Who read this book?'/(lit.) 'Who is the one who read this book?'

28. Views much wider than these and that are consistent with our analysis have been expressed by those working on Tibeto-Burman languages. DeLancey (2002: 56), for example, notes that "[t]he fundamental relativization pattern is the same throughout the family: relativization is a subspecies of clausal nominalization. The modifying clause is nominalized, and then stands in either a genitive or appositive relation to the head noun." Compare Newar examples (3-16) and (7-19) in the text illustrating DeLancey's point. See also Noonan (1997) and (2008).

Cleft

d. *Loq Ali* [*si* *Ø mbace buku=ne*]
ART Ali NMLZR A.read book=this
'It is Mr. Ali who read this book.'/ (lit.) 'Mr. Ali is the one who read this book.'

(5-46) Yaqui (Southern Uto-Aztecan; Alvarez 2012: 89, p.c.)

a. Wh-question
Jabesa [*wa Ø jiosam noktua-me*]
who DEM book read-NMLZR
'Who is reading the book?'/ (lit.) 'Who is the one that is reading the book?'

b. Cleft
Joan [*wa-me Ø yabe-m tea-ka-me*]
John DEM-PL key-PL find-PERF-NMLZR
'It is John who found keys.' (lit.) 'John is the the one who found those keys.'

(5-47) Rainy River Ojibwa (Algonquian; Johns 1982)

a. Wh-question
wenen [[*kaːʔbimpatoːt*]$_{\text{NMLZ}}$]$_{\text{NP}}$
who NMLZR (PST).run.3C
'Who ran?'/(lit.) 'Who is the one that ran?'

b. Cleft
ikwe [[*kaːnagamʌt*]$_{\text{NMLZ}}$]$_{\text{NP}}$
woman NMLZR.sing.3C
'It's the woman who is singing.'/ (lit.) 'The woman is the one who is singing.'

(5-48) Thompson River Salish (Kroeber 1977: 387)

a. Wh-question
*Swe*t [*k* [*wik-t-x^{w}*]$_{\text{NMLZ}}$]
who ART see-TRZ-2S.TS
'Who did you see?'/ (lit.) 'Who is the one that you saw?'

b. Cleft
c'éw'stin [*e* [*n-s-tx^{w}əp*]$_{\text{NMLZ}}$]
soap ART 1SG.POSS-NMLZR-buy
'It is soap that I bought.' /(lit.) 'Soap is what I bought.'

Before turning to the theoretical discussions of these usage patterns of grammatical nominalizations, let us take a quick look at another modification-use of grammatical nominalizations, namely their adverbial use. The adverbial use of grammatical nominalizations typically occur together with a conjunctive particle

or postposition, as in the Mekens and Mẽbengokre examples below (see also van Gijn, this volume), but a straightforward use of nominalizations in adverbial function is also seen in the Americas, as illustrated by some examples below, and elsewhere (e.g. Japanese, Sasak, Gã). As the label for each example below shows, these adverbials denote/refer to such concepts as purpose, condition, time, place, and manner metonymically evoked by the events that the relevant event nominalizations denote (see Van linden, this volume, for a range of adverbial function of nominalizations in a single language).[29]

(5-49) Tarma Quechua (Quechuan; Adelaar 2011)
Purposive: [*yaku picha-q*]-*mi* *away-ya-:*
water clean-NMLZR-AF go-PR-1SG
'I am on my way to clean the water (canal).'
Conditional: [*mana nuqnchik kuga-ta traqtra-sha*]-*m* *kiru-nchi ismu-n*
not we [INCL] coca-ACC chew-NMLZR-AF tooth-4P rot-3SG
'Our teeth rot if we do not chew coca.'

(5-50) Tapiete (Tupí-Guaraní; Ciccone 2008)
Purposive: *she a-ha* *a-heka* [*a-mbɨ'invita-**wërä***] (4)
I 1.SG.AC-go 1SG.AC-look.for 1SG.LOAN-invite-NMLZR
'I go to look for (fish), in order to invite (you).'

(5-51) Kakataibo (Panoan; Zariquiey 2011)
Time: *xu* *ʼikë* *kana* *ʼëx* *Limanu* *kwanakën*
[xu ʼi-kë] kana ʼë=x Lima-nu kwan-akë-n
small be-NMLZR NAR.1SG 1SG=S Lima-LOC go-REM.PST-1/2
'I went to Lima when I was small.'

(5-52) Mekens (Tupí; Galucio 2011)
Temporal/conditional:
kɨrɨ *se-ayt-kwa-t* [*se-akar-ab*]=*ese*
child 3C-cry-TR-PST 3C-fall-NMLZ=LOC
'The child cried when he fell down.'
(lit.) 'The child cried at his own falling.'
Cause (reason)/time:
ōt o-akara [*ōt o-etayap-ka-ab*]=ese
I 1SG-fall I 1SG-slip-TR-MMLZ=LOC
'I fell down because I slipped; I fell down when I slipped.'

29. Recall that these concepts are also denoted by nominalizations that function as the head of NP-arguments (see (3–14)). Nominalizations in adverbial use are adjuncts (*I will try to be home* [*when you arrive*]) rather than arguments (*I want to know* [*when you arrive*]).

(5-53) Mẽbengokre (Macro-Je; Salanova 2011)
Simultaneous event:
[*a-je tep bôr*] *nỳrỳri dja ba ngôj ku'õ*
2-ERG fish roast.PL.NMLZR while FUT I.NOM pot wash
'I'll wash the pots while you roast the fish.'

(5-54) Shoshone (Northern Uto-Aztecan; Dayley 1989: 371)
Simultaneous event:
Nüü [*kahni tünto'e-tü*]$_{\text{NMLZ}}$ *pahekkawa.*
I house climb-NMLZR fell
'Climbing on the house, I fell off.'

(5-55) Thompson River Salish (Kroeber 1997: 381)
Reason:
puyt ku t e [*n-s-qʔaz*]
lie 1S.IS OBL ART 1S.POSS.NMLZR-tired
'I lay down because I got tired.'

Closely related to these adverbial uses of nominalizations is their use in so-called clause-chaining, serial verb, or converb constructions, which are typically translated into English conjoined sentences connected by *and,* as adverbial forms in sequence as *after doing X, after doing Y …* or as participial forms in sequence as *having done X*, *having done Y* …, as seen in the translation of the following Japanese example.

(5-56) Japanese
Taroo wa [*tosyokan ni ik-i*], [*hon o yom-i*], [*siryoo o*
Taro TOP library to go-NMLZR book ACC read-NMLZR material ACC
sirabe-Ø], *syukudai o si-ta.*
check-NMLZR homework ACC do-PST
'Taro, having gone to the library, having read the book, (and) having checked out the material, did the homework.'

The *-i/-Ø* nominalizations here are the grammatical counterparts of the stem nominalizations discussed in Section 2.2, and the structures marked by them, like chained structures in other languages, lack tense and other finite verbal marking, unlike the tense-marked final verb *si-ta* 'do-PST'.

While the chained structures involving stem nominalizations seen in (5-56) is occasionally heard, they are mainly found in written form. As below, the modern colloquial chain structure in Japanese involves additional marking by so-called conjunctive particle *-te*, whose use in chain strctures goes back to Old Japanese (8$^{\text{th}}$ C).

(5-57) *Taroo wa* [*tosyokan ni it = te* (< *ik-i = te*)], [*hon o*
Taro TOP library to go-NMLZR = CON book ACC
yon = de (< *yom-i = te*)], [*siryoo o sirabe = te* (<*sirabe-Ø = te*)],
read-NMLZR = CON material ACC check = CON
syukudai o si-ta.[30]
homework ACC do-PST
'Taro, having gone to the library, having read the book, (and) having checked out the material, did the homework.'

Similar uses of nominalizations as chains are widely observed among the languages of the Americas. For example, Northern Paiute has a nominalizing suffix *-na*, which (a) yields action nominals from verbs, (b) marks object nominalizations in argument nominalization, and (c) which marks event nominalizations functioning as verb complements; (d) it also marks structures chained in the manner of the Japanese examples above. Thornes (2003: 128–129, 443) nicely illustrates these diverse *uses* of *na*-nominalizations in Northern Paiute as below, where his original gloss for *-na* in grammatical nominalizations follows the traditional term "participle" (PTCP).

(5-58) Northern Paiute (Numic, Uto-Aztecan)

a. Action nominal
*i=nossi-na waha-**na***
1=dream-NMLZR tell.of-NMLZR
'…telling of my dreams …' (telling my dreaming)

b. Object argument nominalization (in modification-use)
su=miidɨ [*i=kuhani-**na***] *kai toki kamma*
NOM=meat 1=cook-NMLZR NEG correct taste
'The meat I cooked doesn't taste right.' (lit. "the meat of my cooking")

c. Object complement
u-su [*ka=nɨmɨdzoho u=nagi-kya-**na***] *puni*
3-NOM OBL=people.Masher 3=chase-TRNSL-NMLZR see
'S/he saw the Nemedzoho chasing her/him.'

d. Chained nominalization
*yaisi isu kaiba kussi timatai-**na** yaisi usu pabiʔi*
then this mountain dust rise-NMLZR then that elder.brother
*u-punni-**na** uka kussi-ba yaisi pisa u=supidakwatu*
3=see-NMLZR that.OBL dust.LOC? then well 3=understand
'…then as the dust rose from this mountain, and the elder brother (Wolf) saw it, that dust, then (he, Wolf) understood it (i.e. that it meant).'

30. The parentheses enclose the historically antecedent forms, from which the modern forms arose via phonological changes.

Languages that allow nominalizations without specific nominalizing morphology may use unmarked nominalizations in their chaining constructions. Creek has several nominalizing suffixes, some of which can be used in marking both lexical and grammatical nominalizations. In addition, the language allows grammatical nominalizations that are not specifically marked as such. For example, Martin (2011) analyzes the bracketed structures below as involving no nominalizing morphology.[31]

(5-59) Creek (Martin 2011: 392, 394)

a. *hatâm* [*pô:mi ô:c-iy-a:**n***] *pôn-homp-ick-in*
again we have.FGR-1PL.AG-REF.N 1PL.DAT-eat.LGR-2S.AG-N
om-áɬti:-s
be-FUT-IND
'Then you will eat [what we have]…'

b. *asêy ifá* [*a:-hôyt-a:ti-**t***] *lopéye-i:-t ó:-s*
that dog DIR-stand.FGR-REF-T nice-DUR-T be.FOR-IND
'That dog [standing over there] is friendly.'

We know that the structures in the brackets above have been or are nominalized or nominal because they are marked by the referential markers *-a:n* and *-a:ti*, which mark referential or contrastive nominals, where the former combines with the object case marker -(*i*)*n* and the latter is followed by the subject case marker -(*i*)*t*. It is this type of case-marked nominal structures that form chains in Creek, as seen below, where the case markers take on a switch-reference function, with the subject marker -(*i*)*t* marking same subject across chained structures and the object marker -(*i*)*n* different subject.[32]

(5-60) Creek (Martin 2011: 346)
*a:-oséyy-**ín**,* *halâ:t-ey-**n**,* *an-cíyall-**ín**,*
DIR-come.out-HGR-N hold-FGR-1S.AG-N 1S-DAT-struggle.against-LGR-N
*tí-weyk-éy-**n**,* *hola:n-**ít***
RCP-throw.LGR-1S.AG-N defecate.LGR-T
'[Rabbit] came out [DS], I grabbed him [DS], he struggled against me [DS], I threw him down [DS], and he crapped [SS]…'

31. There is a possibility (if historical) that referential markers *a:n* and *a:ti* seen below are connected with the agentive nominalizer *-a* in the language.

32. The development of switch-reference markers out of nominalizers appears to be a widespred phenomenon. See Jones & Jones (1991: Section 11.2) on Barasano and the contributions to this volume by Zariquiey and Valle & Zariquiey on Kakataibo.

Contrary to the widely used term "clause-chaining", the chained structures are not predicating clauses. They are in fact event nominalizations denoting events that hold temporal relations (sequential or simultaneous) among themselves or with regard to the asserted event.[33] Chained nominalizations may have an overt marker indicating the temporal relations or may not. In the case of Japanese (see (5-56)–(5-57)) and Creek, the order of juxtaposition of nominalizations indicates a sequential temporal relation, whereas in Northern Paiute (5-58), nominalized structures in chain by themsleves denote simultaneously occurring events. Japanese uses the marker *-nagara* for simultaneous events. Northern Paiute, on the other hand, marks sequential events by means of the suffix *-si*, as below, which contrast with the simultaneous events seen in (5-58d).

(5-61) Japanese
Taroo wa [*aruk-i-nagara*] *hon o yonda*.[34]
Taro TOP walk-NMLZR-SIMUL book ACC read
'Taro read while walking.'

(5-62) Northern Paiute (Thornes 2003: 457)
*mi=tɨka-k ɨ-u-**si** nɨmmi tiwau mia-**si** na-noo-ka-**si** oo*
use=eat-APL-PMC-SEQ we.EXCL again go-SEQ MM-carry-TRNSL-SEQ DEM
ka=yamoso-tami
OBL=Ft.Bidwell-toward
'Having allowed us to eat, we went on again, and were hauled on out there to Fort Bidwell …'

We have presented above the general usage patterns of grammatical nominalizations with a special focus on the two uses of argument nominalizations across a wide variety of languages of both South and North America, as well as elsewhere across the world. Having laid out the empirical foundations of the usage patterns of grammatical nominalizations, we are now in a position to launch theoretical discussions. We start with the NP-use of event nominalizations.

5.2 NP-use of event nominalizations: So-called "internally-headed RCs"

One of the major issues pertaining to the use of event nominalizations centers around the construction in (3-3d). Comparison of a similar example and an ordinary relative clause construction in Quechua below illustrates the problem at issue.

---

33. See Section 6.2 on the definitions of clauses, sentences, and nominalizations.

34. Unlike sequential events like (5-56) and (5-57), the *-nagara* form allows only a single simultaneous event per sentence.

(5-63) Bolivian Quechua

a. [*Maria* ***wallpa-ta*** *wayk'u-sqa-n*]*-ta* *mik"u-sayku*
Maria chicken-ACC cook-NMLZR-3-ACC eat-PROG.1PL.EXCL
'We are eating the chicken that Maria cooked.'
(lit.) 'We are eating Maria cooking the chicken.'

b. [[*Maria Ø wayk'u-sqa-n*] ***wallpa***]*-ta* *mik"u-sayku*
Maria cook-NMLZR-3 chicken-ACC eat-PROG.1PL.EXCL
'We are eating the chicken that Maria cooked.'

Following Gorbet's (1974) lead, a large number of researchers (Keenan 1985, Cole 1987, Kuroda 1992, etc.) have analyzed the constructions similar to (5-63a) as "internally-headed relative clauses", assuming (i) that they are relative clauses and (ii) that a head nominal exists within "relative clauses" unlike ordinary relative clause constructions, where a head exists externally in the main clause. A problem with the first assumption is that it is not at all obvious that these structures have the function of relative clauses, which is either to restrict the denotation of the head noun (restrictive relatives) or to identify the head noun in terms of the denotation of a modifying nominalization structure (non-restrictive relatives). That the idiomatic English translations of the relevant structures turn out to be relative clauses is hardly acceptable evidence for the proposed internally-headed RC analysis. Indeed, my Quechua consultant from the Cochabamba village in Bolivia would use the externally-headed RC in (5-63b) over the so-called internally-headed RC in (5-63a) in answering a question such as "What/which chicken are you eating?"

Actually the only evidence that suggests the relative clause status of the so-called internally headed RCs is the fact that similar, but not identical, structures are used for externally-headed relative clauses. There are, however, languages that are said to have internally-headed RCs in the absence of externally-headed RCs (e.g. Yuman languages Diegueño, Jamul Tiipay, and perhaps others, Kutenai, Seri, Parkatêjê), showing a measure of independence of the two.

The second assumption that in these structures an argument internal to the "relative clause" is the argument of the main-clause predicate is also problematic. Those who have studied so-called internally-headed RCs have not looked at constructions like (3-3e), where there is no NP within the "relative clause" that can serve as a main-clause (semantic) argument; accordingly they cannot be analyzed as internally-headed RCs. Notice the exact structural parallelism between so-called internally-headed RCs and the resultative nominalizations below:

(5-64) Bolivian Quechua

a. [*Maria wallpa-ta* *wayk'u-sqa-n*]*-ta* *mik"u-sayku*
Maria chicken-ACC cook-NMLZR-3SG-ACC eat-PROG.1PL.EXCL
(lit.) 'We are eating Maria cooking the chicken.' 'We are eating {the chicken involved in the event of} Maria's cooking a chicken.'

b. [*Maria laranjas-ta ch'irwa-sqa-n*]-*ta* *ujya-ni*
Maria oranges-ACC squeeze-NMLZR-3SG-ACC drink-1SG
(lit.) 'I drink that Maria squeezed oranges.' 'I drink {the result of} Maria's squeezing oranges.'
Cf. **laranjas-ta ujya-ni*
oranges-ACC drink.1SG
'I drink oranges.'

In (5-64b) it is clear that the nominalization-internal argument (*laranjas* 'oranges') is not the semantic argument of the main-clause verb; yet the sentence is perfectly acceptable. Analyzing the form in (5-64a) as an internally-headed relative clause leaves resultative nominalizations of the type seen in (5-64b) unaccounted for.

Those who subscribe to the internally-headed RC analysis have not bothered to expand their data beyond the earlier observations and thus fail to recognize the likely fact that those languages permitting so-called internally-headed RCs allow resultative nominalizations of the type seen above. Besides Bolivian Quechua, in three more languages allowing so-called internally-headed relative clauses for which native speakers were available to the present author, this prediction turned out to be correct, as evident from the following examples from Japanese, Northern Qiang (Tibeto-Burman; China), and Parkatêjê in northern Brazil.

(5-65) Japanese
Ken wa [*Hana ga* ***mikan o*** *sibotte* *kureta*] *no o hitoiki ni nonda*.
Ken TOP Hana NOM orange ACC squeeze.GER gave NPM ACC one.gulp in drank
(lit.) 'Ken drank that Hana squeezed oranges for him in one gulp./Ken drank {the resultant product of} Hana's squeezing oranges for him in one gulp.'
Cf. **Boku wa mikan o nonda.*
I TOP orange ACC drank
'I drank oranges.'

(5-66) Northern Qiang (Tibeto-Burman; courtesy of Chenglong Huang)
a. [*themle-wu* ***tɕytsətʂi*** *ɦa-tʂa-tha-ji*] *lo-qu qa sə-tɕha*.
3PL-AGT orange-juice DIR-squeeze-PART-CSM DEF-CLF 1SG DIR-drink.1SG (CLF = classifier)
'I drank {the orange juice involved in} their squeezing orange juice.'
b. [*themle-wu* ***tɕytsə*** *ɦa-tʂa-tha-ji*] *lo-qu qa s-tɕha.*
3PL-AGT orange DIR-squeeze-PART-CSM DEF-CLF 1SG DIR-drink.1SG
'I drank {the resultant product of} their squeezing oranges.'

Cf. **tɕytsə lo-qu qa sə-tɕha.*
orange DEF-CLF 1SG DIR-drink.1SG
'I drank oranges.'

(5-67) Parkatêjê[35]
wa [*Jõprar te laranja jatenẽ*] (*ata*) *toikõm*
I Jõprar ERG orange squeeze ART drink
(lit.) 'I drink that Jõprar squeezed oranges.'
Cf. **wa laranja toikõm*
I orange drink
'I drink oranges.'

Our claim is that, in both so-called internally-headed RCs and resultative nominalizations of the above type, the semantic arguments of the relevant predicates are not to be sought structure internally in a direct manner, as is done by the proponents of the internally-headed RC analysis, because they are actually something that are evoked metonymically. In fact, many event nominalizations in NP-use lack internal arguments functioning as semantic arguments. For example, the semantic argument of the main-clause predicate in [*John's falling off of the bed*] *happened at 3: 00 am* is the event (of John's falling off of the bed) evoked by the nominalization structure, but the noun *event* is not found anywhere in the relevant structure. Similarly, the semantic object argument of the verb *know* in *I know* [*John is honest*] is a fact (that John is honest), not the state of affairs of John's being honest, but again there is no noun *fact* found in the structure; cf. the synonymy between *I know John is honest* and *I know* ***the fact*** *that John is honest*.

What we are seeing here are cases of the metonymy-meadiated syntax-semantics mismatch, as seen in ordinary metonymic expressions of the following type: *The first violin is sick today*, *I heard three CDs tonight*. The metonymy-based nominalization analysis proposed in this paper treats all these constructions as nominalizations that, like lexical nominalizations discussed earlier, metonymically evoke concepts such as facts, propositions associated with events or state of affairs at large, event protagonists/participants, resultant products, as well as circumstantial matters like time, location, and reason closely associated with an event. In these constructions, nominalization structures function as syntactic arguments as a subject or object precisely because they evoke and stand for thing-like entities just like ordinary nouns do. And it is these metonymically evoked entities that function as semantic arguments of the main predicates. We are arguing that all the structures in (5-68) and (5-69) below display the same syntax-semantic mismatch mediated by metonymy as indicated below, where {…} represents what is evoked metonymically and what functions as semantic arguments.

35. Thanks are due to Marília Ferreira for checking the Parkatêjê forms with native speakers.

(5-68) a. *The first violin* {PLAYER OF} *is sick today*.
b. *I heard three CDs* {MUSICAL SOUNDS OF} *tonight*.

(5-69) a. FACT
English: *Bill knows* [*that John is honest*]$_{\text{NMLZ}}$ {EVENT > FACT}
b. EVENT PROTAGONISTS (so-called internally-headed relative clauses)[36]
Slave
[*lị gah hedéhfe*] *i* {EVENT > PROTAGONISTS} *gháyeyidá*
dog rabbit 3.chased NMLZR 1SG.saw
i. 'I saw the dog that chased the rabbit.'
ii. 'I saw the rabbit that the dog chased.'
c. RESULTANT PRODUCT (resultative nominalization)
Waiwai
[*a-mok-ɾɨ*] {EVENT > RESULT (sounds)} *w-enta*
2-come-AC.NMLZR I-hear+IMM.PST
'I heard you/your coming.'
d. PLACE
Mosetén (Sakel 2004: 95)
*chhiko'-ñi-ti-**dye'*** {EVENT > PLACE}
liquid-put-VD-NMLZR
'places where one washes oneself'

Metonymic meaning extension often works transitively such that one metonymically evoked concept leads to another closely related concept (cf. *the red, white, and blue* > {THE U.S. NATIONAL FLAG} > {U.S.A.} in *Team USA players …representing **the red, white and blue** at the 2014 FIL World Championship …*). It is assumed that event nominalizations of the kind seen above are similar in that the nominalization structures first evoke events portrayed by the relevant structures, which in turn evoke those concepts associated with events, as indicated above. Notice how so-called internally-headed RCs and the resultative nominalization receive a uniform treatment under our analysis, while the internally-headed RC analysis leaves the latter unaccounted for.

36. Our account does not explain why certain languages more readily allow this type of nominalization than others, where the so-called internally headed RCs do not obtain.

(5-70) Quechua

a. [*Maria wallpa-ta wayk'u-sqa-n*] {EVENT > PROTAGONIST}-ta
Maria chicken-ACC cook-NMLZR-3SG -ACC
mik"u-sayku
eat-PROG.1PL.EXCL
'We are eating {the chicken involved in the event of} Maria's cooking a chicken.'

b. [*Maria laranjas-ta ch'irwa-sqa-n*]$_{NMLZ}$ {EVENT > RESULT}-*ta*
Maria oranges-ACC squeeze-NMLZR-3SG -ACC
ujya-ni
drink-1SG
'I drink {the orange juice resulting from} Maria's squeezing oranges.'

As shown by such forms as *I know that John is honest*, *I heard John sing in the shower* and *I saw the man kiss my daughter*, English allows syntax-semantics mismatches too. What is remarkable about those languges permitting so-called internally-headed RCs is the extent to which similar mismatches are allowed, permitting the equivalents to **I recorded John sing in the shower*, **I drank Mary squeeze the oranges* and **I scolded the man kiss my daughter* and **The man kiss my daughter is a flirt.*

5.3 NP-use of event nominalizations: So-called "complement clauses"

A subject complement (5-4a) and an object complement (5-4b) are also known as verb complements, as opposed to noun complements in (5-5). The traditional name for noun complements is "content clause". These traditional descriptions are generally maintained by contemporary researchers as can be seen from Dixon's (2006) and Noonan's (2007) definitions below.

> In many languages certain verbs – notably 'see', 'hear', 'know', 'believe', 'like', and often also 'tell' and 'want' – can take a clause, instead of an NP (noun phrase), as a core argument. This is called a complement clause. (Dixon 2006: 1)

> By complementation, we mean the syntactic situation that arises when a notional sentence or predication is an argument of a predicate. For our purposes, a predication can be viewed as an argument of a predicate if it functions as the subject or object of that predicate. (Noonan 2007: 52)

As indicated by his remark "[l]anguages lacking a full range of complement clause constructions will often employ some kind of nominalization as a complement strategy", Dixon (2006: 37) recognizes a use of nominalizations as complements. The following show that languages of the Americas do indeed use nominalizations as verb complements.

(5-71) Tapiete (Ciccone 2008)
a-mbe'u-po [*yawa yɨ=ware handi awára-wa*]
1.SG.AC-tell-FUT tiger 3-play with fox-NMLZR
'I want to tell that the tiger played with the fox.'
or 'I want to tell about the tiger that played with the fox.'

(5-72) Barasano (Tucanoan; Jones & Jones (1991: 160))
[*wʉ-ri-ka* *ti* *eha-ro-ti-re*] *bãsi-be-a-ha* *yʉ*
fly-PTCPL-hollow 3in arrive-FUT-NMLZR~PROX-O know-NEG-PRS-~3 1SG
'I don't know when the plane is to arrive.' Or 'I don't know whether the plane will arrive or not.'

(5-73) Tariana (Arawakan; Aikhenvald 2009: 201)
wa-dalipa *phiſi* *di-nu-ɾi* *phema-ka-naka*
1PL-towards agouti 3SGNF-come-NMLZR IMP.hear-DEC-PRS.VIS
'One can hear an agouti come towards us.'

(5-74) Shoshoni (Northern Uto-Aztecan; Dayley 1989: 274)
Nüü [*kunai* *wayan**tün**na*] *punikka*.
I wood.O burn.NMLZR.O see
'I see the wood burning.'

(5-75) Thompson River Salish (Kroeber 1977: 381)
ƛ'uɪ x̫ʷox̫ʷst-m'n-xʷ [*k* *s-c'əq'ʷ-t-exʷ*]
NM want-RLT-2S.TS ART NMLZR-write-TRZ-2S.TS
'(and) you want to write it.'

A major issue here is whether what is known as clausal/sentential complements are structures that need to be recognized as distinct from what we have identified above as grammatical event nominalizations. Dixon's (2006) comment quoted above makes it clear that he considers verb complements to be clauses and believes that they are different from nominalizations. The real question here is whether the formal differences between "complement clauses", defined by Dixon as those showing a high degree of structural parallelism with clauses/sentences, and nominalizations make any substantial difference beyond the structural differences.[37]

37. See also the discussions by van Dijk, Haude & Muysken (2011: 3) in their introduction to the recent volume on subordination in South American languages, where these editors endorse Dixon's (2006) division between nominalizations and subordinate clauses by saying that "[n]ominalization in particular is a common strategy that South-American languages use in the same places where other languages have subordinate clauses." Notice the caution taken by Noonan (1985) in the earlier quote, where he says that complements obtain "when a *notional* sentence or predication is an argument of a predicate" rather than asserting that complements are clauses or sentences. (Emphastis added).

This is similar to asking whether the formal difference between the two lexical nominalizations (*a*) *speak**er*** and (*a*) *cook* makes any significant difference beyond the formal difference. If what Dixon considers to be complement clauses and nominalizations were two fundamentally different linguistic units, one would have to ask why they function in a similar way, both denoting events, facts and other abstract substantives, and both standing in a subject and object position that are typically occupied by nouns.[38] This is the crux of the issue that needs to be addressed. Our position is that all those that function as verb and noun complements are nominalizations, whether or not they contain discernible nominalizing morphology or they show structural resemblances to clauses/sentences.

A deeper issue lying beneath the distinction between "complement clauses" and "nominalizations" Dixon draws depends on one's understanding of what nominalization is and upon his analysis. Dixon's understanding of what nominalization is is stated as below:

> 'Nominalization' is used to describe a process (and its result) by which something with the properties of a nominal can be derived from a verb or adjective, or *from a complete clause*. (Dixon 2006: 36; emphasis added)

Unfortunately this understanding does not help us distinguish between what Dixon calls "complement clause" and a nominalization since the latter, according to him, may also have structural properties of a complete clause; and they indeed do (see (5-71) and (5-72) above) like what Dixon considers to be a "complement clause", e.g. *John's playing the national anthem* (*pleased Mary*). One wonders why this example is not a (grammatical) nominalization because the structure certainly has "the properties of a nominal" in denoting an activity (see Dixon 2006: 15) and in syntactically functioning as a subject like an ordinary noun.[39] The two examples Dixon gives in the section on nominalization strategy give the impression that he considers nominalizations those that have a nominalizing morphology, but then one wonders why the *-ing* suffix seen in the English "complement clause" example above is not nominalization morphology.

Dixon would consider the English *that*-construction in the translation of the Jamul Tiipay example below as a complement clause rather than a nominalization,

38. The same question must be posed for those who think that relative clauses are clauses distinguishing themselves from nominalizations (Comrie & Thompson 1985/2007); i.e. why do two fundamentally different structures function alike as noun modifiers if they were?

39. Dixon (2006: 15) discusses distinctions between this example and the expression *John's playing of the national anthem* in terms of the distinction between a complement clause and a nominalization, but the pertinent distinction is between a grammatical nominalization and a lexical nominalization.

perhaps following the Generative Grammar tradition that labels such *that* as a complementizer (COM), a practice followed by many other contemporary grammarians, whereas the original Jamul Tiipay form woud likely be considered a use of nominalization ("a nominalization strategy") because it has the nominalizing morphology *-ch*, which points to a realis state of affairs.

(5-76) Jamul Tiipay (Yuman, Miller 2001: 223)
[[*mi-iima-ch*]$_{\text{NMLZ}}$-pu]$_{\text{NP}}$ *uuyaaw*
2-dance-NMLZR-DEM know
'He knows that you dance.'

However, the early generative grammarian Robert B. Lees, following the more traditional analysis, analyzes *that*-construction in the English translation above as a nominalization, "a factive nominal", which he characterizes as an "abstract object" denoting "an abstract fact, or statement, or [as] information" (Lees 1963: 59ff). And to our mind, Lees is absolutely correct.[40]

Wisely, Miller (2001: Ch.7.2) treats all similar constructions as involving nominalization regardless of whether there is a nominalizing morphology, as in (5-75) above, or not, as in the example below.

(5-77) Jamul Tiipay (Miller 2001: 219)
[[*me-xap*]$_{\text{NMLZ}}$- pu]$_{\text{NP}}$ *nya'wach my-uuwiw*
2-enter DEM we.SJ ½-see.PL
'We saw you come in.'

While the fact that both forms above display an external property of being marked by a demonstrative is a good indication that they are nominals heading an NP, Miller's understanding of nominalization is functional and does not dependent upon morphology.[41]

40. Our major complaint about Lees' work on English nominalizations is that he considers the relevant nominalization structures to be clauses (see Section 6.2). Robert B. Lees was the first PhD in linguistics at MIT, where he trained with Morris Halle and Noam Chomsky. He became a leading generative grammarian in the 1960's, producing a number of influential linguists at the University of Illinois, Urbana-Champaign, including Ronald Langacker.

41. Dixon (2010: 316) also wants to distinguish between relative clauses and nominalizations, as is clear from his remarks: "In some languages, a verbal affix marking a relative clause is homophonous with a nominalizer. This should not be taken to mean that a relative clause is a type of nominalization". Why are certain relative clauses are homophonous with nominalizers if they are distinct constructions? Is it accidental?

5.4 NP-use of argument nominalizations: So-called "headless/free relative clauses"

Argument nominalizations in NP-use, as in (5-78) and (5-79) below, are generally known as "headless relative clauses" or "free-relatives" as if they are a kind of, or a derivative of, relative clauses. As in the case of so-called internally-headed relative clauses, it is not obvious that the relevant structures function as a restrictive or non-restrictive modifier. There is no head noun that sets a domain of denotation which is restricted to a subset or which is identified and commented on by the denotation of the relevant structure.

(5-78) English
You should marry [[*who Ø*^{110} *loves you*]$_{\text{NMLZ}}$^{110}]]$_{\text{NP}}$, *not* [[*who Ø*^{125} *loves your money*]$_{\text{NMLZ}}$^{125}]$_{\text{NP}}$

(5-79) Navajo (Southern Athabaskan; Willie 1989: 435)
[[*Ø*^{15} *'at'ééd yizts'ǫs-yéé*]$_{\text{NMLZ}}$^{15}]$_{\text{NP}}$ *yałtí'*
girl 3sO: 3sS:kiss-NMLZR speaking
'The one who kissed the girl is speaking.'

Actually, there has been little argument for considering these as a type of relative clauses. Most simply assume that to be the case because the same structures are also used as externally-headed relative clauses. This is like putting the cart before the horse to our mind. But there have been, in fact, some proposals to make so-called headless relative clauses align with externally-headed relative clauses. One is a deletion analysis that posits an external head noun, which is then deleted. This is the most popular analysis of headless relative clauses that has been applied to a wide range of languages (e.g. Weber (1989) on Quechua, Huang (2008) on Qiang, Treis (2008) on Kambaata). Another is to posit some kind of phonetically empty pronoun as the head, as proposed by Jelinek (1987).

Besides the fact that the same structure is typically used as ordinary relative clauses, a basic motivation for these analyses appear to be the semantic properties associated with the relevant structures, namely they denote entity concepts. Apparently it has never occurred to the proponents of these analyses that so-called relative clauses *themselves* (e.g. English forms [who loves you], [who loves your money] or Navajo structure ['at'ééd yizts'ǫs-yéé] 'one who kissed the girl' above) have entity denotations associated with them, in the manner described in Section 4, for example. This is largely because these structures had been thought to be clauses or sentences, which certainly do not denote thing-like concepts as ordinary nouns do (see Section 6.2 below). By positing a head noun or a pronoun, the proposed analyses capture the fact that the relevant structures are associated with entity denotations.

The problem with these proposals is that neither the deletion analysis nor the pro analysis is a complete description until they provide an account as to when the putative deletion of the posited head noun applies or how the pro head is actually distributed and interpreted. In offering a complete description, these analyses must refer to context, similar to our analysis. Our argument is that if one has to refer to the context anyway, let the context handle the whole thing. Our analysis maintains that so-called headless relative clauses are grammatical argument nominalizations and that they themselves, not a putative head, have a set of entity concepts as their denotations as described in Section 4. The context of use then determines the referent most appropriate per the Gricean Cooperative Principle. The deletion and pro analyses have been conceived because the proponents of these analyses lacked proper understandings of the nature of grammatical nominalizations and because of their usual practice of analyzing linguistic structures in complete isolation from the context in which they are used.

There are, furthermore, cases in which so-called headless relative clauses (our NP-use of nominalizations) obtain without possible externally-headed RC counterparts. In Russian nominalizations marked by the interrogative pronouns *kotoroyj* 'which' and *chto* 'what' allow both NP-use and modification-use as shown in the examples below.

(5-80) Russian

a. *tot,* [*kotoryj/chto stoit tam*], *eto drug otca*
that WHICH/WHAT stand there it friend father
'The one who is standing there is my father's friend.'

b. *tot chelovek,* [*kotoryj/chto stoit tam*], *eto drug otca*
that man WHICH/WHAT stand there it friend father
'That man who is standing there is my father's friend.'

However, the nominalizations marked by *kto* 'who' allows only NP-use, as indicated by the ill-formed sentence in (5-81b) below.

(5-81) a. (*tot,*) [*kto vymyl ruki*], *mozhet nachatj jestj*
(that) WHO[42] washed hands can start eat
'The one who has washed his hands can start eating.'

b. **Maljchik,* [*kto vymyl ruki*], *mozhet nachatj jestj*
boy WHO washed hands can start eating
'The boy who has washed his hands can start eating.'

42. The glosses WHO, WHAT, etc. are employed to show the etymologies for these nominalizers.

In German whether or not nominalizations marked by *was* 'what' has a modification-use is dialectal. While High German is said to disallow a form like (5-82b) below, Bavarian German or some other southern German dialects may allow it.

(5-82) German
a. *Ich lese* (*das*), [*was Sie empfehlen*]
I read ART WHAT you recommend
'I'll read what you recommend.'
b. %*Ich lese das Buch*, [*was Sie empfehlen*]
I read ART book WHAT you recommend
'I'll read the book which you recommend.'

On the other hand, in the case of those marked by *wer* 'who', a modification-use appears to be generally prohibited. Observe:

(5-83) a. *Ich empfange,* [*wer* (*auch*) *morgen kommt*]
I receive WHO (also) tomorrow comes
'I receive who(ever) comes in tomorrow.'
b. **Ich empfange den Mann*, [*wer morgen kommt*]
I receive ART man WHO tomorrow comes
'I receive the man who comes tomorrow.'

In Spanish *que*-marked nominalizations generally allow both NP-use and modification-use, as in (5-84) below, but those marked by *quien* lack a modification-use.

(5-84) Spanish
a. *Leeré lo* [*que usted recomienda*]
I.will.read ART WHAT you recommend
'I will read what you recommend.'
b. *Leeré el libro* [*que usted recomienda*]
I.will.read ART book WHAT you recommend
'I will read the book which you recommend.'

(5-85) a. *Veré a* [*quien viene mañana*]
I.will.meet to WHO comes tomorrow
'I will meet the one who comes tomorrow.'
b. **Veré al hombre* [*quien viene mañana*]
I.will.meet to.the man WHO comes tomorrow
'I will meet the man who comes tomorrow.'

Thus, all in all, there is little motivation or evidence for analyzing the NP-use of argument nominalizations as (headless) relative clauses. The formal resemblances between so-called headless relative clauses and ordinary externally-headed relative is due to the fact that they represent two different uses of the same nominalization

structures. The fact that so-called headless RCs denote and refer to entity concepts is not due to an external head that is doomed to be deleted or due to a phonetically empty pronoun; rather, it is because grammatical argument nominalizations underlying them themselves have such denotations allowing their referential use in actual discourse (see Section 4).

5.5 Modification-use of argument nominalizations: So-called "relative clauses"

As shown in (5-7), our analysis of relative clause constructions is straightforward. While it captures the observations on the use of nominalizations as relative clauses, it departs from the traditional analysis in several ways. A major claim advanced in this paper is that so-called relative clauses are all nominalizations, whether they contain finite verb forms or they contain specific nominalization morphology. As in the case of verb complements discussed in Section 3.1, those who believe that argument nominalizations and relative clauses are separate structures, which "in certain languages [are] … indistinct" (Comrie & Thompson 1985/2007: 379), must answer (i) why structurally they both have a gap (or a pronoun) in an argument or adjunct position (see Section 6.1), and (ii) why functionally they both modify a noun.

The traditional analysis of RC constructions based on English makes crucial reference to the role of so-called relative pronouns, such as *who* and *which*, that play the double role of indicating the dependency relation between the pronoun and a gap within an RC and of holding the perceived anaphoric relation with a head noun, giving rise to the term "relative pronoun", as in the following representation.

(5-86) *the man*$_i$ [*whom*$_i$ [*you love Ø*$_i$]]

Such an analysis is problematic when applied to other languages in that many, if not most, languages do not use anything like relative pronouns.[43] Most descriptions of RC constructions in a variety of languages label an element marking what looks like an RC as REL or as a relative pronoun. This practice, however, has not been independently justified in most of such descriptions; they simply follow the analysis of (5-86) based on English. To our mind, they are best analyzed as nominalizers, as indicated by our relabeling of them in the examples cited in this paper. Even in English, we can advance an argument for treating *who, which*, etc. as indefinite pronouns used as nominalizers or markers of nominalization (in addition

43. Because of this, some grammarians say that their languages do not have relative clauses (e.g. Jones & Jones 1991: 149). A more accurate way of saying this is that there are no *English-style* relative clause constructions in the relevant languages.

to their use as interrogative pronouns). The use of indefinite pronouns as nominalizers makes good sense because what argument nominalizations denote may be indefinite (e.g. [*Who gets there first*] *gets the prize; You may choose* [*which you find most appealing*]).

Our point is that relativization does not depend on so-called relative pronouns and that the perceived relation between so-called relative pronouns and gaps in RCs can be captured in terms of the role-marking morphology discussed in an earlier section. For example, the German nominalizers *der* and *den*, for example, mark the grammatical role of the entity denoted by an argument nominalization; they are a subject nominalizer and an object nominalizer that mark (or combine with) a subject nominalization and an object nominalization, respectively, as below.

(5-87) German

a. *Ich treffe* [*den Mann,* [***der*** [*Ø morgen kommt*]$_{\text{SUB.NMLZ}}$]$_{\text{NP}}$
I meet ART man SUB.NMLZR morning comes
'I meet the man who comes tomorrow.'

b. *Ich treffe den,* [***den*** [*du mir Ø vorgestellt hast*]$_{\text{DO.NMLZ}}$]
I meet ART DO.NMLZR you me introduce have
'I meet the to one whom you introduced to me.'

A requirement in languages with role-marking nominalization morphology like German is that the morphology correctly indicates the type of argument nominalization involved. That is, a subject nominalizer must combine with a subject argument nominalization with a gap in subject position, as in (5-87a), and an object nominalizer with an object argument nominalization with a gap in object position, as in (5-87b). The role of the English nominalizer *whom* is exactly the same as that of the German DO nominalizer *den* (except for the additional gender information coded in the latter). All languages with role-indicating morphology examined in Section 3.2 have similar requirements (see Section 9.1 for an important implication of these points in the analysis of relative clause constructions).

While many languages are similar to German in having role-indicating morphology, there are many others that do not; accordingly, such morphology, like so-called relative pronouns, is not an essential feature of RC constructions in general, as can be seen in the Japanese and the Toba pattern below.

(5-88) Japanese

a. [[*Ø hon o yomu*]$_{\text{SUB.NMLZ}}$ *kodomo*]$_{\text{NP}}$
book ACC read.PRS child
'a child who reads a book'

b. [[*kodomo ga Ø yomu*]$_{\text{OBJ.NMLZ}}$ *hon*]$_{\text{NP}}$
child NOM read.PRS book
'a book which a child reads'

(5-89) Toba (Messineo & Porta 2009: 57)[44]

a. *so* [*ʃijaGawa* [*Ø i-waGan a-so qaʔaɲole*]$_{\text{SUB.NMLZ}}$]$_{\text{NP}}$
DD man 3A-hit FEM.DD young.lady
'the man who hit the young lady'

b. so [*ʃijaGawa* [*a-so qaʔaɲole i-waGan Ø*]$_{\text{OBJ.NMLZ}}$]$_{\text{NP}}$
DD man FEM.DD young.lady 3A-hit
'the man whom the young lady hit'

The examples of RC constructions above indicate that a minimal requirement, the essential feature of RC constructions, is that they involve as a modifier an argument nominalization with a gap (or a pronoun as in Thai, Modern Hebrew and some others) in an argument (or an adjunct) position. The relevant argument nominalizations may or may not involve morphology indicating the grammatical role of the entity denoted by the nominalization. These considerations suggest the following analysis of RC constructions.

(5-90) a. Japanese
[[*Ø*^{i} *hon o yomu*]$^{i}_{\text{SUB.NMLZ}}$ *kodomo*]$_{\text{NP}}$
book ACC read.PRS child
'a child who reads a book'

b. Chinese
[[*Ø*^{i} *zài nàr diào lǐyú*]$^{i}_{\text{SUB.NMLZ}}$ *= de*] *háizi*]$_{\text{NP}}$
PROG there fish carp = NMLZR child
'a child who is fishing carp there'

c. German
[*der Mann* [[*der* [*Ø*^{i} *dich liebt*]$^{i}_{\text{SUB.NMLZ}}$]]$_{\text{NP}}$
the man SUB.NMLZR you love
'the man who loves you'

d. English
[*the man* [*whom* [*you love Ø*^{i}] $^{i}_{\text{OBJ.NMLZ}}$]]$_{\text{NP}}$
OBJ.NMLZR

e. Thai
yaa [*thîi* ***man***^{i} *mòtʔaayúʔ*]$^{i}_{\text{SUB.NMLZ}}$
medicine NMLZR 3SG expire
'the medicine which has expired'

The above analysis embodies the idea that nouns have a denotation index in the manner of [dog]^{i} that points to a set of concepts that they denote. Nominalizations as nominals share this property, as indicated above. In the case of argument

44. Toba has the non-role marking nominalization marker *maʒi*, which can be used at the beginning of an argument nominalization.

nominalizations, these denotation indices bind a variable in the form of a zero or pronominal argument, thereby indicating the grammatical role that the entity denoted by the nominalization structure stands for. In (5-90a), (b), (c), (e) above, the nominalizations denote an entity that stands for the subject role. In (5-90d) the nominalization denotes an entity that stands for the object role.

Notice that in our analysis there is no role that so-called relative pronouns play with regard to the gap in the nominalization or with regard to the head noun. This is a desired consequence of our analysis, which analyzes a restrictive relative clause construction as involving two independent nominals, each with its own denotation set. Restricting the denotation of the head noun means specifying its subset by the denotation of the modifying argument nominalization. Thus, the only requirement for the modifying nominalization with respect to the head noun in an RC construction is that the former denotes entities that intersect with those denoted by the head noun, as in Figure 3 below. Our analysis is highly compatible with the treatment of a restrictive relative clause construction in Formal Semantics, which would define the denotation of such a construction as the intersection of two sets of entities; e.g. {x| x is a man} ∩ {x| you love x} ("the intersection of the set of all x such that x is a man and the set of all x such that you love x"), where x's are two independent variables.[45]

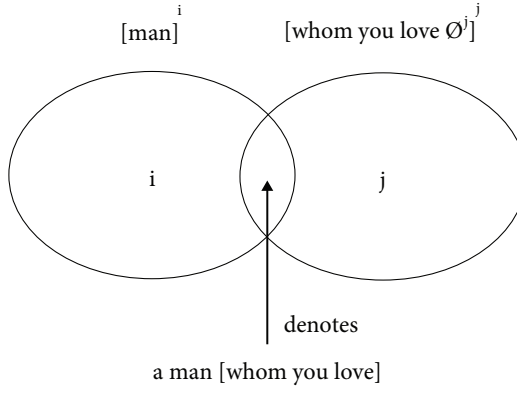


Figure 3. Denotation of restrictive relative clause

As is clear from the exposition above, so-called subject relative clause construction is simply a combination of a head noun and a subject argument nominalization with a gap (or rarely a pronoun) in subject position (5-90c), (e), and so-called

45. The Formal Semantic analysis would have a difficult time in deriving the second set for the modification involving event nominalizations without a gap (see Section 3.1), which would not yield to an analysis calling for operator movement, as in the case of the generative analysis of wh-relatives in English.

object RC construction a combination of a noun head and an object argument nominalization with a gap or a pronoun in object position (5-90d). The relativization process per se involves no movement or deletion of a pronoun or a noun coreferential with the head noun. It simply brings together, or merges, a head noun and a grammatical argument nominalization that restricts the denotation of the head noun in the manner shown in Figure 3.

In addition to the distinction between the restrictive and the non-restrictive relative clause constructions, there are relevant constructions that deviate from the canonical RC constructions discussed above. First, many languages of the world do not seem to make a clear distinction between restrictive and non-restrictive RCs as in English, where the latter is set off by a brief pause in speech and a comma in writing. Nevertheless, there is a functional difference between the two even if the two constructions are formally alike. Instead of restricting the denotation of the head noun to its subset, the denotation of a modifying argument nominalization in non-restrictive RC constructions identifies that of the head noun either under strict identity, that the two denotations denote an identical entity, or as an instance of of the latter. For example, in the construction *the man, **who you have decided to marry against my advice**,* (*is a real crook*), the modifying argument nominalization likely denotes a single-member set. In such a case the total identity obtains between the denotation of a single-member set of the head noun and that of a single-set member of the modifying argument nominalization. On the other hand, in a construction like *the man*, ***who you love dearly***, (*is a real crook*), it is likely that the denotation set of the modifying nominalization contains multiple entities, denoting all those that the addressee loves dearly. In this case, the denotation of the modifying nominalization identifies that of the head noun as an instance; namely, that the denotation of the head noun *man* is an instance of those that the addressee loves dearly. In both these cases, the modifying nominalizations identify the denotation of the head noun in terms of alternative and more elaborate ways of determining it, resulting in the commonly-held observation that a non-restrictive relative clause provides additional information about the denotation of the head noun.

The canonical RC constructions form a noun phrase with a head nominal and a modifying argument nominalization. Many languages, including English, however, allow the modifying nominalization to be separated from the NP containing the head noun, as in the examples below (see Fleck, this volume, for more examples).

(5-91) Tapiete (Gonzaléz 2005: 232)
hama angu'a nohe-ha [*wɨ-ro-po-ha-p ɨ-wa*] *hoka*
then drum pull out-IMPER 3-COM-dance-IMPER-INST-NMLZR DEM
ha'e (cf. 5-9)
(s)he
'then they take out the drum to dance, that is', (lit.) '…the drum with which they dance …'

(5-92) Harakmbut (Van linden this volume)
arakmbut-ta i'-uk-i [*henpu wa-mba-baeri*]*-ta*
person-ACC 1SG-search-1.IND string.bag NMLZR-VRL-make-ANIM-ACC
'I am looking for the person who makes string bags.'

These are reminiscent of what Hale (1976) refers to as the adjoined relative clause in Australian languages, although it is not clear whether or not these and similar constructions in other parts of the world are "typically …separated from the main clause by a pause" (Hale 1976: 78).

Relevant to the interpretations of these "adjointed relative clauses" is the question of whether or not the modifying nominalizations themselves form a noun phrase and hence are referential. Our analysis of the canonical RC constructions, as shown in (5-90) above, juxtaposes the head noun and the modifying argument nominalization similar to an appositive construction of the type, *John, the butcher*. However, juxtaposition by itself does not mean that the two nominal constituents are in the appositive relation seen in the type given here that involves identification of the referent of one NP by another referential NP. Our RC analysis maintains that the modifying argument nominalization is not a noun phrase and hence, while it denotes an entity, it does not refer to a discourse entity, just as nouns in a noun compound such as $[[goat]_N\ [cheese]_N]_N$ do not refer individually (or even collectively as a compound noun) – the entire noun phrase containing the compound, as in *We ate* $[[[goat]_N\ [cheese]_N]_N]_{NP}$ refers. Likewise, the nominalization structure identified as $[\ldots]_{NMLZ}$ in this paper does not refer by itself; it only refers when it heads an NP.[46]

The discussion above is relevant in the interpretations of the two similar constructions in Spanish below.

46. Our field has seen many loose uses of the term "referential/referring expression" in reference to nouns and nominalizations (Croft 1991, Shibatani 2009, Cristofaro this volume) that ignore an important distinction between denotation and reference (see Section 4 on this distinction).

(5-93) Spanish

a. [*El hombre* [*que está bailando*]$_{\text{NMLZR}}$]$_{\text{NP}}$ *es mi tío.*
the man NMLR is dancing is my uncle
'The man who is dancing is my uncle.'

b. [*El hombre*]$_{\text{NP}}$ [*el* [*que está bailando*]$_{\text{NMLZR}}$]$_{\text{NP}}$ *es mi tío.*[47]
the man the NMLR is dancing is my uncle
'The man, the one who is dancing, is my uncle.'

In (5-93b), the nominalization marked by the definite article forms a noun phrase and is therefore referential, while the counterpart in (5-93a) does not form an NP and does not refer.

Indeed, a contrast similar to that observed in the Spanish examples above is seen in a number of languages that have a special marker for an NP-use of grammatical nominalizations like the Spanish articles, e.g. *el* in (5-93b). Toba has demonstrative determiners that mark an NP-use of grammatical nominalizations (see Section 6). These demonstrative determiners, like the Spanish articles, do not mark nominalizations when they are in modification-use, as seen in (5-93a) above and (5-94a) below. Determiner marking in these languages, therefore, signals an NP-use of nominalizations, where they are playing a referential function. Observe the parallel pattern between (5-93) and (5-94).

(5-94) Toba (courtesy of Cristina Messineo)

a. *ajem si-kjen* [*so* [[*ʃijaGawa*] [*Ø-neta-ge da cako*]$_{\text{NMLZ}}$]$_{\text{N'}}$]$_{\text{NP}}$
I 1A-greet DD man 3-be-DIR DD Chaco
'I greeted the man who lives in Chaco.'

b. *ajem si-kjen* [[*so* [*ʃijaGawa*]]$_{\text{NP}}$ [***so*** [*Ø-neta-ge da cako*]$_{\text{NMLZ}}$]$_{\text{NP}}$
I 1A-greet DD man DD 3-be-DIR DD Chaco
'I greeted the man, the one who lives in Chaco.'

It is this kind of contrast that must be investigated in order to understand the true nature of the "adjoined" type seen in (5-91) and (5-92), or even the ordinary juxtaposed one, for that matter.[48]

It is by now clear that the traditional term "relative clause" is a misnormer since what it designates is an argument nominalization in modification-use. The term can now be understood as a label for it. Such a label, however, is misleading since it suggests that nominalizations are clauses. We show below that there

47. As in the simple *John, the butcher* type, the constituency of the two juxtaposed NP's is not clear; i.e. whether or not they form a larger NP constituent.

48. In Japanese, the referential appositive nominalization occurs after the head noun (e.g., *kono hon,* [[*boku ga kinoo katta*] *no*], 'this book, the one that I bought yesterday'), while the ordinary RC is prenominal ([*boku ga kinoo kata*] *kono hon* 'this book that I bought yesterday').

is ample crosslinguistic evidence that nominalizations are not clauses, let alone sentences. But first a clearer understanding is in order of what sentences, clauses, and nominalizations are.

6. Sentences, clauses, and nominalizations

An early recognition of the use of grammatical nominalizations as noun modifiers is found in the context of Japanese grammar in Yoshio Yamada's monumental grammar, *Nihonbunpōron* (Theory of Japanese Grammar) published in 1908. More recently, Matisoff (1972) recognized a connection between nominalizations and relative clauses in the Tibeto-Burman language Lahu. He also mentions similar morphological connections among nominalizations, relative clauses, and possessive constructions in Lahu, Mandarin Chinese, and Japanese (see Section 7). Matisoff's observation has been followed by others specializing in Tibeto-Burman languages, such as DeLancey (1986) and Noonan (1997), and more recently by DeLancey (2002) and Noonan (2008).

Many other recent studies on nominalizations and relative clauses such as those contained in Yap, Grunow-Hårsta & Wrona (2011) and Comrie & Estrada-Fernández (2012) clearly recognize the use of nominalizations as relative clauses, but for some unclear reason and without any justification they continue to use the term "relative clause", (i) as if some structures identifiable as relative clauses exist apart from argument nominalizations, but "in certain languages relativization is indistinct from nominalization" (Comrie & Thompson 1985/2007: 379) or (ii) as if nominalizations somehow turn into clauses under modification-use. Yamada (1908: 1462) simply states that we may call a grammatical nominalization used for noun modification an "adnominal *clause*" without offering the reason for it.

6.1 Tense and nominalization

Without clear definitions of clauses and sentences (and nominalizations, for that matter!) on the part of those who believe that relative clauses (our grammatical argument nominalizations) are clauses, it is difficult to know the true rationale underlying their belief about the clausehood of RCs. However, one observation that has been made is that RCs in some languages may stand as sentences; hence RCs are sentences that have been made dependent clauses by embedding them.[49]

49. Rice (1989: 25), in her otherwise excellent grammar, tells us that "[a] relative clause is a sentence that modifies a noun". Compare this with Nevis, Pesetsky & Rodrigues's (2009: 366) characterization of event nominalizations: "a verb may merge with *a sentence*, as in *Mary thinks*

Such a possibility arises when the predicate in an RC has a tensed verb or more broadly a finite verbal form associated with sentences. For example, Comrie and Horie (1995: 68) tell us that "[w]hat precedes the head noun [as in (6-1a) below, for example,] is a well-formed sentence in its own right", as can be seen from the fact that it can stand as a sentence; see (6-1b).

(6-1) Japanese
a. [$Ø_1$ *kinoo* $Ø_2$ *katta*] *hon*
yesterday bought book
'the book that (I) bought yesterday'
b. $Ø_1$ *kinoo* $Ø_2$ *katta.*
yesterday bought
'(I) bought (it) yesterday.'
(As an answer to the question "When did you buy the book?")

Comrie and Horie are equating the gaps found in the RC in (6-1a) with the anaphoric gaps found in sentence (6-1b). This, however, is a mistake. Anaphoric gaps can be filled by full noun phrases, albeit perhaps redundantly, but one of the gaps in the RC/argument nominalization in (6-1a) cannot. Compare (6-1) with (6-2) below:

(6-2) a. [$Ø_1$/ ***boku*** *ga* *kinoo* $Ø_2$/ ****sono*** ***hon*** *o* *katta*] *hon*
I NOM yesterday that book ACC bought book
'the book [that I bought Ø /*that book]'
b. $Ø_1$/ ***Boku*** *wa* *kinoo* $Ø_2$/ ***sono*** ***hon*** *o* *katta.*
I TOP yesterday that book ACC bought
'I bought that book yesterday.'

(6-2a) is as bad as its English translation with the full noun phrase in object position. In other words, the two gaps in RC (6-1a) are different from the two gaps in sentence (6-1b). Object argument nominalizations must have an obligatory gap ($Ø_2$) in object position in both English and Japanese, while the latter may contain an anaphoric gap in other positions.[50] Clauses and sentences, on the other hand, have no such constraint. Argument nominalizations are thus different from clauses and sentences in both English and Japanese.[51]

[*that the world is round*]" and "…a noun can merge with *a sentence*, as it does in (*the*) *claim* [*that the world is round*],…" (Emphasis added).

50. Japanese, as in some other languages, allow a resumptive pronoun in a position lower in the grammatical relation hierarchy.

51. Comrie and Horie (1995) recognizes this difference in footnote 5, page 75, but does not deal with this most crucial issue in comparing the structures of RCs/argument nominalizations and

A similar conclusion obtains with the Mayan language K'ichee' spoken in Guatemala, whose argument nominalizations modifying a noun qua RCs contain a finite verb form and appear to be able to stand as sentences, as shown in (6-3b) below.

(6-3) K'ichee' (Larson & Norman 1979: 357; the grammaticality judgement courtesy of Telma Can Pixabaj)

a. *lee ixoq lee [x-Ø-u-ch'ay lee achih]*
the woman NMLZR ASP-3SG.ABS-3SG.ERG-hit the man
'the woman whom the man hit/the woman who hit the man'

b. *x-Ø-u-ch'ay lee achih*
ASP-3SG.ABS-3SG.ERG-hit the man
'S/he hit the man/The man hit him/her.'

However, just like the case of Japanese above, nominalizations qua RCs are different from sentences. The latter can have a full array of arguments appearing as full noun phrases, while the former must contain a gap. Observe:

(6-4) a. *lee ixoq lee [x-Ø-u-ch'ay*
the woman NMLZR ASP-3SG.ABS-3SG.ERG-hit
***lee ixoq**/Ø lee achih]*
the woman/Ø the man
'the woman whom [the man hit *the woman/Ø]' or 'the woman who [*the woman/Ø hit the man]'

b. *x-Ø-u-ch'ay **lee ixoq** lee achih*
ASP-3SG.ABS-3SG.ERG-hit the woman the man
'The woman hit the man/the man hit the woman.'

The difference between argument nominalizations qua RCs and clauses/sentences seen here also obtains in those languages that may contain a pronoun instead of a gap in argument nominalizations/RCs. Thus, the pronoun in subject position of a subject nominalization cannot be replaced by a full noun, as shown in (6-5b) below, which is just as bad as its English translation.

(6-5) Thai

a. *thəə mây khuan kin yaa [thîi **man** mòtʔaayúʔ]*
2SG not should eat medicine NMLZR 3SG expire
(lit.) 'You should not take the medicine which **it** has expired.'

b. **thəə mây khuan kin yaa [thîi **yaa** mòtʔaayúʔ]*
2SG not should eat medicine NMLZR medicine expire
(lit.) '*You should not take the medicine [which **the medicine** has expired].'

(complement) clauses/sentences. Other related papers by Comrie (Comrie 1996 and 1998a, b) repeat similar views about Japanese and other languages without even mentioning this issue.

Similar examples can be adduced from a diverse array of languages whose argument nominalizations contain verbal forms similar to those occurring in sentences, with tense and other finite features or without any of them, as in isolating Asian languages like Thai above.

The reluctance to recognize these nominalizations as such is rooted in the fact that they may contain formal finite features such as tense, aspect, and person marking, characteristics of sentences as in the examples above. However, there is nothing that prevents nominalizations from having these features since the information they carry can be highly valuable in distinguishing types of entities they denote. For example, *what*-nominalizations in English make a crucial difference in what they denote depending on the tense information they contain; e.g. *what I was buying* vs. *what I am buying;* also cf. *John's purchasing of a house last year* vs. *John's purchasing of a house next year*. Indeed, in many languages nominalizing morphology itself may incorporate tense information (and/or aspectual, as well as evidential values in some languages), as shown in (6-6)–(6-8), or may allow a separate tense expression within nominalized structures, as in the Oceanic examples given in (6-9) and (6-10) below.[52]

(6-6) Korean

a. [*cikum pap-ul mek-**nun***] *kes*
now meal-ACC eat-PRS.NMLZR NPM
'one who is eating a meal now'
(cf. [[*cikum pap-ul mek-**nun***] *ai*] 'a child who is eating a meal now')

b. [*ecey pap-ul mek-**un***] *kes*
yesterday meal-ACC eat-PST.NMLZR NPM
'one who ate a meal yesterday'

c. [*pap-ul mek-**ul***] *kes*
meal-ACC eat-FUT.NMLZR NPM
'one who will eat a meal'

(6-7) Hixkaryana (Derbyshire 1999: 48–49)

a. Event/Action nominalizer-Past tense: *-th~~ɨ~~r~~ɨ~~*
i-wanota-th~~ɨ~~r~~ɨ~~ komo
3-sing-AC.NMLZR COL
'their singing (in the past)'

52. See Fleck (this volume), Peña (this volume), Valle & Zariquiey (this volume), and Jones & Jones (1991) for additional cases of time-bound nominalization morphology in Matses, Wampis, and Barasano. Also compare these cases with so-called present and past participles in English forms, *a breaking chair/a broken chair*.

b. Nominalizer of the S (Protagonist of intransitive event)/O (Patientive protagonist of transitive event)-Past tense: *-saho*
i̶-manho-saho *uɾo/omoɾo/moki*
IMPERS-dance-S.NMLZR 1/2/3PRO
'I (am)/you (are)/he (is) the one who danced.'

(6-8) Imbabura Quechua (Cole 1982)
a. [*Marya Ø riku-shka*[53]] *runa*
Maria see-PST.NMLZR man
'the man whom Maria saw'
b. [*Marya Ø riru-k*] *runa*
Maria see-PRS.NMLZR man
'the man whom Maria sees'
c. [*Marya Ø riku-na*] *runa*
Maria see-FUT.NMLZR man
'the man whom Maria will see'

(6-9) Xârâcùù (Oceanic; Moyse-Faurie 2016: 182)
È kê pwî [*êê-mwata* ***na*** *rê anyââ*]
3.SG eat.tubes banana NMLZR-grate PST POSS mommy
'He is eating bananas which have been chopped up by his mother.'

(6-10) Marquesan (Oceanic; Moyse-Faurie 2016)
[*Te* ***i*** *ite-tina na tunane tata eka te tihe te kui*
SPEC PST see-NMLZR PAUC brother nearly reach SPEC come SPEC mother
i una], *atahi kokoti na tunane te ouoho no Hina*.
LOC top then cut PAUC brothers SPEC hair POSS Hina
'When the brother saw that the mother had nearly reached the top, (then) they cut Hina's hair.'

Tense is intimately connected with a sentence because the latter asserts the truth of a predication made by a clause as obtaining at a specific time. Since nominalizations do not perform this kind of illocutionary function, they typically lack an expression of tense. On the other hand, tense indication one way or another adds some vital information about what is denoted by nominalizations. There is thus no need to assume that nominalizations cannot be marked for tense, and we should not uncritically assume that tense-marked structures are clauses or sentences.

53. The Imbabura *-shka* corresponds to the object argument nominalizer/event nominalizer *-sqa*, and *-k* to the subject argument nominalizer *-q* in Bolivian Quechua. The connection between object argument nominalizer and past tense is seen elsewhere; e.g. Kakataibo (Zariquiey 2011).

6.2 Functional definitions of clauses, sentences, and nominalizations

Those who believe that argument nominalizations (used in RC constructions) and other types of grammatical nominalizations are clauses/sentences are victims of the formal orientation in linguistics that attempts to characterize the nature of grammatical constructions in terms of formal properties such as internal structural properties. As seen above, grammatical nominalizations (partially) share internal structures with clauses and sentences. But these structure-internal formal properties are like the skeletal structures that a roast turkey shares with a live one. Just as a roast turkey and a live bird are functionally very different and are accordingly treated differently (e.g. we do not keep them in the same cage), grammatical constructions such as clauses, sentences, and nominalizations must similarly be defined functionally and distinguished according to the functions they perform. Shibatani (2017, 2018a) offers the functional definitions of clauses, sentences, and nominalizations along the following lines, in terms of different kinds of speech act they perform:

- **Clauses predicate**: By uttering a structure like [John is honest] as a clause, a speaker ascribes the relational property denoted by a predicate phrase to the referent of the subject noun phrase.
- **Sentences perform illocutionary acts**: By uttering a structure like [John is honest] as a declarative sentence, a speaker asserts the truth of the predication made by the clause of the same structure. By uttering a structure like [Is John honest] as an interrogative sentence, a speaker asks whether the predication made by the clause [John is honest] is true or not.
- **Nominazations denote** (things and thing-like entity concepts): By uttering a structure like [John is honest] as a nominalization, a speaker evokes and establishes a form-meaning connection between the structure and a metonymically motivated meaning, e.g. a fact, related to a state-of-affairs (or broadly an event).

In other words, a clause is a grammatical structure associated with a speech act of ascribing a verbal property to the referent of a subject nominal. Sentences, on the other hand, perform different kinds of speech act, namely illocutionary acts, such as asserting that the predication made by a clause is true (declarative sentences), questioning whether or not the predication is true (yes-no questions), ordering (imperative sentences), warning, promising, etc. Notice that predication and assertion are two distinct types of speech act, as clearly shown in the case of yes-no questions. Nominalization structures neither predicate nor assert; they instead denote. That is, by uttering a noun or nominalization, a speaker evokes a mental connection between a string of sounds with an entity concept. Many such mental connections are permanent, as in the case of nouns and lexical nominalizations,

while grammatical nominalizations evoke nonce connections between the forms and concepts.

Besides the structural properties that nominalizations may share with clauses and sentences, as in our examples here, meaning similarities among them is likely another reason that many believe that nominalizations are clauses or sentences. For one thing, nominalizations share meaning similarities with clauses and sentences in terms of presuppositions that they are associated with. The event nominalization [(that) John recklessly shoots tresppassers] in a sentence like *Bill knows* (*that*) *John recklessly shoots trespassers* presupposes the proposition "John recklessly shoots trespassers". Likewise, the argument nominalization [who John shot] qua a relative clause as in *I knew the man who John shot* presupposes the proposition "John shot someone/something". How one arrives at these presuppositions from the nominalization structures is an interesting and difficult question. One should not presume that meaning resides in structures in a straightforward manner. A presuppositional proposition can also be derived from a sentence that does not contain a clausal structure corresponding to the form of a presupposition in question; e.g. *John is a good cook* presupposes "John cooks"; *The nurse attended the crying baby* presupposes "The baby was crying".

Perhaps a more fundamental issue here is a distinction between linguistic meaning and propositional meaning, the latter of which obtains when a structure (a linguistic form, a series of speech sounds) is used in performing speech acts of predicatation and assertion. A structure can have linguistic meaning apart from these speech acts, as in the case of those metalinguistic statements used as examples in linguistics discussions, such as *John is honest*, *John sent a Christmas gift to his grandmother*, which have full and coherent meanings similar to actual clauses and sentences, while they lack a referential property and truth value, i.e. propositional meanings of clauses and sentences. In short, a nominalization may share meaning similarities with a clause/sentence at the level of linguistic meaning, however such a thing is arrived at.[54] And it is at this level that the internal structure of nominalizations plays an important role in determing their meanings and their similarity to those of clauses/setentences.

The event nominalization of the form [that [John drank beer excessively]] accordingly shows a high resemblance in linguistic meaning to the clause/sentence *John drank beer excessively*, whereas those that show only partial structural similaries, such as [John's drinking beer excessively], [drinking beer excessively], [drinking

54. We have not made much progress on this from the early phase of Generative Grammar, where an attempt was made to arrive at linguistic meaning by an algorithm that compositionally built up the meaning on the basis of phrase structrues/markers. Fillmore's Case Grammar was also an attempt to capture meaning similarities across different syntax structures.

excessively], and [drinking], share partial meaning similarities to the relevant clause/sentence according to their structural complexity. The difference in structural complexity of this type correlates with the generality in meaning, from the most specific for [that [John drank beer excessively]] to the least specific for [drinking], but it does not make one nominalization functionally "more/less nominalized" than others.

Being nominal, nominalizations may head an NP and function as arguments of clauses and sentences. They do not stand alone like sentences in their capacity as nominal structures. However, nominalizations may become clauses and sentences when they become used to perform illocutionary acts, just as a noun can be used as a sentence issuing a warning (*Fire!*), for example (see next section). Conversely, sentences/clauses do not function as NP arguments. The only case in which they function as arguments is when used as a direct quotation; e.g. *Muhammad Ali said/boasted/wrote, "I am the greatest of all!"* In this way, grammatical constructions – sublexical morphemes, words, as well as larger phrasal units – are defined in terms of their functions, not by their formal or meaning similarities to other structures, though these provide supporting evidence for treating alike structures bearing the same function.

From our perspective, the terms "clause" and "nominalization" are mutually exclusive. The terms "nominalized clause" and "clausal nominalization" used widely in the field make sense only in one reading, namely in referring to nominalizations that share structural similarities with clauses but which are not clauses, not in the sense of the nominalizations that are clauses, which we believe do not exist. In the face of the proposed terms such as "grammatical nominalization" and "event nominalization" in this paper, it is not clear whether ambiguous terms such as "clausal nominalization" and "nominalized clause" have any theoretical status, besides the difficulty in precisely determining the degree of structural similarity that a nominalization must have in order for it to qualify as a clause.

6.3 Insubordination

A discussion of the distinction between clauses and sentences, on the one hand, and nominalizations, on the other, cannot be complete without touching on the use and development of nominalizations as sentences. The term insubordination or desubordination refers to a phenomenon in which a dependent structure, or a structure that does not stand as a sentence by itself, comes to be used as a stand-alone sentence (see Evans 2007). There are several clear cases in which nominalizations, which usually do not stand as sentences, develop into sentences. We discuss here two such cases; one, where event nominalizations as a whole get reanalyzed as sentences, eventually replacing sentences marked by a finite verb, and the other, where a sentence develops out of nominalizations combined with "auxiliary verbs" via a dropping of the latter.

In modern central dialects of Japanese, including Tokyo Japanese, the finite verb (known as *shūshi-kei* 'conclusive form') and the nominalized verb form (known as *rentai-kei* 'adnominal form'), which were distinct in Old Japanese (8th C) for many verbs, are identical, except for the copula, whose finite form is *da* and the nominalized form *na*. The merger of the finite and the nominalized verb form resulted from the use of nominalizations as sentences, which had been seen from the time of Old Japanse through Early Middle Japanese, and which was completed during the Late Middle Japanese period (13th–17th C). The following, from the *Genji Monogatari* (*The Tale of Genji*) first published in 1008, is a well-known example of the use of an event nominalization as a sentence.

(6-11) *Suzume no ko o Inuki ga nigasi-tu**ru**.*
sparrow GEN child ACC Inuki GEN let.go-PERF.NMLZR
(lit.) 'Inuki's letting my baby sparrow go – (shucks!).'

The perfective ending in the above example is in nominalized form, whose finite form is *-tu*, which would end usual declarative sentences. The use of the event nominalization as a sentence above was made possible because it performed the illocutionary act of evincing a feeling of disgust, regret, lamentation, or surprise. When the use of forms like this continued to expand, their illocutionary forces eroded, and they have eventfully been reanalyzed as ordinary declarative sentences, replacing those that ended in the finite verb form. This change brought about a reinterpretation of the genitive form modifying a nominalization (*Inuki ga*) as a subject. Stand-alone event nominalizations of the above type with a special illocutionary force have been reported elsewhere, where nominalizers involved are analyzed as a "stance-marker" (see Yap & Grunow-Hårsta (2010)).

Japanese nominalizations, being nominal, cannot predicate over a subject referent unless they combine with a copula or some tense-carrying "auxiliary verbs", many of which are grammaticalized versions of verbs, such as *suru* 'do', *yaru* 'give', *iru* 'exist/be' and their honorific variants.[55] Compare the following forms.

(6-12) Tokyo Japanese
a. *Taroo ga tegami o kai-ta.*
Taro NOM letter ACC write-PST
'Taro wrote a letter.'
b. *Taroo ga tegami o kak-i nasat-ta.*
Taro NOM letter ACC write-NMLZR do.HON-PST
'Taro wrote (HON) a letter.'

55. There is a good possibility that nominalization chains discussed earlier are precursors of these constructions. See (5-56) and (5-57).

Sentence (6-12b) contains the nominalized form *kak-i* 'writing', which by itself cannot predicate. It can, however, combine with the honorific auxiliary *nasar-* 'do', which inflects for tense. *Nasaru* has the somewhat archaic-sounding imperative form *nasare*, which in Western Japan dialects is pronounced as *nahare*, as in (6-13a), below. Now, Western Japan dialects, but not Tokyo Japanese, may drop this imperative auxiliary and let the form ending in nominalization carry an imperative force; hence turning a nominalization into a sentence, as in (6-13b).

(6-13) Western Japan dialects (Osaka, Kyoto)
a. *Tegami o kak-i nahare.*
letter ACC write-NMLZR do.HON
'Write a letter (honorific)/(lit.) Do writing a letter (honorific).'
b. *Tegami o kak-i!*
letter ACC write-NMLZR
'Write a letter (plain)!'

Tokyo Japanese, on the other hand, allows a dropping of a related polite imperative auxiliary in the form *nasai* 'do (HON)', when the nominalized form is honorified by the use of the prefix *o-*, as seen below.[56]

(6-14) Tokyo Japanese
a. *Tegami o o-kak-i nasai.*
letter ACC HON-write-NMLZR do.POL
'Write a letter (polite).'
b. *Tegami o o-kak-i!*
letter ACC HON-write-NMLZR
'Write a letter (polite)!'

A more widespread pattern in Japanese dialects involves combinations of the conjunction-marked nominalization *V-i=te*, which also combines with a variety of inflecting auxiliaries, including the grammaticalized forms of the verbs of giving marking a benefactive sentence, as in (6-15a) below. The dropping of the auxiliary verbs results in nominalization-based imperatives, as in (6-15b).

(6-15) Tokyo Japanese
a. *Tagami o ka-i=te* (< *kak-i=te*) *kure/kudasai.*
letter ACC write-NMLZR=CON give.me/give.me.POL
'Write a letter for me/Write me a letter.'
b. *Tegami o ka-i=te!*
letter ACC write-NMLZR=CON
'Write a letter for me/Write me a letter!'

56. Cf. *tegami* 'letter': ***o****-tegami* 'letter (honorific)'.

Finally, Japanese also allows a dropping of the copula *da* and its honorific variant *desu*, as well as the interrogative particle *ka,* allowing an event nominalization marked by the NP-use marker *no* to stand as a sentence. Compare;

(6-16) Tokyo Japanese
a. *Taroo ga kinoo kita no desu ka?*
Taro NOM yesterday came NPM COP Q
(lit.) 'It is that Taro came yesterday?' 'Did Taro come yesterday?'
b. *Taroo ga kinoo kita no?*
Taro NOM yesterday came NPM
(lit.) 'It is that Taro came yesterday?' 'Did Taro come yesterday?'

Since the copula *da*/*desu* can drop by itself, a declarative sentence identical in form to the interrogative sentence (6-17b) also obtains; the interrogative form (6-16b) is differentiated from the declarative one (6-17b) by a rising intonation.

(6-17) Tokyo Japanese
a. *Taroo ga kinoo kita no da/desu.*
Taro NOM yesterday came NPM COP/COP.POL
(lit.) 'It is that Taro came yesterday.' 'Taro came yesterday.'
b. *Taroo ga kinoo kita no.*
Taro NOM yesterday came NPM
(lit.) 'It is that Taro came yesterday.' 'Taro came yesterday.'

Like the other nominalization-based sentences, (6-17b) is usable only as a response to a question such as (6-16a), (6-16b), or in response to someone demanding an explanation for one's action. The sentential use of nominalizations, in other words, is most prevalent in face-to-face inter-personal situations, where the context allows the hearer to figure out the illocutionary force intended by the speaker.

The discussion on the insubordination phenomenon in Japanese above bears on the issues surrounding the treatment of the Pirahã nominalization raised by Everett (2005, 2009). In these papers Everett reverses his earlier nominalization treatment (Everett 1986) of the *sai*-marked structures seen below.

(6-18) Pirahã (Mura-Matanawi; Brazil)
a. *hi ob-áaxáí* [*kahaí kai-**sai***]
3 see/know-INTNS arrow make-NMLZR
'He really knows how to make arrows.'
b. *kóxoí soxóá xibíib-i-haí* [*tiobáahai bíio kai-**sai***]
Kóxoí already order-PROX-REL.CERT child grass do-NMLZR
'Kóxoí already ordered the child to cut the grass.'

Everett (2005, 2009) now identifies the suffix *-sai* as a marker of old information, and reanalyzes a sentence like (6-18a) as a paratactic coordination, as in the manner below.

(6-19) (*Hi*) *xob-áaxáí*. (*Hi*) *kahaí kai-sai*.
(3) see-well (3) arrow make-OLD.INFO
'He is really smart/very talented. (That is with respect to the fact that) he makes arrows.' (Everett 2009: 410)

Everett's motivation for his reanalysis of the Pirahã nominalizer *-sai* as an old information marker lies in his observations and conclusion below:

> "The second verb above [(6-19)], 'to make', is a bare root followed by *-sai*. This looks like nominalization until we see that: (i) the verb can take a full range of inflection …; (ii) that the sentence in which *-sai* appears can also appear as a main clause [(6-20) below]. "
>
> "If both clauses refer to topical information, both can bear the *-sai* suffix. If *-sai* were a nominalizer, however, we would not expect it to appear on both clauses since, presumably, a nominalized clause would not be a stand-alone sentence (cf. **John running the store*, **Rome's destruction of Carthage*)." (2009: 410)

(6-20) a. *Kóhoi xob-áaxáí xáagí-sai*.
name see-well permanent:to.be-OLD.INFO
'Kóhoi really knows his stuff.'
b. *Kóhoi hi kahaí kai-b-íigí-sai*.
name 3 arrow make-MOVE:DOWN-CONT-OLD.INFO
'He is finishing making arrows.' (Everett's glosses)

Clearly Everett does not know that grammatical nominalizations may have fully inflected verb forms (e.g. [that [John might have been arrested]]). Neither is he aware of the phenomenon of insubordination, which by itself does not invalidate a nominalization analysis.[57]

6.4 Evidence that nominalizations are not clauses or sentences

There are some compelling pieces of evidence pointing to the nominal nature of grammatical nominalizations that help distinguish them from clauses and sentences. Below we examine the two quintessentially nominal phenomena of plural and classifier marking.

57. The reason that *sai*-marked nominalizations are associated with the notion of old information has likely to do with their associated presuppositions. See the contributions to this volume by Bruil, Gipper & Yap, Machado & Peña for additional cases and discussions of insubordination.

6.4.1 *Plural marking*

Languages that have plural marking on nouns may mark grammatical argument nominalizations similarly since both may denote countable entities. Observe the following Bolivian Quechua forms.

(6-21) Bolivian Quechua
a. *wasi* 'house': *wasi-kuna* 'houses'
llank'a-q 'worker': *llank'a-q-kuna* 'workers'
b. [*wallpa-ta* *wayk'u-q*]-***kuna***
chicken-ACC cook-SUB.NMLZR-PL
'ones who are cooking a chicken'
c. [[*wallpa-ta* *wayk'u-q*]$_{NMLZ}$-***kuna***] *warmi-**kuna***]$_{NP}$
chicken-ACC cook-SUB.NMLZR-PL woman-PL
'women who are cooking a chicken'

The examples (6-21b.c) show that the subject grammatical nominalization involved plays a denoting function, just like a simple noun *wasi* 'house' in (6-21a), rather than the predication or the assertion function of a clause and a sentence. Notice, however, that a Quechua sentence, as in some other languages, may contain a verb marking plurality of an NP referent within a sentence, as in the following sentence.

(6-22) *Waki-n runa humu-n-**ku**.*
some-3 man come-3-PL
'Some of the men come.'

Crucially, the plural morpheme marking verbs differs from that marking nominals, although there is an obvious similarity in form.

Similar plural marking of grammatical nominalizations is seen in a fair number of languages, as the following data show.

(6-23) Capanawa (Panoan; Peru; Loos 1999)
[*ʔoá* *tsaʔot-ai*]$_{NMLZ}$ -***bo*** *his-**ɨ*** (236)
there sit-PRS -PL see-IMER
'Look at those (who are) sitting over there.'
Cf. [*ʔani hiwi mebi* *taʂpat-ai*]$_{NMLZ}$ *teʂpan anin ʔiso* *honiti* (236)
big tree branch bifurcate-PRS fork LOC monkey hide-PRS
'A monkey is hiding in the fork of a branch that bifurcates from a large tree.'

(6-24) Nheengatu (Tupí-Guaraní; Brazil; Cruz 2011 and p.c.)
a. *re-su* *re-mu-tawari* *kau* [*re-yu-mu-kuaku*
2SG.A-go 2SG.A-CAUS-tobacco DEM 2SG.A-R/R-CAUS-be.fasting
wa]=***ita*** *u-mbau* *arã*
NMLZR=PL 3SG.A-eat PROS
'You are going to bless those whom you made fast.'

b. *Ai-te paa nhaã pedasu itá=ita [maxi posu upe wa]=**ita***
3SG-FOC REP DEM piece stone=PL leper well LOC NMLZR=PL
'(They say that) he becomes those stones that are in the well of lepers.'

(6-25) Yaqui (Alvarez 2012 and p.c.)
a. [*in jinu-ka-'u*]-***m*** *sikili*
1SG.GEN buy-PERF-NMLZR-PL red
'Ones I bought are red.'
b. *U-me bisikleeta-m* [*in jinu-ka-'u*]-***m*** *sikili*
DET-PL bicycle-PL 1SG.GEN buy-PERF-NMLZR-PL red
'The bicycles that I bought are red.'

(6-26) Salve (Rice 1989: 83)
[*nịwą́ kedaw'í] i* ***ke*** *gogháyeyida*
long 3PL.sat NMLZR PL I SG.saw.3PL
'I met ones who stayed a long time.' (Hare dialect)

Turkish also allows plural marking on grammatical argument nominalizations, but does not permit doubling of plural marking on both the head noun and the modifying nominalization, as in (6-27c) below.

(6-27) Turkish (Göksel & Kerslake 2005: 449 and Yu Kuribayashi p.c.)
a. [*Opera-yı sev-me-yen*]$_{\text{NMLZ}}$-***ler***-*e şaşıyorum.*
opera-ACC like-NEG-NMLZR-PL-DAT surprised.1SG
'I am surprised at those who don't like opera.'
b. [[*Opera-yı sev-me-yen*]$_{\text{NMLZ}}$ *kişi-**ler***]-*e şaşıyorum.*
opera-ACC like-NEG-NMLZR person-PL-DAT surprised.1SG
'I am surprised at the people who don't like opera.'
c. *[[*Opera-yı sev-me-yen*]$_{\text{NMLZ}}$-***ler***] *kişi-**ler***]-*e şaşıyorum.*
opera-ACC like-NEG-NMLZR-PL person-PL-DAT surprised.1SG
'I am surprised at the people who don't like opera.'

Tapiete grammatical argument nominalizations, in addition to plural marking, show another nominal feature, foreign to clauses and sentences, namely, diminutive marking, as below.

(6-28) Tapiete (Golzárez 2005 and Coccine 2008)
a. *o-che-wa-**reta***
3AC-sleep-NMLZR-PL
'(the ones) who are sleeping'
b. *karai-re* [*tumpa i-ñe'ë mbe'u i-a-**reta***]
white.man-PL god 3.POSS-language tell be-NMLZR-PL
'the gringoes (white men) who were announcing (reciting) the Bible'

c. *hau-wa-**mi***
1:eat-NMLZR-DIM
'what little I eat'

d. *ko ñ-a'engu-mba-**mi***
DEM 3IN-be.deaf-NEG.NMLZR-DIM
'this (one) who is not a little deaf'

Perhaps these examples above are not surprising since all of them involve clear nominalizing morphology, indicating that the relevant forms are nominals. Therefore, the fact that even those languages that involve no nominalizing morphology in their grammatical nominalizations allow plural marking on them is significant in showing that nominalizations may contain verbal forms similar to finite verbs and that nominalization is independent from morphological marking. The case in point is the use of a plural marker on grammatical nominalizations with a finite verb form in Toba, which, according to Cristina Messineo (p.c.), is observed in the speech of younger speakers, as in the following example.

(6-29) Toba (courtesy of Cristina Messineo)
*hilos [na [nqat-te-wek]-**pi***]
threads DD take-PROG-out-PL
'los hilos, los está sacando', 'the threads, the ones (s/he) is taking out'

Finally, Piapoco, spoken in the eastern plains of Colombia, has a nominalizer that combines number and gender information, as below, where gender (± M) is indicated only in singular forms.

(6-30) Piapoco (Arawak; Colombia; Klumpp & Burquest 1983)

a. *yà-a-wa [i-té-**eyéi**-ca yà-ana*] (395)
3M-go-ASP 3M-carry-[+pl]-ASP 3M-limb
'(The ones) who carry the animal's leg go.'

b. *ábiba asìeli [yà-amè-**eri** sísade Cadá néese*] (390)
other man 3M-arrive-[+M/−pl] from.there Cada from
'the other man who arrived from Cada'

6.4.2 *Classifier marking*

The Piapoco data bring us to the next nominal feature that reflects the entity-denoting property of nominalizations, namely classifiers. Japanese numeral classifiers, mostly Chinese loans, occur in several syntactic positions. Two common patterns are shown below, where the numeral classifier *san-satu* [three-CLF.BOUND] 'three bound (things)' occurs prenominally (6-31a) and as an adverb away from the modified noun (6-31b).

(6-31) a. *Ken wa* ***san-satu*** *no hon o kinoo motte kita.*
Ken TOP three-CLF GEN book ACC yesterday carry.GER came
'Ken brought three books yesterday.'
b. *Ken wa hon o kinoo* ***san-satu*** *motte kita.*
Ken TOP book ACC yesterday three-CLF carry.GER came
'Ken brought three books yesterday.'

Grammatical nominalizations in Japanese do not seem to readily allow prenominal numeral classifiers, but they can be quantified by adverbial numeral classifiers, indicating that grammatical nominalizations denote entities rather than predicate or assert like clauses and sentences.

(6-32) a. **Boku wa* [***san-satu*** *no* [*Ken ga motte kita*]$_{\text{NMLZ}}$] *no o*
I TOP three-CLF GEN Ken NOM carry.GER came NM ACC
kinoo yonda.
yesterday read
(lit.) 'I read yesterday three of what John brought.'
b. *Boku wa* [*Ken ga motte kita*]$_{\text{NMLZ}}$ *no o kinoo*
I TOP Ken NOM carry.CON came NPM ACC yesterday
san-satu *yonda.*
three-CLF read
'I read yesterday three of what Ken brought.'

It is interesting to notice that the adverbial quantifier in (6-32b) has the partitive interpretation of reading three of what Ken brought, rather than quantifying what is denoted by the grammatical nominalization. But the point is that the choice of an adverbial quantifier is determined by the denotation of the grammatical nominalization, as the comparison between (6-32b) and the following clearly shows.

(6-33) *Boku wa* [*Ken ga motte kita*]$_{\text{NMLZ}}$ *no o kinoo*
I TOP Ken NOM carry.CON came NPM ACC yesterday
san-bon *nonda.*
three-CLF drank
'I drank yesterday three (bottles) of what Ken brought.'

The choice of adverbial classifiers indicates different types of things that Ken brought. The use of *satu* in (6-32b) indicates that what Ken brought were books or book-like bound materials, while the use of *hon/bon* in (6–33) indicates that what Ken brought were contained in cylindrical containers such as bottles.

Interestingly Chinese allows the pattern in (6-34a) disfavored in Japanese. Observe:

(6-34) Mandarin Chinese

a. *sān-běn shū*
three-CLF book
'three books'

b. *sān-zhī niǎo*
three-CLF bird
'three birds'

c. *Sān-běn* [*wǒ mǎi-de*]$_{\text{NMLZ}}$ hěn guì.
three-CLF I buy-NMLZR very expensive
(lit.) 'Three what [books] I bought were very expensive.'
Cf. [*wǒ mǎi-de*]$_{\text{NMLZ}}$ *shū*
I buy-NMLZR book
'book that I bought'

d. *sān-zhī* [*wǒ mǎi-de*]$_{\text{NMLZ}}$
three-CLF I buy-NMLZR
(lit.) 'three what [animals] I bought'

Forms (6-34c, d) show that the grammatical nominalizations *wǒ mǎi-de* 'what I bought' may denote a variety of things evoked by this structure, and depending on what they actually denote, different classifiers are chosen in quantifying the denoted objects, such as books or book-like materials as in (6-34c) and animals as in (6-34d).

That grammatical nominalizations denote, rather than predicate or assert, is also clearly seen from the use of classifiers in Thai, which allows optional classifier marking of grammatical argument nominalizations. Observe.

(6-35) Thai (courtesy of Kingkarn Thepkanjana)

a. *khruu* [*lăaj khon*]
teacher many CLF.PERSON
'many teachers'

b. *măa* [*sìi tua*]
dog four CLF.BODY
'four dogs'

c. (*chán chɔ̂ɔp*) [[*thîi khwĕɛ nay tûu*]$_{\text{NMLZ}}$]$_{\text{NP}}$
I like NMLZR hang in closet
'(I like) the one hanging in the closet.'
(Answer to the question "Which skirt do you like?")

c′. (*chán chɔ̂ɔp*) [***tua*** [*thîi khwĕɛ nay tûu*]$_{\text{NMLZ}}$]$_{\text{NP}}$
I like CLF NMLZR hang in closet
'(I like) the one hanging in the closet.'

d. (*chán chɔ̂ɔp*) [*kràprooŋ* [***tua*** [*thîi khwĕɛ nay tûu*]$_{\text{NMLZ}}$]]$_{\text{NP}}$
I like skirt CLF NMLZR hang in closet
'(I like) the skirt hanging in the closet.'

e. **Tua* [*kràprooŋ khwěɛ nay tûu*]
CLF skirt hang in closet
'A skirt hangs in the closet.'

Notice that a clause/sentence is never marked by a classifier, as indicated by the ungrammatical form in (6-35e) above.

While classifier makring is largely optional in Thai in the sense that relevant structures may or may not be marked by a classifier depending on how specific one wants to be with regard to what a nominalization denotes, some languages use classifiers, instead of a general nominalizer, in marking nominalizations. Cantonese is one such language, in which grammatical nominalizations may be marked by classifiers, the choice of which depends on what they denote.[58] Observe;

(6-36) Cantonese (Matthews & Yip 1994: 111, 112, p.c.)

a. [*nī* ***dī***] *yú*
this CLF fish
'this fish'

b. [*sām* ***go***] *hohksāang*
three CLF student
'three students'

c. [[[*Ngóhdeih hái Faatgwok sihk*] ***dī***]$_{\text{NMLZ}}$ *yéh*]$_{\text{NP}}$ *géi hóu-sihk ga.*
we in France eat CLF food quite good-eat NM
'The food that we ate in France was pretty good.'

d. [[[*Gaau léih tàahn kàhm*] *gó*] ***go***]?
teach you play piano that CLF
'The one who teaches you (to play the) piano?'

Asia is not the only area in which nominalizing classifiers occur. The Amazon Basin is another area where classifiers play important grammatical roles, including their use as numeral classifiers and for marking grammatical nominalizations, again underscoring the point that grammatical nominalizations denote entities, which can be classified according to their nature. Observe the following data from Bora, where the classifier *hà* marks an argument nominalization denoting an object like a shelter or with a sheltering function and *kpà* a slab-like object.

(6-37) Bora (Witotoan; Colombia, Peru, Brazil; Thiesen & Weber 2012)

a. *ó-a^{x}t^{jh}ùmɨ-^{ʔ}* [*a^{s}í-ːβ^{j}ɛ̀*]-***hà*** (382)
I see-<t> burn-sIn-<shelter>
'I saw a house that was burning.' (lit.) 'I saw one (shelter-like thing) that was burning.'

58. The nominalizing function of classifiers is discussed fully in Section 8.

b. *ò-k^{h}è t-à:k^{h}ìù [è:-kpà [é:-há tʃì^{L}:ɲè í^{x}*
1-OBJAN you.IMP-give that-<slab> that-<shelter> below
*k^{jh}à]-**kpà**]βìù* (388)
be-<slab>-THM
'Give me the plank that is under the house.' (lit.) 'Give me the (slab-like thing) one (slab-like thing) which is under the (shelter-like thing).'

Like grammatical nominalizations in other languages, these nominalizations marked by classifiers also have a modification-use. First observe (6-38a) below, in which a gender-based classifier marks argument nominalizations, a phenomenon very common among Amazonian languages. Argument nominalizations like this can modify a head noun, as in (6-38b).

(6-38) a. **[*Ø hóáà-k^{h}è uɾ^{s}kpá:pò*]-*:pè*** *ts^{h}á:-ʔì* (379–380)
John-objAn teach-<SG.MSC> came-<t>
'(The one-MSC) who taught John came.'

b. *ó á^{x}t^{h}ùmɨ́-^{ʔ} [ò:ʔí-:p^{j}è]-khè* **[*Ø ò-k^{h}è ɨ́^{sʔ}tó*]-*:pè*-khè** (381)
I see-<t> dog-<SG.MSC> I-OBJAN bite-<SG.MSC>-OBJAN
'I see the dog that bit me.

The Bora patterns above may seem quite exotic, but, as a matter of fact, the gender-based classifier system is widespread among Indo-European languages, and several languages incorporate it in their nominalization markers as Bora does. The Indo-Aryan language Marathi, for example, has argument nominalizations involving verbal forms traditionally labeled participle, which mark the gender of a denoted person. Compare the following examples.

(6-39) Marathi (Indo-Aryan; courtesy of Prashant Pardeshi)

a. [*dāgine ʦor-lel-**i***] *qām-č-ī nokriṇ āhe*
jewels steal-PST.NMLZR-**FSG** we-GEN-FSG female.servant be.PRS
'The one (FEM) who stole the jewels is our maid.'

b. [*dāgine ʦor-lel-**i***] *bāī ām-č-ī nokriṇ āhe*
jewels steal-PST.NMLZR-**FSG** woman we-GEN-FSG female.servant be.PRS
'The woman who stole the jewels is our maid.'

(6-40) a. [*dāgine ʦor-lel-**ā***] *ām-č-ā nokar āhe*
jewels steal-PST.NMLZR-**MSG** we-GEN-MSG male.servant be.PRS
'The one (MASC) who stole the jewels is our male servant.'

b. [*dāgine ʦor-lel-**ā***] *māṇus ām-č-ā nokar āhe*
jewels steal-PST.NMLZR-**MSG** man we-GEN-MSG male.servant be.PRS
'The man who stole the jewels is our male servant.'

The Marathi argument nominalizations above also show that clauses/sentences do not underlie them. In Marathi, as in other similar Indo-Aryan languages, verbs agree in gender and number with an unmarked (or nominative) nominal argument. In the past tense, an A nominal is marked ergative and a P nominal is left unmarked, and in such a case a verb agrees with the P nominal, as shown below:

(6-41) a. *mol.karṇi-ne dāgine ʦor-le*
maid-ERG jewels.**MPL** steal-PST.**MPL**
'The maid stole the jewels.'
b. *nokar-ne dāgine ʦor-le*
male.servant-ERG jewels.**MPL** steal-PST.**MPL**
'The male servant stole the jewels.'

Comparison between these and the nominalized forms in (6-39) and (6-40) makes it clear that they do not underlie the latter, in which nominalized forms do not agree with the unmarked P argument.

German distinguishes three gender classes of masculine, feminine, and neuter, and, like ordinary nouns, grammatical argument nominalizations are distinguished according to these classes depending on what they denote. Observe:

(6-42) German
a. *Ich kenne den* [***der*** [*Ø morgen kommt*]].
I know ART.MSC.ACC MSC.SUB.NMLZR tomorrow comes
'I know the one (MSC) who comes tomorrow.'
b. *Ich kenne die* [***die*** [*Ø morgen kommt*]].
I know ART.FEM.ACC FEM.SUB.NMLZR tomorrow comes
'I know the one (FEM) who comes tomorrow.'
c. *Ich kenne das* [***das*** [*Ø morgen kommt*]].
I know ART.NEUT.ACC NEUT.SUB.NMLZR tomorrow comes
'I know the one (NEUT) who comes tomorrow.'

The German nominalizers clearly combine information about the gender class and about the grammatical relation that the denoted entities are identified with. In fact, German has a double marking system, whereby nominalizations have the structure internal marking discussed above as well as the external marking by the articles indicating the gender class of the entity denoted by the nominalizations, as observed in (6-42) above.[59] In some languages, external classificatory articles of the German type are the only clue to the nature of the entities denoted by grammatical argument nominalizations.

59. The articles marking these argument grammatical nominalizations, based on demonstrative pronouns, differ from those marking ordinary nouns, though there is a great deal of overlap in form and function between the two sets.

In Toba, all nominals in NP-use are marked by what Messineo (2003) calls nominal classifiers (*clasificadores nominales*), which encode configurational and deictic, as well as number and gender information regarding the denotation of the following nominal, as below.[60]

(6-43) Toba (see Messineo (2003: 145ff) for details)

a. ***na*** *pioq*
DD dog
PROXIMAL
'this dog'

b. ***yi-wa*** *pioq*
DD-PAUC dog
HORIZONTAL
EXTENDED
'two or three dogs lying down'

c. ***a-so*** *waaka*
FEM-DD COW
DISTAL
'a COW'

Just like the German articles mentioned above, these demonstrative determiners indicate the nature of the entity denoted by grammatical nominalizations, as seen below.

(6-44) Toba (courtesy of Cristina Messineo)

a. *s-acʔek* ***a-so*** [(*ntonigiʃi*) [*ʔaw-ʔot* *Ø ʃikajt*]$_{\text{NMLZ}}$]$_{\text{NP}}$
1A-eat FEM-DD (tortilla) 2A-make yesterday
DISTAL
'I ate what you made yesterday.' 'I ate the tortilla that you made yesterday.'

b. [***na-wa*** [*Ø chigoqchigiŋa yi* *Espinillo*]$_{\text{NMLZ}}$]$_{\text{NP}}$
DD-PL 3.come.from DD Espinillo
PROXIMAL
t-r-ayge *da* *Salta ko'ollaGa*
3.go<PAUC> DD Salta PST
'Those who came from the Espinillo went to Salta.'

These demonstrative determiners (DDs) (and the demonstratives not discussed here) show two things. Syntactically, grammatical nominalizations are nominal and they are marked by either a demonstrative determiner (or a demonstrative) in

60. Cristina Messineo (p.c.) now would call these "demonstrative determiners" (DDs) and "demonstratives" (DEMs). I gloss the examples with this new terminology.

their NP-use, just like any ordinary nouns. In other words, determiners function as NP-use markers for both ordinary nouns and grammatical nominalizations.[61] Semantically, grammatical nominalizations, especially argument nominalizations, denote concrete entities, whose physical properties including number and gender are marked by DDs (or demonstratives), again just like any ordinary nouns.[62] These facts would not be easily explained if grammatical nominalizations were clauses and sentences that do not denote substantives. Notice also that grammatical nominalizations are perfectly compatible with finite verb forms in Toba as well.

We conclude this section by pointing out that English grammatical argument nominalizations also classify their denotations in terms of the human/non-human distinction similar to the animate/inanimate distinction that Newar nominalizers mark. The case in point is the distinction between *who(m)* and *which*, the former marking a human denotation by an argument nominalization and the latter a non-human denotation. Observe:

(6-45) a. *You may marry* [[***who*** [*you like* Ø]]$_{\text{NMLZ}}$]$_{\text{NP}}$
b. *You may choose* [[***which*** [*you want* Ø]]$_{\text{NMLZ}}$]$_{\text{NP}}$

This again shows that what we consider to be grammatical nominalizations are denoting rather than predicating or asserting like clauses and sentences.

7. Nominal-based nominalizations

Perhaps the most innovative proposal made in Shibatani's work on nominalization is to reanalyze the genitive or possessive construction as a nominal-based nominalization.[63] There are several motivations for this radical departure from the traditional analysis. First of all, what forms like *his* and *John's* denote are those things with which the person referred to is intimately connected, as in the case of ordinary metonymic expressions, such as things that are possessed permanently or temporarily or things to which the person is connected as an author or a theme (as in the case of the theme of a photo).

61. See Section 8.1 on the development of these NP-use markers as nominalizers.

62. Movima (a language isolate spoken in Bolivia; Haude 2006) and Musqueam (Suttles 2004) and related Salish languages have demonstrative systems similar to Toba.

63. We are dealing with the genitive of "possession" in this article, not other kinds of gentive, such as, for example, the Russian genitive used in leu of the partitive case.

Secondly, the precise references of the "possessive" forms are determined by context, again as in the case of ordinary metonymic expressions and verbal-based grammatical nominalizations discussed above.

Observe:

(7-1) A: *Which car do you like?*
B: *I prefer **John's** over **Bill's**.*

(7-2) A: *Is this the book that Bill brought?*
B: *No, that's **John's**. I saw **Bill's** on the dining table.*

By the Gricean Cooperative Principle, we would interpret *John's* and *Bill's* in (7-1) to be referring to the cars intimately connected with the referents of *John* and *Bill*, while in (7-2) *John's* and *Bill's* would be likely understood to be referring to the books belonging to the referents of *John* and *Bill*.[64] The relevant forms above represent NP-use of N-based nominalizations. These, like V-based grammatical nominalizations studied above, also have a modification-use, as shown below.

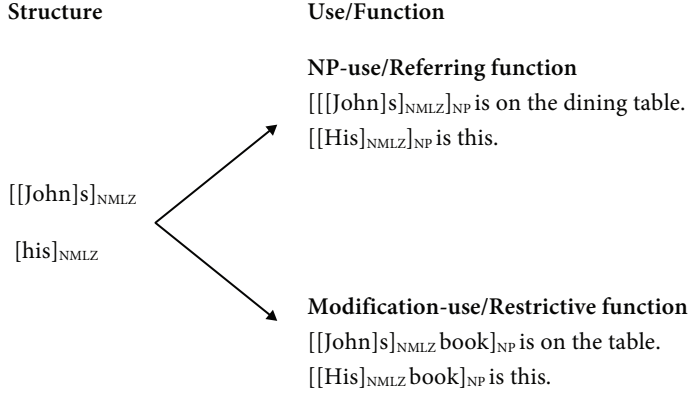


Figure 4. Two uses of N-based nominalization

The newly proposed nominalization analysis does away with the genitive case for possessive constructions altogether as well as the parts of speech of "possessive pronouns/pronombres posesivos" and "possessive adjectives/adjetivos posesivos" recognized in traditional grammar. The former (*his, mine*, etc.) are no more than instances of the NP-use and the latter (*his, my*, etc.) those of the modification-use

64. These nominal-based nominalizations in NP-use are described variously as "free genitives" (e.g. Stolz et al. 2008: 390), "headless adnominal" (e.g. Noonan 2008: 130), and "headless possessive" (Koptjevskaja-Tamm 1995; van der Voort 2009). See also Dryer (2008). The same arguments against such treatments can be raised as those against treating the NP-use of argument nominalizations as "headless relative clauses" in Section 5.4.

of N-based nominalizations. Traditional grammar makes the same mistake as those who recognize relative clauses apart from the modification-use of grammatical argument nominalizations.

Besides the logical consistency between the analysis of N-based nominalizations outlined above and that of V-based nominalizations, Shibatani (2017, 2018a) offers morphological evidence unifying these two types of grammatical nominalization. The NP-use of V-based nominalizations in Modern Japanese requires *no*-marking very much similar to the *one*-marking found in Modern English. Compare the forms of the two uses of V-based nominalizations in the Japanese forms and their English translations below.

(7-3) Japanese

a. [[*suki na*]$_{\text{NMLZ}}$]$_{\text{NP}}$ ***no*** *to kekkon sinasai.*
like COP NPM with marriage do.IMP
'Marry **one** [who [you like]$_{\text{NMLZ}}$]$_{\text{NP}}$

b. [[*suki na*]$_{\text{NMLZ}}$ *hito*]$_{\text{NP}}$ *to kekkon sinasai.*
like COP person with marriage do.IMP
'Marry [a person [who [you like]]$_{\text{NMLZ}}$]$_{\text{NP}}$

Shibatani traces the *no*-marker above to the Classical Japanese *no* that marks the NP-use of the N-based nominalization (or the genitive form), as seen in an example like the one below.

(7-4) *Hitozuma to* [*wa ga*[65]]
man's.wife and I GEN
no *hutatu omouni hanarekosi sode wa awaremasereru.*
NPM two think leave.behind sleeve TOP exceedingly.sad
'As I think about both a man's wife and **mine**, the sleeves left behind are exceedingly sad.' (*Yoshitadashū,* 10th C)

The extension of the *no*-marking from the NP-use of N-based nominalization (aka the genitive/possessive), as in (7-4) above to that of the V-based nominalization, as in (7-3a) started in the early 17th century. Shibatani's point is that this extension of the *no*-marking from one domain to another indicates that the two domains are recognized as a unified phenomenon.

Crosslinguistic investigations reveal a large number of cases where both N-based nominalizations and V-based nominalizations take the same morphological marking. There are two patterns of marking here. One is the Japanese

65. Classical Japanese had two genitive particles (or nominalizers for nouns), *no* and *ga*. The *no* particle that marks the NP-use of grammatical nominalization is related to the nominalizer/genitive *no*. In some other dialects (e.g. Kōchi, Toyama) *ga* is used as the marker of the NP-use of grammatical nominalizations.

pattern, where only the NP-use of both N-based and V-based nominalizations are marked the same way, and the other, perhaps more compelling pattern is where both N-based and V-based nominalizations involve identical markings in all usage patterns.

The Korean use of the particle *kes* is similar to the Japanese *no*-marking, where only NP-uses of N-based and V-based nominalizations are marked identically, as below.

(7-5) Korean

a. NP-use of V-based nominalization
Na-nun [[[*apeci-ka cwu-n*]$_{\text{NMLZ}}$-***kes***]$_{\text{NMLZ'}}$]$_{\text{NP}}$-*ul ilk-ko-iss-ta*.
I-TOP father-NOM give-NMLZR-NPM-ACC read-GER-be-IND
'I am reading what father gave (me).'

a′. NP-use of N-based nominalization
[[[*emeni-uy*]$_{\text{NMLZ}}$-***kes***]$_{\text{NMLZ'}}$]$_{\text{NP}}$-*un ku chaek i-ta*.
mother-NMLZR-NPM-TOP that book COP-IND
'Mother's is that book.'

b. Modification-use of V-based nominalization
[[*apeci-ka cwu-n*]$_{\text{NMLZ}}$ *chaek*]$_{\text{NP}}$]
father-NOM give-NMLZR book
'the book that father gave (me)'

b′. Modification-use of N-based nominalization
[[*emeni-uy*]$_{\text{NMLZ}}$ *chaek*]$_{\text{NP}}$
mother-NMLZR book
'mother's book'

Notice that the nominalizers themselves are different for the V-based (*-n*) and N-based (*-uy*) nominalization, yet the results of these processes are treated alike in their NP-use, as seen in (7-5a) and (7-5a′).

Many languages of the world (e.g. a variety of Japanese dialects, Ryukyuan, Telugu and some other Dravidian languages) show a similar marking pattern for the NP-use of both N-based and V-based grammatical nominalizations (see Section 6.1 for the Telugu pattern). While the ultimate origins of the Japanese particle *no* and Korean *kes* cannot be known, many languages recruit as markers of NP-use of nominalizations a noun meaning "thing", as closely documented in a variety of Ryukyuan languages by Shibatani and Shigeno (2013).[66] The Kwa language

66. Many Korean scholars think that *kes* was originally a noun with the meaning of "thing", but there is no evidence for it. The "thing" reading they associate with *kes* actually comes from the nominal denotation of the nominalizations they mark. Like Japanese, Middle Korean did not have the *kes* marking, yet those nominalizations without *kes* have exactly the same "thing" reading as their modern counterparts with *kes*. See Lee (1975).

Gã of Ghana uses *nɔ́*, deriving from a noun meaning "thing", and *mɔ̀*, which means "person" as a noun, as markers of NP-use of nominalizations – the former when a non-human is denoted and the latter for a human denotation. While in Gã the origins of these markers are transparent, Campbell (2017: Chapter 6) presents strong evidence that they are grammaticalized and do not mean "thing" or "person" when they occur with the NP-use of nominalizations. For example, *nɔ́* as a noun takes a definite article but *nɔ́* as a marker of NP-use never does, and *nɔ́*, meaning inanimate "thing" as a noun, can mark both human and animate referents in the NP-use of N-based nominalizations for a smaller group of speakers.[67]

(7-6) Gã (Kwa; Campbell 2017: 111, 538, 550, 557)

a. NP-use of V-based nominalization
[***nɔ́*** [*nɪ́* *ɪ́=súmɔ́ɔ́ɔ̀*]] *jí* *àníháó*
NPM NMLZR 1SG=like.NEG COP laziness
'What I don't like is laziness.'

a′. NP-use of N-based nominalization
shɪ́ [[*Ellen*] ***nɔ́***]=! *ɔ́* [*Pàpá Tèí* ***nɔ́***]=! *ɔ́* *lɛ́=!ɛ́*
but Ellen NPM=TOP Papa Tei NPM=TOP 3SG.OBJ=TOP
ɪ́=nyɛ́ɛ́ɛ́ *má-!yá*
1SG-able.NEG 1SG.FUT-go
'But as for Ellen's and Papa Tei's I couldn't attend them.'

b. Modification-use of V-based nominalization
[*àtàlé* [*nɪ́* *àmɛ̀=sùmɔ̀*]]
dress NMLZR 3PL=like
'the dress that they like'

b′. [[*Elma*] *bɪ̂*][68] *jí* *lɛ̀*
Elma child COP 3SG.OBJ
'She's Elma's child.'

---

67. Notice that these markers of the NP-use of nominalizations are the result of *appropriations of lexical nouns* meaning "thing", "person", "matter", "place" etc. as a grammatical marker, rather than resulting from their lexical use through a gradual grammaticalization process. This is so because the lexical use of these nouns continues in parallel with the grammatical use, as in many Ryukyuan languages, where constructions marked by *mun*(*u*) meaning "thing" lexically can be ambiguous (e.g. *wa: munu* 'my thing/mine'). In the Taketomi Ryukyuan this ambiguity has been resolved by changing the grammatical version as *ũnu* (*ba: munu* 'my thing' vs. *ba: ũnu* 'mine'). See Section 8 below and Shibatani & Shigeno (2013) on the use of *mun*(*u*) in verbal-based nominalizations.

68. Notice that Gã, as in many languages, does not have an overt nominalization marker for N-based nominalizations.

The recruiting of a noun meaning "thing" as a marker of NP-use of nominalizations is also seen in the Panoan language Kakataibo (Zariquiey 2011).

Notice at this juncture that we are dealing with two kinds of nominalization markers[69] – one, a marker of NP-use glossed as NPM, which, as in the examples above, marks only the NP-use of nominalizations, and the other, a nominalizer (or nominalizing morphology) glossed as NMLZR, which marks nominalizations and which appears wherever they are used, e.g. in both NP-use and modification-use of nominalizations; see the nominalizers *-n* for V-based nominalizations and *-uy* for N-based nominalizations in the Korean examples in (7-5) above. It is mistaken to identify NP-use markers (NPM's), such as the Japanese particle *no* and Korean *kes* (as well as Gã *nɔ́*) above, as nominalizers (as in, e.g. Horie 2008; Frellesvig 2010; Yap, Grunow-Hårsta & Wrona 2011).[70]

Both Japanese *no* in the central dialects and Korean *kes* are a modern development and the earlier (e.g. Middle Japanese and Middle Korean) NP-use of nominaliztions was not marked by them, but they are now both obligatory as NP-use markers in the modern languages. Likewise, the English NP-use marker *one* seen in the translation of (7-3a) above is a modern development. Grammatical nominalizations marked by *wh*-forms were freely used as NP-heads without the *one*-marking in Middle English, and it is still optional in Modern English in certain contexts, as in many of the examples in this paper.[71]

The difference between NP-use markers and nominalizers can be clearly seen by comparing the marking pattern of Korean *kes* in (7-5) above and that of the Mandarin *de* nominalizer below, where *de* occurs in both NP- and modification-use.

69. We use the term "nominalization marker" as a cover term for both markers of nominalizations in NP-use, like Korean *kes* and Gã *nɔ́*, and nominalizers, like Korean *-n* and *-uy* and Mandarin *de*, which occur in both contexts of NP- and modification-use.

70. Whether or not a morphology marking nominalization is an NP-use marker or a nominalizer cannot be determined without a systematic investigation of usage patterns of nominalization. It is, therefore, dangerous to rely on others' descriptions of nominalization markers, as done in many research papers, since most past works on this topic have not done the necessary systematic investigation called for.

71. The *Middle English Dictionary* compiled by the University of Michigan (https://quod.lib.umich.edu/m/med/) contain forms such as *And* [*which falleþ on þat furste flur*] *schal beo Quene* and [*Who aske this*] *Leseth his asking trewely*, both of which would require marking by *one* in Modern English. Sinhala is another language that has recruited the numeral *eka* 'one' as a marker of the NP-use of both event and argument nominalizations, where the marking is obligatory.

(7-7) Mandarin Chinese

a. NP-use of V-based nominalization
Nǐ méi yǒu [[*wǒ xǐhuān Ø*] =***de***]$_{\text{NMLZ}}$]$_{\text{NP}}$
you not have I like =NMLZR
'You don't have what I like.'

a′. NP-use of N-based nominalization
Zhèi běn shū shì [[[*wǒ*]=***de***]$_{\text{NMLZ}}$]$_{\text{NP}}$
this CLF book COP I=NMLZR
'This book is mine.'

b. Modification-use of V-based nominalization
[[*wǒ xǐhuān Ø*] =***de***]$_{\text{NMLZ}}$ *yī fu*]$_{\text{NP}}$
I like =NMLZR clothes
'the clothes that I like'

b′. Modification-use of N-based nominalization
[[*wǒ*]=***de***]$_{\text{NMLZ}}$ *shū*]$_{\text{NP}}$
I=NMLZR book
'my book'

While the marking of both N-based and V-based nominalizations by the same nominalizing morphology has been noticed by many, including Matisoff (1972), there has been no answer as to why a V-based nominalization and a "possessive construction" are marked the same way.[72] As for Mandarin Chinese, Li & Thompson (1989) recognize two different *de*, one for nominalizing verbs (p. 575ff) and the other termed "associative" *de* (p. 113ff) for N-based expressions, as if we are dealing with two different particles that are accidentally similar in form.[73] Such a treatment cannot explain why a similar "accident" happens in so many languages across the globe. Sposato (2012), in his description of relative clauses of the Miao language Xong, opts for Li & Thompson's term in describing one type of V-based argument nominalization and N-based nominalizations, both marked by what appear to be interchangeable markers *naond* and *nangd*, leaving unanswered the question why relative clauses and possessive constructions are marked the same.[74]

---

72. DeLancey (1986: 1), maintaining a narrow, verb-centered view of nominalization, finds it "odd that a dependent noun [of possessive constructions] would be marked as nominalized".

73. After completing this manuscript, I learned that Zhu (1982) had offered an analysis highly similar to ours, recognizing N-based nominalizations along with V-based ones in terms of the Mandarin particle "de" under discussion.

74. It is probable that the Xong pattern results from contact with Chinese. Sposato recognizes other functions these markers play, such as marking adverbs and functioning as a sentence final emphatic marker. The development of nominalizers into these functions is not at all rare (see Yap & Grunow-Hårsta (2010)).

(7-8) Xong (Miao-Yao (Homong-Mien); southern China; Sposato 2012)

a. [*Wud jangs* ***nangd***][75] *nis ndut-lid ndut-ghueax.* (58)
[3SG plant ASSOC] COP tree-plum tree-peach
'What he planted were plum trees and peach trees.'

b. [*Wel hauk* ***naond***] *jud jix raut.* (57)
1SG drink ASSOC alcohol NEG good
'The alcohol that I'm drinking is no good.'

c. *Ob-naind nis* [*wel* ***naond***]. (59)
NOM-this COP 1SG ASSOC
'This is mine.' (NOM=nominalizing prefix or general nominal prefix)

d. [*dab-guoud* ***naond***] *zhoux.mioux* (59)
AN-dog ASSOC ear
'the dog's ear' (AN=animal prefix)

A wide range of Tibeto-Burman languages (e.g. Lahu, Burmese) show a pattern similar to the Chinese (Sinitic) and Xong (Miao-Yao) pattern above, marking both V-based argument nominalizations and N-based nominalizations the same way. While the data clearly cry out for uniform treatment of these marking patterns, the nominalization solution was unavailable to those who have puzzled over the relevant data because of the past beliefs that nominalization applies only to verbs or non-nominal elements. We have shown in the beginning of this paper that these beliefs are in fact groundless; and what we see in Chinese and Xong above (as well as Matisoff's earlier observations on Lahu) confirm this.

Turning to languages outside the Sino-Tibetan sphere, it is not easy to find comparable cases among languages of the Americas largely because most descriptive grammars fail to record the NP-use of N-based nominalizations comparable in form to *mine* and *the child's*, as in ***Mine/The child's*** *is this one*. However, at least two languages are found that show the use of the same marking pattern for both V-based and N-based nominalizations. Also, there are a fair number of languages that treat V-based and N-based nominalizations similarly in terms of classifier marking, as discussed below.

The first, a more straightforward, case found in the Mosetenan language Mosetén in Bolivia concerns the nominalizers *-tyi'* (M) and *-si'* (F), which classify their denotations according to their gender. While Sakel (2004) treats them as linkers, she

75. This and the example in (7-8c) show that the term/gloss "associative" is quite infelicitous if what is referred to by this term were to "indicate[s] that two noun phrases [connected by it] are 'associated' or 'connected' in some way" (Li & Thompson 1989: 113), because these examples do not involve two noun phrases connected by *nangd* or *naond*. On the other hand, if the term is understood to be referring to a nominalizer marking structures that denote entities metonymically associated with them, then it makes a good sense. But this is not what is intended by Li & Thompson's use of the term.

clearly recognizes their nominalization function, citing forms like the ones below representing both V-based lexical and grammatical nominalizations. Also observed below are N-based grammatical nominalizations utilizing the same nominalizers.

(7-9) Mosetén (Sakel 2004: 97, 106, 107, 111)

a. V-based nominalization

i. *mi'* *jäe'mä* [*jï'-wë-ti-ksi-**tyi'***] (NP-use)
3M.SG DM CA-move-DT-3P.O.M.S-NMLZR
'the one who made them come'

ii. *Mi'* *soñi'* [*yo-dye'-**tyi'***] *i-khan bae'-i äej-ä-te*
3M.SG man R-NMLZR-NMLZR M-IN live-VI.M.S kill-VI-3M.O
jïirï-ty ïtsïkï. (Modification-use)
one-M jaguar
'The man who lives in here killed a jaguar.'

b. N-based nominalization

i. *phen-mi'* [*jäe'mä iits* [*Kose-**si'*** *jike*
woman-3M.SG DM DE.M Jose-NMLZR PS
Toreno-win]]… (NP-use)
Tureno-C
'his wife, the late Jose Tureno's'

ii. *mö'* [[*aka'* *jäechbäe'-sï*] *Martin-**si'***] (Modification-use)
3F.SG house red-NMLZR Martin-NMLZR
'Martin's red house'

While Sakel (2004: 106) translates (7-9b.i) above as "the wife of the late Jose Tureno" in a modification pattern, a literal interpretation seems to be the one suggested for this form, where the possessor form is understood to be in an NP-use in apposition with the preceding noun phrase "his wife". This seems correct in view of the occurrence of the demonstratives in (7-9b.i). As seen in (7-9a.ii) and (7-9b.ii), modifying nominalizations are not marked by demonstratives.

The other case marking N-based nominalization by the same nominalizer that marks V-based nominalization concerns the marking by the Tapiete nominalizer *-wa*, seen earlier (5-9). The situation is not as straightforward as the Mosetén case above. The NP-use of possessive forms involves the extra morpheme *a'ámpo*, which Gonzaléz (2005: 243) simply glosses as "POSS", as below.

(7-10) Tapiete (Gonzaléz 2005: 114, 243)[76]
*mbo'u she sh-a'ámpo-**wa*** (NP-use)
send I 1SG.POSS-POSS-NMLZR
'Pass me mine.'

76. See the Newar forms in (7–91) below that also show a doubling of nominalization markers.

sh-anka (Modification-use)
1SG-head
'my head'

Turning to languages outside the Americas, first observe the comparable marking pattern in the Niger-Congo language Yoruba in West Africa, where the nominalizer for V-based nominalizations has high tone, whereas the counterpart for N-based ones has mid tone.

(7-11) Yoruba (Ajiboye 2005)
a. V-based nominalization
Mo ri eyi [***tí*** *Kúnlé ni*]
I see this NMLZR Kunle own
'I saw the one that Kunle owns.
Cf. [*ère* [***tí*** *Kúnlé ni*]] (90)
statue NMLZR Kunle own
'the statue that Kunle owns'
b. N-based nominalization
Mo ri [***ti*** *Kúnlé*] (107)
1SG see NMLZR Kunle
'I saw Kunle's.'
Cf. [*ère* [***ti*** *Kúnlé*]]
statue NMLZR Kunle
'Kunle's statue'

The Kushitic language Kambaata marks nominalization by a final accent in both V-based and N-based nominalizations.

(7-12) Kambaata (Cushitic; Ethiopia; Treis 2008)
a. V-based nominalization
i. [[*dagujj-ó*] *adab-áa*]
run-3M.PVO.NMLZR boy-M.ACC
'the boy who ran'
ii. [[*xuujj-o-sé*] *adab-áa*]
see-3M.PVO-3F.OBJ.NMLZR boy-M.ACC
'the boy who saw her'
b. N-based nominalization
i. [*ann-í*] *hiz-óo*
[N-M.GEN.NMLZR] N-M.ACC
'father's brother'
ii. [*ann-i-sé*] *hiz-óo*
[N-M.GEN-POSS.NMLZR] N-M.ACC
'her father's brother'

iii. [*ann-aakk-a-sé*] *hiz-óo*
[N-PL-F.GEN-POSS.NMLZR] N-M.ACC
'her fathers' brother'

The Austronesian language Lamaholot in eastern Indonesia makes use of another kind of suprasegmental feature to mark nominalization of both verbs and nouns, namely nasality, which is likely connected to the Proto-Austronesian genitive determiner **ni*. Observe:

(7-13) Lamaholot (Austronesian; Nagaya 2011: 194, 200)
a. V-based nominalization
go hope [*meʔə̃*] / [*topi meʔə̃*].
1SG buy red.NMLZR / hat red.NMLZR
'I bought the red one/red hat.'
b. N-based nominalization
go gute Hugo nəʔẽ (*hepe*).
1SG take Hugo 3SG.NMLZR knife
'I will take Hugo's (knife) [lit. 'Hugo his (knife)'].'

The Indo-Aryan language Nepali, likely influenced heavily by surrounding Tibeto-Burman languages, marks both V-based argument nominalizations and event nominalizations the same way as N-based nominalizations, as shown below.

(7-14) Nepali (courtesy of Madhav Pokharel)[77]
a. V-based event nominalization
[[*u Dhilo aa-e*]=***ko***] *durbhaagya bha-yo.*
s/he late come-PF-NMZLR unfortunate BE-PFV
'That s/he came late was unfortunate.'
b. V-based argument nominalization
[[*Madhav le ma laai di-e*]=***ko***] *ma paDh-dai chu*
Madhav ERG I DAT GV-PF-NMLZR I read-PROG am
'I am reading what Madhav gave me.'
Cf. [[[*Madhav le ma laai di-e*]=***ko***] *kitab*]
Madhav ERG I DAT GV-PF-NMLZR book
'the book that Madhav gave me'
c. N-based nominalization
[*Madhav*=***ko***] *ma paDh-dai chu*
Madhav=NMLZR I read-PROG am
'I am reading Madhav's.'

77. To appreciate a systematic different between nominalizers (NMLZRs) and NP-use markers (NPMs), compare these Nepali examples with a NMLZR and the Telugu examples in (8–1) with an NPM.

Cf. [[*Madhav=**ko***] *kitab*]
Madhav=NMZLR book
'Madhav's book'

A comparable pattern is seen in Modern Hebrew, in which *she* marks a similar range of nominalizations as in Nepali (also see Shibatani & bin Makashen (2009) for another Semitic language Soqotri).

(7-15) Modern Hebrew (courtesy of Ana-Marie Hartenstein)
a. V-based nominalization
i. *Ani yodaat* [***she*** [*ata lo bemet rofe*]]
I know NMLZR you no real doctor
'I know that you are not really a doctor.'
ii. *zo* [***she*** [*Yoav raa etmol*]] *xi xavera sheli*
this.FEM NMLZR Yoav saw yesterday is friend my
'The one Yoav saw (her) yesterday is a friend of mine.'
b. N-based nominalization
[*Ha-kova* [***she**-l* *Moshe*]] *shachor, aval* [***she**-l*
ART-hat NMLZR-DAT? Moshe black but NMLZR-DAT?
Yakov] *chum*.
Yakov brown.
'Moshe's hat is black but Yakov's is brown.'

Next, those languages that use classifiers as nominalizers may mark both V-based and N-based nominalizations by classifiers, as in Cantonese below.

(7-16) Cantonese (Matthews & Yip 1994: 108, 111, 112)
a. V-based nominalizations marked by classifiers
i. [*Ngóhdeih hái Faatgwok sihk*] ***dī*** *yéh géi hóu-sihk ga*.
we in France eat CLF food quite good-eat PRT
'The food we ate in France was pretty good.'
ii. [[*Gaau léih tàahn kàhm*] *gó*] ***go***?
teach you play piano that CLF
'The one who teaches you piano?'
b. N-based nominalizations marked by classifiers
i. [*léih (gó)* ***dī***] *pàhngyáuh*
you (that) CLF friend
'those friends of yours'
ii. [*léih* ***go***] *pàhngyáuh*
you CLF friend
'your friend'

iii. *Lī dēoi* *hai* [ngóh ***dī*** *phàngyáuh*], [*léih gó* ***dī***]
these (lit. This pile) COP I CLF friend you that CLF
hóeng gópihn. (courtesy of Haowen Jiang)
LOC there
'These are my friends, and yours are over there.'

Similar use of classifiers is also seen among Amazonian languages, as shown by the Tucano language Barasano (also known as Pãnerã) in Colombia.

(7-17) Barasano (Tucano; Jones & Jones 1991: 61,150)
a. V-based nominalization
[*hũʉ̃* [*ō* *kãhi-ri-**kʉ̃***] *ãbo-a-ha* *yʉ̃*
hammock there hang-NMLZR-CLF want-PRS-3 1SG
'I want the hammock that is hanging there.'
b. N-based nominalization
[*hũʉ̃* [*ĩ-ya-**gʉ̃***]]
hammock 3MASC.SG-NMLZR-CLF
'his hammock'

Barasano has different nominalizers for V-based (*-ri*) and N-based (*-ya*), as do many other languages. However, the results of the nominalizations are treated alike, as indicated by use above of the same classifier marking *kʉ̃/gʉ̃*, which is for a long hammock. Both V-based and N-based forms have NP-use such that (7-17a), without the head noun *hũʉ̃*, would mean "I want one (hammock-thing) hanging there" and (7-17b) "his (hammock-thing)".

In the related language Tuyuca, classifier marking is optional for the N-based nominalization in modification-use, while it is obligatory in the NP-use, as seen below.[78]

(7-18) Tuyuca (Tucano; courtesy of Janet Barnes)
a. V-based nominalization
[*nɨ́ká* [*bako-á-ri-**gɨ́***]]
leg to.have.been.bitten-RECENT-SG.NMLZR-CLF
<cylindrical.shape,long.and.solid>
'the leg that was bitten'
b. N-based nominalization
i. $[[yɨ́ pakɨ́\text{-}ya\text{-}\textbf{ró}]_{\text{NMLZ}}]_{\text{NP}}$
my father-NMLZR-CLF:2D.FELXIBLE
'my father's'(as in "They are my father's/My father's are those.")

78. In Bora, only nominal-based nominalizations appear to be marked by classifiers only in their NP-use.

ii. [[*yɨɨ pakɨ-ya*(-***ro***)]$_{\text{NMLZ}}$ *sirúra*]$_{\text{NP}}$
my father-NMLZR(-CLF:2D.FLEXIBLE) trouser
'my father's trousers'

This is the pattern that we find in the Tibeto-Burman language Newar, which has classifier-based nominalizers. Observe the data below, where an N-based form has its own nominalizer (*-yā*), but it further takes the nominalizer marking V-based nominalization (*-mha*), indicating that N-based nominalizations are treated like V-based nominalizations.

(7-19) Newar (courtesy of Kazuyuki Kiryu)

a. [[*ana dan-ā cwā*=***mha***] *macā*] [*rām=yā*]=***mha*** *kha:*.
there stand-CM exist.ND=NMLZR child Ram=NMLZR=NMLZR COP
'The child standing over there is Ram's.'

b. [*ana dan-ā cwā:*=***mha***] [*rām=yā*(=***mha***)] *macā*] *kha:*.
there stand-CM exist.ND=NMLZR Ram=NMLZR(=NMLZR) child COP
'The one standing there is Ram's child.'

Finally, Bantu noun-class marking, which also has a nominalizing function,[79] marks both V-based and N-based nominalizations, as shown by the Chichewa examples below.

(7-20) Chichewa (Mchombo 2004 and p.c.)

a. V-based nominalization in NP-use
[[***chi***-*méné ndí-ná-gúla*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *chi-ná-lí ch-ódúla*.
7-NMLZR I-PST-buy 7-PST-be 7-expensive
'What I bought was expensive.'

a′. V-based nominalization in modification-use
[*chi-péwá* [***chi***-*méné ndí-ná-gúla*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *chi-ná-lí ch-ódúla*.
7-hat 7-NMLZR I-PST-buy 7-PST-be 7-expensive
'The hat that I bought was expensive.'

b. N-based nominalization in NP-use
[[***ch***-*ángá*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *chí-ma-sangaláts-á a-lenje*.
7-my 7-HAB-please-FV 2-hunters
'Mine pleases hunters.'

b′. N-based nominalization in modification-use
[*chi-péwá* [***ch***-*ángá*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *chí-ma-sangaláts-á a-lenje*.
7-hat 7-my 7-HAB-please-FV 2-hunters
'My hat pleases hunters.'

79. Cf. Digo forms; *ku-fwits-a* [15-hide-fv] 'hiding', *m-ris-a* [1-feed-fv] 'herdsman', *chi-tsek-o* [7-laugh-fv] 'laughter', *chi-digo* [7-Digo]'Digo language/culture'. Nicolle (2013)

All in all, there is ample evidence that languages around the globe also nominalize nouns and noun phrases.[80] Our reanalysis of the so-called genitive case as an N-based nominalizer not only captures the parallel patterns exhibited by V-based and N-based nominalizations we have examined above, but also offers a new insight into the meaning relationship between the "possessor" and the "possessum". Those various "possessive" relations (*John's head*, *John's father*, *John's letter*, *John's hospital*, etc.) are attributed to the metonymic inferences associated with a nominalization process creating the so-called possessive form, *John's*. Modification in possessive constructions means restricting the denotation of the head noun (e.g. *book*) with that of the N-based nominalization (e.g. *John's*), with the latter denoting things with which John is associated and specifying a subset of the former. The denotation of an entire possessive construction (e.g. *John's book*) is an intersection of the denotation of a head noun and that of an N-based nominalization, in the same way as the so-called restrictive relative clause restricts the denotation of the head noun (see Section 5.5).

Traditional grammars treat the genitive case/possessive form as an inflection similar to a grammatical case such as nominative and accusative. The genitive, however, differs from grammatical case in that it either modifies another noun (*his/John's car*) or stands in NP positions where case inflected forms cannot freely stand (*His/John's/*Him is here*; *I saw his/John's/*he*), showing that it is different from ordinary case forms, the syntactic positions of which are fixed – a nominative form in subject position, an accusative form in object position. The genitive form, in contrast, is not bound to any particular syntactic position, and even to the modifier position in possessive constructions, as the foregoing discussions amply demonstrate.[81] More importantly, the denotation of the nominal in different case forms remains constant, with case inflections adding only grammatical meanings. For example, the nominative *I* marks a subject function, and the accusative *me* an object function, while maintaining their denotation constant, namely the speaker. This is not so with the genitive form; *mine* does not denote the speaker but rather something metonymically related to the speaker. The same applies to languages using particles to mark case, as in Japanese forms *boku ga* (I NOM) and *boku o* (I ACC) vis-à-vis *boku no* (*o mite*) '(look at) mine'.

The nominalization analysis proposed above treats the genitive case form as derivationl, as nominal-based nominalization, similar to the derivations of *pig* > *piglet*, *village* > *villager*, which derive new nouns with new denotations, unlike the inflections such as *he*/*him*, *pig*/*pigs*, which do not change the denotations,

80. See next section on forms such as *my/mine* that are associated with specific functions.

81. Cf. $[[[\textit{the Queen of England}]_{\text{NP}}\textbf{'s}]_{\text{NMLZ}}\ \textit{hat}]_{\text{NP}}$

but which differ only in grammatical meanings (subject/object, singular/plural). The same applies to Japanese, Korean and other languages that make use of particles. While Japanese particles *ga* (nominative), *o* (accusative), *ni* (dative) are case particles, the so-called genitive case particle *no* is not. Like the English possessive clitic *'s*, it is derivational; *John no* does not denote/refer an individual named John but things that are metonymically related to that person.

8. Structure, use, and form

One of the most interesting facts to observe in crosslinguistic research is the way languages respond to unity and divergence of function in terms of linguistic forms. The functional unity underlying nominalization phenomena is the creation of nominal structures denoting entity concepts. The divergences stem from several factors. One is the difference in input, i.e. verbal-based or nominal-based. The outputs of the nominalization process are of several types. We have distinguished between lexical and grammatical nominalizations, the former whose form-meaning relationships are entrenched and whose grammatical properties are highly similar to ordinary nouns, and the latter non-lexical grammatical structures created for the nonce. Of the grammatical nominalizations, there are event nominalizations and argument nominalizations. These nominalizations play different functions depending on their use, in particular whether they head a noun phrase, where they play referential function, or they modify a head noun, either restricting the denotation of the head noun (so-called restrictive relative clauses) or identifying the denotation of the head noun (so-called non-restrictive or appositive relative clauses). Careful observations of the data discussed above indicate that languages respond to these functional similarities and divergences in different ways. An interesting question to be raised is whether general crosslinguistic patterns emerge on the basis of which we might be able to make predictions about change in form over time.

As for the distinction between lexical and grammatical nominalizations, languages often make a clear formal distinction between the two. Yet, we have seen that a fair number of languages do formally express the functional unity between the two by marking both types in a morphologically uniform way. Indeed, in some cases the form is ambiguous allowing either a lexical or grammatical interpretation. For example, the Central Alaskan Yup'ik forms in (2-2) and the Mayrinax Atayal form in (3-12) can be interpreted either lexically (in the sense of the word, e.g. "child" or "singer") or grammatically (in the sense of "one who is small" or "one who sings"). In situations like this, it is likely that grammatical nominalizations give rise to lexical nominalizations, where a form denoting an entity in an analytic manner has been applied to an entity whose meaning is not entirely

compositional, as in the case of designating a person who sings routinely or whose singing constitutes a professional activity. The opposite direction of development, where the marking of lexical nominalizations has been extended to grammatical nominalizations needs to be documented.[82]

Turning to the distinctions between nominal-based and verbal-based nominalizations and between event and argument nominalizations, many languages of the world make clear formal distinctions in them. But, again, a fair number of languages from different parts of the globe formally express functional unity by morphologically marking them in a similar way. We recognize two patterns of formal identity across these types of nominalization. One pattern expresses the fundamental functional unity that binds all these types of nominalization (i.e. a marker indicating that they are all nominalizations), using the same nominalizing morphology for all of them, as in Nepali (see (7-14)) and Modern Hebrew (7-15), among others. The other, perhaps more widespread pattern responds to the commonality in their use function by marking the same all these types of nominalization (only) when they are in NP-use. This can be seen most clearly in the Telugu forms below, where the particle *di* marks the shared referential function of nominalizations in NP-use in a uniform manner, as below. Some other Dravidian languages, Korean (7-5), Gã (7-6), and a variety of Ryukyuan (Shibatani & Shigeno 2013) and Japanese dialects (Shibatani 2017) show this marking pattern.

(8-1) Telugu (courtesy of K. V. Subbarao)

a. N-based nominalization in NP-use
*idi naa-**di**.*
this my-NPM
'This is mine.'
Cf. *naa pustakam*
my book
'my book'

b. V-based event nominalization in NP-use
[*vāḷḷu vacc-in-a*]-***di*** *naaku telusu.*
they.NOM come-PST-NMLZR-NPM to me known
'I know that they came.'
Cf. [*vāḷḷu vacc-in-a*] *sangati*
they.NOM come-PST-NMLZR news
'the news that they came'

82. The English grammatical nominalizations of the form [singing the national anthem] appears to be an innovation based on the lexical counterparts involved in [singing of the national anthem] and the like. The parallel grammatical nominalizations do not obtain in other Germanic languages like German and Swedish.

c. V-based argument nominalization in NP-use
neenu [*vāḷḷu icc-in-a*]-***di*** *cadiveenu*.
I.NOM they NOM give-PST-NMLZR-NPM read
'I am reading what they gave (me).'
Cf. [*vāḷḷu icc-in-a*] *pusutakam*
they.NOM give-PST-NMLZR book
'the book which they gave (me)'

Many languages make a clear formal distinction between N-based and V-based nominalization, and between V-based event nominalization and V-based argument nominalization. English and many others have special forms (known as the genitive case or possessive form) for N-based nominalization distinct from those for V-based nominalizations (traditionally referred to as gerundive, participial, infinitive, or adnominal). A comparison of Portuguese and Spanish pronoun-based nominalizations reveals how different languages respond differently to the functional demand at two different levels. Portuguese, like Italian, maintains the formal uniformity in favor of formally expressing the functional unity underlying the pronoun-based nominalizations – that the relevant forms are of the same substance regardless of their use; NP-use: *Olhe para o* [***meu***] 'Look at mine'; Modification-use: *Olhe para* [***meu*** *livro*] 'Look at my book'.

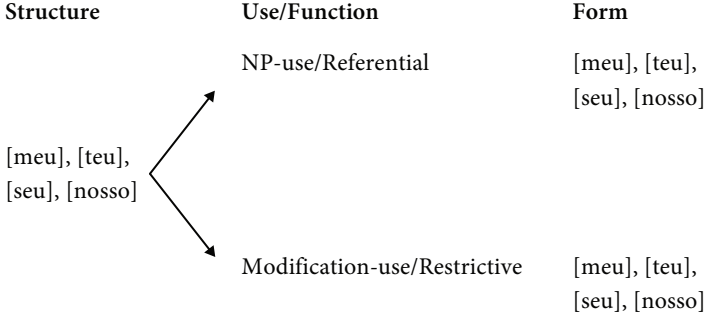


Figure 5. Portuguese pronoun-based nominalizations for first- (*meu*), second- (*teu*), third-person (*seu*) singular masculine froms and first-person plural masculine form (*nosso*)

On the other hand, Spanish, like French (and English), differentiates forms according to their use/function so as to express formally the difference in the usage and function; MP-use: *Mira el* [***mío***] "Look at mine"; Modification-use: *Mira* [***mi*** *libro*] "Look at my book" (see next page).

When a language develops usage-based specialized forms, they are likely to acquire new grammatical properties, just as the Spanish NP-use form *mío* or its English counterpart *mine* cannot modify a noun. Recall that Northern Paiute

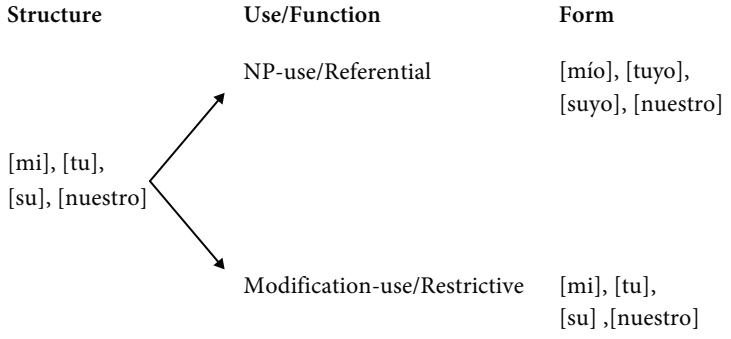


Figure 6. Spanish pronoun-based nominalizations for first- (*mi*), second- (*tu*), third-person (*su*) singular masculine forms and first-person plural masculine form (*nuestro*)

event nominalizations marked by *-na* allows an adverbial function, including chaining constructions (see (5-58)). The forms performing this chaining function denoting/referring to simultaneous events are identical with those functioning as verb complements, and so the *-na* forms bearing the adverbial and V-complement NP function look alike. Northern Paiute has developed a specialized form marked by *-si* for events holding a sequential temporal relation (see (5-62)). Such a specialized form would not be able to function as an NP-head functioning as a syntactic argument.

A similar but a subtler pattern is observed about the specialized adverbial nominalizations marked by the conjunctive particle *-te*/*-de* in Japanese (see (5-57)). While they show a measure of syntactic nominality in patterning like ordinary nouns in the formation of noun-modifying forms, they are unlike ordinary nouns or straightforward event nominalizations in that they cannot head an argument NP. Compare:

(8-2) Japanese
  a. Noun and nominalization-based adverbial in modification-use
    i. [[[*eiga*] ***no***] *kaeri-miti*
      movie NMLZR return-road
      (lit.) 'a return-road of a movie'/ 'a return road taken after a movie'
    ii. [[[*eiga o mi=te*] ***no***] *kaeri-miti*
      movie ACC watch=CON NMLZR return-road
      (lit.) 'a return road of having watched a movie'/ 'a return road taken after having watched a movie'
  b. Noun and nominalization-based adverbial in NP-use
    i. [[*Eiga*]]$_{NP}$ *wa omosiroi*.
      movie TOP fun
      'A movie is fun.'

ii. *[[*Eiga o mi=te*]] *wa omosiroi.*
movie ACC watch.NMLZR=CON TOP fun
(lit.) 'Having watched a movie is fun.'

Caution must be taken in that even those forms that function as NP-heads may not behave alike since NPs themselves can function differently. The Portuguese pronoun-based nominalizations discussed above, such as *meu* 'my/mine (MASC)' and *minha* 'my/mine (FEM)', display different syntactic properties depending on different uses of NPs they head. When they head an NP functioning as a syntactic argument, they require a definite article, but when they head an NP functioning as a nominal predicate, article marking is optional. Compare:

(8-3) Portuguese
a. [*O meu*]$_{NP}$ *é aquele carro.*
the.MSC 1.SG.MSC.NMLZR is that car
'Mine is that car.'
b. Aquele carro *é* [(*o*) *meu*]$_{NP}$
that car is (the.MSC) 1.SG.MSC.NMLZR
'That car is mine.'

8.1 From NP-use markers to nominalizers

As a way of showing the importance of N-based nominalizations, the remainder of this subsection examines crosslinguistic marking patterns in grammatical nominalizations. Shibatani's study of Ryukyuan and Japanese dialects (Shibatani & Shigeno 2013) shows that forms in NP-use tend to be more complex than the ones seen in modification-use, the former with an explicit marker for the referential function that the forms in NP-use bear (cf. Spanish *mi* 'my' vs. *mío* 'mine' and their English equivalents). Over time, however, the complex forms in NP-use spread to the modification context, leveling the formal difference between the two uses. Shibatani and Shigeno (2013) also show that once a language achieves uniformity in formal marking across different uses of the same type of nominalization, it begins to differentiate forms according to the difference in use, by adding an extra marker to the forms in NP-use. This cyclic development seems to be one way for a language to negotiate with the opposing forces for formal uniformity (revealing the underlying unity) and divergence (expressing the difference in usage/function).

Shibatani and Shigeno's work demonstrates that the NP-use of N-based nominalization is the site where innovations start out, which then spread to the modification-use of N-based nominalizations, as well as to the NP-use of V-based nominalization, and then to their modification-use, as in the manner below.

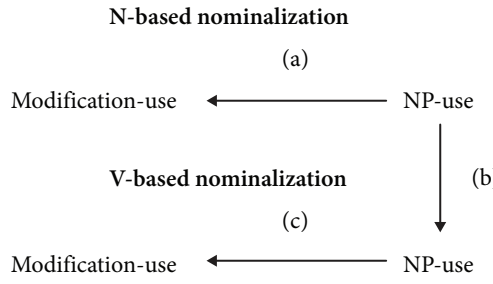


Figure 7. Directions of spread of nominalization markers

The (a) pattern of development is clearly seen in a recent innovation in Ūyama Okinawan on the main island of Okinawa in the Ryukyuan archipelago. There are two NP-use markers, *si* and *mun*(*u*), seen in both N-based and V-based nominalizations that have a wide geographical distribution in the archipelago. The fact that *si* and its phonological variants occur only in peripheral areas such as southern Amami, Yoron, and Kumejima in the north and Ishigaki and Taketomi in the south indicate that it is older than *mun*(*u*), whose occurrence with N-based nominalizations is seen in both central Ryukyuan islands and peripheral areas. The fact that while the etymology of *si* is unknown, *mun*(*u*) is trasparently connected to the noun *mun*(*u*) 'thing' corroborates the observation that the *mun*(*u*) marking is a newer development.[83] The marking patterns exhibited by *si* and the more widespread use of *mun*(*u*) for N-based nominalizations are shown below.

(8-4) Yoron Ryukyuan (Yoron Island, Kagoshima prefecture, Japan)
a. NP-use of N-based nominalization
[*sinsee nu*]=***si*** *ja are*
teacher NMLZR=NPM TOP that
'The teacher's is that.'
b. Modification-use of N-based nominalization
[*sinsee nu*] *hasa*]
teacher NMLZR umbrella
'teacher's umbrella'

(8-5) Agena Okinawan (Okinawa Island, Okinawa prefecture, Japan)
a. NP-use of N-based nominalization
ʔure: [*sinsi: ga*] ***mun***=*do:*.
that teacher NMLZR NPM=COP
'That is the teacher's.'

83. The use of the marker *-si* is attested in *Omorosōshi*, a collection of Ryukyuan songs compiled between 1531 and 1623, while the beginning use of *mun*(*u*) is unclear, largely due to the difficulty in determining whether *mun*(*u*) is functioning as a noun with the meaning of "thing" or as an NP-use marker.

b. Modification-use of N-based nominalization
ʔure: [[*sinsi: ga* *kuc*u]=*do:*.
that teacher NMLZR shoe=COP
'Those are the teacher's shoes.'

The following data show that the (a) pattern of development illustrated in Figure 7 is seen in the Ūyama dialect of Okinawan, where the *mun*-marking is spreading to the modification context, as seen in (8-6c) below, where it has replaced the *N* nominalizer.

(8-6) Ūyama Okinawan (Okinawa Island)
a. *Unu kucu ja* [*waN*] ***mun***=*do.*
that shoe TOP I.NMLZR NPM=COP
'Those shoes are mine.'
b. [[*waN*] *kucu*]
I.NMLZR shoe
'my shoes'
c. [[*wa* ***mun***] *kucu*]
I NMLZR shoe
'my shoes'

The replacement of the *N* nominalizer by *mun* has not yet developed to its full extent in that only first and second person pronominals can be marked by *mun* in the modification context. The fact that the neighboring dialect on Tsuken Island does not show this pattern of spread of the NP-use marker *mun* to the modification context shows that the above is a true innovation in Ūyama Okinawan.

A similar development illustrating the (a) pattern in Figure 7 is seen in Creek, which involves the noun root etymologically meaning 'thing' in the formation of the N-based nominalizations in NP-use, as below.

(8-7) Creek (Martin 2011: 144)
i. *ca-ná:ki* 'mine'
1SG-THING
ii. *ci-ná:ki* 'yours'
2-THING
iii. *i – ná:ki* 'his/hers/theirs'
3-THING
iv. *po-ná:ki* 'ours'
1PL-THING
v. *ca-ná:ki-t* *ô:-s*
1S.PAT-THING-T be.FGR-IND
'It's mine.'

These forms contrast with those in the basic modification pattern, which marks the modified noun by personal prefixes directly, as below.

(8-8) a. relational noun
  i. *ca-hácko* 'my ear'
  ii. *ca-cá:ta* 'my blood'
 b. non-relational noun
  i. *am-ifa* 'my dog'
  ii. *am-mí:kko* 'my chief'

However, Creek allows what Martin (2011: 137–138) calls periphrastic possession using the NP-use forms in (8-7). Observe:

(8-9) a. *toł-sakká:ka* [***ca-ná:ki***]
  eye.glasses 1S.PAT-THING
  'my glasses'
 b. *ifá* [***ca-ná:ki-n***] *î:st* *o:w-éy-s*
  dog 1S.PAT-THING-N take.SG.FGR-T be.1SG-1S.AG-IND
  'I'm holding my dog.'

According to Martin (2011: 138), some speakers prefer these periphrastic possessive forms for some non-relational nouns over the prefixed forms, as in (8-8), indicating that a shift from prefixed forms to periphrastic forms may be happening, as a way of leveling the formal difference between the NP-use form (8-7) and the modification-use form (8-8); i.e. from the pattern ***ca-ná:ki*** 'mine': ***am***-*ifa* 'my dog' to the pattern ***ca-ná:ki*** 'mine': *ifá* ***ca-ná:ki*** 'my dog'.[84]

The (b) pattern of spread of NP-use markers in Figure 7 shows the importance of N-based nominalizations, whose innovations spread to V-based nominalizations. Ryukyuan languages also illustrate this pattern of development. As noted above, the innovative replacement of the *si* marker by *mun*(*u*) in the NP-use of N-based nominalizations among Ryukyuan languages appears almost complete except for those peripheral islands mentioned above. This replacement pattern has spread to the NP-use of V-based nominalizations in a wide range of Ryukyuan languages except for the Okinawan dialects on the main island of Okinawa, where while *mun*(*u*) has replaced *si* in the NP-use of N-based nominalization, that of V-based nominalizations retain the older *si*-marker. Compare the following sets of data from Agena Okinawan and Irabu Ryukyuan of the Miyako archipelago in

84. A similar use of the noun meaning "thing" as an NP-use marker for N-based nominalizations is also seen in Thai, where the marker *khɔ̌ŋ* with the lexal meaning of "thing" obligatarily marks NP-use. The spread of this marking to the modification context is again secondary in that it remains optional.

the south; in the latter the *munu*-marking has spread from the NP-use of N-based nominalization to that of V-based nominalizations.[85]

(8-10) Agena Okinawan (courtesy of Shigehisa Karimata)

a. N-based nominalization

i. *ʔure:* [*sinsi: ga*] ***mun****=do:*. (NP-use)
that teacher NMLZR NPM=COP
'That is the teacher's.'

ii. *ʔure:* [[*sinsi: ga*] *kucu*]*=do:*. (Modification-use)
that teacher NMLZR shoe=COP
'Those are the teacher's shoes.'

b. V-based nominalization

i. [*ʔama=nkai tattʃo:****se:*** (< *-****si*** *ja*)] *wa:*
there-LOC stand.PROG.NPM.TOP (NM TOP) I.NMLZR
ikigangwa. (NP-use)
son
'The one standing there is my son.'

ii. [[*ʔama=nkai tattʃo:-ru*] *ttʃo:* (< *ttʃu ja*)] *wa:*
there=LOC stand.PROG-NMLZR person.TOP person TOP I.NMLZR
ikigangwa. (Modification-use)
son
'The person who is standing there is my son.'

(8-11) Irabu Ryukyuan (Miyako archipelago; Okinawa Prefecture, Japan; Shimoji 2008 and p.c.)

a. N-based nominalization

i. *uri a* [*ba ga*] ***munu***. (NP-use)
that TOP I NMLZR NPM
'That is mine.'

ii. *uri a* [[*ba ga*] *zin*] (Modification-use)
that TOP I NMLZR money
'That is my money.'

b. V-based nominalization

i. [*kama n taci-ur*] ***munu*** *u ba ga biki-vva.* (NP-use)
there LOC stand-PROG NPM TOP I NMLZR male-child
'The one standing there is my son.'

85. In addition to the *munu*-marked form in (8-10a), Irabu retains older nominalization forms involving the particle *si* as well as the even older verbal nominalizer, known as *rentai-kei* 'ad-nominal form' in Japanese linguistics.

ii. [[*kama n taci-ur*] *ffa*] *u ba ga*
threre LOC stand-PROG child TOP I NMLZR
biki-vva. (Modification-use)
male-child
'The child who is standing there is my son.'

Finally, we shall examine crosslinguistic marking patterns that show the development pattern (c) in Figure 7, namely the spread of a marker from the NP-use of V-based nominalization to the modification-use. Recall that Toba allows V-based nominalizations without any nominalization marker, which, however, must be marked by a demonstrative determiner when they head an NP, similar to the Portuguese pronoun-based nominalization in NP-use seen earlier. In the modification-use, however, a demonstrative determiner is not used, as in (8-12b) below.

(8-12) Toba (courtesy of Cristina Messineo)
a. NP-use
[***so*** [*neta'age da Chaco*]$_{\text{NMLZ}}$]$_{\text{NP}}$ *i-waGan so Juan*
DD 3.exist.DIR DD Chaco 3-hit DD Juan
'The one who lives in Chaco hit Juan.'
b. Modification-use
[[*so ſijaGawa* [***Ø*** [*neta'age da Chaco*]$_{\text{NMLZ}}$]$_{\text{NP}}$ i-*waGan so Juan*
DD man 3.exist.DIR DD Chaco 3-hit DD Juan
'The man who lives in Chaco hit Juan.'

The demonstrative determiner marking in NP-use has not yet been extended to the modification context above, where the Ø marker indicates its absence in (8-12b).[86]

When we turn to K'ichee', we see that the determiner marking of V-based nominalization in NP-use has been extended to the modification context, as seen below.

(8-13) K'ichee' (courtesy of Telma Can Pixabaj)
a. *x-Ø-inw-il* ***lee*** [*ixoq*]
ASP-3SG.ABS-1SG.ERG-see the woman
'I saw the woman.'
b. *x-Ø-inw-il* ***lee*** [*x-Ø-u-ch'ay lee*
ASP-3SG.ABS-1SG.ERG-see the/NMLZR ASP-3SG.ABS-3SG.ERG-hit the
achih]$_{\text{NMLZ}}$
man
'I saw the one whom the man hit/I saw the one who hit the man.'

86. Presumably it is possible to use a DD in the place of Ø in (8-12b), but it would result in a structure different from a restrictive relative clause construction. See Section 5.5.

c. *x-Ø-inw-il* [*lee ixoq* [***lee***
ASP-3SG.ABS-1SG.ERG-see the woman the/NMLZR
[*x-Ø-u-ch'ay* *lee achih*]$_{NMLZ}$]
ASP-3SG.ABS-3SG.ERG-hit the man
'I saw the woman whom the man hit/I saw the woman who hit the man.'

The use of *lee*, or its dialectal variants, in the modification context as in (8-13c) does not seem entirely obligatory at present, though its use appears highly favored according to the investigation of its status by a K'ichee' specialist known to the present author. In other words, the determiner *lee* is in a final stage of becoming a nominalizer/relativizer, so that V-based nominalizations become formally uniform in both the contexts of NP-use and modification-use, as in (8-13b) and (8-13c). Compare these with the Toba forms in (8-12), where the forms of V-based nominalizations are distinguished according to the usage pattern. K'ichee' would eventually reach the stage where the determiner *lee* becomes an obligatory nominalizer/relativizer as in German, which has developed nominalizers out of demonstrative pronouns. Observe the following where *der* marking a V-based nominalization is obligatory in both NP-use (8-14a) and modification-use (8-14b).

(8-14) German
a. [*Der* [***der*** *morgen kommt*]$_{NMLZ}$]$_{NP}$ *ist*
ART.MSC.SUB SUB-NMLZR.MSC tomorrow comes is
mein Freund.
my friend
'The one who (MSC) comes tomorrow is my friend.'
b. [*Der Mann* [***der*** *morgen kommt*]$_{NMLZ}$]$_{NP}$ *ist*
ART.MSC.SUB man SUB-NMLZR.MSC tomorrow comes is
mein Freund.
my friend
'The man who comes tomorrow is my friend.'

The above discussions identify three sources of nominalization markers, namely (i) particles whose origins are unclear as in the cases of the Korean *kes* and the nominalizers for N-based nominalization (aka genitive particles) *no* (and *ga*) in Japanese, (ii) a lexical source in terms of nouns meaning "thing" or others meaning "person" (Gã *mɔ̀* 'person'), "matter" (Japanese *koto*, not discussed in this paper), "place" (Thai *thîi*)), and (iii) determiners/demonstratives. At least some of these clearly start out as NP-use markers, and some of them have already become nominalizers, marking all occurrences of grammatical nominalizations as such.

Besides the nominalization markers above, many languages of the world use noun classifiers as makers of grammatical nominalizations, some of which have become nominalizers. First of all, the widely held understandings of classifiers

miss the mark and fail to recognize their nominalization function.[87] Even those who recognize the nominalizing function of classifiers do not appear to properly grasp its scope due to a limited understanding of what nominalization is all about.[88] Because of this, we spend the next several pages detailing the nominalization function of classifiers and showing that classifier-marked nominalizations behave exactly like ordinary nominalizations in both NP- and modification-use. The marking patterns in these two functions also parallel between classifier-marked nominalizations and ordinary nominalizations, as we shall see. The essential difference between ordinary nominalizers and (nominalizing) classifiers is simply that the former derive nominals with a new denotation without classifying them, the latter both nominalize and classify derived nominals into different categories based on the nature of the new denotation, such as shape, size, consistency, animacy, gender, and function. The nominalization function of classifiers is clearly seen when they derive new nouns (lexical nominalizations), as in the Hmong and Barasano examples below.

(8-15) White Hmong (Riddle 1989; Jaisser 1987: 173)
- a. V-based
  - i. *said* 'see' > ***tus*** *said* 'supervisor'[89]
  - ii. *sau* 'write' > ***tus*** *sau* 'writer'
- b. N-based
  - i. *dej* 'water' > ***tus*** *dej* 'river'
  - ii. *ntawv* 'paper' > ***phau*** *ntawv* 'book'
  - iii. *ntawv* 'paper' > ***tsab*** *ntawv* 'letter'

87. Observe the following; "[a] classifier denotes some salient perceived or imputed characteristic of the entity to which an associated noun refers (or may refer)." (Allen 1977: 285). "Numeral classifiers are morphemes that only appear next to a numeral, or a quantifier; they may categorize the referent of a noun in terms of its animacy, shape, and other inherent properties."(Aikhenvald 2006: 466) Numeral classifiers actually apply to verbs, nouns, demonstratives as well as grammatical nominalizations; see below). Compare these definitions with the following characterization of Barasano classifiers by Jones & Jones (1991: 49): "Barasano has an extensive system of noun classifiers, which provide concordance (agreement) within the noun phrase, and *are used to form referring expressions that head noun phrases*." (Emphasis added).

88. Payne (1985: Chapter 4), who clearly recongnizes the nominalistion function of classifiers in Yagua, talks about the derivational (nominalizing) function and the inflectional function of classifiers, the latter of which is actually no more than modification-use of classifier-marked nominalizations (see (8-17b) and (8-18b)).

89. *tus* = classifier for human beings, animals, things that closely affect people, and things that come in short length; *phau* = stacks of things and things piled up on each other; *tsab* = written messages.

(8-16) Barasano (Jones & Jones 1991)

a. V-based

i. *bōa-**gʉ*** 'a male worker'
work-PROX.MASC

ii. *bue-**go*** 'a female student'
study-PROX.FEM

b. N-based

i. *riri-**rahe***
clay-CYLINDRIC
'clay cylinder (used as cooking pot support)'

ii. *kōbe-**hãi***
metal-FLAT.THIN
'machete'

iii. *héá-**bʉ***
fire-CYLINDRIC CONTAINER
'firearm'

So-called numeral classifiers are numeral-based grammatical nominalizations that derive nominal structures denoting entities characterized by quantity and then classified. E.g.,

(8-17) Thai (Iwasaki & Ingkaphirom 2005: 69, 75)

a. NP-use
[*săm tua*] *nán hăay pay năy*
three CLF that disappear ASP where
'Where have the three (e.g. dogs) gone?'

b. Modification-use
ní khăay [*măa* [*săm tua*]]
Ni sell dog three CLF
'Ni sold three dogs.'

(8-18) Ainu (Bugaeva 2012: 811)

a. NP-use
cóka [*tu-**n***] *ci=ne* *na,* [*tu-**p***] *en=kor-e*
1PL.EXC two-HUM 1PL.(EXC).A=COP FIN two-THING 1SG.O=have-CAUS
yan
IMP.POL
'There are two of us (lit. 'we are two humans'), so give us two pieces.'

b. Modification-use[90]
   i. [*menoko* [*tu-**n***]]
      woman two-HUM
      'two women'
   ii. [*chise* [*tu-**p***]]
      house two-THING
      'two houses'

(8-19) Barasano (Jones & Jones 1991: 50, 59, 113)
a. NP-use
   *yuk**ʉ** ha sure* *~~ʉ~~se* *y**ʉ*** [*idia-**ro***] *Sabīdo barari* *Luka*
   tree hit cut.off.many weave 1s three-CLF Sabino four (fences) Luke
   [*h**ʉa-se***] Arike [h***ʉa-se***] *~~ʉ~~se-ka-h**ʉ*** *y**ʉa***
   two-CLF Eric two-CLF weave-FAR.PST-~3 1x
   'Cutting down trees, weaving, I (wove) three (length of) fences, Sabino (wove) four, Luke (wove) two, (and) Eric wove two.'
b. Modification-use
   i. [*hũ**ʉ**-re* [*h**ʉ**-**ʉ**-re*]] *ãbo-a-ha* *y**ʉ***
      hammock-OBJ two-CLF-OBJ want-PRES-~3 I
      'I want two hammocks.'
   ii. [[*h**ʉa-rahe***] *kõbe-prahe-ri*]
      two-CYLINDRIC metal-cylinder-PL
      'two metal cans'

As already seen, classifiers also nominalize pronouns, ordinary nouns and noun phrases giving rise to so-called "relational/possessive" classifiers (Aikhenvald 2006), which are actually nominal-based nominalizations known as the genitive/possessive form discussed in Section 7. The following Cantonese example shows both modification- and NP-use of N-based nominalizations marked by a classifier.

(8-20) Cantonese (courtesy of Haowen Jiang)
[*keoi*^{5} *gaan*^{1} *fong*^{2}] *daai*^{6} *gwo*^{3} [*ngo*^{23} *gaan*^{1}]
3SG CLF room big exceed I CLF
'His room is bigger than mine.'

As below, possessions are marked by classifiers in Barasano, and these can also have a modification-use (8-21b), though the head nouns are typically unexpressed since the classifiers and context sufficiently narrow down the denotation/reference of the entire noun phrase (Jones and Jones 1991: 62).

90. In prenominal position, numerals may directly modify the head noun; e.g. *sine menoko* 'one woman'.

(8-21) Barasano (Jones & Jones (1991: 62))

a. i. y*ʉ* ya-ro
1SG NMLZR-CLF
'mine' (e.g. seat)

ii. *ĩ* *ya-ga*
3MSG NMLZR-HOLLOW
'his hollow thing', 'his thing having a hole' (e.g. needle)

iii. so *ya-tuti*
3FSG NMLZR-STACK
'her stack' (e.g. book)

b. i. [*ĩ* *ya-g*~~*ʉ*~~] *jũ*~~*ʉ*~~
he NMLZR-MASC hammock
'his hammock'

ii. [*y*~~*ʉ*~~ *y-*~~*ʉ*~~] *bãk-*~~*ʉ*~~
I NMLZR-MASC progeny-MASC
'my son'

Classifiers may also derive new nominal structures from grammatical nominalizations. The Thai argument nominalization (8-22a) below may denote/refer to all kinds of things, including a cat, a shirt, a skirt, a desk, a knife, a notebook, as well as a book.[91]

(8-22) Thai (courtesy of Kingkarn Thapkanjana)

a. [*thîi* *chán súuu mûawaannií*] *phɛɛŋ* *mâk*
NMLZR I buy yesterday very expensive
'What I bought yesterday was very expensive.'

b. [*tua* [*thîi* *chán súuu mûawaannií*]] *phɛɛŋ* *mâk*
CLF NMLZR I buy yesterday very expensive
'What I bought yesterday was very expensive.'

c. [*lêm* [*thîi* *chán súuu mûawaannií*]] *phɛɛŋ* *mâk*
CLF NMLZR I buy yesterday very expensive
'What I bought yesterday was very expensive.'

The form in (8-22b) with the classifier *tua* (< 'body') is a new nominal structure that denotes only things such as a cat, a shirt, and a skirt, a desk, which have some sense of association with a body. On the other hand, (8-23c) with *lêm* refers to sharp, slender objects and book-like bound objects, such as a knife, a book and a notebook. This pattern parallels that found in the classifier marking of N-based

91. Cf. Hmong pattern: [*uas dawb*] *yog kuv tus* ([NMLZR white] is I CLF) 'The one (e.g. skirt, pencil, etc.) that is white is mine', ***daim*** [(*uas*) *dawb*] *yog kuv tus* (**CLF** [NMLZR white] in I CLF) 'The one (e.g. skirt, *pencil) that is white is mine', ***tus*** [(*uas) *dawb*] *yog kuv tus* 'The one (e.g. pencil, *skirt) that is white is mine' (curtesy of Martha Ratliff).

nominalizations; [*khɔ̌ɔŋ chán*] 'mine (all kinds of things)', *tua* [*khɔ̌ɔŋ chán*] 'mine (e.g. a cat, a shirt, a skirt, a desk)', *lêm* [*khɔ̌ɔŋ chán*] 'mine (e.g. a knife, a book, a notebook)'.

A similar pattern is also seen Barasano, as below, where classifier marking derives a structure denoting more specific objects out of a grammatical nominalization with a general denotation.

(8-23) Barasano (Jones & Jones 1991: 146)

a. *ĩ sua-ka-ti*
3MS weave-FAR.PST-~PROX.NMLZR[92]
'things he wove'

b. *ĩ sua-ka-ti-bʉ*
3MS weave-FAR.PST-~PROX.NMLZR-BASKET
'the basket (I saw) him weave' or 'his woven basket'

Classifiers, thus, play various nominalizing functions across languages. On the other hand, the use of classifiers differs considerably from one language to another but a general trend is seen, similar to the development of nominalization markers examined above. In general, the pattern of classifier marking in the modification context is equivalent to or less developed/grammaticalized than that in the context of NP-use. In Thai classifiers are generally "optional" in the sense that a given form may or may not be marked by a classifier, though the use of a classifier changes meaning (cf. (8-22)). A major exception is numerals that need to be nominalized by a classifier in order to function as an entity-denoting nominal (as opposed to denoting numbers and numerals). Compare the numeral-based nominalizations, where classifier marking is obligatory in both NP- and modification-use, and the demonstrative-based ones below, in which classifier marking is optional in both contexts.

(8-24) Thai (cf. (8-17))

a. NP-use
[*săm* *(*tua*)] *nán hăay pay nǎy*[93]
three CLF that disappear ASP where
'Where have those three (e.g. dogs) gone?'

b. Modification-use
ní khăay [*măa* [*săm* *(*tua*)]]
Ni sell dog three CLF
'Ni sold three dogs.'

92. The suffix *-ti* is a time-bound nominalizer associated with the far past tense marker *-ka*.

93. The notation *(…) indicates an obligatory element, as opposed to (…) indicating an optional element.

(8-25) Thai (Iwasaki & Ingkaphirom (2005: 64) and Shoichi Iwasaki p.c.)

a. NP-use
[(*khan*) *níi*] *thawràay*
CLF this how.much
'How much is this?', 'How much is this.CLF (e.g. car)?'

b. Modification-use
rót [(*khan*) *níi*] *thawràay*
car CLF this how.much
'How much is this car?'

Unlike Hmong and Cantonese discussed below, Thai does not use a classifier to nominalize pronouns and ordinary nouns in forming so-called possessive constructions. Instead, they have recruited the noun *khx̌ɔŋ* 'thing', which, when used as the nominalizer for N-based nominalizations, is grammaticalized to the extent that it can denote people and animals as well. This nominalizer is obligatory in the NP-use of N-based nominalizations, but optional in the modification-use, as shown below.

(8-26) Thai (Iwasaki & Ingkaphirom (2005: 13, 69)

a. NP-use
[*(*khx̌ɔŋ*) *khun*] *yùu nǎy*
NMLZR you stay where
'Where is yours?'

b. Modification-use
náŋsx̌ʉ̌ [(*khx̌ɔŋ*) *nók*]
book NMLZR Nok
'Nok's book'.

In the case of Thai V-based argment nominalizations, a classifier is optional in both NP- and modification-use. Thus, parallel to the NP-use pattern seen in (8-22) above, a classifier may or may not occur in the modification context, as below.

(8-27) Thai

a. *kràproon* [(*tua*) [*thîi* *chán súuu mûawaannníi*]] *phɛɛŋ* *mâk*
skirt CLF NMLZR I buy yesterday very expensive
'The skirt that I bought yesterday was very expensive.'

b. *náŋsx̌ʉ̌* [(*lêm*) [*thîi* *chán súuu mûawaannníi*]] *phɛɛŋ* *mâk*
book CLF NMLZR I buy yesterday very expensive
'The book that I bought yesterday was very expensive.'

White Hmong uses classifiers obligatorily in marking not only numerals, but also demonstratives and N-based nominalizations (possessive constructions).

(8-28) White Hmong (courtesy of Martha Ratliff)

a. *Kuv muaj* [*peb tus*]
I have three CLF
'I have three of them (e.g. pencil).'

b. *Kuv muaj* [*tus no*]
I have CLF this
'I have this one (e.g. pencil).'

c. *Kuv muaj* [*kuv tus*]
I have I CLF
'I have mine (e.g. pencil).'

Like all these forms, V-based grammatical nominalizations in Hmong need to be marked by a classifier in NP-use even when marked by the nominalizer *uas*. For example;

(8-29) White Hmong (courtesy of Nerida Jerkey)
[[*(***tus**) [*uas hais*]] *yog Maiv Yaj*
CLF NMLZR say COP Mai Ya
'The one who told (this story) is Mai Ya.'

The above fact is consistent with the fact that in Hmong classifiers function like a determiner marking all definite or specific nouns. Grammatical nominalizations in modifification-use, however, cannot be marked by a classifier, and the following is ungrammatical; removing the second classifier *daim* renders the sentence grammatical.

(8-30) White Hmong (courtesy of Martha Ratlif)
Daim tiab* [*daim*** [*uas kuv niam ntxhua*]] *yog kuv daim.*
CLF skirt CLF NMLZR I mother wash COP I CLF
'The skirt that my mother is washing is mine.'

Notice that the nominalizer *uas* is omissible in White Hmong under certain circumstances, sometimes preferably, e.g. when short nominalizations are involved.

(8-31) White Hmong (courtesy of Martha Ratliff)

a. [*Daim* [*uas dawb*]] *yog kuv daim.*
CLF NMLZ white COP I CLF
'The white one (e.g. a skirt) is mine.'

b. [*Daim* [*dawb*]] *yog kuv daim.*
CLF white COP I CLF
'The white one is mine.'

c. **Daim tiab* [***daim*** *dawb*] *yog kuv daim.*
CLF skirt CLF white COP I CLF
'The white skirt is mine.'

While in (8-31b) the classifier *daim* appears to be functioning like a nominalizer, a verbal nominalization marked by it cannot appear in a modification context, as shown in (8-31c), which must replace the classifier *daim* with the nominalizer *uas* to become grammatical.

The non-verbal based nominalizations marked by classifiers (see (8-28) above) may modify a nominal. While some report (e.g. Nerida Jerkey, p.c.) tells us that they cannot modify those that are already marked by a classifier, as in (8-32a)–(8-32c), our own Hmong consultant permits doubling of classifier marking, as in (8-32d)–(8-32e).

(8-32) White Hmong

a. [[*ob lub*] [*(*lub*) *rooj*]]
two CLF CLF table
'two tables'

b. [[*kuv lub*] [*(*lub*) *rooj*]]
I CLF CLF table
'my table'

c. [[*kuv* *(*lub*)] [[*ob lub*] *rooj*]] (cf. 8-33a)
I CLF two CLF table
'my two tables'

d. [[*daim nplooj*] [(*daim*) *no*]]
CLF leaf CLF this
'this leaf'

e. [[*kuv daim*] [(*daim*) *ntawd*]]
I CLF CLF that
(lit.) 'my that one', 'that one of mine'

In any event, it is clear that classifier-marked nominalizations are more limited in the modification context than in the NP-use context, where classifier marking is obligatory.

Compared to Thai and Hmong situations above, Cantonese classifiers may mark grammatical nominalizations very consistently in both NP- and modification-use, showing that in this language classifiers are beginning to supplant the general nominalizer *kɛ*^{33} (cf. (b) and (c) below), corresponding to the Mandarin nominalizer *de*.

(8-33) Cantonese (courtesy of Haowen Jiang; Matthews & Yip 1994: 111, 112)

a. *ngo*^{23} *kin*^{33} *tou*^{35} *sam*^{55} *pun*^{35} (*ʃy*^{55})
I see reach three CLF book
'I see three (books).'

b. *nei*^{55} *pun*^{35} (*ʃy*^{55}) *hei*^{22} *ngo*^{23} *kɛ*^{33}
this CLF book COP I NMLZR
'This (book) is mine.'

c. *keoi*^{5} *gaan*^{1} (*fong*^{2}) *daai*^{6} *gwo*^{3} *ngo*^{23} *gaan*^{1}
3SG CLF room big exceed I CLF
'His (room) is bigger than mine.'

d. [[*gaau*^{3} *lei*^{5} *taan*^{4} *kam*^{4}] *go*^{2}] *go*^{3}*?*
teach you play piano that CLF
'The one who teaches you piano?'

d′. [*ngo*^{5}*dei*^{6} *hai*^{2} *faat*^{3}*gwok*^{3} *sik*^{6}] *di* *je*^{5} *gei*^{2} *hou*^{2}*-sik*^{61}
we in France eat CLF food quite good-eat
'The food we ate in France was pretty good.'

Classifiers have not yet taken over the general nominalizer *kɛ*^{33} in Cantonese, where the latter is still a default nominalizer. The use of classifiers in nominalization function, except for numerals and demonstratives, is said to be part of colloquial speech and there are also contexts in which classifiers and *kɛ*^{33} are not interchangeable. For example, *kɛ*^{33} in (8-33b) cannot be replaced by *pun*^{35}; but see (8-33c), where the classifier marking *ngɔ*^{2} *gaan*^{1} (I CLF) 'mine' is permitted. Like Hmong above, Cantonese also disfavors the use of a CLF-marked modifier when the head nominal is also marked by a classifier. Thus, while the classifier-marked modifier is permited in (8-33b), (8-33c) above, (8-34a) below is not possible, requiring a modifier marked by the default nominalizer, as in (8-34b).[94]

(8-34) Cantonese (courtesy of Haowen Jiang)

a. *[[*ngɔ*^{23} *pun*^{35}] [[*nei*^{55} *pun*^{35}] *ʃy*^{55}]][95]
I CLF this CLF book
(lit.) 'my this book', 'this book of mine'

b. [[*ngɔ*^{23} *kɛ*^{33}] [[*nei*^{55} *pun*^{35}] *ʃy*^{55}]]
I NMLZR this CLF book
(lit.) 'my this book', 'this book of mine'

94. According to Haowen Jiang (p.c.), there are other southern Chinese dialects (e.g. Jixi Huitseu Mandarin) in which classifier marking is more advanced than in Cantonese, allowing forms that are not possible in the latter. The crosslinguistic pattern of classifier-marking seems to suggest the following hierarchy, where the numenrals are the the easiest to nominalize by classifiers: NUM > DEM > V-based/(Pro)NOUN NMLZ.

95. The avoidance of the use of classifiers in this modification context is reminiscent of the prohibition of the use of a plural form of V-based nominalizations in the modification context when the head noun is marked plural (see (6–27)).

In contrast to Hmong and Cantonese, the Newar animate (*-mha*) and inanimate (*-gu:*) classifiers have fully given rise to the nominalizers *-mha* and *-gu* (NB, a short *u*), and the classifiers (e.g. *-gu:*) now nominalize only numerals. That is, in Modern Newar only numerals are nominalized by *-gu:* and other appropriate classifiers (e.g. *-ga:* for cars), and all else must use the nominalizer *-gu*, as below:

(8-35) Newar (courtesy of Kazuyuki Kiryu)

a. *jĩ:* *cha-gu:/ni-gu:* (*saphu:*) *khane du*
1SG.ERG one-CLF/two-CLF book see can.ND
'I can see one/two (books).'

b. *thwa*[96] (*saphu:*) *bwã*
this book read.IMP
'Read this (book).'

b'. *thu=gu* (*saphu:*) *bwã.*
this=NMLZR book read.IMP
'Read this (book).'

c. *wa gāri Rām=yā=gu* (*kha:*)
that car Ram=NMLZR=NMLZR COP
'That car is Ram's.'

c′. *wa Rām=yā*(*=gu*) *gāri* (*kha:*)
that Ram=NMLZR=NMLZR car COP
'That is Ram's car.'

d. *thaũ:* [*tuyu=gu* (*wāsa:*)] *phi:* *nu*
today white=NMLZR clothes wear.FC HORTATIVE
'Today, let's wear the white (clothes).'

As can be seen in (8-35c′), nominalizer marking is optional in the modification-use of N-based nominalizations in Newar, which is marked by the N-based nominalizer *yā*.

Classifier languages of South America generaly show classifier usage-patterns similar to the Asian counterarts studied above. In particular, many languages use classifiers as nominalizers with some restrictions of their use in the modification context. Yagua, a Peba-Yaguan language spoken primarily in northeastern Peru, obligatorily mark numerals and demonstratives with classifiers in both NP- and modification-contexts, while the marking of adjective-based nominalizations is optional in the modification context. Observe:

(8-36) Yagua (Payne 1985: 131, 174, 176, 179, 189)

a. … *tą́ą́yų* *tá-**duu**-gu**ii**-dee-tée* *vurya-jiriy-tée*
why one-TUBE-one-DIM-EMPH we.INCL-grab-EMPH
'…why did we get one (flute)?'

96. This is likely to be an older nominalized form.

a′. *tá-**juu**-guɨɨ-dee tuváriy vada*
one-EGG-one chicken egg
'one chicken egg'

b. *Néé sa-rúpąą-núúy-tyée jíʔíta jiy-**nu** day-rà*
NEG 3SG-fail-IMPERF-EMPH JITTA this-ANIM:SG DAY-INAN
'This one did not fail to get it.'

b′. *jiy-**nu** vánu*
this-ANIM.SG man
'this man'

b″. *jiy-**see** núúy-**see***
DEM-STICK write-STICK
'this pencil'

c. **ray-jụnúuy-rà jąąmu*
1SG-see-INAN big
(lit.) 'I see big.'

c′. *ray-jụnúuy-rà jąąmu-**dasiy***
1SG-see-INAN big-THIN.POLE
'I see the big blowgun', 'I see the big pole (or anything big that is pole-like).'

c″. *nay rápụụy-**ra***
stranger worthless-CLF.NEUT
'worthless stranger'=mestizo

c‴. *sunupana rúnay*
anatto red
'red anatto'

As shown in (8-36a, a′), Yagua infixes a classifier for numeral-based nominalizations in both NP and modification uses. Example (8-36b″) shows that classifier marking is retained in the modification context for demonstrative-based nominalizations. Notice further that modification by a classifier-marked demonstrative can modify a classifier-marked head, which was optional in Hmong and was avoided in Cantonese, showing that leveling of the demonstrative forms in the two use contexts is more advanced in Yagua. The usage pattern of adjective-based nominalizations by classifiers differs slightly from this. That the NP-use of adjective-based nominalizations requires classifier marking, like numeral-based and demonstrative-based forms, is clearly indicated by the ungrammatical form in (8-36c) and the grammatical form in (8-36c′). In the modification context, adjective-based nominalizations may be marked by a classifier, as in (8-36c″), or may not, as in (8-36c‴), indicating that classifier marking for adjective-based nominalizations in Yagua is not fully grammaticalized in the modification context.

As for V-based nominalizations, Yagua involves a nominalization marker, which typically consists of a demonstrative, a classifier, and the ending *-tìy*. Observe:

(8-37) Yagua (Payne 1985: 106, 109)

a. *naañ-ą junu-rà* [*jiy-ra-tìy rá-raniy*]
1DLEXICL-IRR cut-INAN DEM-CLF.NEUT-TIY INAN-stand
'We are going to cut this which is standing.'

b. *ray-mutívyey-rà jimyichara* [*jiy-ra-tìy sa-tą́ą́ryụ́y*
1SG-cook-INAN food DEM-CLF.NEUT-TIY 3SG-buy
Tomása-rà]
Tom-INAN
'I cooked the food that Tom bought.'

Besides classifier-marked nominalizers involving demonstratives, Yagua seems to allow the nominalizer ending *-tìy* to attach directly to a pronoun (see Payne (1985: Chapter 4) for such examples).

Barasano, an Eastern Tocanoan language spoken in southeastern Colombia, uses classifying nominalizers very extensively. Unlike Yagua, where N-based nominalizations (so-called possessive/genitive constructions) are not marked by classifiers, and where V-based nominalizations (so-called relative clauses) involve a special marker, Barasano uses classifiers, which are numerous in number and which vary in form, to mark all these types of nominalization. Classifier-marked numerals denote objects classified according the classifier marking them, as in (8-38a). These can be used as modifiers as in (8-38a′). But many classifiers being specific enough, modification structures like these are far less commonly used; the NP-use of classifier-marked numerals does the job. That is, instead of (8-38a′.ii), the one without the head noun is most likely used (cf. (8-38c.iii)).

(8-38) Barasano (Jones & Jones 1991: 50, 57, 60, 61, 62, 66, 78, 86)

a. i. *hʉ̵a-hãi*
TWO-FLAT.THIN
'two flat, thin objects'

ii. *hʉ̵a-rã*
TWO-ANM.PL
'two living things'

a′. i. *hʉ̵a-hãi sudi-hãi*
TWO-FLAT.THIN cloth-FLAT.THIN
'two pieces of cloth'

ii. *hũʉ̵-re hʉ̵-ʉ̵-re ãbo-a-ha yʉ̵*
hammock-OBJ TWO-HAMMOCK want-PRS-~3 I
'I want two hammocks.'

b. i. *adi-re ãbi-a-bʉ yʉ*
these-OBJ pick.up-MOT-PST~3 I
'I brought these.'
ii. *adi-tuti-re ãbi-a-ha bʉ*
this-STACK-OBJ pick.up-MOT-FUT.IMPV.PROX you
'You take this stack.thing (book).'

b′. i. *ti gũbu-re bʉha …*
that log-OBJ find
'finding that log'
ii. *ʉye-godo-a ti-a-re ãbo-a-ha yʉ*
oil-CLEARED.OUT-HOLLOW that-HOLLOW-OBJ want-PRS-~3 I
'I want that empty oil can.'

c. i. *yʉ ya-ro*
I NMLZR-THING
'my thing (e.g. seat)'
ii. *ĩ ya-ga*
he NMLZR-HOLLOW
'his hollow thing/thing having a hole (e.g. needle)'
iii. *haibẽ ĩ ya-tuti yã-a-ha ti*
Jim he NMLZR-STACK be-PRS-~3 it
(lit.) 'That is Jim's (stack stuff).' 'That is Jim's book.'

c′. *hũʉ ĩ ya-gʉ*
hammock he NMLZR-HAMMOCK
'his long hammock'

As in (8-38b, b′) above, a demonstrative may or may not be marked by classifiers in both NP- and modification-use, whereas N-based nominalizations require classifier marking in both NP- and modification-use, as in (8-38c, c′).

V-based nominalizations in Barasano are formed in two ways. A first method is to attach nominalizing classifiers directly to verb roots and stems as in the examples below, involving the general classifiers *-se* and *-ro*, meaning "thing", "time", "place", and time-bound nominalizer (*-ka-ti* 'far past, non-proximal'), as well as more specific nominalizing classifiers that distinguish animacy-number (*-rã* 'inanimate singular') and gender-number (*-go* 'feminine singular', *-gu* 'masculine singular').[97]

(8-39) Barasano (Jones & Jones 1991)
a. *wʉ-se*
fly-CLF
'flying things (airplanes)'

97. Some of these can be further marked by a classifier; e.g. compare (8-40b) and *yã-ro-hũ* (be-NMLZR-CLF.PLACE) 'the place where'.

b. *ĩ yã-ro*
he be-CLF
'when/where he is'

c. *ĩ suka-ka-ti*
he weave-FAR.PST-~PROX.NMLZR
'things he wove'

d. *bue-ria-rã*
study-PST.CONJ-ANP
'ones who probably studied'

e. i. *bõa-gũ*
work-MASC.SG
'a male worker'
ii. *bue-go*
study-FEM.SG
'a female student'

f. *ĩ suka-ka-ti*
he weave-FAR.PST-~PROX.NMLZR
'things he wove'

All these allow both NP- and modification-use, as illustrated below.

(8-40) NP-use of V-based nominalizations (Jones & Jones 1991: 85, 148, 150, 169)

a. *kahi* [*idi-re*] *idi-ka-ti bũ*
coca drink-CLF drink-FAR.PST-Q you
'Did you drink the halcinogenic drink?'

b. *ĩ-re* [*yʉ̃ ĩsi-boa-ka-ti-re*]
he-OBJ I give-BUT-FAR.PST-~PROX.NMLZR-OBJ
rea-koã-yu-ĩ
move.away-FF-INFER-3 MSG
'He threw away what I had given him.'

c. [*yʉ̃ yi-boa-rã*], *y*ũ-re *kũdi-beti-s-ã-ĩdã*
say-BUT-ANP I-OBJ respond-NEG-PRS.PROX-3p they
'The ones I talked to aren't responding to me.'

d. [*ado eha-go-re*] *bãsi-be-a-ha yʉ̃*
here arrive-FEM.SG-OBJ know-NEG-PRS.PROX-~3 I
'I don't know the feminine one who just arrived here.'

(8-41) Modification-use of V-based nominalizations (Jones & Jones 1991: 6, 143, 151

a. [*oko* [*kedi-se*]] (143)
water fall-CLF
(lit.) 'falling-thing water', 'rain'

b. *to kõ-ro yã-ka-bã* [*rõbi-a* [*ʉ̃e-o-rã*]]
that count-NMLZR be-FAR.PST-3p female-p dirty-CAUS-ANP
'There were that many women who were dirtying (the river).'

c. [*gʉbo sudi tẽdi [bʉ̃-re yʉ sẽdi-ka-ti*]]
foot clothes tennis.shoes you-OBJ I ask-FAR.PST-~PROX.NMLZR
hua-ri bʉ
pickp.up-Q you
'Did you get the tennis shoes that I asked you for?'

d. [*sĩg-o [ĩdã rãka bue-go*]] *yʉ-re ãbo-a-bõ so*
one.AN-3FEM.SG they with study-FEM.SG I-OBJ want-PRS-FEM.SG she
'A girl who studies with them wants me (for her husband).'
(lit.) 'A female who is a female student with them, she wants me.'

A second way to form V-based nominalizations in Barasano involves the suffix *-ri*, which is recognized as a nominalizer but is glossed as PTCPLE (participle) by Jones & Jones (1991: 43). While the other nominalizers discussed above appear to derive nominalizations with concrete denotations like things and persons, *-ri* derives event nominalizations as well, which is probably why Jones and Jones (1991) glosses it as PTCPLE. The event-denoting *-ri* nominalizations are widely used as adverbials including the use as chain structures discussed in Section 5.1. When they derive forms denoting concrete objects, their NP-use typically calls for classifier marking, though it may be possible to use them without a classifier, as in the following example.

(8-42) Barasano (Jones & Jones 1991: 43)
[*uka-ri*] *kʉti-go yã-ka-bõ*
write-NMLZR have-FEM.SG be-FAR.PST-3FEM.SG
'She had spots (as a characteristic).' (lit.) 'She was a female possessor of writings/a female one who had writings.'

The more commonly observed usage pattern of *-ri* nominalizations has additional classifier marking as in the following examples.

(8-43) Barasano (Jones & Jones 1991: 24, 135, 150)

a. [*sʉ̃be-ri-hãi*] [*sũa-ri-hãi*] *ãbo-a-ha*
green-NMLZR-FLAT.THIN red-NMLZR-FLAT:THIN want-PRS-~3
'I want either a green cloth or a red one.'

b. [*yʉ ĩa bʉha-ri-hãi*] *yĩ-beti-bʉsa-a-ha ti*
I see find-NMLZR-FLAT.THIN dark-NEG-very.much-PRS-~3 it
'The cloth I found isn't very dark.' (lit.) 'What flat, thin thing I found, it isn't very dark.'

c. [*õ kãhi-ri-kʉ*] *ãbo-a-ha yʉ*
there hang-NMLZR-HAMMOCK want-PRS-~3 I
'I want that hammock which is hanging there.' (lit.) 'I want that hanging hammock-thing.'

As expected, all these forms allow a modification-use as well, as below.

(8-44) Barasano (Jones & Jones 1991: 63, 66, 150)

a. [*hũʉ̃* [*õ* *kãhi-ri-kʉ̃*] *ãbo-a-ha* *yʉ̃*
hammock there hang-NMLZR-HAMMOCK want-PRS-~3 I (cf. (8-43c))
'I want that hammock which is hanging there.'

b. *gahe* [*gũbu* [*yao-ri-kʉ̃*]]
another log long-NMLZR-LOG
'another long log'

c. [*sũa-ri-hãi* [*yoa-ri-hãi*]] *ãbo-a-ha* *yʉ̃*
red-NMLZR-FLAT.THIN long-CLF.FLAT.THIN want-PRS-~3 I
'I want a long red piece of cloth.' (lit.) 'I want a red thin, flat thing that is long thin, flat.'

Interestingly and most relevant to the main point of the discussion in this subsection, *-ri* nominalizations, which appear to be normally marked by a classifier in NP-use, need not be marked by a classifier in the modification context, as below, showing that the marking pattern in the modification context lags behind that in the NP-use context.

(8-45) Barasano (Jones & Jones 1991: 21, 111, 144, 152)

a. *gahe gũbu* [[*yoa-ri*] *gũbu*] (cf. (8-44b))
other log long.NMLZR log
'another long log'

b. [[*bõa-ri*] *bãs-o*][98]
work-NMLZR human-FEM.SG
(lit.) 'working female human', 'a woman/girl worker'

c. [[*sĩa-ri*] *bãs-ʉ̃*] *yã-a-bĩ* *ĩ*
kill.NMLZR human-MASC.SG be-PRS-3MASC.SG he
(lit.) 'He is a killing/killer male human.' 'He is a killer.'

d. [[*ĩe-ri*] *hai-gʉ̃*] *yã-a-bĩ*
dirty-NMLZR big-MASC.SG be-PRS-3MASC.SG
(lit.) 'He is a dirty big male.' 'He is very dirty.' 'He has a lot of dirt.'
Cf. [*hai-gʉ̃*] *yã-a-bĩ*
big-MASC.SG be-PRS-3MASC.SG
(lit.) 'He is a big male.' 'He is big.'

98. There is some issue regarding identification of the head in these forms, but in view of the author's observantions that "[t]he descriptive modifier generally precedes the head noun in the noun phrase" (Jones & Jones 1991: 4), the literal readings we provide seem to be more faithful interpretations of the data than the idiomatic translations by the authors. Barasano is an OV and postpositional language.

Tuyuca, closely related to Barasano, has a classifier system similar to Barasano. According to Jane Barnes (p.c.), classifier marking is obligatory when N-based nominalization (aka gentives/possessives) is used as an NP-head, but it is optional in the modification context, similar to the pattern seen in Newar (see 8-35c, c′). For example,

(8-46) Tuyuca (Jane Barnes, p.c.)
a. [*yi̱i̱ paki̱-ya-**ró***]
my father-NMLZR-2D.FLEXIBLE
'my father's' (as in "They are my father's/ My father's are those.")
b. [[*yi̱i̱ paki̱-ya*(-***ro***)] *sirúra*]
my father-NMLZR-2D.FLEXIBLE trouser
'my father's trousers'

The above study of classifier marking on nominalization structures in the two usage contexts of NP-use and modification-use also corroborate the pattern of development of nominalization markers recognized by Shibatani and Shigeno (2013) (see Figure 7 (p. 129). Specifically, in both N-based and V-based nominalizations, nominalization markers start out in NP-use, as markers of referential use of nominalizations. As in (a) and (c) in Figure 7, these markers then spread to the modification context, eventually becoming nominalizers, marking nominalization structures as such regardless of their usage contexts.[99] The observed patterns of developet of nominalization markers allow us to draw the following generalizations.

(8-47) Generalizations on the forms of nominalizations
The form of nominalizations in modification-use is equal to or simpler in formal complexity than that in NP-use. In particular, if a modifying nominalization involves nominalization morphology, the same marker is found in its NP-use, either in the contemporary language or in a historically earlier form or in its dialects.

The caveat about a historically earlier form of the language or its dialects in the above generalization is necessary since a new marker is likely to be introduced first in the context of NP-use either replacing the old marker or incrementally, once the uniformity in form is achieved in the two usage contexts.

We might be able to draw some kind of generalization from the pattern of development depicted as (b) in Figure 7, but since what is involved here are two different types of nominalization, N-based and V-based, rather than two different uses of a single type of generalization, it might be difficult to do so. In many

99. A complete study of the forms of nominalizations must examine the marking patterns in nominalizations in the adverbial function as well.

languages N-based and V-based nominalizatins have distinct morphology, and we need more cases attesting the pattern of development depicted in Figure 7(b).

9. Summary and implications

After summarizing the discussions above in general terms, the final Subsection 9.1 discusses the implications of the present study for both descriptive and theoretical studies.

Past studies on nominalization tended to focus on lexical nominalizations because they typically have clear morphological marking. We showed in the beginning that across different languages the same lexical nominalization morphology may apply to units larger than words, suggesting the existence of grammatical nominalizations. The field has been slow to recognize grammatical nominalizations because many do not have clear nominal morphology or the forms involved have the same verbal form as clauses and sentences. We have argued that the notion of nominalization is neither morphological nor syntactic, but functional. Crosslinguistic investigations reveal clearly that formally different structures cohere in their semantics and usage patterns, supporting this view. At the same time, such studies provide crucial evidence, including morphological support, that is hard to find when dealing with single languages such as Japanese and English.

We have argued that traditional studies fail to distinguish between structures and their use, a failure that has led to the recognition of what is no more than different uses of the same basic structures as independent constructions. We have argued strongly that relative clauses are simply uses of grammatical argument nominalizations. So-called internally-headed RCs are event nominalizations in NP-use, which evoke various concepts metonymically related to the events such as the abstract concepts of facts and propositions or concrete concepts such as event protagonists and resultant products. What are known as headless relative clauses are instances of the NP-use of argument nominalizations, which also have a modification-use giving rise to what are known as relative clause constructions with a modified head noun.

The reanalysis of the genitive case or the possessive form as a nominal-based nominalization reveals that nothing like possessive pronouns and possessive adjectives exist as separate parts of speech. Similar to the case of relative clauses, they are no more than two uses of N-based nominalizations. We have provided ample crosslinguistic evidence in support of this new analysis, including classifier marking, which nominalizes and then classifies the entities denoted.

Table 2 below shows how disparate constructions and those thought to be related yet independent construction types in the past studies receive a uniform

analysis in terms of the notion of grammatical nominalizations of different types under different usage patterns:

Table 2. Comparison of past treatments and nominalization-based analysis

Past treatments	Nominalization-based analysis
a. Verb complement	Event nominalization: NP-use
b. Noun complement	Event nominalization: (Adnominal) Modification-use
c. Internally-headed RC	Event nominalization: NP-use
d. Headless/free RC	Argument nominalization: NP-use
f. Relative clause (RC)	Argument nominalization: (Adnominal) Modification-use
g. Adverbial clause, Converb, Clause-chain	Event nominalization: (Adverbial) Modification-use
h. Headless/Free genitive	Nominal-based nominalization: NP-use
i. Genitive/Possessive	Nominal-based nominalization: (Adnominal) Modification-use
j. (Numeral) classifier	Nominal/Numeral/Verbal-based nominalization

9.1 Implications

The crosslinguistic study of nominalizations and their roles in grammar presented above has some far-reaching implications for both descriptive practice and theoretical debates. Indeed, the philosopher of language Zeno Vendler, who also worked on nominalization, noted that "the grammar of nominalizations is a centrally important part of linguistic theory" (Vendler 1967: 125).[100] Of the various implications to many parts of grammar, we limit our discussions here to those pertaining to the understanding and analysis of relative clauses because of the high impact that this research topic has had in the field over the past fifty years.

We start with the widely received definitions of relative clauses and their evaluations in the light of our analysis of nominalization in this paper. The following two definitions perhaps represent the general understanding in the field of what relative clause constructions are:

> a construction consisting of a nominal or common noun phrase …which may be empty …and a subordinate clause interpreted attributively modifying the nominal. The nominal is called the head and the subordinate clause the RC. The

100. In this book of high relevance to linguists, Vendler recognizes grammatical nominalizations and that they may denote abstract entities like a fact, a proposition, as well an event. Though he does not invoke the notion of metonymy, his discussions make it abundantly clear that many nominals, both ordinary nouns and nominalizations, denote entities that are evoked metonymically.

> attributive relation between head and RC is such that the head is involved in what is stated in the clause. (Lehmann (1986: 664))

> A relative clause (RC) is a subordinate clause which delimits the reference of an NP by specifying the role of the referent of that NP in the situation described by the RC. (Andrews (2007: 206))

Both these definitions are based on the traditional clause-based analysis of English-type relative clauses that recognizes an involvement of so-called relative pronouns that link the head noun with an argument position in the RC structure (our grammatical argument nominalization), as if the denotation/referent of the head noun is directly represented in the RC structure (see (5-86)). Such descriptions, however, fail to capture the notion of "restrictive" modification properly, which involves two independent sets of entity denotations. If the head NP is coreferential with an NP in the relative clauses as in the traditional analysis, we are dealing with a single referent; and the notion of restrictive modification does not fall out from such an analysis. Our analysis of RC constructions does not recognize as essential the English-type relative pronouns, which are not found in most languages outside the Indo-European sphere. More importantly it is not really the case that the denotation/referent of the head noun "is involved in what is stated in the clause" or that their role is specified "in the situation described by the RC". Aside from the point that so-called RCs do not "state" or "describe" like clauses and sentences, the denotations of the two nominals involved in RC constructions, a head noun and a modifying argument nominalization, actually denote two distinct sets of denotation. The only and important requirement for the restrictive RC construction is that the denotation of the argument nominalization be able to specify a subset of that of the head noun, as in the manner shown in Figure 3. Our analysis not only translates straightforwardly to the Formal Semantics treatment (see page 84), but also captures in a very simple manner the intent behind the definition of the restrictive RC construction by Keenan and Comrie below, which suffers from the fact that there is no evidence that a truth-bearing sentence underlies an RC. Argument nominalizatons qua RCs do not assert; they only presuppose.

> [an RC construction] specifies a set of objects (perhaps a one-member set) in two steps: a larger set is specified …and then restricted to some subset of which a certain sentence, the restricting sentence, is true. The domain of relativization is expressed in surface structure by the head NP, and the restricting sentence by the restricting clause, which may look more or less like a surface sentence depending on the language. (Keenan & Comrie 1977: 63-64)

Let us now turn to the single most influential paper on the topic of relative clauses, namely Keenan & Comrie (1977). Our new analysis first calls for a reassessment of

the NP-Accessibility Hierarchy that plays the central role in the Keenan-Comrie analysis of RCs. Our analysis suggests that grammatical relations actually have nothing to do with relative clause formation per se, which is viewed as bringing together a grammatical argument nominalization and a head noun to form a larger NP constituent without ever asking whether what is being relativized on is Subject, Object, or Oblique. A so-called subject relative clause is simply a modification of a noun by a subject nominalization, and a so-called object RC is no more than bringing an object nominalization and a head noun together under the modification function. Under the proposed analysis of RC constructions, there is no process involved that "accesses" an argument position, as in the traditional generative analysis, which creates a gap in an argument position as part of the relativization process. In our analysis, a gap in the modifying structure of an RC construction is a property of an argument nominalization.

This does not invalidate a hierarchy of grammatical relations like the one posited by Keenan and Comrie. Indeed, such a hierarchy is plausible for argument nominalizations, since some languages, such as Yup'ik, allow only argument nominalizations of the absolutive argument. Apparently some dialects of K'ichee' are like Yup'ik, while other dialects allow argument nominalizations pointing to both absolutive and ergative roles, as shown in the examples cited in this paper (see Larsen & Norman 1979). Those Austronesian languages (many Formosan and Philippine languages as well as Malagasy) maintaining the proto-Austronesian four-way focus contrast allow argument nominalizations of various types, such as subject nominalization, object nominalization, locative, and beneficiary (see the Malagasy-German comparison in the following discussion). Those that have reduced the focus contrast to two (AF and PF), as in many languages of Indonesia (Bahasa Indonesia, Javanese, Balinese, Sasak, etc.), allow only subject and object nominalizations; obliques must be first made applicative objects before they can be the target of argument nominalization. The English and German gerundive argument nominalization applies only to subjects; e.g. *the man* [***Ø holding a book in hand***], **the book* [***the man holding Ø in hand***]. Our point is that while argument nominalizations refer to grammatical relations, the relativization process itself does not, contrary to the claim advanced by Keenan and Comrie. Indeed, the relevance of relational hierarchies makes much more sense in the metonymy-based analysis of nominalization than the clause-based analysis of relativization. Since the absolutive/ergative and the subject/object arguments, as central argument types, code most salient event protagonists intimately associated with a wide range of event types, they are the easiest to evoke metonymically.

Our new analysis of RCs also has significant implications for the description of RCs in individual languages. In particular, it shows that those Austronesian languages (e.g. Malagasy and Tagalog), which are claimed to obey the subject-only

constraint on relativization actually relativize on any argument as does German, which is said to relativize down to the genitive position in the Accessibility Hierarchy. Keenan and Comrie demonstrate that relativization on a subject (9-1b) is possible in an Actor focus construction, but an object in such a construction cannot be relativized (9-1c). For an object to be relativized, it must be made into a subject by turning an AF construction to a Patient focus construction (9-2a).

(9-1) Malagasy AF construction

a. *n-**i**-kapoka ilay alika t-aminy hazokely **ilay lehilahi***
PST-**AF**-hit DEF dog PST-with stick **DEF man**
'The man hit the dog with a stick.'

b. *n-a-hita ilay lehilahy* (*izay*) [*n-**i**-kapoka ilay alika t-aminy*
PST-AF-see DEF man NMLZR PST-**AF**-hit DEF dog PST-with
hazokely **Ø**] *aho*
stick 1SG (AF-NMLZR + SUB NMLZ)
'I saw the man [who Ø hit the dog with a stick].'

c. **n-a-hita ilay alika* (*izay*) [*n-**i**-kapoka* **Ø** *t-aminy hazokely*
PST-AF.see DEF dog NMLZR PAST-**AF**-hit PST-with stick
ilay lehilahi] *aho*
DEF man 1SG
'I saw the dog [that the man hit Ø with a stick].'

(9-2) Malagasy PF construction

a. *no-kapoh-**in**-ilay lehilahy t-aminy hazokely **ilay alika***
PST-hit-**PF**-DEF man PST-with stick **DEF dog**
'The man hit the dog with a stick.'

b. *n-a-hita ilay alika* (*izay*) [*no-kapoh-**in**-ilay lehilahy t-aminy*
PST-AF-see DEF dog NMLZR PST-hit-**PF**-DEF man PST-with
hazokely **Ø**] *aho*
stick 1SG (PF-NMLZR + OBJ NMLZ)
'I saw the dog that Ø was hit by the man with a stick.'

Assuming PF and other non-AF constructions to be passive, Keenan and Comrie conclude that only subjects can be relativized on in Malagasy, instantiating a language in which the subject-only constraint on relativization obtains (Keenan & Comrie 1977; Comrie & Keenan 1979). As it turns out, the real reason why (9-1c) is ungrammatical is not because what has been relativized on (the gap position) is object position, but because the construction has an incompatible combination of AF marking and object nominalization. Recall from the earlier discussion on another focusing Austronesian language, Mayrinax Atayal (see (3-12)), that focus marking in Austronesian has a role-marking function for argument nominalizations, where AF marking in the verb marks a subject nominalization, PF marking

an object nominalization, LF marking a locative nominalization, and CF marking a beneficiary or an instrumental nominalization. AF marking, therefore, can combine only with a subject nominalization, as in (9-1b), and cannot combine with an object nominalization, as in (9-1c). An object nominalization must be marked by the PF marker in the verb, as in (9-2b),[101] not by the AF marker as in (9-1c).

Languages with role-marking nominalizers all behave this way, such that a subject/agent nominalizer must mark a subject nominalization, an object/patient nominalizer an object nominalization, and so forth. Indeed, the Malagasy pattern is paralleled by German, which also has role-marking nominalizers, similar to AF and PF markers in focusing Austronesian languages. Observe:

(9-3) German subject nominalization

a. *Der Junge sieht den Hund.*
ART boy sees ART dog
'The boy sees the dog.'

b. [*der Junge* [***der*** [*Ø den Hund sieht*]]]
ART boy **SUB.NMLZR** ART dog sees (SUB.NMLZR + SUB NMLZ)
'the boy who sees the dog'

c. *[*der Hund* [***der*** [*der Junge sieht Ø*]]
ART dog **SUB.NMLZR** ART boy sees (SUB.NMLZR + DO NMLZ)
'the dog that the boy sees'

d. [*der Hund* [***den*** [*der Junge sieht Ø*]
ART dog **OBJ.NMLZR** ART boy sees (DO.NMLZR + DO NMLZ)
'the dog that the boy sees'

The reason that (9-3c) is bad is not because German cannot relativize on an object, but because it has the incompatible combination of a subject nominalizer and an object nominalization, as in (9-1c) for Malagasy. The phrase in (9-3d) is grammatical because the object nominalizer marks an object nominalization, just like the Malagasy form (9-2b). We see an exact parallelism between Malagasy and German. Indeed, Malagasy can relativize on any argument and adjunct that German can as long as the marking pattern is consistent. Just to drive the point home, another parallelism between the two languages using an oblique nominalization and its use as a modifier (relative clause) is shown below.

101. Like many other languages, e.g. Mongolian, Turkish, Japanese, Yaqui, and Quechua, object nominalizations in focusing Austronesian languages may have an agent in the genitive form.

(9-4) German oblique/source nominalization

a. NP-use
Ich treffe den [[***von dem*** *ich das Buch bekommen habe*]$_{\text{NMLZ}}$]$_{\text{NP}}$
I meet ART **from IO.NMLZR** I the book receive.PP have
'I meet the one from whom I received the book.'

b. Modification-use
Ich treffe den [*Mann* [***von dem*** *ich das Buch bekommen habe*]$_{\text{NMLZ}}$]$_{\text{NP}}$
I meet ART man **from IO.NMLZR** I the book received have
'I meet the man from whom I received the book.'

(9-5) Malagasy oblique/source nominalization

a. NP-use
Ho hita-ko ilay [*n-**indrama**-ko* (*an'ilay/ilay*) *boky*]
FUT see-1SG.GEN the PST-**borrow.CF**-1SG.GEN (the/the) book
'I will see the one from whom I borrowed the book.'

b. Modification-use
Ho hita-ko ilay lehilahy [*n-**indrama**-ko* (*an'ilay/ilay*) *boky*]
FUT see-1SG.GEN the man PST-**borrow.CF**-1SG.GEN (the/the) book
'I'll see the man from whom I borrowed the book.'

The parallelism between German and Malagasy is again clear – (9-4b) and (9-5b) are grammatical RC constructions because the nominalizers mark correct nominalization structures in both cases. If the nominalizers in these examples were the subject nominalization marker *der* for German or the subject nominalizing AF form for Malagasy, both would be ungrammatical. As long as nominalizers and nominalization structures are kept consistent, both languages can nominalize any argument down to the genitive in the relational hierarchy, and the resulting nominalizations can be used as modifiers (relative clauses). We can observe the same thing in English, which has an object nominalizer that uniquely marks human object nominalizations, and which, therefore, cannot combine with a subject nominalization, as in (9-6b) below.

(9-6) English object and subject nominalizations

a. *Marry* [*who/whom* [*you love Ø*]] (object nominalization)
a′. *Marry a man* [*who/whom* [*you love Ø*]]
b. *Marry* [*who/*whom* [*Ø loves you*]] (subject nominalization)
b′. *Marry a man* [*who/*whom* [*Ø loves you*]]

As is clear from the above, the proposed nominalization-based analysis of relative clause constructions yields a very different result from the traditional clause/sentence-based analysis by Keenan and Comrie (1977), Comrie and Keenan (1979) and others. This is true of all focusing Austronesian languages such as Atayal and other Formosan languages, Tagalog and other Philippine languages, Malay/Indonesian, Sasak, Sumbawa, and others (see Shibatani (2008) on Sasak and Sumbawa).

There are many other important theoretical and descriptive implications engendered by the proposed analysis of nominalization, but let us close this already lengthy paper by touching on the problems of the power of a grammatical theory. The analysis of so-called relative clauses and complement clauses as nominalizations rather than as clauses allows a much tighter theoretical framework for syntactic analysis; namely only structures of equal or lower rank can be embedded within a given structure. Current theories, which consider relative clauses and complement clauses as clauses, allow clauses to be embedded under a structure lower in rank such as NP and VP. Such theories allowing any type of embedding are too powerful and hence are weak theories.

As these brief comments suggest, the new analysis of nominalizations proposed in this paper opens up many interesting new developments in both descriptive and theoretical arenas.

Acknowledgements

This chapter is a vastly expanded version of an earlier paper, Shibatani (2018a), which contrasts Japanese nominalizations with those of the world's languages. The present paper, on the other hand, draws data from the Americas and elsewhere and offers more detailed discussions on both conceptual and empirical issues surrounding nominalization phenomena. I owe a great deal of indebtedness to many scholars who have helped me gather and interpret the nominalization data contained in this paper, but especially to Akio Ogawa (German), Roberto Zariquiey (Spanish), Cristina Messineo (Toba), Jane Barnes (Tuyuca), Doris Payne (Yagua), Marília Ferreira (Parkatêjê and Portuguese), Telma Can Pixabaj and Mareike Sattler (K'ichee'), Kingkarn Thepkanjana and Shoichi Iwasaki (Thai), Haowen Jiang (Mandarin Chinese, Cantonese), Sung-Yeo Chung (Korean), Stephen Matthews and Virginia Yap (Cantonese), Nerida Jarkey, Martha Ratliff, and Elizabeth Riddle (White Hmong), Kazuyuki Kiryu (Newar), Prashant Pardeshi (Marathi), Miki Nishioka (Hindi), and Albert Alvarez (Yaqui). I am also grateful to Francesc Queixalós, David Fleck, and Roberto Zariquiey for many useful discussions on nominalization. I am also grateful to David Fleck and Haowen Jiang, who read an earlier version carefully and helped spot typos and make stylistic improvments. The research and preparation of the work reported here was in part supported by the International Joint Research Promotion Program of Osaka University (PI: Sung-Yeo Chung) and the project "Noun Modifying Expressions" (PI: Prashant Pardeshi) of the National Institute for Japanese Language and Linguistics (NINJAL).

List of less commonly used abbreviations

AF	actor focus	MM	middle marker
ANP	animate plural	MOT	motion
BUT	contraexpectation	MSG	masculine singular
C	time reference 'before', 'completed'	MPL	masculine plural
CA	causative	NAR	narrative register
CF	circumstantial focus	ND	neutral disjunct
CL	clitic	NPM	NP-use marker
CLF	classifier	NMLZ	nominalization
CM	concatenated form	NMLZR	nominalizer
COLL	collective	MOD	modal clitic
CON	conjunctive particle	O	object
CONJ	conjecture	OBJ.An	animate object
DE	demonstrative pronoun	OM	object marker
DIR	directional	P	possessor
DL	dual	PF	patient focus/affirmative evidential
DM	'uh', preform	PFV	perfective
DT	associated motion marker	PR	progressive
EVD	evidential	PROX	proximate
EX	existential	PRT	particle
F	feminine	PS	optional past tense marking
FGR	falling tone grade	REF	referential
FSG	feminine singular	REL	relative case (Yup'ik)
FV	final vowel	S	subject
GER	gerundive	SIN	single action
HAB	habitual	<t>	verb-terminating classifier
HON	honorific	THM	theme (grammatical relation)
INC	inceptive aspect marker	TOP	topic marker
IR	irealis	TR	transitivizer
LF	locative focus	TRZ	transitivizer
LGR	lengthened grade	VI	verbal stem marker
LIN	linker	~	non-

References

Adelaar, Willem. 2011. Participial clauses in Tarma Quechua. In van Gijn, Haude & Muysken (eds), 267–280.

Aikhenvald, Alexandra Y. 2006. Classifiers and noun classes: Semantics. In *Encyclopedia of Language and Linguistics*, 2nd edn, Keith Brown (ed.), 463–471. Oxford: Elsevier. https://doi.org/10.1016/Bo-08-044854-2/01111-1

Aikhenvald, Alexandra Y. 2009. Complement clause types and complementation strategies in Tariana. In Robert M. W. Dixon & Alexandra Aikhenvald (eds), 178–203.

Ajiboye, Oladipo Jacob. 2005. Topics on Yoruba Nominal Expressions. PhD dissertation, University of British Columbia.

Allan, Keith. 1977. Classifiers. *Language* 53(2): 285–311. https://doi.org/10.1353/lan.1977.0043

Allan, Keith. 2001. *Natural Language Semantics*. Oxford: Blackwell.

Alvarez, Albert. 2012. Relative clauses and nominalizations. In Comrie & Estrada-Fernández (eds), 65–96.

Andrews, Avery. 2007. Relative clauses. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, Timothy Shopen (ed.), 206–236. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.004

Bills, Garland, Vallejo, Bernardo & Troike, Rudolph. 1971. *An Introduction to Spoken Bolivian Quechua*. Austin TX: University of Texas Press.

Blake, Barry J. 1979. *A Kalkatungu Grammar* [Pacific Linguistics Series B, No. 57]. Canberra: Research School of Pacific Studies, Australian National University.

Bugaeva, Anna. 2012. Southern Hokkaido Ainu. In *The Languages of Japan and Korea*, Nicolas Tranter (ed.), 799–845. London: Routledge.

Campbell, Akua Asantewaa. 2013. Grammatical nominalizations in Akan and Gã. Ms, Department of Linguistics, Rice University.

Campbell, Akua Asantewaa. 2017. A Grammar of Gã. PhD dissertation, Rice University.

Caponigro, Ivano, Torrence, Harold & Cisneros, Carlos. 2013. Free relative clauses in two Mixtec languages. *International Journal of American Linguistics* 79(1): 61–96. https://doi.org/10.1086/668608

Ciccone, Florecia. 2008. Relative clauses and nominalization in Tapiete (Tupí-Guaraní). Term paper for the seminar "Typology and Functional Linguistics: Case studies on relative clauses, nominalization, serial verbs and complex predicates" by Masayoshi Shibatani, University of Buenos Aires, Argentina.

Cole, Peter. 1982. *Imbabura Quechua*. Amsterdam: North Holland.

Cole, Peter. 1987. The structure of internally headed relative clauses. *Natural Language and Linguistic Theory* 5: 277–302. https://doi.org/10.1007/BF00166587

Comrie, Bernard. 1996. The unity of noun modifying clauses in Asian languages. *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics*, 1077–1088. Salaya, Thailand: Mahidol University.

Comrie, Bernard. 1998a. Attributive clauses in Asian languages: Toward an areal typology. In *Sprache in Raum und Zeit, In memoriam Johannes Bechert*, Winfried Boeder, Christoph Schroeder, Karl Heinz Wagner & Wolfgang Wildgen (eds), 51–60. Tübingen: Gunter Narr.

Comrie, Bernard. 1998b. Rethinking the typology of relative clauses. *Language Design* 1: 59–86.

Comrie, Bernard. 2006. Syntactic typology: just how exotic ARE European-type relative clauses?', In *Linguistic Universals*, Ricardo Mairal and Juana Gil (eds), 130–154. Cambridge: CUP.

Comrie, Bernard & Estrada-Fernádez, Zarina (eds). 2012. *Relative Clauses in Languages of the Americas: A Typological Overview* [Typological Studies in Language 102]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.102

Comrie, Bernard & Horie, Kaoru. 1995. Complement clauses versus relative clauses: Some Khmer evidence. In *Discourse Grammar and Typology: Papers in Honor of John W. M. Verhaar* [Studies in Language Companion Series 27] Werner Abraham, T. Givón & Sandra A. Thompson (eds), 65–75. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.27.07com

Comrie, Bernard & Keenan, Edward L. 1979. Noun phrase accessibility revisited. *Language* 55: 649–664. https://doi.org/10.2307/413321

Comrie, Bernard & Thompson, Sandra A. 1985/2007. Lexical nominalization. In *Language Typology and Syntactic Description*, Vol. 3: *Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 334–381. Cambridge: CUP.

Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Chicago IL: University of Chicago Press.

Cruz, Aline. 2011. Fonologia e gramática do Nheengatú. PhD dissertation, Free University Amsterdam.

Crystal, David. 1980. *A Dictionary of Linguistics and Phonetics*. Oxford: Wiley-Blackwell.

Davis, Henry. 2010. A unified analysis of relative clauses in St'át'imcets. *Northwest Journal of Linguistics* 4(1): 1–43.

Dayley, Jon P. 1989. *Tümpisa (Panamint) Shoshone Grammar*. Berkeley CA: University of California Press.

DeLancey, Scott 1986. Relativization as nominalization in Tibetan and Newari. Ms, presented at the 19th International Conference on Sino-Tibetan Languages and Linguistics.

DeLancey, Scott. 2002. Relativization and nominalization in Bodic. In *Proceedings of the Twenty-eighth Annual Meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics*: *In Honor of Prof. James A. Matisoff*, Patrick Chew (ed.), 55–72. Berkeley CA: BLS.

Derbyshire, Desmond. 1999. Carib. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 23–64.

Dixon, Robert M. W. 2006. Complement clauses and complementation strategies in typological perspective. In *Complementation: A Cross-linguistic Typology*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 1–48. Oxford: OUP.

Dixon, Robert M. W. 2010. *Basic Linguistic Theory*, Vol 2. Oxford: OUP.

Dixon, Robert M. W. & Aikhenvald, Alexandra Y. (eds). 2009. *The Amazonian Languages*. Cambridge: CUP.

Dryer, Matthew S. 2007. Noun phrase structure. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, Timothy Shopen (ed.), 151–205. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.003

Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and Empirical Foundations*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.

Everett, Daniel. 1986. Pirahã. In *Handbook of Amazonian Languages*, Vol. 1, Desmond Derbyshire & Geoffrey Pullum (eds), 200–326. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110850819.200

Everett, Daniel. 2005. Cultural constraints on grammar in Pirahã: Another look at the design features of human language. *Current Anthropology* 46: 621–646. https://doi.org/10.1086/431525

Everett, Daniel. 2009. Pirahã culture and grammar: A response to some criticisms. *Language* 85: 405–442. https://doi.org/10.1353/lan.0.0104

Fillmore, Charles. 1976. Frame semantics and the nature of language. In Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech. Volume 280: 20-32.

Fillmore, Charles. 1982. Frame semantics. In *Linguistics in the Morning Calm*. 111-137. Seoul: Hanshin Publishing Co.

Frellesvig, Bjarke. 2010. *A History of the Japanese Language*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511778322

Galucio, Ana Vilacy. 2011. *Subordinate adverbial constructions in Mekens*. In van Gijn, Haude & Muysken (eds), 25–44.

Gerdts, Donna B. & Hinkson, Mercedes Q. 2004. Salish numeral classifiers: A lexical means to a grammatical end. *Language Typology and Universals* 57: 247–279. https://doi.org/10.1524/stuf.2004.57.23.247

Givón, T. 2009. *Genesis of Syntactic Complexity*. Amsterdam: John Benjamins.

González, Hebe. 2005. A Grammar of Tapiete (Tupi-Guarani). PhD dissertation, University of Pittsburgh.

Gorbet, Larry. 1974. Relativization and Complementation in Diegueño: Noun Phrases as Nouns. PhD dissertation, University of California, San Diego.

Göksel, Aslı & Kerslake, Celia. 2005. *Turkish: A Comprehensive Grammar*. London: Routledge. https://doi.org/10.4324/9780203340769

Graczyk, Randolph. 2007. *A Grammar of Crow*. Lincoln NE: University of Nebraska Press.

Guerrero, Lillián. 2012. On relative clauses and related constructions in Yaqui. In Comrie & Estrada-Fernández (eds), 97–126.

Gutièrrex-Bravo, Rodrigo. 2012. Relative clauses in Yucatec Maya: Light heads vs. null heads. In Comrie & Estrada-Fernández (eds), 253–268.

Hale, Ken. 1976. The adjoined relative clause in Australia. In *Grammatical Categories in Australian Languages*, Robert M. W. Dixon (ed.), 78–105. Atlantic Highlands NJ: Humanities Press.

Hale, Ken & Platero, Paul. 1973. Aspects of Navajo anaphora: Relativization and pronominalization. *Diné Bizaad Náníl'iih* 1: 9–28.

Huade, Katharina. 2006. A Grammar of Movima. PhD dissertation, Radboud University Nijmegen.

Horie, Kaoru. 2008. The grammaticalization of nominalizers in Japanese and Korean: A contrastive study. In *Rethinking Grammaticalization: New Perspectives* [Typological Studies in Language 76], María José López-Couso & Elena Seoane (eds), 169–187. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.76.09hor

Huang, Chenglong. 2008. Relativization in Qiang. *Language and Linguistics* 9(4): 735–768.

Huang, Lillian M. 1995. *A Study of Mayrinax Syntax*. Taipei: The Crane Publishing Co.

Hutchison, John P. 1981. *The Kanuri Language: A Reference Grammar*. Madison WI: University of Wisconsin Press.

Iwasaki, Shoichi & Ingkaphirom, Preeya. 2005. *A Reference Grammar of Thai*. Cambridge: CUP.

Jacobson, Steven A. 1984. *Yup'ik Eskimo Dictionary*. Fairbanks AK: Alaska Native Language Center, University of Alaska.

Jaisser, Annie. 1987. Hmong classifiers. *Linguistics of the Tibeto-Burman Area*10 (2): 169–176.

Jelinek, Eloise. 1987. Headless relatives and pronominal arguments: A typological perspective. In *Native American Languages and Grammatical Typology. Papers from a Chicago Linguistics Society Parasession*, Paul D. Kroeber & Robert E. Moore (eds). Bloomington IN: Indiana University Linguistics Club.

Jensen, Cheryl.1999. Tupí-Guaraní. In Robert M. W. Dixon & Alexandra Aikhenvald (eds), 125–164.

Johns, Alana. 1982. A unified analysis of relative clauses and questions in Rainy River Ojibway. In *The Proceedings of the Thirteenth Algonquian Conference*, William Cowan (ed.), 161–168. Ottawa: Carleton University.

Jones, Wendell & Jones, Paula. 1991. *Barasano Syntax* [Studies in the Languages of Colombia 2]. Arlington TX: Summer Institute of Linguistics and the University of Texas.
Keenan, Edward. 1985. Relative clauses. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, Timothy Shopen (ed.), 141–170. Cambridge: CUP.
Keenan, Edward & Comrie, Bernard. 1977. NP accessibility and universal grammar. *Linguistic Inquiry* 8: 63–100.
Klumpp, James & Burquest, Donald A. 1983. Relative clauses in Piapoco. *International Journal of American Linguistics* 49(4): 388–399. https://doi.org/10.1086/465801
Koptjevskaja-Tamm, Maria. 1995. Possessive and relational forms in Chukchi. In *Double Case: Agreement by Suffixaufnahme*, Frans Plank (ed.), 301–321. Oxford: OUP.
Kroeber, Paul. 1997. Relativization in River Thompson Salish. *Anthropological Linguistics* 39: 376–422.
Kuroda, S.-Y. 1992. *Japanese Syntax and Semantics: Collected Papers*. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-2789-9
Lakoff, Geroge & Johnson, Mark. 2003. *Metaphors We Live By*. Chicago IL: University of Chicago Press. https://doi.org/10.7208/chicago/9780226470993.001.0001
Langacker, Ronald. 1987. *Foundations of Cognitive Grammar*, Vol. I. Stanford CA: Stanford University Press.
Langacker, Ronald. 1991. *Foundations of Cognitive Grammar*, Vol. II. Stanford CA: Stanford University Press.
Larsen, Thomas & Norman, William.1979. Correlates of ergativity in Mayan grammar. In *Ergativity: Towards a Theory of Grammatical Relations*, Frans Plank (ed.), 347–370. New York NY: Academic Press.
Lee, Sung-nyong. 1975. A study of the word Kes in Middle Korean. *Cintan Hakpo* 39: 105–138.
Lees, Robert. 1963. *The Grammar of English Nominalizations*. The Hague: Mouton.
Lefebvre, Claire & Muysken, Pieter. 1988. *Mixed Categories: Nominalizations in Quechua*. New York NY: Springer.
Lehmann, Christian. 1986. On the typology of relative clauses. *Linguistics* 24: 663–680. https://doi.org/10.1515/ling.1986.24.4.663
Li, Charles N. & Thompson, Sandra A.1989. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley CA: University of California Press.
Loos, Eugene. 1999. Pano. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 227–250.
Malchukov, Andrej. 2004. *Nominalization, Verbalization: Constraining a Typology of Transcategorial Operations*. Munich: Lincom.
Martin, Jack. 2011. *A Grammar of Creek (Muskogee)*. Lincoln NE: University of Nebraska Press. https://doi.org/10.2307/j.ctt1df4h3s
Matisoff, James. 1972. Lahu nominalization, relativization, and genitivization. In *Syntax and Semantics*, Vol. 1, John Kimball (ed.), 237–258. New York NY: Seminar Press.
Matthews, Stephen & Yip, Virginia. 1994. *Cantonese: A Comprehensive Grammar*. London: Routledge.
Mchombo, Sam. 2004. *The Syntax of Chichewa*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511486302
Messineo, Cristina. 2003. *Lengua Toba* (*Guaycurú*). *Aspectos gramaticales y discursivos*. Munich: Lincom.
Messineo, Cristina & Porta, Andrés. 2009. Cláusulas realtivas en Toba (Guaycurú). *International Journal of American Linguistics* 75(1): 49–68. https://doi.org/10.1086/598206

Miller, Amy. 2001. *A Grammar of Jamul Tiipay*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110864823

Mithun, Marianne. 2012. Questionable relatives. In Comrie & Estrada-Fernández (eds), 269–300.

Miyaoka, Osahito. 2012. *A Grammar of Central Alaskan Yupik (CAY)*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110278576

Montgomery-Anderson, Brad. 2008. A Reference Grammar of Oklahoma Cherokee. PhD dissertation, University of Kansas.

Moore, Denny. 2012. Relative clauses in Gavião of Rondônia. In Comrie & Estrada-Fernández (eds), 243–252.

Moyse-Faurie, Claire. 2016. Referential markers in Oceanic nominalized constructions. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada Fernández (eds), 171–203. Amsterdam: John Benjamins.

Nagaya, Naonori. 2011. The Lamaholot Language of Eastern Indonesia. PhD dissertation, Rice University.

Nevis, Andrew, Pesetsky, David & Rodrigues, Cilene. 2009. Pirahã exceptionality: A reassessment. *Language* 85: 355–404. https://doi.org/10.1353/lan.0.0107

Nicolle, Steve. 2013. *A Grammar of Digo: A Bantu Language of Kenya and Tanzania*. Dallas TX: SIL International.

Noonan, Michael. 1985. Complementation. In *Language Typology and Syntax Description*, Vol. 2: *Complex Constructions*, Timothy Shopen (ed.), 42–141. Cambridge: CUP.

Noonan, Michael. 1997. Versatile nominalizations. In *Essays on Language Function and Language Type. Dedicated to T. Givón*, Joan Bybee, John Haiman & Sandra A. Thompson (eds), 373–394. Amsterdam: John Benjamins. https://doi.org/10.1075/z.82.21noo

Noonan, Michael. 2008. Nominalization in Bodic languages. In *Rethinking Grammaticalization: New Perspectives* [Typological Studies in Language 76], María José-López-Cruso & Elena Seoane (eds), 219–237. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.76.11noo

Payne, Doris. 1985. Aspects of the Grammar of Yagua: A Typological Perspective. PhD dissertation, University of California, Los Angeles.

Payne, Thomas. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: CUP.

Rice, Keren. 1989. *A Grammar of Slave*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110861822

Riddle, Elizabeth. 1989. White Hmong noun classifiers and referential salience. Paper presented at the Association of Asian Studies, inc. annual meeting, Washington DC.

Riddle, Elizabeth. 1993. The relative maker *uas* in Hmong. *Linguistics of the Tibeto-Burman Area* 16(2): 57–68.

Rodrigues, Aryon. 1999. Macro-Jê. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 165–206.

Ross, John Robert. 1973. Nouniness. In *Three Dimensions of Linguistic Theory*, Osamu Fujimura (ed.), 137–257. Tokyo: The TEC Corporation.

Sakel, Jeanette. 2004. *A Grammar of Mosetén*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110915280

Salanova, Andrés Pablo. 2011. *Relative clauses in Mẽbengokre*. In van Gijn, Haude & Myusken (eds), 45–78.

Schachter, Paul & Otanes, Fe T. 1972/1983. *Tagalog Reference Grammar*. Berkeley CA: University of California Press.

Seki, Lucy. 2000. *Gramática do Kamaiurá, Língua Tupi-Guarani do Alto Xingu*. Campinas: Editora UNICAMP and São Paulo State Official Press.

Searle, John. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: CUP.

Shibatani, Masayoshi. 2008. Relativization in Sasak and Sumbawa, Eastern Indonesia. *Language and Linguistics* 9(4): 865–916.

Shibatani, Masayoshi. 2009. Elements of complex structures, where recursion isn't: The case of relativization. In *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution* [Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 163–198. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.07ele

Shibatani, Masayoshi. 2017. Nominalization. In *Handbook of Japanese Syntax*, Masayoshi Shibatani, Shigeru Miyagawa & Hisashi Noda (eds), 271–332. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9781614516613-009

Shibatani, Masayoshi. 2018a. Nominalization in crosslinguistic perspective. In *Handbook of Japanese Contrastive Linguistics*, Prashant Pardeshi & Taro Kageyama (eds), 345–410. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9781614514077-013

Shibatani, Masayoshi. 2018b. Nominalization. In *Handbook of Japanese Linguistics*, Yoko Hasegawa (ed.), 431–462. Cambridge: CUP.

Shibatani, Masayoshi & Awadh bin Makhashen, Khaled. 2009. Nominalization in Soqotri, a South Arabian language of Yemen. In *Linguistics of Endangered Languages: Contributions to Morphology and Syntax*, W. Leo Wetzels (ed.), 9–31. Leiden: Brill.

Shibatani, Masayoshi & Shigeno, Hiromi. 2013. Amami nominalizations. *International Journal of Okinawan Studies* 4(1): 107–138.

Shimoji, Michinori. 2008 A Grammar of Irabu, A Southern Ryukyuan Language. PhD dissertation, Australian National University.

Sposato, Adam. 2012. Relative clauses in Xong (Miao-Yao). *Journal of the Southeast Asian Linguistics Society* 5: 49–66.

Starosta, Stanley, Pawley, Andrew & Reid, Lawrence A. 1982. The evolution of focus in Austronesian. In *Papers from the 3rd International Conference on Austronesian Linguistics*, Vol. 2, Amran Halim, Lois Carrington & Stephen A. Wurm (eds), 145–170. Canberra: Australian National University.

Stolz, Thomas, Kettler, Sonja, Stroh, Cornelia & Urdze, Aina (eds). 2008. *Split Possession: An Areal-linguistic Study of the Alienability Correlation and Related Phenomena in the Languages of Europe* [Studies in Language Companion Series 101]. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.101

Suttles, Wayne. 2004. *Musqueam Reference Grammar*. Vancouver: UBC Press.

Tamura, Yuki-Shige. 2017. Grammaticalization and nominalization in Central Alaskan Yup'ik. Presentation at the Nominalization Festival III, Osaka University, 8 July.

Thiesen, Wesley & Weber, David. 2012. *A Grammar of Bora with Special Attention to Tone*. Dallas TX: SIL International.

Thornes, Tim. 2003. A Northern Paiute Grammar with Texts. PhD dissertation, University of Oregon.

Thornes, Tim. 2012. The nominalization-relativization connection in Northen Paiute. In Comrie & Estrada-Fernández (eds), 147–170.

Treis, Yvonne. 2008. Relativization in Kambaata from a typological point of view. In *Interaction of Morphology and Syntax: Case Studies in Afroasiatic* [Typological Studies in Language 75], Zygmunt Frajzyngier & Erin Shay (eds), 161–206. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.75.08tre

van der Voort, Hein. 2009. Possessive expressions in the Southwestern Amazon. In *The Expression of Possession*, William B. McGregor (ed.), 343–388. Berlin: Mouton de Gruyter.

van Gijn, Rik, Haude, Katharina & Muysken, Pieter. 2011. *Subordination in South America: An overview*. In van Gijn, Haude & Myusken (eds), 1–24.

van Gijn, Rik, Haude, Katharina & Muysken, Pieter (eds). 2011. *Subordination in Native South American Languages*, [Typological Studies in Language 97]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.97

Van Valin Jr., Robert D. 1977. Aspects of Lakohta Syntax. PhD dissertation, University of California, Berkeley.

Vendler, Zeno.1967. *Linguistics in Philosophy*. Ithaca NY: Cornell University Press.

Watkins, Laurel J. 1984. *A Grammar of Kiowa*. Lincoln NE: University of Nebraska Press.

Weber, David. 1989. *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley CA: University of California Press.

Willie, Mary A. 1989 Why there is nothing missing in Navajo relative clauses? In *Athapaskan Linguistics: Current Perspectives on a Language Family*, Eung-Do Cook and Karen Rice (eds), 407–437. Berlin: Mouton de Gruyter.

Yamada, Yoshio. 1908. *Nihonbunpōron (Theory of Japanese grammar)*. Tokyo: Hōbunkan.

Yaowapat, Natchanan & Prasithrathsint, Amara. 2009. A typology of relative clauses in mainland Southeast Asian languages. *The Mon-Khmer Studies Journal* 38: 1–23.

Yap, Foong Ha & Grunow-Hårsta, Karen. 2010. Non-referential uses of nominalization constructions: Asian perspectives. *Language and Linguistics Compass* 3(1): 1–22.

Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick. 2011. Introduction. Nominalization strategies in Asian languages. In Yap, Grunow-Hårsta & Wrona (eds), 1–57. https://doi.org/10.1075/tsl.96

Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick (eds), 2011. *Nominalization in Asian Languages: Diachronic and Typological Perspectives* [Typological Studies in Language 96]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.96

Zariquiey, Roberto. 2011. A Grammar of Kashibo-Kakataibo. PhD dissertation, La Trobe University.

Zhu, Dexi . 1982. *Yǔfǎ jiǎngyì (Lecture notes on grammar)*. Beijing: Shangwu Yinshuguan.

CHAPTER 3

Nominalization in cross-linguistic diachronic perspective

Sonia Cristofaro
University of Pavia

While more and more data are now available on languages rich in nominalizations, such as those of Latin America, the literature on nominalization is mainly synchronically orientented. The paper discusses several pieces of diachronic evidence about the origins of nominalization cross-linguistically. This evidence challenges the idea, widely held in the functional-typological literature, that the use of nominalizations reflects a non-default treatment of particular expressions, and that this motivates the distinguishing properties of nominalizations vis-a-vis other constructions. Diachronic evidence also points to possible motivations for the fact that nominalizations fail to consistently display the same structural properties, both cross-linguistically and within individual languages, and fail to be consistently used in the same contexts from one language to another.

1. Introduction

Over the past decades, more and more data have become available on languages rich in nominalizations, such as those of Latin America. This has led to renewed interest in these constructions in functionally and typologically oriented research. In this framework, nominalizations have mainly been investigated in relation to subordinate clauses, word formation, and parts of speech classes (Koptjevskaja-Tamm 1993, Croft 1991 and 2001, Hengeveld 1992, Cristofaro 2003, Malchukov 2004, Comrie and Thompson 2007). In addition, a variety of phenomena have been described that are related to nominalizations cross-linguistically. For example, the reanalysis of constructions involving nominalizations can give rise to new alignment, TAM, voice, and word order patterns. The ellipsis of a main predicate taking a nominalized complement can lead to patterns where the latter is used independently to convey the meaning originally associated with the construction as a whole, for example background information, various types of modal meanings, exclamations, or hot news (insubordination: Evans 2007, Mithun 2008, Cristofaro 2016).

https://doi.org/10.1075/tsl.124.03cri

These patterns have been described for many languages of Latin America (see, for example, Gildea 1998 and many of the papers collected in van Gijn, Haude, and Muysken 2011 and Comrie and Estrada-Fernández 2012), and are discussed in several papers in this volume (Bruil, Cahon, Gipper and Yap, Machado, Peña). They basically involve a number of diachronic processes whereby constructions involving nominalizations can give rise to new ones. Comparatively little attention, however, has been devoted to how nominalizations arise in the first place. Theoretical studies of nominalization as a general phenomenon usually only refer to the synchronic properties of different nominalization types, for example in terms of argument structure, presence vs. absence of particular categorial distintions (e.g. TAM distinctions), or the entity type denoted by the construction, e.g. agents, patients or actions (Koptjevskaja-Tamm 1993, Malchukov 2004, Comrie and Thompson 2007). Research on nominalization in individual languages has collected evidence about the origins of nominalizers, the dedicated morphemes sometimes used to mark nominalizations, but the relevant data are unsystematic and have not been integrated into theoretical treatments of nominalization in general.

The goal of this paper is to show that, while overall scanty, the available diachronic evidence on the origins of particular nominalization types cross-linguistically poses various challenges for a number of traditional assumptions about nominalization as a general phenomenon. In particular, this evidence challenges an idea, widely held in functionally and typologically oriented approaches, that the use of nominalizations reflects a non-default treatment of particular expressions, and that this motivates the distinguishing properties of nominalizations vis-a-vis other constructions. Diachronic evidence also points to possible motivations for various phenomena not fully accounted for in traditional views of nominalization, such as the fact that nominalizations fail to consistently display the same structural properties, both cross-linguistically and within individual languages, and the fact that they fail to be consistently used in the same contexts from one language to another.

A full understanding of nominalization, then, requires evidence about the historical origins of individual nominalization types in particular languages, in addition to data on the synchronic properties of these constructions. As this type of evidence is currently generally lacking for the languages of Latin America, the relevant issues will be illustrated based mainly on other languages. As will be shown in Sections 4–5, however, these issues have general implications for a number of structural and distributional properties of nominalizations that are cross-linguistically widespread and can be observed in many languages of Latin America. In this respect, it is hoped that the paper will point to new issues in the description and analysis of nominalizations in individual languages, particularly, as is the case for Latin America, languages where these constructions have been investigated in some detail, but mainly in a synchronic perspective.

2. Some traditional assumptions about nominalization

The constructions identified as nominalizations in the literature are ones where lexical roots denoting processes, states or properties display at least some of the structural properties usually associated in the language with lexical roots denoting things, persons, or places. These properties include, for instance, case or gender marking, determiners, or possessive marking on the arguments notionally corresponding to A, P, or S arguments (that is, following a standard practice in typology, the two argument of transitive clauses and the only argument of intransitive clauses).[1]

Due to the presence of these properties, for example, relative clauses have been identified as nominalizations in several Tibeto-Burman languages (DeLancey 1999, among others). Relative clauses in many languages of Latin America are also often analyzed in this way. A case in point is Epps' (2008, 2009) analysis of relative clauses in Hup, a Nadahup language of Amazonia where the relative clause verb can take case and number affixes. In (1), for example, 'steal' carries an object affix.

Hup (Nadahup)

(1) *tɨh=tæ̃hʔín-ăn-ɑ̂h,* [tɨh *toh-ʔé-p=ɑ̂y-***ă̰n**-*ɑ̂h*]$_{NMLZ}$
3SG = wife-OBJ-DECL 3SG steal-PERF-DEP=FEM-OBJ-DECL
'to the woman he had stolen' (Epps 2009: 292)

Similar analyses of relative clauses in other languages of Latin America are provided, for example, in Weber (1983), da Silva Facundes (2000), and several papers in Comrie and Estrada-Fernández (2012), van Gijn, Haude, and Muysken (2011), Chamoreau and Estrada-Fernández (2016), and this volume.

Many current approaches to nominalization assume that this phenomenon originates from a non-default treatment of particular expressions, and that this motivates the distinguishing properties of nominalizations vis-a-vis other constructions. When used in the description of individual languages, for example, the

1. This description is intentionally neutral as to the status of the relevant constructions in terms of parts of speech distinctions. Nominalizations are traditionally defined as constructions 'turning something into a noun' (Comrie and Thompson 2007: 334). The distinction between nouns and other parts of speech is, however, a problematic issue, which many linguists argue can only be resolved in a language-specific and construction-specific way (Croft 2001, among others). Also, the various constructions identified as nominalizations in the literature usually only display some of the properties that can be regarded as distinctive for nouns in the language. For these reasons, these constructions are described here in terms of specific combinations of structural properties and types of conceptual entity denoted by a lexical root (processes, states, or properties as opposed to things, persons, or places), rather than in terms of an opposition between nouns and other parts of speech.

notion of nominalization typically involves an underlying assumption that nominalizations are special constructions used when verbs or adjectives are exceptionally being assimilated to nouns, so that they display at least some properties of the latter. In a number of functionally oriented theoretical approaches, nominalizations are assumed to reflect the fact that particular expressions are being used in a non-default function. For example, Heine and Kuteva (2007: 107) suggest that nominalizations are used when some expression encodes a non-default construal of particular conceptual entities, in the sense that non-time stable, dynamic phenomena (of the type usually encoded by lexical roots denoting processes) are construed as time-stable, thing-like phenomena (of the type usually encoded by lexical roots denoting things, persons and places). Similar ideas are developed in a number of models of parts of speech proposed in typology and cognitive linguistics (Hopper and Thompson 1984 and 1985, Langacker 1987 and 1991, Croft 1991 and 2001, Hengeveld 1992). In these models, different parts of speech classes are defined by default combinations of lexical roots on the one hand and discourse functions or cognitive profiles on the other. Nominalizations and other constructions such as predicate nominals or predicate adjectives encode non-default combinations, which are possibly perceived by speakers as peripheral members of the class. For example, Hopper and Thompson (1984, 1985) and Croft (1991, 2001) assume that a speaker's mental representation of different parts of speech classes, namely nouns, verbs and adjectives, has a prototype structure with central and peripheral members. Prototypical nouns obtain when lexical roots denoting persons, things or places are used in discourse in order to refer to an entity. Prototypical verbs obtain when lexical roots denoting processes or states are used to predicate something about an entity, while prototypical adjectives obtain when lexical roots denoting properties are used to modify some other expression. Nominalizations are used when lexical roots denoting processes, states or properties occur in referring function.[2] These combinations are unexpected and less frequent than the prototypical ones where roots denoting processes or states are used in predicating function, or roots denoting properties are used in modifying function. As a result, they may be signaled through nominalizers. Also, the construction may not display the full array of properties found when the relevant roots occur in their prototypical function, for example (for roots denoting processes or states)

2. It should be noted that this view contrasts with several descriptions of individual languages where particular constructions are identified as nominalizations even though they are not used for reference. Relative clauses, for example, are analyzed as nominalizations in many languages despite that they are traditionally regarded as performing a modifying, rather than a referring function (though see Shibatani and Makhashen 2009 and Álvarez González, this volume, for an analysis of relative clauses as referring expressions).

TAM or person distinctions, but it may display properties normally associated with prototypical nouns (case or gender marking, determiners, possessive marking on arguments).[3]

The idea that the use of nominalizations reflects a non-default treatment of particular expressions is based on the synchronic structural properties of the relevant constructions, for example presence of nominalizers or absence of particular categorial distinctions. To the extent that particular factors are assumed to motivate nominalization, however, those factors should in principle play a role in the diachronic processes leading to the development of the relevant constructions in individual languages. In what follows, it will be argued that in many cases the available evidence about these processes does not actually support the idea that nominalizations originate from a non-default treatment of particular expressions, and that this motivates their distinguishing properties vis-a-vis other constructions. In particular, these properties often reflect the properties of particular source constructions that give rise to the nominalization, rather than an opposition between the default and non-default uses of particular expressions.

3. Nominalization and the origins of nominalizers

Diachronic evidence about the origins of nominalizations mainly pertains to a number of recurrent cross-linguistic processes that give rise to nominalizers. In particular, nominalizers have been shown to typically develop from semantically generic expressions such as 'thing', 'matter', 'one', 'person', 'that', and the like, which occur in referring function in different types of source constructions and evolve into nominalizers as these constructions are reanalysed.

In many cases, for example, the source construction is one where the referring expression is modified by an expression denoting a process or a state, that is, 'the one/the person Verbing (something)' (or 'the Verbing one/person'), 'matter/thing (of) Verbing' (or 'Verbing matter/thing'), 'place for Verbing' (or 'Verbing place'). Alternatively, this expression can be used in predicating function, e.g. 'one Verbs'. Over time, the construction maintains its global meaning, but the specific meaning of the referring expression is bleached, so that its referring function is transfered onto the construction as a whole, and the original referring expression survives as a marker of this function, that is, a nominalizer.

3. Langacker (1987) and (1991) and Hengeveld (1992) propose similar models, but Langacker defines parts of speech in terms of prototypical cognitive profiles, rather than discourse function, while Hengeveld's model does not involve prototypicality.

While diachronic evidence about the etymology of nominalizers is generally lacking for the languages of Latin America, Gipper and Yap (this volume) suggest, for example, that this may have been the origin of the nominalizer *=ti* attested in the Bolivian language Yurakaré, which is structurally similar to the demonstrative pronoun *ati*. Similarly, Moore (1989) shows that in Gavião, a Tupian language of Brazil, the nominalizers *méne* and *mát* also functions as pronouns, as can be seen for *méne* in (2a) and (2b). This suggests that the relevant nominalizations developed from constructions where an expression denoting a process or a state modified the pronoun, for example, for (2a), 'that of hunting, the cause' or the like.

Gaviã o (Tupi)

(2) a. *[gakoráá* ***méné****]*$_{NMLZ}$ *tígí*
hunt NMZR cause
'cause to hunt' (Moore 1989: 314)

b. ***méne*** *sot-ka teé b'o tá-máà*
that bad-make CONTIN FOC 3PL-AUX.PAST
'They messed that up.' (Moore 1989: 311)

Outside Latin America, these processes have been postulated for several languages, for example in the Tibeto-Burman family. A case in point are Classical and Lhasa Tibetan, where the nominalizer *-rgyu*, illustrated in (3) and (4), is historically derived from the noun *rgyu* 'matter, substance' (Beyer 1992: 296).

Classical Tibetan (Sino-Tibetan)

(3) *[ña-la dgos-****rgyu****]*$_{NMLZ}$
I-to need-NMZR
'something for me to need' (Beyer 1992: 297)

Lhasa Tibetan (Tibeto-Burman)

(4) *['di'i skad = cha dris =* ***rgyu****]*$_{NMLZ}$ *gus = zhabs med-pa red*
this-GEN question ask-NMZR polite not.be-PERF/DISJUNCT
'It's not polite to ask about this.' (DeLancey 2003: 284)

The etymology of the nominalizer suggests that the original structure of these constructions may have been 'my needed substance', 'the matter of asking', or the like.

In Lotha Naga, the nominalizer *-ò* probably developed from a demonstrative pronoun (Herring 1991: 66). Thus, a relative clause such as the one in (5), 'the boy who will come tomorrow' may have been, originally, 'the boy, that one coming tomorrow' (or, literally, 'the boy, coming tomorrow that one'; more on this in Section 5 below).

Lotha Naga (Tibeto-Burman)

(5) *ēpóeróró [ocüà rō sa-**ò**]*$_{NMLZ}$
boy tomorrow come VM-NMZR
'the boy who will come tomorrow' (Herring 1991: 61)

In Qiang, the nominalizer *-m* is historically derived from the noun *mi* 'person' (LaPolla 2003: 223–9), so that the construction in (6), literally 'the hat wearer', must have been, originally, 'the person wearing a hat', as in the English translation.

Qiang (Tibeto-Burman)

(6) *[tawə-ta-**m** le-ze]*$_{NMLZ}$
hat-wear-NMZR DEF-CL
'the person wearing a hat' (LaPolla 2003: 224)

The Niger-Congo language Supyire has a range of nominalizers transparently related to lexical items, for example 'person', 'thing', 'time', or 'place' (Carlson 1994: 107–119). These are illustrated by the constructions in (7a)–(7d) below, which are plausibly derived, respectively, from structures of the type 'beg person' (for 'person that begs'), 'thing make noise' (for 'thing makes noise', 'thing that make noise', or 'thing to make noise'), 'thing separate' (for 'thing separates', 'thing that separate', or 'thing to separate'), 'time to pay taxes', 'place where one lies down'.[4]

Supyire (Niger-Congo)

(7) a. *[ŋáára-**fóo**]*$_{NMLZ}$
beg-NMZR
'beggar'
(cfr. *foo* 'owner, possessor, person in charge': Carlson 1994: 115–6)

b. *[**ya**-tin-ŋɛ]*$_{NMLZ}$
NMZR-make.noise-G2
'musical instrument' (cf. *yaaga* 'thing': Carlson 1994: 112)

c. *Ndé la à [py ˋŋàmi-pìì **kà**-laha-ní]*$_{NMLZ}$ *kè*
DEM it PERF be twins-DEF NMZR-let.go-G3 REL
'that which caused the separation of the twins (from each other)'
(cf. *kyaa* 'thing': Carlson 1994: 112–3)

d. *Kà lànmpú-ŋi [tèè-kaan-ní]*$_{NMLZ}$ *sì nɔ̀*
and taxes-DEF NMZR-give-G3 NARR arrive
'Then the time to pay taxes arrived.' (cf. *tèrè* 'time': Carlson 1994: 113)

4. In cases where multiple readings are provided for the source construction, this is because the latter is compatible with all of these readings, and there is no evidence to decide for any of them.

e. *[***ta***-sinaga]*$_{NMLZ}$
NMZR-lie.down
'bedroom, place where one lies down'
(cf. *tɛʔɛ́* 'place' in the related language Cebaara: Carlson 1994: 110)

In Mojave, the nominalizer *ʔč-*, used in agent nominalizations and illustrated in (8a), is related to the indefinite pronoun *ʔč* 'something' illustrated in (8b). An expression such as 'bird' in (8a), then, must have been originally 'something flies', 'something that flies' (Munro 1976: 229).

Mojave (Hokan)

(8) a. *[ʔč-iyer]*$_{NMLZ}$
NMZR-fly
'bird' (Munro 1976: 229)

b. *ʔč* *isva:r*
something sing
'sing something' (Munro 1976: 229)

A different path leading to the development of nominalizers from originally referring expressions has been proposed by Estrada-Fernández (2008, 2012) for Pima Bajo, a Uto-Aztecan language of Mexico. In this language, a nominalizer *-kɨg*, used in relative clauses and illustrated in (9), may have evolved from a demonstrative element *higai* 'that one'.

Pima Bajo (Uto-Aztecan)

(9) *ɨg* *okosi* *[in = nɨir-*~~**kɨg**~~*]*$_{NMLZ}$ *ɨg* *gɨ'id*
DET.SUBJ woman 1SG.NONSUBJ = see.PERF-REL DET.SUBJ big
'The woman I saw is big.' (Estrada-Fernández 2012: 134)

Estrada-Fernández argues that, while the nominalization originated from the combination of the demonstrative with a co-occurring expression denoting a state or a process, the two were originally part of two distinct clauses in the source construction, with the demonstrative occurring in what becomes the main clause in the relative clause construction. For example, the original structure of a sentence such as the one in (9) would have been 'the woman I saw, that one is big', or the like. Over time, the demonstrative became attached to the preceding lexical root and evolved into a nominalizer. In this scenario, contrary to the other cases described above, the elements that give rise to the nominalization combine as an epiphenomenal result of linear adjacency, rather than because they stand in a specific relationship vis-a-vis each other.

The developmental processes postulated for nominalizers cross-linguistically have several consequences for traditional assumptions about nominalizations and their structural properties, as described in Section 2. For one thing, these processes

show that nominalizations need not originate from some non-default treatment of particular expressions. Rather, they can develop as semantically generic expressions are used in their standard referring function. In the source construction, these expressions are accompanied by a modifying or predicating expression, or are eventually combined with some adjacent expression modifying some other element in the sentence, as in the Pima Bajo case illustrated in (9) above. Over time, the referring function is transfered onto the construction as a whole, so that the modifying or predicating expressions become directly associated with this function. This, however, is a side effect of the referring expression losing its original meaning, rather than an effect of a non-default treatment of the modifying or predicating expressions in themselves.

In this scenario, contrary to traditional assumptions (see e.g. Hopper and Thompson 1984 and 1985, Langacker 1987 and 1991, Croft 1991 and 2001), the use of nominalizers cannot be regarded as a way to signal that particular expressions are being treated in a non-default way. Nominalizers develop from elements that are originally used in their standard function, that is, as referring expressions, and are there because their meaning ('person', 'one', 'matter', 'place' and the like) provides a specific contribution to the overall meaning of the construction. These elements survive in the construction even when their meaning is not transparent any more, at which stage they function as semantically generic morphemes that identify the construction, that is, nominalizers. This development, however, is a result of a process of semantic bleaching and grammaticalization, rather than the fact that speakers make a conceptual distinction between the default and non-default uses of particular expressions and signal this distinction through special morphology.[5]

Other structural properties of nominalizations also need not be a result of a non-default treatment of particular expressions. For example, absence of categorial distinctions typically associated with predication, such as TAM distinctions, is often assumed to originate from the fact that particular expressions are used for reference rather than predication (Hopper and Thompson 1984: 737–8, among others). In the source construction, however, these expressions are actually used for predication or modification, not reference, and alternative explanations are sometimes available for the absence of particular categorial distinctions in the resulting nominalization.

5. In such cases, the fact that elements denoting processes or states are used in modifying function in the source construction can be regarded as a non-default use of these elements, because they are usually used for predication (see Deutscher 2009 for similar remarks). This, however, is irrelevant to the issue of the function of nominalizers. The construction does not originally involve any nominalizers, and when the referring expression evolves into a nominalizer, the latter does not signal modification anyway.

For example, in constructions of the type of the Lhasa Tibetan one in (4) or the Supyire one in (7d), absence of TAM distinctions directly follows from the fact that the relevant expressions do not describe a specific occurrence of some state of affairs, but are rather used to present the entity denoted by the referring expression as an instance of a particular type ('the matter of asking' as opposed to some other matter, rather than a specific instance of an asking event; 'a place for lying down' as opposed to some other type of place, rather than a place where somebody lies down on some specific occasion). Of course, this type of explanation is related to the specific properties of particular source constructions that give rise to the nominalization, so it may or may not be applicable to different nominalization types, and should be tested against actual diachronic data about the origins of the relevant constructions. The general point is, however, that, if particular nominalizations originally consist of the combination of a referring expression and a modifying or predicating expression, absence of particular categorial distinctions cannot be explained in terms of an a priori assumption that the modifying or predicating expression is exceptionally being used for reference. Rather, this phenomenon should be investigated in light of what specific categorial distinctions are missing in the nominalization, and the original structure and semantics of the source construction.

Similar observations apply to another distinguishing property of nominalizations, the fact that expressions denoting processes, states or properties display morphology typically associated with ones denoting things, persons or places (such as case or gender markers, determiners, or possessive marking on arguments). This is generally taken as a result of the fact that the former expressions are being assimilated to the latter, possibly, as assumed in functionally oriented approches, because they exceptionally encode the same type of conceptual construals or are exceptionally being used for the same discourse function (reference). Nominalizers, however, are derived from expressions denoting things, persons or places, so it is possible that the presence of the relevant morphology in the nominalization is originally due to the presence of these expressions, rather than some non-default treatment of expressions denoting processes, states of properties (in fact, as mentioned earlier, these expressions are not used in referring function in the source construction). For example, in the Qiang construction in (6), the definite classifier could originally have applied to the element *mi* 'person' that provided the source for the nominalizer. In the Pima Bajo sentence in (9), the nominalizer evolved from a demonstrative, and the A argument in the relative clause is indexed through possessive morphology. The use of this morphology could be due to the fact that the entity denoted by the demonstrative was originally possessed by the notional A argument of the construction, that is, 'the woman, my seen one', as has been proposed for similar structures in other languages, e.g. several Cariban languages (Gildea 1998) and West Greenlandic (Fortescue 1995).

Another case in point is provided by the Mojave relative clause construction illustrated in (10).

Mojave (Hokan)

(10) *hatčoq [ʔ-u:ta:v-n^{y} -č]$_{\text{NMLZ}}$ ən^{y} əʔi:l^{Y}-pč*
dog 1-hit-DEM-SUBJ black-TNS
'The dog I hit is black.' (Munro 1976: 194)

In this construction, the lexical head of the nominalized relative clause is case-marked according to its role in this clause, while the verb in the relative clause carries a demonstrative affix followed by a case marker that indexes the role of the head in the main clause. For example, in (10), 'dog' is zero marked because it is the P argument of the relative clause verb, while this verb carries the subject case marker because 'dog' occurs as an S argument in the main clause. Given the SOV structure of the language, this construction could in principle have evolved from one of the type 'dog I hit, that one is black' (for 'I hit the dog, that one is black'), where the demonstrative occurs in what becomes the main clause in the relative clause construction and is case-marked according to its role in this clause. The nominalization could be a result of the demonstrative combining with the preceding verb, in which case the case marking on the nominalized verb would be the one originally applying to the demonstrative.[6]

This type of explanation too may or may not actually apply to different nominalizations, including the cases just discussed. In particular, individual properties (for example, the markers used for particular arguments) may or may not be actually compatible with the original structure of the source construction and the meaning of the element that gave rise to the nominalizer. Also, the relevant morphology can occur in the absence of nominalizers. For example, Trio, a Cariban language of Suriname, has various types of nominalizations where expressions denoting states or processes are directly combined with morphology normally

6. While this possibility is not discussed in the literature on Mojave and related languages displaying similar relative clause constructions (see, for example, Langdon 1970, Gorbet 1976, or Miller 2001), this process is similar to the one postulated by Estrada-Fernández (2008, 2012) for Pima Bajo. A similar process is also reconstructed by Heine and Reh 1984 for the Niger-Congo language Ewe. In this language, sentences involving relative clauses involve two distinct relative clause markers, e.g. 'Woman REL came yesterday REL is no longer here' for 'The woman who came yesterday is no longer here'. The two relative markers originated, respectively, from a postposed demonstrative and a definite article in a construction of the type 'Woman that, the yesterday having come one, is no longer here' (for 'that woman, the one who came yesterday, is no longer here'). While the resulting construction is not regarded as a nominalization by Heine and Reh, this process resembles the ones described here for Pima Bajo and Mojave in that the relative clause originates from the combination of adjacent elements originally belonging to different clauses.

associated with expressions denoting things, persons or places, including case affixes, plural marking, and possessive person markers (Carlin 2004: 351–61).

Trio (Cariban)

(11) a. *[wewe-ton tuna-ton ihkërën-ma-**ke**]*$_{NMLZ}$
tree-PL water-PL destroy-INCH.STAT-INSTR
'because the trees and rivers are being destroyed' (Carlin 2004: 353)

b. *[president **i**-w-ëepï-se=**to**]*$_{NMLZ}$ *n-a-i*
President 3POSS-1TR-come-DESID=PL 3 > 3-be-NCERT
tï-pata-pona
3POSS.COREF-village-DIR
'They want the president to come to their village.' (Carlin 2004: 499)

In such cases, the relevant morphology plausibly applied from the beginning to the expression denoting a state or a process. In line with traditional views of nominalization, then, its use should be assumed to be related to the function of this expression, rather than the presence of some other expression in the source construction. As long as the nominalization involves a nominalizer, however, the use of this morphology could in principle be related to the original presence of an expression denoting a person, a thing or a place, which later gave rise to the nominalizer. In general, then, this phenomenon cannot be accounted for in terms of an apriori assumption that some expression is being treated in a non-default way (in the sense of being assimilated to some other expression). Rather, this is an issue that needs to be investigated on a case-by-case basis, in light of the specific properties of the relevant nominalization types (for example, presence vs. absence of nominalizers) and the properties of the source construction.

4. Diachrony and the structural diversity of nominalizations

Nominalizations display considerable structural diversity, both cross-linguistically and within individual languages. As the details are extensively discussed in the literature (Koptjevskaja-Tamm 1993, Malchukov 2004, Comrie and Thompson 2007, Cristofaro 2003 and 2007, Yap and Wrona 2011), only a few representative examples from languages of Latin America will be discussed here for illustrative purposes.

In general, some major parameters of structural variation are whether or not the nominalization is marked by a nominalizer, what categorial distinctions are encoded in the construction, presence vs. absence of different types of morphology usually associated with referring expressions, and the encoding of arguments (what arguments are encoded overtly, whether or not individual arguments are encoded in the same way as in non-nominalized constructions).

In Trio, for example, some nominalizations have no nominalizers, as in (11) above and (12a), while others display a variety of nominalizers, for example *-ne*, as in (12b). Argument roles are usually indicated through person indexation in the language. In constructions without nominalizers, however, notional A arguments are not indexed, whereas P and S arguments are indexed by possessive person affixes, rather than the person affixes used in non-nominalized constructions. This can be seen from (12a), which also shows that A arguments can be encoded as goal NPs. As can be seen from (12b), on the other hand, nominalizations in *-ne* denote notional A arguments, and P arguments are indexed through possessive person affixes. Both nominalization types can take past tense markers (Carlin 2004: 351–69).

Trio (Carib)

(12) a. *[ë-eta-se]*$_{NMLZ}$ *w-a-e* ***i-ja***
2POSS-hear-DES 1 >3.1TR-be-CERT 3-GOAL
'I want him to listen to you.' (Carlin 2004: 356)

b. *j-i-ponopï-rëken [j-ene-**ne**-npë-ton]*$_{NMLZ}$
1-POSS-TR-tell-only 1POSS-see-NMZR-PAST-PL
'The people who knew (saw) me told me.' (Carlin 2004: 368)

In Apuriña, the arguments of nominalizations in *-inhi* are encoded in the same way as in non-nominalized constructions, as can be seen from the treatment of the first person argument in (13a) and (13b). The nominalization can take aspect markers, such as the progressive marker in (13a).

Apurinã (Arawakan)

(13) a. *[aiko* ***nota*** *sa-nanu-t-inhi-mokaru]*$_{NMLZ}$
house 1SG go-PROGR-VRBLZR-NMZR-GOAL
'my being going to the house'[7] (da Silva Facundes 2000: 608)

b. ***nota*** *muteka*
1SG run
'I run.' (da Silva Facundes 2000: 247)

In Huallaga (Huánaco) Quechua (Weber 1983, 1989), nominalizations in *-q* and *-sha-* are used in relative clauses. Nominalizations in *-q* can only be used to relativize A and S arguments, as in (14a), while nominalizations in *-sha-* can also be used to relativize other grammatical roles, for example P arguments, as in (14b). In the relative clause, the relativized arguments are not encoded overtly, while non-relativized arguments are marked for case in the same way as in non-nominalized

7. The element glossed as 'verbalizer' is a formative that must be added to various bases in order for these bases to combine with several types of grammatical elements, including for example directional, causative, and progressive markers (da Silva Facundes 2000: 305–25).

constructions, but are indexed by possessive person prefixes. This can be seen from the accusative marker on the P argument 'you' in (14b) and the possessive person indexes for the A argument in (14a) and the P argument in (14b). While these two nominalizations types cannot encode tense, they are inflected for aspect, as can be seen from the imperfective marker in (14b).

Huallaga (Huánuco) Quechua (Quechuan)

(14) a. *Runa [maqa-**sha-yki**]$_{NMLZ}$ sha:-yka-mu-n*
man hit-NMZR-2POSS come-IMPFV-afar-3
'The man whom you hit is coming.' (Weber 1989: 281)

b. *[Qam-ta maqa-**shu-q**]$_{NMLZ}$ sha:-yka-mu-n runa*
you-OBJ hit-2POSS-NMZR come-IMPFV-afar-3 man
'The man who hit you is coming.' (Weber 1989: 281)

In Hixkaryana, nominalizations in *-nye* cannot have overtly encoded A or S arguments, while P arguments are indexed through possessor prefixes. Apart from person, none of the inflectional distinctions normally allowed to verbs in the language (tense, aspect, mood and voice) is encoded in the construction.

Hixkaryana (Carib)

(15) *[r-ompamnohɨ-**nye**]$_{NMLZ}$*
1POSS-teach-NMZR
'the one who teaches me' (Derbyshire 1979: 167)

Traditional views of nominalization in general, as outlined in Section 2, provide no explanation for the structural diversity of different nominalization types. If particular structural properties of individual nominalizations are manifestations of some general phenomenon, namely some non-default treatment of particular expressions, then it is not clear why these properties should fail to consistently appear from one nominalization type to another, both cross-linguistically and within individual languages. For example, the idea that nominalizers are used to signal that particular expressions are used in a non-default function is weakened by the fact that many nominalizations do not display nominalizers, as shown by the Trio construction in (12b). Likewise, if absence of particular categorial distinctions or use of particular morphology reflect the fact that particular expressions are being treated as referring expressions, then one needs to account for why not all nominalizations display these particular properties, as shown by the Apuriña, Huallaga (Huánaco) Quechua and Hixkaryana constructions in (13)–(15).

Structural diversity is, however, expected in a diachronically oriented approach where the properties of individual nominalizations originate from properties of particular source constructions, rather than reflecting some general phenomenon. For example, nominalizers will be found in a nominalization if the source construction

involves elements that grammaticalize into nominalizers, and they won't be found if the source construction involves no such elements. Similarly, as detailed in Section 3, absence of particular categorial distinctions or presence of morphology normally used for referring expressions may be a consequence of the properties of particular source constructions that give rise to the nominalization. These phenomena, then, need not manifest when the nominalization originates from a different source. While these hypotheses need to be investigated on a case-by-case basis, they point to a new research approach to nominalization, one in which the structural properties of individual nominalizations are assessed in the perspective of possible source constructions for that particular nominalization, rather than in the perspective of some more general phenomenon independent of these constructions.

5. Diachrony and the distribution of nominalizations

The use of nominalizations as opposed to non-nominalized constructions is not consistent cross-linguistically. From one language to another, the same contexts may or may not allow the use of nominalizations in apparently arbitrary fashion.

A typical environment for nominalizations are, for example, complement clauses (Cristofaro 2003, Noonan 2007, among others). In many languages, however, the use of nominalizations is limited to complements of particular types of main predicates, and these are not the same from one language to another. This can be observed, once again, in several languages of Latin America. In Mosetén, for example, nominalizations are used in complements of manipulative predicates, that is, predicates describing a process of causation or attempted causation ('make', 'ask to', 'order' and the like), as can be seen from (16a) below. However, 'finish' verbs, illustrated in (16b), take non-nominalized complements.

Mosetén (Mosetean)

(16) a. *Yäe ködye-yë [sob-a-k-**dye**' öi-yä' phen]*$_{NMLZ}$
1SG beg-1SG/2SG visit-VM-AP-NMZR F-L.F-AD woman
'I beg you to visit this woman.' (Sakel 2004: 432)

b. *Äej-ä-i phe-ya-ki jike ya-ksi aj*
stop-VM-M.SUBJ talk-VM-AP.M.SUBJ PAST say-3PL.O.M.SUBJ yet
phi-ke-dye-si' …
run-VM-BEN-L.F
'When they had finished to talk he said to them concerning the race …'
(Sakel 2004: 431)

Pilagá displays the opposite pattern, that is, manipulative verbs cannot take nominalized complements, ((17a)), but 'finish' verbs can ((17b)).

Pilagá (Guaykuruan)

(17) a. *hayem se-na(t)-pega da' t'-ont-aʂan*
1SG 1-say-ASP COMP 3-work-VD
'I asked (him) to work.' (Vidal 2001: 359)

b. *soʂote [y-imat di' l-onta-naʂak]$_{NMLZ}$*
before 3-finish CL 3POSS-work-NMZR
'He finished his work.' (Vidal 2001: 356)

Individual languages also often use different nominalizations in different contexts. For example, as can be seen from the sentences in (14) above, Huallaga (Huánaco) Quechua uses different nominalizations in different types of relative clauses: nominalizations in *-sha* are usually used to relativize items other than A and S arguments, while these arguments are relativized through nominalizations in *-q*. A similar situation is found with complement clauses. As illustrated in (18), nominalizations in *-sha* are also used in complements of utterance verbs, while complements of perception verbs require nominalizations in *-q*.

Huallaga (Huánuco) Quechua

(18) a. *Chawra maman-shi willapaq wamran-ta [marka-chaw*
then his:mother-REPORT she:tells:him her:son-DAT town-LOC
*tiya-**sha**-n-ta]$_{NMLZ}$*
live-NMZR-3POSS-ACC
'Then his mother told her son that she had lived in a town.'
(Weber 1983: 89)

b. *muskishkaa [kamcha-ta rupa-yka-**q**-ta]$_{NMLZ}$*
I:smelled toasted.corn-ACC burn-IMPFV-NMZR-ACC
'I smelled that the corn was burning.' (Weber 1983: 95)

These distributional patterns are not accounted for by traditional views of nominalization. In these views, nominalization reflects the fact that particular expressions are assimilated to other expressions, possibly because they are exceptionally used for reference. In principle, one would expect this phenomenon to be triggered by particular properties of the contexts of use of the relevant expressions, for example properties leading to these expressions being used for reference rather than predication or modification. In this case, however, it is not clear why nominalization is not consistently attested across the same range of contexts cross-linguistically. Also, if particular structural properties of nominalizations are motivated by a non-default treatment of particular expressions, then, to the extent that particular contexts lead to this treatment, it is not clear why those contexts should not allow any nominalization displaying the relevant structural properties. For example, in classical definitions of complement clauses (Noonan 2007, among others), these are clauses functioning as arguments of a main predicate, hence they can be assumed

to be performing a referring, rather than a predicating function. In traditional views, this provides a motivation for the use of nominalizations in these clauses. In this case, however, it is not clear why this use should be limited to particular complement clause types in some languages, nor why these should vary arbitrarily from one language to another. On a similar note, if nominalizers are used to signal a non-default treatment of particular expressions, as traditionally assumed, then any nominalizer will perform this function, so it is not clear why particular contexts should not allow the use of particular nominalizers as opposed to others.

These issues, however, can be at least partly accounted for by looking at the available diachronic evidence about the origins of nominalizations. While this evidence may not make it possible to shed light on specific individual cases (such as the ones described above), it shows that whether or not particular nominalizations can be used in particular contexts is related to the original meaning of the source construction.

This is illustrated in (19) for Qiang. In this language, the nominalizer *-m*, derived from the noun *mi* 'person' and discussed in regard to Example (6) above, is used to relativize subjects and recipients. Locations, on the other hand, are relativized through constructions involving a different nominalizer, *-s*, derived from a noun meaning 'place' or 'earth' (LaPolla and Huang 1996: 223–9).

Qiang (Sino-Tibetan)

(19) a. *upu [tʈi-ʈhə-topu-**m**-le:]*$_{NMLZ}$ *ʈʈəu-la zi*
uncle wine-drink-like/love-NMZR-DEF.CL home-LOC exist
'The uncle who likes drinking liquor is at home.' (LaPolla 2003: 228)

b. *[qa-wu-panə-dele-**m**]*$_{NMLZ}$ *mi*
1SG-AGT-thing-give-NMZR person
'the person to whom I gave something' (LaPolla and Huang 1996: 227)

c. *[qa-lu-**s**-ta]*$_{NMLZ}$
1SG-come-place-NMZR-LOC
'the place that I came from' (LaPolla and Huang 1996: 224)

As repeatedly pointed out in the literature (DeLancey 1986 and 1999, LaPolla and Huang 1996, Noonan 1997, Gildea 1998, Givón 2012, among others), a possible source for nominalized relative clauses are constructions where the lexical items that ultimately give rise to the nominalizer occur themselves as heads of a relative clause, e.g. 'the person who likes drinking liquor', 'the place that I came from'.[8] In the relative clause, the role of these elements will plausibly be related to their

8. Nominalized relative clauses arise from these constructions as the relevant lexical items evolve into nominalizers and the construction as a whole is used in apposition to other nouns, leading to new relative clause constructions where these nouns functions as heads. Constructions such as the ones in (19a-b), for example, would originally have been, literally, 'the uncle, the person

meaning, for example, nouns meaning 'person' will occur as agents or recipients in the relative clause, and nouns meaning 'place' will occur as locations. This will give rise to restrictions in the distribution of the resulting nominalizers across different types of relative clauses, in the sense that nominalizers arising from items normally occurring in particular roles in the relative clause will be restricted to those roles, at least initially.

In other cases, while particular uses of a nominalizer do not directly reflect its original meaning, they are plausibly derived from this meaning through processes of context-driven inference, as described in grammaticalization studies and studies of language change in general (Heine 2003, Traugott and Dasher 2005, among many others). In Qiang, for example, the nominalizer *-sa* is also used to relativize instruments, as in (20).

Qiang (Sino-Tibetan)

(20) a. *[laupin-tʂhopu-**s**]*$_{NMLZ}$ *səpe*
tubercolosis-treat-NMZR medicine
'medicine used to treat tubercolosis' (LaPolla and Huang 1996: 226)

b. *[stua-haʨcə-**s**]*$_{NMLZ}$ *tʂuats*
food-eat-NMZR table
'the table used for eating food' (LaPolla and Huang 1996: 226)

While no direct connection can be established between the notion of instrument and the original locative meaning of the nominalizer, some contexts are compatible with both, for example, in (20b) a table used for eating food is actually a table where food is eaten. Such contexts, then, may have determined the extension of the nominalization from the relativization of locations to that of instruments.

In Supyire, as mentioned in regard to Example (7) above, nominalizations in *ta-*`are used to denote locations, in accordance with the locative origin of the nominalizer. In addition, they are also used in temporal and purpose clauses, as illustrated in (21).

Supyire (Niger-Congo)

(21) a. *Uru u à pyi mìi shyéré-ŋi [wyɛ́rɛ́-ŋi*
he(EMPH) he PERF be my wittness-DEF money-DEF
***tá**-kan-ge e]*$_{NMLZ}$
LOC.NMZR-give-DEF at
'It was he who was my witness when the money was given.'
(Carlson 1994: 111)

who likes drinking liquor' and 'the person to whom I gave something, the person' (DeLancey 1999, LaPolla and Huang 1996, Noonan 1997, Givón 2012).

b. *Canŋ kà mì máha ŋ-kare dú-gé e [fàa*
day IND I PAST IP-go stream-DEF to fish
***tá**-cya-ge e]*$_{NMLZ}$
LOC.NMZR-seek-G2.SG to
'One day I went to the stream to catch fish.' (Carlson 1994: 111)

While temporal and purpose clauses do not specifically involve the notion of location, the use of the nominalization in these clauses may have developed in contexts involving this notion. If the nominalization denotes the place of an action, then speakers may infer that that it refers to the time of this action, that is, expressions of the type 'at the place where X takes place' can be reinterpreted as 'when X takes place', as in (7b). Motion towards the place of an action can be reinterpreted as motion in order to perform that action, that is, 'to the place where X takes place' is reinterpreted as 'to achieve X', as in (7c). These are in fact instances of a well-known grammaticalization process whereby spatial expressions develop new, more abstract meanings through processes of context-induced inference (Heine, Claudi, and Hünnemeyer 1991, among several others).

Old Chinese has a nominalizer *zhe* derived from a semantically generic noun that originally conveyed various meanings related to individuation, e.g. 'the one that', 'people/items possessing a certain feature'. Nominalizations in *zhe* can be used in contexts relatively consistent with these meanings, such as relative clauses, as well as contexts apparently unrelated to the notion of individuation, such as conditional clauses. The latter use is likely to have developed from the relative clause use through processes of inference in contexts such as the one in (22), which are compatible both with a relative clause interpretation and a conditional interpretation (Yap and Wang 2011).

Old Chinese (Sino-Tibetan)

(22) *[shun **zhe**]*$_{NMLZ}$ *cuo zhi*
obey **NMZR** leave him/them
'Those who obeyed/ If anyone obeyed, (he) left them alone.'
(Xun Zi: Yap and Wang 2011: 74)

Epps (2009) describes a similar process for Hup. In this language, nominalizations in *-Vp* (where *V* is a vowel copying the vowel of the preceding sillable) can be used as relative clauses and to convey a variety of adverbial meanings. Epps (2009) submits that the relative clause use is the original one. This use could be related to the original function of the nominalizer, as there is some evidence that the latter might have evolved from a topic marker, and several languages display formal parallels between relativization and topicalization. The adverbial clause use originated through the reanalysis of relative clauses lending themselves to an adverbial interpretation, of the type in (23).

Hup (Nadahup)

(23) *[ʔám = yɨʔ key-níh-ɨ̃p]*$_{NMLZ}$, *ʔăn* *bʾɔ̆t-an*
2SG = TEL see-be.like-NMZR 1SG.OBJ manioc.field-DIR
wɨdway-ʔý-áh
arrive.go.out-VENT-DECL
'Someone who looked like you/ Looking like you, (he) came to me in the manioc field.' (Epps 2009: 299)

Diachronic evidence also shows that the contexts that do not allow the use of particular nominalizers are ones incompatible with, or less directly related to the original semantics of the nominalizer. As shown by examples (24) and (25) below, for example, the Qiang nominalizer *-s* and the Supyire nominalizer *ta-* are not used in complements of 'want' verbs (in the two languages, these complements are not nominalized).

Qiang (Sino-Tibetan)

(24) *the: tɕəu kə ɕtɕaq-lu*
3SG home go heart-come
'She wants to go home.' (LaPolla and Huang 1996: 230)

Supyire (Niger-Congo)

(25) *Mìì lá mpyi u ú ˊŋ-káré*
my desire was he SUBJNCT IP-go
'I wanted him to go.' (Carlson 1994: 430)

In both of these cases, the meaning of the sentence is relatively incompatible with the original locative meaning of the two nominalizers, as witnessed by the oddity of sentences such as 'She wanted the place where she goes or 'I wanted the place where he goes (as opposed, for example, to (7b), 'He was my witness when/at the place where the money was given').

The facts just described provide a natural diachronic explanation for the distribution of particular nominalizations across different context, and suggest that this distribution may not be related to some non-default treatment of particular expressions. Rather, individual nominalizations will be used in contexts more directly related to the meaning of the source construction, and will not be used in contexts unrelated, or less directly related to this meaning. This is in accordance with classical assumptions in grammaticalization studies and historical linguistics in general, particularly the idea that the distribution of individual constructions is at least partly determined by their original meaning (see e.g. the notion of persistence proposed in Hopper and Traugott 2003).

Languages also display a number of cases where the distribution of some nominalization is in contrast with, or cannot be clearly related to the semantics of the source. In Classical Tibetan, for example, nominalizations marked by the

nominalizer *-rgyu* can be used to refer to humans despite that the nominalizer is derived from a noun meaning 'substance, matter', as mentioned in regard to (3) above. This is illustrated in (26).

Classical Tibetan (Sino-Tibetan)
(26) bla-ma *[oū-**rgyu**]*$_{NMLZ}$
lama come-NMZR
'the lama to come' (Beyer 1992: 296)

In several languages, nominalizers derived from locative expressions can be used to relativize not only locations, but also apparently unrelated grammatical roles, namely P arguments. This is the case with the Qiang nominalizer *-s*, as well as the Middle Chinese nominalizer *suo*, derived from a noun meaning 'place'.

Qiang (Sino-Tibetan)
(27) *[tɕile-(ŋuəɳi) pə-**s**]*$_{NMLZ}$ *pies ŋuə*
1PL-TOP buy-NMZR meat COP
'What we need (to buy) is meat.' (LaPolla and Huang 1996: 234)

Middle Chinese
(28) *[min zhi **suo** shi]*$_{NMLZ}$ *da di dou fan huo geng*
people GEN NMZR eat basically beans cuisine beans soup
'What people eat is basically cuisine and soup made of beans.'
(Zhan Guo Ce: Yap and Wang 2011: 83)

While in such cases the semantics of the source does not seem to provide an explanation for particular uses of the nominalization, this does not rule out that there could still be a link between the two that is not immediately apparent, for example through analogy or processes of context-induced inference of the type of those described for (20)–(23) above. These cases, then, do not invalidate the idea that the original meaning of a nominalization plays a key role in shaping its distribution. Rather, they point to the need for further research on the often highly particularized factors that may lead to particular nominalizations being extended from one context to another.

6. Concluding remarks

The available diachronic evidence about the development of nominalizations cross-linguistically poses some major challenges for traditional assumptions about nominalization in general. At least some nominalizations do not originate as special constructions used when particular expressions are being assimilated to others. Rather, they develop as some referring expression in the source construction loses its specific meaning, so that the properties of this expression are transfered

onto the construction as a whole. This implies that particular distinguishing properties of the nominalization may be a result of the original structure of the source construction, rather than some special treatment of particular expressions. This, however, need not be the case for all of the constructions that can be regarded as instances of nominalization under traditional criteria. This suggests that these criteria do not actually capture a unified phenomenon. Rather, they identify a series of constructions that originate through different mechanisms and may be motivated in terms of different principles. This provides an explanation for the structural diversity of nominalizations, as well as the variation in their cross-linguistic distribution across different contexts.

All this has significant implication for research on nominalization in individual languages. This research usually focuses on the structural synchronic properties of particular nominalizations, and sometimes the status of the nominalization in terms of part of speech distinctions, for example to what extent the construction can be regarded as an instance of a noun. Most analyses, however, do not deal with issues such as why the nominalization displays particular structural properties as opposed to others, or why it is used in particular contexts as opposed to others. This is the case with most existing treatments of nominalizations in the languages of Latin America, even comprehensive ones such as Weber 1983 for Huallaga (Huánaco) Quechua or da Silva Facundes 2000 for Apuriña.

In order to gain a full understanding of nominalization phenomena, then, it is essential for new research on this topic, particularly research on less described languages, to concentrate on the possible sources of individual nominalization types, as well as the specific diachronic processes that trigger the extension of particular nominalizations from one context to another. This can provide crucial clues as to why the nominalization displays particular structural properties, for example why it is or is not marked by nominalizers, why particular arguments are encoded as possessors, or presence vs. absence of particular inflectional distinctions. The contexts of occurrence of individual nominalizations also often form a complex network best understood in diachronic perspective.

While direct diachronic information on these issues may be difficult to obtain, significant progress can be made through internal reconstruction, intragenetic comparison, or simply by making hypotheses about possible connections between the various uses of a nominalized clause, much in the vein of works such as Epps 2009 or Yap and Wang 2011. As also discussed in Cristofaro 2012, this type of research has a bearing not only on nominalization phenomena in themselves, but also on a number of more general issues such as why speakers use different constructions in different types of subordinate clauses, the origins of non-finite verb forms, and what evidence do we actually have for prototype models of parts of speech and grammatical categories in general.

Abbreviations

>	acting on	GOAL	goal
1TR	one-argument transitive verb	INCH.STAT	inchoative stative
ACC	accusative	IND	indicative
AD	adessive	INSTR	instrumental
AGT	agentive marker	IP	intransitive verb prefix
AP	antipassive	L	linker
ASP	aspect	LOC	locative
AUX	auxiliary	M	masculine
BEN	benefactive	NARR	narrative
CERT	certainty	NCERT	non-certainty
CL	classifier	NMLZ	nominalization
COMP	complementizer	NMZR	nominalizer
CONTIN	continuing	NONSUBJ	non-subject
COP	copula	PAST	past
DAT	dative	POSS	possessive
DECL	declarative	REL	relative
DEF	definite	REPORT	reportative
DEM	demonstrative	SUBJ	subject
DES	desiderative	SUBJUNCT	subjunctive
DET	determiner	TEL	telic
DIR	directional	TNS	tense
DISJ	disjunctive	TOP	topic
EMPH	emphatic	TRANS	transitive
F	feminine	VD	valency derivation
FOC	focus	VENT	venitive
G2	gender2	VM	verbal marker
G3	gender 3	VRBLZR	verbalizer
GEN	genitive		

References

Beyer, Stephan V. .1992. *The Classical Tibetan language*. Albany: State University of New York Press.

Carlin, Eithne B. 2004. *A Grammar of Trio, a Cariban Language of Suriname*. Frankfurt am Main: Peter Lang.

Carlson, Robert. 1994. *A Grammar of Supyire*. Berlin and New York: Mouton de Gruyter. https://doi.org/10.1515/9783110883053

Chamoreau, Claudine & Estrada-Fernández, Zarina (eds). 2016. *Finiteness and Nominalization*. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.113

Comrie, Bernard & Estrada-Fernández, Zarina (eds). 2012. *Relative clauses in languages of Americas*. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.102

Comrie, Bernard & Thompson, Sandra A. 2007. Lexical nominalization. In *Language Typology and Syntactic Description*. 2nd Edition, Volume 3: *Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 334–381. Cambridge: CUP. https://doi.org/10.1017/CBO9780511618437.006

Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.

Cristofaro, Sonia. 2007. Deconstructing categories: Finiteness in a functional-typological perspective. In *Finiteness: all over the clause*. Irina Nikolaeva (ed.), 91–114, Oxford: OUP.

Cristofaro, Sonia. 2012. Cognitive explanations, distributional evidence, and diachrony. *Studies in Language* 36, 645–670. https://doi.org/10.1075/sl.36.3.07cri

Cristofaro, Sonia. 2016. Routes to Insubordination: a Cross-Linguistic Perspective. In *Insubordination*, Evans, Nicholas R. & Watanabe, Honoré (eds), 393–422. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.115.15cri

Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Chicago and London: The University of Chicago Press.

Croft, William. 2001. *Radical Construction Grammar*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

da Silva Facundes, Sidney. 2000. The Language of the Apurinã People of Brazil (Maipure/Arawak). PhD dissertation, State University of New York, Buffalo.

DeLancey, Scott. 1986. Relativization as nominalization in Tibetan and Newari'. Ms, University of Oregon. <http://tibeto- burman.net/nominalizationworkshop.html>

DeLancey, Scott. 1999. Relativization in Tibetan. In *Topics in Nepalese Linguistics*, Yogendra P. Yadava & Warren W. Glover (eds), 231–249. Kathmandu: Royal Nepal Academy.

DeLancey, Scott. 2003. Lhasa Tibetan. In *The Sino-Tibetan Languages*, Graham Thurgood & Randy J. LaPolla (eds), 270–88. London and New York: Routledge.

Derbyshire, Desmond C. 1979. *Hixkaryana*. Lingua Descriptive Studies. Amsterdam: North Holland.

Deutscher, Guy. 2009. Nominalization and the evolution of subordination. In *Syntactic Complexity*, Talmy Givón & Masayoshi Shibatani (eds), 199–214. Amsterdam and Philadelphia: John Benjamins https://doi.org/10.1075/tsl.85.08nom

Epps, Patience. 2008. *A Grammar of Hup*. Berlin and New York: Mouton de Gruyter. https://doi.org/10.1515/9783110199079

Epps, Patience. 2009. Escape from the noun phrase: From relative clause to converb and beyond in an Amazonian language. *Diachronica* 26, 287–318. https://doi.org/10.1075/dia.26.3.01epp

Estrada-Fernández, Zarina. 2008. *Cláusolas relativas en pima bajo. Línguas Indígenas Americanas* 8, 69–83.

Estrada-Fernández, Zarina. 2012. From demonstrative to relative marker to clause linker: Relative clause formation in Pima Bajo. In Bernard Comrie & Zarina Estrada-Fernández (eds), 127–146. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.102.06est

Evans, N. (2007). Insubordination and its uses. In *Finiteness: all over the clause*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.

Fortescue, Michael. 1995. The historical source and typological position of ergativity in Eskimo languages. *Etudes/Inuit/Studies* 19, 61–75.

Gildea, S. (1998). *On reconstructing grammar: Comparative Cariban morphosyntax*. Oxford: OUP.

Givón, Talmy. 2012. Towards a diachronic of relative clause. In Bernard Comrie and Zarina Estrada-Fernández (eds), 1–26. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.102.01giv

Gorbet, Larry P. 1976. *A Grammar of Diegueño Nominals*. New York: Garland.

Heine, Bernd. 2003. Grammaticalization. In *The Handbook of Historical Linguistics*, Brian D. Joseph and Richard D. Janda (eds), 576–601. Oxford: Blackwell. https://doi.org/10.1002/9780470756393.ch18

Heine, Bernd, Claudi, Ulrike & Hünnemeyer, Friederike. 1991. *Grammaticalization*. Chicago: University of Chicago Press.

Heine, Bernd & Kuteva, Tania. 2007. *The genesis of grammar : a reconstruction*. Oxford: OUP.

Heine, Bernd & Reh, Mechtild. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.

Hengeveld, Kees. 1992. *Non-verbal predication*. Berlin and New York: Mouton de Gruyter. https://doi.org/10.1515/9783110883282

Herring, Susan C. 1991. Nominalization, relativization, and attribution in Lotha, Angami, and Burmese. *Linguistics of the Tibeto-Burman Area* 14, 55–72.

Hopper, Paul J. & Thompson, Sandra A. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60, 703–752. https://doi.org/10.1353/lan.1984.0020

Hopper, Paul J. & Thompson, Sandra A. 1985. The iconicity of the universal categories noun and verb. In *Iconicity in Syntax*, John Haiman (ed.), 151–183. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.6.08hop

Hopper, Paul J. & Traugott, Elizabeth C. 2003. *Grammaticalization*. 2nd edition. Cambridge: CUP. https://doi.org/10.1017/CBO9781139165525

Koptjevskaja-Tamm, Marja. 1993. *Nominalizations*. London and New York: Routledge.

Langacker, Ronald W. 1987. Nouns and verbs. *Language* 63, 53–94. https://doi.org/10.2307/415384

Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar. Vol. II*: *Descriptive Applications*. Stanford: Stanford University Press.

Langdon, Margaret. 1970. *A Grammar of Diegueño. The Mesa Grande Dialect*. University of California Publications in Linguistics 66. Berkeley and Los Angeles: University of California Press.

LaPolla, Randy J. 2003. Qiang. In *The Sino-Tibetan Languages*, Graham Thurgood & Randy J. LaPolla (eds), 573–87. London and New York: Routledge.

LaPolla, Randy J. & Huang, Chenglong. 1996. *A Grammar of Qiang*. Berlin and New York: Mouton de Gruyter.

Malchukov, Andrej. 2004. *Nominalization/verbalization: Constraining a typology of transcategorial operations*. München and Newcastle: Lincom Europa.

Miller, Amy. 2001. *A Grammar of Jamul Tiipay*. Berlin and New York: Mouton de Gruyter. https://doi.org/10.1515/9783110864823

Mithun, Marianne. 2008. The extension of dependency beyond the sentence. *Language* 83, 69–119.

Moore, Denny. 1989. Gavião Nominalizations as Relative Clauses and Sentential Complements Equivalents. *International Journal of American Linguistics* 55, 309–325. https://doi.org/10.1086/466122

Munro, Pamela. 1976. *Mojave Syntax*. New York: Garland.

Noonan, Michael. 1997. Versatile Nominalizations. In *Essays in language function and language type*, Bybee, Joan, Haiman, John & Thompson, Sandra A. (eds), 373–394. Amsterdam: John Benjamins. https://doi.org/10.1075/z.82.21n00

Noonan, Michael. 2007. Complementation. In *Language Typology and Syntactic Description*. 2nd *Edition*, Volume 2: *Complex constructions*, Timothy Shopen (ed.), 52–150. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.002

Sakel, Jeannette. 2004. *A Grammar of Mosetén*. Berlin and New York: Mouton de Gruyter. https://doi.org/10.1515/9783110915280

Shibatani, Masayoshi & bin Makhashen, Khaled A. 2009. Nominalization in Soqotri, a South Arabian language of Yemen. In *Endangered languages*: *Contributions to Morphology and Morpho-syntax*, Wetzels, W. Leo (ed.), 9–31. Leiden: Brill.

Traugott, Elizabeth C. & Dasher, Richard B. 2005. *Regularity in Semantic Change*. Cambridge: CUP.

van Gijn, Rik, Haude, Katharina & Muysken, Pieter (eds). 2011. *Subordination in Native South American Languages*. John Benjamins. https://doi.org/10.1075/tsl.97

Vidal, Alejandra. 2001. Pilagá Grammar. PhD dissertation, University of Oregon.

Weber, David J. 1983. *Relativization and nominalized clauses in Huallaga (Huanaco) Quechua*. University of California Publications in Linguistics 103. Berkeley and Los Angeles and London: University of California Press.

Weber, David J. 1989. *A Grammar of Huallaga (Huánaco) Quechua*. University of California Publications in Linguistics 112. Berkeley and Los Angeles and London: University of California Press.

Yap, Foong Ha & Wang, Jiao. 2011. From light noun to nominalizer and more: The grammaticalization of *zhe* and *suo* in Old and Middle Chinese. In *Nominalization in Asian Languages: Diachronic and typological perspectives*, Yap, Foong-Ha, Grunow-Hårsta, Karen & Wrona, Janick (eds), 61–108. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.96.02yap

Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick. 2011. Introduction. Nominalization strategies in Asian languages. In *Nominalization in Asian Languages: Diachronic and typological perspectives*, Yap, Foong-Ha, Grunow-Hårsta, Karen & Wrona, Janick (eds), 1–57. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.96.01yap

PART II

Areal studies on nominalization in South America

CHAPTER 4

Case markers as subordinators in South American indigenous languages

Rik van Gijn
University of Zurich

Nominalization (in different forms and guises) is one of the most common subordination strategies in South American indigenous languages. A frequently used nominalization strategy is to use case markers to indicate the semantic or structural relationship of the nominalized clause to the proposition expressed in the main clause. This paper aims to establish the extent to which these extensions of case markers to subordinate clauses are non-random, and to separate the role of language contact, genealogical retention, and universal preferences in explaining the distribution of case markers as subordinators in South American indigenous languages.

1. Introduction

South American Indigenous languages show recurrent patterns in their subordination strategies (van Gijn et al. 2011). Recurring constructions, found across language families, are for instance the use of bound subordinators, internally headed relative clauses, multi-verb constructions, clause-chaining, switch reference, and nominalization. Van Gijn (2014) showed that nominalization as a subordination strategy is found significantly more often in South American languages than would be expected on the basis of global patterns. One of the particularly common constructions identified in van Gijn (2014) is the formation of adverbial clauses by using case markers (defined as functional elements that indicate the relationship a noun or NP bears to its head) to indicate the semantic relation of the adverbial clause to the proposition expressed in the main clause.

The present contribution is intended as a follow-up study to van Gijn (2014), aimed at achieving a more detailed perspective on case-marked subordinate clauses in South American languages. More specifically, it aims to establish the extent to which the observed patterns are non-random, and can be accounted for in terms of language contact, genealogical retention, or perhaps more general explanatory principles relating to human cognitive or communicative preferences.

https://doi.org/10.1075/tsl.124.04gij

The paper is organized as follows: in Section 2 I introduce some more background to the issue of case marking in subordinate clauses, outline the leading questions of the paper, and introduce the language sample used for this paper. Section 3 is dedicated to a comparison of the case systems of the sample languages. Section 4 describes the patterns found of case markers used in adverbial clauses, which are discussed in terms of different possible accounts in Section 5. In Section 6, finally, the general conclusions for the paper are drawn.

2. Preliminaries

It is a widely observed phenomenon that case markers, generally associated with noun phrases, are often also used in (functional equivalents of) subordinate clauses. This is found in unconnected geographical areas and across many different language families (e.g. Blake 2001; Heine 2008). This raises the question of what the driving forces behind this connection between NP case marking and subordinate clause case marking are. Three types of answers seem to be likely candidates:

1. Genealogical retention
   The functional extension of case markers to mark subordinate clauses is a common grammaticalization path and can be stated in terms of general grammaticalization principles (see e.g. Heine & Kuteva 2007; Heine 2008), but the specifics of this grammaticalization path may of course differ from one situation to another. One of the potentially determining factors for the distribution of case markers in different types of subordinate clause is genealogical retention. Although case systems seem to suffer rather different fates in different situations (see Kulikov 2006, 2008), some case systems seem to be particularly time-stable. We also know that some aspects of case marking, like patterns of syncretism (see Baerman & Brown 2005) have a substantial genealogical component. It is, therefore, conceivable that specific extensions of case markers to subordinators were established before the languages of a family started to diverge, and that they simply retained this extension.
2. Contact-induced diffusion
   It has long been recognized (see e.g. Weinreich 1953; Thomason & Kaufman 1988) that grammatical (especially morphologically bound) material is less easily borrowed than lexical material. It is less clear, however, how easily more abstract structural dimensions of systems may spread through contact (see e.g. Matras & Sakel 2007; Johanson 2008). It has furthermore been observed that some aspects of case marking, including the presence of case and the number of cases per language, but also some formal parameters like fusion, seem to

have (macro-)areal tendencies (Bickel & Nichols 2008). It may therefore also be the case that language contact is the main factor responsible for the occurrence of case markers in subordinate clauses.

3. Functional pressures
   The fact that case markers are often found on subordinate clauses may in itself be regarded as resulting from functional pressures. One of the explanatory principles invoked in Cristofaro (2003) to explain patterns of subordination is the likelihood for a subordinate event to be construed as a thing (versus a process). Following Langacker (1987a, b), Cristofaro argues that entities are processed differently than e.g. actions or properties. The distinguishing cognitive feature of nouns is "that they designate sets of entities that are scanned summarily as a unitary whole (things)" (Cristofaro 2003: 159), whereas actions are prototypically processed in phases that occur sequentially in time. In subordination constructions, in Cristofaro's approach, the main event imposes its processing profile over the entire construction, leaving the dependent events to be scanned without such a sequential profile, which makes them cognitively more like things. Moreover, subordinate clauses typically perform discourse and syntactic functions associated with nouns (argument of a verb, reference, etc.) which also makes them more like nouns, and therefore more likely to acquire nominal characteristics like case marking (see Croft 1991; Malchukov 2006).

The three answers, in their pure form, predict different distributional patterns: the genealogical factor predicts similar extensions of case markers to subordinate clauses within but not across families, whereas the areal factor predicts the opposite pattern. The functional factor would predict overall, continent-wide preferences. Of course, the three answers are not mutually exclusive, and different factors may have conspired to produce the actual patterns.

The goal of this paper is to establish a) the bandwidth of case marking in South America in terms of presence of case and the types of cases that are present, b) the extent to which the case functions that are present in South American languages have been extended to marking subordinate verbs, and c) which factors might have played a role in shaping the distributions that we find. In order to achieve these goals I have looked at a sample of 60 South American languages, representing 26 families and 10 isolate languages. The approximate location of the sample languages is given in Map 1, the designations of the numbers can be found in Table 1.

Map 1. The sample

Table 1. The sample languages, their iso-codes, affiliations and main sources

#	Name	iso	affiliation	Main source(s)
1	Ika	ARH	CHIBCHAN	Frank 1985
2	Warao	WBA	ISOLATE	Romero-Figueroa 1997
3	N Embera	EMP	CHOCOAN	Mortensen 1999
4	Panare	PBH	CARIBAN	Payne & Payne 2013
5	Yanam	SHB	YANOMAMAN	Goodwin-Gómez 1990
6	Puinave	PUI	ISOLATE	Girón 2008
7	Emérillon	EME	TUPIAN	Rose 2011
8	Páez	PBB	ISOLATE	Jung 2008
9	Trio/Tiriyó	TRI	CARIBAN	Meira 1999
10	Cubeo	CUB	TUCANOAN	Morse & Maxwell 1999
11	Awa Pit	KWI	BARBACOAN	Curnow 1997
12	Hup	JUP	MAKUAN	Epps 2008
13	Desano	DES	TUCANOAN	Miller 1999
14	Tariana	TAE	ARAWAKAN	Aikhenvald 2003
15	Imbabura Qu	QVI	QUECHUAN	Cole 1982
16	Dâw	KWA	MAKUAN	Andrade-Martins 2004
17	Hixkaryana	HIX	CARIBAN	Derbyshire 1985
18	Tsafiki	COF	BARBACOAN	Dickinson 2002
19	Miraña	BOA	BORAN	Seifart 2005, Thiesen & Weber 2012
20	Yagua	YAD	PEBA-YAGUAN	Payne 1985
21	Kokama	COD	TUPIAN	Vallejos 2010
22	Matses	MPQ	PANOAN	Fleck 2003
23	Urarina	URA	ISOLATE	Olawsky 2006
24	Aguaruna	AGR	JIVAROAN	Overall 2007
25	Timbira	XRI	MACRO-GE	Popjes & Popjes 1986, Alves 2004
26	Shipibo-K	SHP	PANOAN	Valenzuela 2003
27	Jarawara	JAA	ARAWAN	Dixon 2004
28	Apurinã	APU	ARAWAKAN	Facundes 2000
29	Kakataibo	CBR	PANOAN	Zariquiey 2011
30	Karitiana	KTN	TUPIAN	Storto 1999, 2011; Everett 2006
31	Huallaga Qu.	QUB	QUECHUAN	Weber 1989
32	Yaminahua	YAA	PANOAN	Faust & Loos 2002
33	Karo	ARR	TUPIAN	Gabas Jr. 1999

(*continued*)

Table 1. (*continued*)

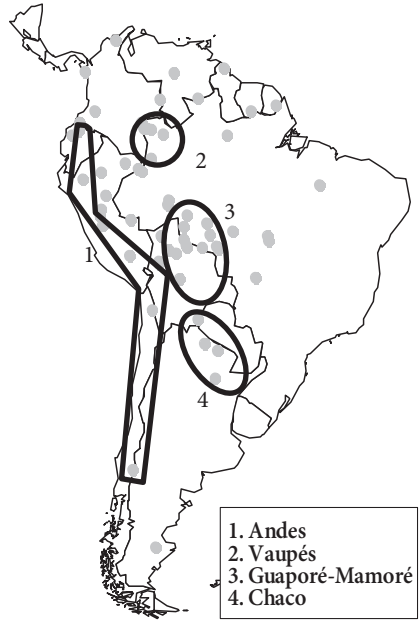
#	Name	iso	affiliation	Main source(s)
34	Yanesha'	AME	ARAWAKAN	Duff-Trip 1997
35	Wari'	PAV	CHAPACURAN	Everett & Kern 1997
36	Rikbaktsa	RKB	MACRO-GE	Silva 2011
37	Kwazá	XWA	ISOLATE	van der Voort 2004
38	Ese Ejja	ESE	TACANAN	Vuillermet 2012
39	Trumai	TPY	ISOLATE	Guirardello 1999
40	Kanoê	KXO	ISOLATE	Bacelar 2004
41	Kamaiurá	KAY	TUPIAN	Seki 2000
42	Itonama	ITO	ISOLATE	Crevels 2012
43	Sabanê	SAE	NAMBIKWARAN	Antunes 2004
44	Mekens	SKF	TUPIAN	Galucio 2001
45	Baure	BRG	ARAWAKAN	Danielsen 2007
46	Mamaindê	WMD	NAMBIKWARAN	Eberhard 2009
47	Cavineña	CAV	TACANAN	Guillaume 2008
48	Movima	MZP	ISOLATE	Haude 2006
49	Cuzco Qu.	QUZ	QUECHUAN	Lefebvre & Muysken 1988; Muysken p.c.
50	Mosetén	CAS	MOSETENAN	Sakel 2004
51	Leko	LEC	ISOLATE	van de Kerke 2009
52	Bororo	BOR	MACRO-GE	Crowell 1979; Nonato 2008
53	Yurakaré	YUZ	ISOLATE	van Gijn 2006
54	Aymara	AYR	AYMARAN	Hardman 2001
55	Tapiete	TPJ	TUPIAN	González 2005
56	Wichí	WLV	MATACOAN	Terraza 2009
57	Pilagá	PLG	GUAYCURUAN	Vidal 2001
58	Mocoví	MOC	GUAYCURUAN	Grondona 1998
59	Mapudungun	ARN	ISOLATE	Smeets 2008
60	Tehuelche	TEH	CHONAN	Fernandez Garay 1998, 2004

In terms of genealogy, the sample is built up as indicated in Table 2:

Table 2. Genealogical units in the sample

Isolates	12	Arawan	1
Tupian	6	Aymaran	1
Arawakan	4	Boran	1
Panoan	4	Chapacuran	1
Cariban	3	Chibchan	1
Macro-Ge	3	Chocoan	1
Quechuan	3	Chonan	1
Barbacoan	2	Jivaroan	1
Guaycuruan		Matacoan	1
Makuan	2	Mosetenan	1
Nambikwaran	2	Peba-Yaguan	1
Tacanan	2	Yanomaman	1
Tucanoan	2		

Proposed linguistic areas (Sprachbünde) are indicated in Map 2: the Andes (Torero 2002), the Vaupés (Aikhenvald 2002), the Guaporé-Mamoré (Crevels & van der Voort 2008), and the Chaco (Comrie et al. 2010).



Map 2. Linguistic areas

3. Case systems in South American languages and their extensions to subordinate clauses

In this section I will first review case marking in general and then zoom in on the use of case markers as markers of interclausal relations in complex clauses. In order to structure the discussion, I have made a number of subdistinctions within possible case systems. First I distinguish between core, or structural, case on the one hand (ergative, accusative, genitive), and peripheral, or semantic, case on the other. This distinction is not without its problems, and we will see that South American languages often have polysemous case markers with both core and peripheral functions. For the sake of exposition, I have treated these multifunctional markers as instances of homonymy. By doing so, I do not intend to make a synchronic statement about these markers, but rather to try to group the behavior of the different core functions versus peripheral functions of case markers since they are quite different and may also be expected to lead to different types of grammaticalizations in subordinate clauses. I will mention where relevant when core verbal case markers or genitive case markers have peripheral functions as well. Peripheral functions are subdivided into dative/benefactive[1] (Section 3.3), instrument/comitative[2] (Section 3.4), and spatial cases (Section 3.5). Two further categories are 'oblique' (Section 3.6), for those markers that have more than one of the peripheral functions just mentioned, and 'other' (Section 3.7) for those markers to which none of the above categories apply. First, however, I will assess the extent to which the languages in the sample have case at all (Section 3.1).

3.1 Presence of case

Case markers in this paper are defined as functional elements that indicate the relationship a noun or NP bears to its head. I take a broad view on case markers, where I disregard fusion as a criterion, but rather look at grammatical wordhood, following Bickel & Nichols (2007). Excluded are those markers that either take obligatory agreement or govern case, since they form grammatical words of their own. Map 3 shows the languages with case markers (black) and the languages

1. Here I have not taken into account the fact that in some languages dative behaves as a core case. Instead I gave preference to the semantic connections between dative and benefactive markers.

2. This grouping is based on their frequent formal/functional connection cross-linguistically (Stolz et al. 2005, 2006: 23–25).

without case markers (grey).[3] As can be seen, case as a category is very widespread in South American languages, and is found in Andean as well as Amazonian languages. The languages that do not have case use one (or both) of two strategies: verbal strategies and/or adpositional strategies. Both these strategies are in fact widespread, but most languages use them in addition to case marking.

Map 3. Presence of case

Sample languages with case

Tariana, Apurinã, Yanesha', Baure, Jarawara, Aymara, Awa Pit, Tsafiki, Miraña, Trio/Tiriyó, Hixkaryana, Ika, Embera, Northern, Tehuelche, Mocoví, Aguaruna, Timbira, Rikbaktsa, Bororo, Hup, Dâw, Mapudungun, Mosetén, Sabanê, Nasa Yuwe, Matses, Shipibo-Konibo, Kakataibo, Yaminahua, Yagua, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Ese Ejja, Cavineña, Cubeo, Desano, Karitiana, Karo, Mekens, Emérillon, Kokama, Kamaiurá, Tapiete, Warao, Puinave, Urarina, Kwaza, Trumai, Kanoê, Movima, Leko, Yurakaré, Yanam

Sample languages without case

Panare, Wari', Pilagá, Wichí, Mamaindê, Itonama

An example of a language that resorts to verbal strategies seemingly exclusively is Itonama, an isolate language spoken in northeast Bolivia (reference number 42). Crevels (2012: 248) mentions that the language has no case markers or adpositions. What it does have, is person marking, applicatives, incorporation, and spatial and directional markers on the verb, which can perform the functions normally associated with case.

Other sample languages that show similar behavior are the Matacoan language Wichí (Terraza 2009), neighboring the Guaykuruan language Pilagá (Vidal 2001), and Nambikwaran Mamaindê (Eberhard 2009). The pattern of "verbal case marking" is exemplified for Wichí in Example (1).

3. Languages without case markers in the sample are Panare, Wari', Pilagá, Wichi, Mamaindê, and Itonama.

(1) Wichi [MATACOAN] Terraza 2009: 220
*n-p'u-**hu** wahat n-k^{y}oti*
1-roast-**APPL** fish 1POSS-grandfather
'I roast fish for my grandfather.'

The other type of alternative strategy is exemplified by the Chapacuran language Wari', which has a single preposition, which cannot be counted as a case marker because it shows agreement. This preposition is also used to mark certain subordinate clauses.

(2) Wari' [CHAPACURAN] Everett & Kern 1997: 22
*param 'ina-em **pain** [ca mao wa]$_{NMLZ}$*
desire 1SG:REA.NONFUT-2SG **PREP: 3NEUT** IRR.NONFUT go INF
'I want you to go.'

Inflected adpositions are a common feature in Tupian (Rodrigues & Cabral (2012) and Cariban languages (Derbyshire 1999), though not necessarily to the exclusion of case markers. Verbal relation marking through e.g. applicatives, motion and position markers is also a very common strategy, which seems to have more areal characteristics (see e.g. Wise 2002; Guillaume & Rose 2010; van Gijn 2015). In the remainder of this paper, only the languages with case markers that fall within the definition given above will be considered.

3.2 Core case

An issue that has come up in the discussion about Andean versus Amazonian profiles is the presence of core case and, related to that, the alignment pattern of the language. Andean languages are reputed to have rather clear accusative alignment patterns in main clauses, marked by case. Alignment patterns in Amazonian languages tend to be either ergative-based, or to exhibit split marking in one of the roles. Moreover, case markers to indicate structural relations (both in the clausal and in the noun phrase context) are said to be rare in Amazonian languages (see e.g. Dixon & Aikhenvald 1999: 8). In this paper I define core case in a shallow way as case markers that encode ergative, accusative, or genitive (nominative and absolutive marking is too uncommon to yield any useful patterns).[4]

Map 4 shows the presence in the sample languages of one or more of these three cases. As can be seen in Map 4, structural case is certainly not uncommon in Amazonian languages, but it does seem to be more present across the board in Andean languages.

4. I stay on the conservative side and do not include zero markers, since their empirical status is often problematic.

Map 4. Presence of core case

Languages without core case

Tariana, Apurinã, Baure, Jarawara, Panare, Trio, Hixkaryana, Wari', Pilagá, Mocoví, Mapudungun, Wichí (Mataco), Mamaindê, Nasa Yuwe, Yagua, Karitiana, Karo, Mekens, Emérillon, Kokama, Tapiete, Urarina, Itonama, Movima, Leko, Yurakaré

Languages with core case

Yanesha', Aymara, Awa Pit, Tsafiki, Miraña, Ika, Embera, Northern, Tehuelche, Aguaruna, Timbira, Rikbaktsa, Bororo, Hup, Dâw, Mosetén, Sabanê, Matses, Shipibo-Konibo, Kakataibo, Yaminahua, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Ese Ejja, Cavineña, Cubeo, Desano, Kamaiurá, Warao, Puinave, Kwaza, Trumai, Kanoê, Yanam

For the subset of languages with core case, Maps 5 and 6 show languages with ergative and accusative case markers, respectively. The distribution of both types of case markers suggests that areal factors may play a role.

The languages with ergative case markers are Ika, Northern Embera, Yanam, Puinave, Timbira, Shipibo-Konibo, Kakataibo, Yaminahua, Ese Ejja, Trumai, and Cavineña. Ergative alignment is considered to be a recessive feature of languages (Nichols 2003), i.e. it tends not to be time-stable, either within families or within areas. Nevertheless, ergative case marking is stable in the Panoan (e.g. Loos 1999: 240) and Tacanan (Guillaume & Rose 2011: 464) language families.

Most of the ergative markers in the sample have several possible functions, which is consistent with the observation that ergative case markers most commonly develop from other case markers (McGregor 2009: 499). In Panoan languages, the ergative marker is generally homophonous with other case markers, e.g. in Matses, Shipibo and Kakataibo, the ergative marker *-n* is also used to mark instrumental, temporal, and genitival functions, as well as locative in Shipibo. In Yaminahua, the ergative markers seems to be less extended, marking ergative and vocative. The Ika ergative case marker *-se* additionally marks locative, Yanam *-n* also marks instrumental. Northern Embera *-(p)a* additionally marks ablative, Puinave *-at* indicates ablative, and a more general oblique, and in Timbira, ergative *tɛ* also marks genitive. The connection between ergative case markers and

Map 5. Presence of ergative case

Languages with ergative case

Ika, Embera, Northern, Timbira, Matses, Shipibo-Konibo, Kakataibo, Yaminahua, Ese Ejja, Cavineña, Puinave, Trumai, Yanam

other case functions is frequent cross-linguistically (Heine & Kuteva 2002: 180, Heine 2008: 467) and may suggest that the ergative in these languages may have arisen through reinterpretation of other case roles in e.g. nominalized or passive constructions (but note the unusual situation in some Panoan languages where the ergative also marks both instrumental and genitive[5]).

Languages in the sample with an accusative marker are Cubeo, Awa Pit, Hup, Desano, Imbabura Quechua, Dâw, Tsafiki, Miraña, Aguaruna, Huallaga Quechua, Rikbaktsa, Kwaza, Sabanê, Cuzco Quechua, and Bororo. A number of these languages have a conditionally appearing accusative marker, for instance in the Tucanoan languages of the sample, an object marker is used only for animate objects (Cubeo) or specific objects (Desano). Similar or stronger constraints

5. I thank Roberto Zariquiey for bringing this to my attention.

Map 6. Presence of accusative case

Languages with accusative case

Awa Pit, Tsafiki, Miraña, Aguaruna, Rikbaktsa, Bororo, Hup, Dâw, Sabanê, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Cubeo, Desano, Kwaza

exist in Kwazá, Hup, Awa Pit, Miraña, Aguaruna, Rikbaktsa,[6] and Sabanê.[7] This means that accusative case markers that conform to the inflectional prototype of appearing automatically, without conditioning (Corbett 2006) are almost exclusively found in the Andes.

Maps 7 and 8 show the extension of the ergative and accusative case markers (respectively) to subordinate clauses. Disregarded here are headless relative clauses that fulfill a referential function and receive the appropriate case marker according to the role the relativized argument plays in the main clause, exemplified in (3), since they do not constitute cases where the functionality of the case marker is carried over to event-denoting bases.

6. The accusative case marker *-tɨ:* in Rikbaktsa is constructionally conditioned, as it appears only in so-called "periphrastic structures" which contains an auxiliary that can only be marked for subject (Silva 2011: 112).

7. The exact function of the "object marker" *-k(a)* in Sabanê is unclear, further research may show that the marker should not be analyzed as an accusative marker.

Map 7. Ergative case in subordinate clauses

Languages with ergative case in subordinate clauses

Timbira, Ese Ejja, Cavineña, Trumai

(3) Yaminahua [panoan] Faust & Loos 2002: 147
[mani pei-pefe-a]$_{NMLZ}$ ***-tõ*** *mexteteke pi-i*
banana leaf-carry.on.shoulders-ptc-**erg** twigs eat-prog
'The lizard with wings (lit.: he who carries banana leaves on his shoulder) eats twigs.'

Instances of NPs containing a relative clause, which are marked by a phrase-final clitic are not taken into account either. This situation type is examplified in (4) from Trumai.

(4) Trumai [isolate] Guirardello 1999: 412
[di nïchïts ka'chï pata-t' ke]$_{NMLZ}$***=k*** *mi'ïrau kiṭï hai-tl*
woman now walk arrive-nlz rel=**erg** necklace give 1sg-dat
'The woman who just arrived gave me a necklace.'

As can be seen in Map 7, most ergative cases do not function as subordinators. There are a few potential cases of extensions, to be discussed below. That ergative case markers do not extend to mark subordinate clauses is not unexpected given

Map 8. Accusative case in subordinate clauses

Languages with accusative case in subordinate clauses

Aguaruna, Bororo, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Cubeo (potential)

the cross-linguistic dispreference for clausal transitive subjects in transitive clauses (Hopper & Thompson 1980). The fact that the use of the accusative case to mark subordinate clauses is predominantly found in the Andes is not unexpected either, given the conditional appearance of the accusative marker in many Amazonian languages, where conditions are often related to animacy.

The potential extensions of the ergative marker are found in the Tacanan languages Cavineña and Ese Ejja, in the isolate language Trumai, and in Timbira [MACRO-GÊ]. Cavineña presents the clearest case: the ergative marker = *ra* is also found on purpose of motion and (though a minor pattern) reason clauses (5).

(5) Cavineña [TACANAN] Guillaume 2008: 715, 719

a. *tudya i-ke kwinana-wa wira=**ra***
then 1SG-FM emerge-PERF urinate=**PRP.MOT**
'I went outside to urinate.'

b. *e-tsaka uje-da ju-ya aje-**ra***
NPREF-legs painful-ASUF be-IMPFV walk=**REAS**
'My legs hurt from walking.'

In Trumai, reason clauses are marked with the marker *-ak*, which is close to the ergative marker *-(a)k/-ek*. The reason marker *-ak* and the ergative marker *-Vk* seem to be historically related, especially considering conspicuous though somewhat opaque patterns of allomorphy: the ergative marker has allomorph *-ts* used for the first person, the reason clause marker has allomorph *iets'*. The marker *-ka* is used in combination with a third person absolutive clitic on the verb. Although the diachronic specifics of the relation between the ergative marker and the reason marker (including the pattern of allomorphy) remain somewhat enigmatic, Guirardello (1999: 406) hypothesizes that the use of the reason marker may have been an extension of the function of the ergative marker to encode the causer in a causative construction (although synchronically the markers have to be regarded as different). Trumai purpose clauses are marked with *(a)hak*, possibly also related to the ergative marker (though with less confidence). In Timbira there is a possible extension of the ergative/genitive marker to reason clauses, discussed below in this section.

In Ese Ejja the ergative case marker possibly forms part of a number of subordinators. Vuillermet (2012: 599–600) argues that some of the switch-reference markers of the language have been partially formed on the basis of an absolutive (ø) versus ergative (*=a*) opposition where the latter marks coreference between the intransitive subject of the dependent clause and the transitive subject of the main clause, and the former between the absolutive argument of the dependent clause and the intransitive subject of the main clause (Table 3).[8] These absolutive-ergative oppositions are suggested to derive from attributively used participial constructions displaying case agreement with their head noun (ibid.).

Table 3. Correspondences of absolutive and ergative cases in ese Ejja subordinators (Vuillermet 2012)

	$S/P_{DEP} = P_{MAIN}$	$S_{DEP} = A_{MAIN}$
before	-ximawa = ø	-ximawa = a
condition	=ø = xemo	=a = xemo
reason	=ø = xejojo	=a = xejojo

Similar patterns may be found in the Panoan languages, where the /n/-like form of the ergative may appear in parts of the switch-reference paradigm (see Valenzuela 2003), but the distribution is less clearly governed by the role of the pivotal participant in either of the connected clauses.

8. The zero marking is of course problematic (and therefore not considered in the database) but the functional correspondence does seem to speak for the ergative case marker as part of the S = A markers.

Summarizing, there is little unequivocal evidence for grammaticalization paths from ergative to subordinator in the sample, but for the languages that potentially show such a grammaticalization there seems to be at least a connection between ergative and reason clauses, and possibly purpose clauses. There are too few data points to be able to say anything definite about whether these grammaticalizations are driven by genealogical, areal, or general pressures, but it is probably not a coincidence that Tacanan and Panoan languages show evidence of this path, and it is likely that the switch-reference systems and their interaction with transitivity in these languages have played a major role in the extensions of the ergative markers.

Moving on to the accusative markers, Quechuan languages show a straightforward connection of the accusative case to marking nominalized clausal complements:

(6) Cuzco Quechua [QUECHUAN] Lefebvre & Muysken 1988: 18
[papa mikhu-y]$_{NMLZ}$***-ta*** *muna-n*
potato eat-INF-**ACC** want-3
'He wants to eat potatoes.'

A similar construction exists in Aguaruna.

(7) Aguaruna [JIVAROAN] Overall 2007: 428
naŋkama-a-u-ai *[anɨntaĩ tsupi-hu-ta]*$_{NMLZ}$***-na***
begin-HIAF-REL-COP: 3:DECL heart cut-APPL-ACT.NLZ-**ACC**
'He began to cut (the boa's) heart.'

In addition, there are relativized complements:

(8) Aguaruna [JIVAROAN] Overall 2007: 534
dɨka-a-ma-ha-i *[amɨ wai-tu-ka-mau]*$_{NMLZ}$***-na****-ka*
know-IMPFV-REFL-1SG-DECL [2SG see-1SG.OBJ-INTS-NONA/S:REL]-**ACC**-FOC
'I know you (who) saw me.'

Bororo can also mark its complements with the object marker *-ji*:

(9) Bororo [BORORoAN] Nonato 2008: 147
a-jorödü-re *[boe* *e-wogu-re* *dü]*$_{NMLZ}$***-ji***
2SG-see-ASSERT Bororo 3PL-fish-ASSERT COMP-**ACC**
'You saw that the Bororos fished.'

The constructions in these languages are comparable: they show an extension of the accusative marker to marking clausal complements. In all of the cases this extension is facilitated by the presence of some kind of subordinator or nominalizer, which makes these uses of the accusative marker comparable to the nominalized

relative clauses mentioned above, which were not taken into consideration. The marker *-re* in Cubeo [TUCANOAN] is possibly found as an element in a number of different-subject forms of the switch-reference system, e.g. for 'when' clauses (*-e-re*), simultaneous (*-e-reka*), where the initial *-e* is a nominalizer (Morse & Maxwell 1999: 161–9). However, given that the case marker *-re* has functional extensions into the spatial realm (locative, ablative) it is unclear which functional connection is responsible for the extension to adverbial clauses. Kamaiurá [TUPÍ] deserves a special mention because, in the analysis of Seki (2000), the language exhibits a 'nuclear case marker' *-a,* which does not distinguish between different types of core case, but instead marks an argument as belonging to the core. This marker can also be used to mark clausal complements:

(10) Kamaiura [TUPÍ] Seki 2000: 171
1-potar=ete *i-jo-taw-**a***
1SG-want=really 3-go-NLZ-**NUC**
'I want him to go.'

Summarizing, in a number of languages the accusative markers extend to marking clausal complements, though in combination with some type of subordinator or nominalizer.

The third structural case marker considered here is the genitive. Maps 9 and 10 show the presence of the genitive case marker and its distribution as a subordinator, respectively.

As can be seen on Map 9, the genitive is particularly widespread in western South America, though by no means exclusively in the Andes. The genitive is a feature of several families, like e.g. Quechuan, Aymaran, Panoan, Tacanan, Barabacoan. Map 10 shows that in only very few cases the genitive extends to mark subordinate clauses, and as we will see, most of these cases are somewhat problematic.

In Mosetén [MOSETENAN], possessive constructions mark the possessor with one of the so-called linker suffixes (*-tyi'* 'masculine' or *-si'* 'feminine') depending on the gender of the head noun. These markers can also be used to form relative clauses and simultaneity clauses (in the form of converbs). However, in its nominal use, the linker suffixes have a broader extension: they function as relation markers within the noun phrase, also for instance between adjectives and nouns.

(11) Mosetén [MOSETENAN] Sakel 2004: 106
*jaem'-**tyi'*** *mintyi'*
good-**LK.M** man
'a good man'.

Map 9. Presence of genitive case

Languages with genitive case

Yanesha', Aymara, Awa Pit, Tsafiki, Ika, Tehuelche, Timbira, Dâw, Mosetén, Matses, Shipibo-Konibo, Kakataibo, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Ese Ejja, Cavineña, Warao, Trumai, Kanoê

In Imbabura Quechua, the benefactive marker *-paj* codes genitives as well. This marker is found as a constitutive element of the same-subject purposive marker *-ngapaj* (*-nga* is third person future). Since other Quechuan languages have extended the cognate benefactive *-paj* (this marker is different from the genitive marker in most Quechuan languages) to purpose clauses, it seems likely that, in Imbabura Quechua, it was the benefactive function that allowed for the extension as well.

In Awa Pit, genitive is marked by *-pa*, which has an allomorph *-wa* after a vowel. Sequential clauses are marked with *-tpa* (after V-final stems) or *-tawa* (after C-final stems). This marker seems to contain the postposition that is used for genitive (Curnow 1997: 271). However, Awa Pit has a homophonous locative/allative postpostion *-pa*/*-wa* which is probably a more likely source for the extension to sequential clauses (see Section 3.5).

Map 10. Genitive case in subordinate clauses

Languages with genitive case in subordinate clauses (potential)

Awa Pit, Timbira, Mosetén, Imbabura Quechua

In Timbira, finally, the marker *-te*[9] marks reason clauses. In Alves (2004), this marker also encodes ergative and genitive in NPs. However, in Popjes & Popjes' (1986) analysis, this marker encodes experiencers of habitual states, and subjects of transitive clauses in paste-tense clauses.

(12) Timbira [MACRO-GÊ] Popjes & Popjes 1986: 165
jaco me capi te pĩ here jakep ame to [ajpẽn cahhyr
Jaco and Capi ERG/PAST wood twig cut 3PL INST REC beat
pram]$_{NMLZ}$ ***te***
want **REAS**
'Jaco and Capi cut twigs because they wanted to beat each other with them.'

In summary, there seems to be no uncontroversial evidence for an extension of genitive markers to marking subordinate clauses in the languages of the sample.

9. I follow the spelling in Popjes & Popjes (1986) here. In Alves (2004) this marker is spelled *-tɛ*.

3.3 Dative cases

The category of dative case is broadly conceived of here as those case markers that encode beneficiaries, recipients, maleficiaries, etc. unless they also mark locative, instrumental, or spatial relations, in which case they have been classified as 'obliques' (see below). Maps 11 and 12 show the distributions of dative cases in general (Map 11) and as subordinators (Map 12).

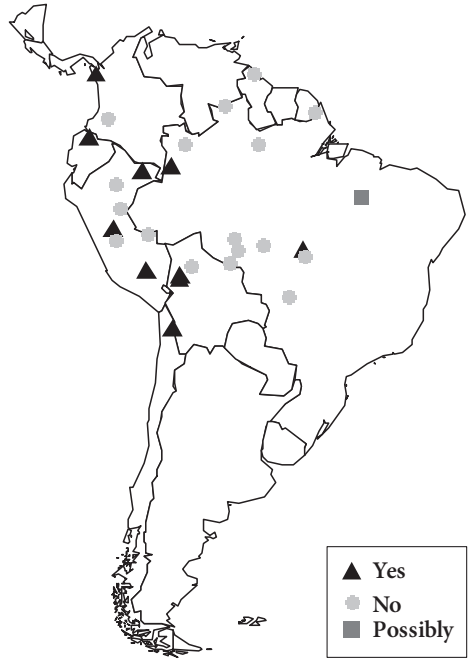
Map 11. Presence of dative case

Languages with dative case

Yanesha', Aymara, Miraña, Hixkaryana, Embera, Northern, Timbira, Rikbaktsa, Bororo, Dâw, Mosetén, Nasa Yuwe, Shipibo-Konibo, Yaminahua, Yagua, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Cavineña, Karo, Mekens, Emérillon, Kamaiurá, Warao, Urarina, Kwaza, Trumai, Leko, Yanam

The extension of a dative marker seems to be common in the Andean languages. It is a pattern found throughout the Quechuan family, where the benefactive marker *-paq* (or regional variants thereof) can also be used to mark purpose clauses. The extension of the dative/genitive marker in Imbabura Quechua to purpose clauses was mentioned above. Other Quechuan varieties generally distinguish between the genitive and benefactive, with only the latter expanding to purpose clauses, strengthening the argument for a grammaticalization path benefactive →

Map 12. Dative case in subordinate clauses



Languages with dative case in subordinate clauses

Timbira (possibly), Aymara, Miraña, Embera, Northern, Mosetén, Yagua, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Trumai, Leko

purposive for Imbabura Quechua as well. Example (13) shows the use of the marker *-paq*, in combination with the irrealis nominalizer *-na*:

(13) Huallaga Quechua [QUECHUAN] Weber 1989: 206
*wañu-chi-ma:-na-n-**paq** parla-ku-sha*
die-CAUS-1OBJ-NLZ-3POSS-**BEN** converse-REFL-3PERF
'They agreed to kill me.'

This construction is possibly related to the Central Aymaran construction. As is well known, Quechuan and Aymaran languages have had a long history of contact and structural convergence (see e.g. Adelaar & Muysken 2004).

(14) Central Aymara [AYMARAN] Hardman 2001: 213
*ch'uq pall-ja-fiani [p"iry apa]$_{NMLZ}$ **-ñataki***
potato sort-PART-4>3FUT fair carry-**OBLG.PRP**
'We'll sort potatoes to take to market.'

Further potential areal extensions are found in the foothill languages Mosetén (15) and Leko (16):

(15) Mosetén [MOSETENAN] Sakel 2004: 438
*khäkï katyi'-in jäe'mä dyam~dyam jedye'-in jäe'mä saeks-i-**dye**-si'*
because HSAY-PL filler little~RED thing[F]-PL FILLER eat-VSM-**BEN**-LK.F
'Because there are only a few things to eat.'

(16) Leko [ISOLATE] van de Kerke 2009: 324
*dira hoyno-tha hal-ate uywas-ich-**moki***
four pig-DIM buy-PAST.1 raise-INF-**BEN**
'I bought four pigs to raise.'

Further north, Northern Embera likewise marks purpose clauses with a marker that can also be used to encode beneficiaries.

(17) Northern Embera [CHOCOAN] Mortensen 1999: 121
k^{h}uriwa ete-de wã-tua mɯ̃ wárra-rã mɯ̃-a [hu tawa-i]$_{NMLZ}$
Guatín get-LOC go-IMPFV 1SG offspring-PL 1SG-ABL breast give.drink-IRR
k^{h}ãrẽã
BEN
'Guatín, go get my children so that I may nurse them!'

The path beneficiary to purpose is clearly the most common for dative-like markers, and particularly associated with the Andes and adjacent areas. Nevertheless, a few alternative grammaticalization paths are shown by other languages. The beneficiary marker *-llii* in Miraña can mark reason clauses, and the dative marker *-(i)va* in Yagua can additionally mark 'until'-clauses. Trumai dative markers extend to certain complement relations (complements of verba dicendi, verbs of liking and perception verbs), and there is a possible diachronic connection between dative and temporal conditional clauses (Guirardello 1999: 192–3). In Timbira, the dative/benefactive marker *-mã* is homonymous with the topic marker *mã* (except that the latter is not bound), which in turn seems to have developed into a third-person different-subject marker in semantically versatile complex sentences (Alves 2004: 146). Although the grammaticalization chain beneficiary > topic > different subject marker is speculative, it is an imaginable development, perhaps via a cleft construction.

Table 4 summarizes the uses of dative/benefactive case markers in subordinate clauses in the languages of the sample.

Table 4. Extensions of dative-like case markers to subordinate clauses

Language	Case marker	Extension
Timbira (P)	*mã* BEN/DAT	Perception complements (DS)
Embera, Northern	*k^{h}ãrẽã* BEN	Purpose
Imbabura Quechua	*-paj* BEN	Purpose (*-ngapaj*)
Miraña	*-llii* BEN	Reason
Yagua	*(i)va* DAT	Until
Huallaga Quechua	*-paq* BEN	Purpose
Trumai	*-ki* DAT	Complementation (perception, fear, liking, communication)
	-(a)s/(i)s DAT	Temporal/conditional
Cuzco Quechua	*-paq* BEN	Purpose, conditional
Mosetén	*-dye* BEN	Purpose
Leko	*-moki* BEN	Purpose
Aymara	*-taki* DAT	Purpose

3.4 Comitatives and instruments

As shown in Map 13, comitative and instrumental case markers are extremely common in South American languages. Quite often, there is a single marker for both functions, confirming the functional connection between the semantics of these case roles. It is not very common, however, for an instrumental or comitative marker to grammaticalize into a subordinator (Map 14). Only a handful of examples in the sample languages show this connection, without there being any obvious genealogical or areal pattern. In terms of recurring grammaticalization paths, the most common pattern is for the comitative or instrumental case marker to encode simultaneity/manner or overlap (when) clauses. This link is found in Bororo, Desano,[10] Kwazá, Mosetén, Trumai, and possibly in Kakataibo, Yaminahua and Kokama.[11] The use of the comitative marker in Bororo simultaneous clauses (pos-

10. The use of the comitative marker *bẽrã* to mark temporal clauses is constructionally rather restricted, as it is used only for constructions with a time word and a nominalized form of the verb *wa* 'to go' (Miller 1999: 151).

11. Some of the subordinate markers in these languages show elements that may be linked to the comitative marker, like instrument marker = *pu(pe)* in Kokama which is possibly present in the subordinators *-npu* (sequential, *-n* is a nominalizer) and = *puka* ('when', =*ka* is a locative), see Vallejos (2010). In Yaminahua one of the comitative markers, *-ya*, as well as the instrument marker *-ña* may be part of the subordinator *yanã* 'when'. Kakataibo has a subordinator *-këbëtan* (different subject simultaneous) which probably contains the A participant comitative *-bëtan*.

sibly with a causal reading) and the instrument marker in Kwazá 'when' clauses are given in (18) and (19), respectively.

Map 13. Presence of comitative/instrument case

Languages with com/ins case

Tariana, Apurinã, Awa Pit, Miraña, Trio/Tiriyó, Hixkaryana, Ika, Tehuelche, Aguaruna, Timbira, Rikbaktsa, Bororo, Dâw, Mosetén, Nasa Yuwe, Matses, Shipibo-Konibo, Kakataibo, Yaminahua, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Ese Ejja, Cavineña, Cubeo, Desano, Karo, Mekens, Emérillon, Kokama, Kamaiurá, Warao, Puinave, Urarina, Kwaza, Trumai, Leko, Yurakaré, Yanam

Map 14. Comitative/instrument case in subordinate clauses

Languages with com/ins case in subordinate clauses

Emerillon, Hixkaryana, Desano, Trumai, Mosetén, Bororo, Kakataibo (possibly), Kokama (possibly), Yaminahua (possibly), Kwazá

(18) Bororo [bororoan] Nonato 2008: 79
e-jagare-re *[tu-okwage-i ji-dü]*$_{NMLZ}$ ***apo***
3pl-be.happy-assert 3a-eat-inf thm-comp **com**
'They were happy eating it [the corn].'

(19) Kwazá [isolate] Van der Voort 2004: 508
[hako'ri duky-'tõi ũi-e-'nã-tsy-wy]$_{NMLZ}$ ***-ko***
moon other-cl:eye lie-again-fut-ger-time-**inst**
'in the next month when it will be new moon'.

Another connection, observed in Emerillon and Hixkaryana, is between instrument and reason.

(20) Emerillon [TUPÍ-GUARANÍ] Rose 2011: 335)
[aman-a-r-aʔɨr-a-te o-ʔar-a-r]$_{NMLZ}$***-ehe*** *eaɲ o-kakuwa.*
rain-REF-REL-SON-REF-FOC 3C-fall-REF-RELN-**REAS** quickly 3C-grow
'Because it is the son of the rain who is born, he grows very quickly.'

Table 5 summarizes the different extensions of the comitative and/or instrumental markers in the sample languages.

Table 5. Extensions of comitative and instrumental case markers to interclausal relation markers

Language	Case marker(s)	Extensions	Comments
Emerillon	*-ehe* (comitative)	reason	
Hixkaryana	*ke* (instrument)	reason	
Desano	*bẽrã* (comitative)	when	Constructionally limited.
Trumai	*tam* (comitative)	simultaneity	
Mosetén	*tom* (comitative)	simultaneity	
Bororo	*apo* (comitative)	simultaneity	Possibly extensions to reason.
Kokama (Potential)	*=pu(pe)* (instrument)	when, succession	Extensions are not entirely certain.
Yaminahua (Potential)	*-ya* (comitative), *-ña* instrument	when	Extensions are not entirely certain. Status of *-ya* as a comitative case marker is moreover doubtful (R. Zariquiey, p.c.).
Kakataibo (Potential)	*-bëtan* (comitative A)	-kebëtan	
Kwazá	*-ko* (instrument)	when	

3.5 Spatial cases

Spatial case markers are clearly the most common type of case markers in South American languages. Moreover, they are the case markers that most frequently extend to mark subordinate verbs and clauses for their relation to some superordinate clause. Both facts are visualized in Maps 15 and 16, respectively.

Map 15. Presence of spatial case

Languages with spatial case

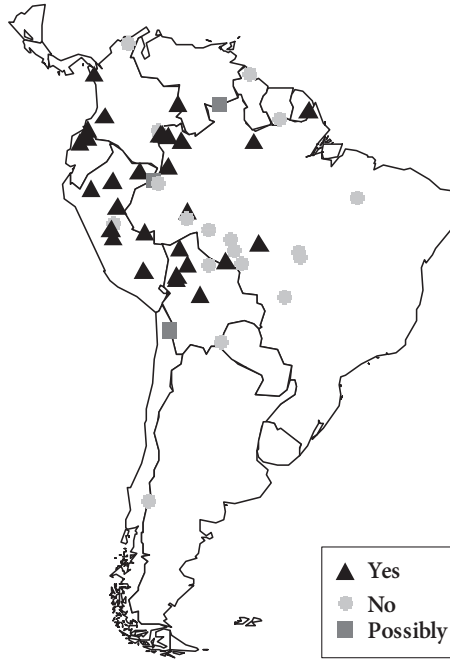
Tariana, Apurinã, Yanesha', Baure, Jarawara, Aymara, Awa Pit, Tsafiki, Miraña, Trio/Tiriyó, Hixkaryana, Ika, Embera, Northern, Aguaruna, Timbira, Rikbaktsa, Bororo, Hup, Dâw, Mapudungun, Mosetén, Sabanê, Nasa Yuwe, Matses, Shipibo-Konibo, Kakataibo, Yaminahua, Yagua, Imbabura Quechua, Huallaga Quechua, Cuzco Quechua, Ese Ejja, Cavineña, Cubeo, Desano, Karitiana, Karo, Mekens, Emérillon, Kokama, Kamaiurá, Tapiete, Warao, Puinave, Urarina, Kwaza, Trumai, Leko, Yurakaré, Yanam

I have not counted location clauses because they do not genuinely present semantic extensions, although they may be indicative of a rather flexible distinction between nouns and verbs, especially if no nominalization is required before the case marker can be applied, as in (21).

(21) Baure [ARAWAK] Danielsen 2007: 407
bueno tiow [noiy ver eto-pi-a-po]$_{NMLZ}$ -yi=ro
bueno CLEFT there PERF finish-words-LK-PFV.REFL-LOC = 3SGM
'Well, this is where already the words were finished.'

Table 6 sketches the different extensions of the spatial case markers to contexts of clause combinations.

Map 16. Spatial case in subordinate clauses



Languages with spatial case in subordinate clauses

Embera, Yanam (potentially), Puinave, Emerillon, Paez, Awa Pit, Hup, Desano, Tariana, Imbabura Quechua, Dâw, Hixkaryana, Tsafiki, Miraña, Yagua, Kokama (potentially), Urarina, Aguaruna, Shipibo (potentially), Huallaga, Yaminahua, Yanesha', Jarawara, Rikbaktsa, Ese Ejja, Mekens, Cavineña, Cuzco Q, Moseten, Leko, Yurakaré, Aymara (potentially)

Table 6. Extensions of spatial case markers to interclausal relation markers

Language	Case marker(s)	Extensions	Comments
Embera	*-de* LOC	simultaneity, when; comparison (purpose of motion)	
Yanam (potential)	*-ha* various spatial	temporal (*-pɪha*)	
Puinave	*-a* DIR, and *-'u* ADH	temporal, conditional, reason, complements	
Emerillon	*-upi* PERL	comparative (simultaneity)	

(*continued*)

Table 6. (*continued*)

Language	Case marker(s)	Extensions	Comments
	koti DIR	cognition	
Paez	*-te* LOC	temporal (DS) > concessive	
	-xũ ABL	succession	
	-na DIR	comparative	
Awa Pit	*-kima* LIM	until (temp)	
	-pa/-wa LOC/DIR/GEN	succession	
Hup	*-an* various spatial	simultaneity	Many relator-like spatial elements can have temporal interpretations.
Desano	*kore* before	precedence	
	pi?ri after	succession	Unclear if the postposition has a spatial meaning as well.
Tariana	*-se* various spatial	sequence simultaneity	
Imbabura Quechua	*-manda* ABL	reason	
	-kaman LIM	until	*-ngakaman* is used.
Dâw	*xáx* 'among'	simultaneity	
Hixkaryana	*hona* DIR	purpose	
	way 'to, by'	manner	
Tsafiki	*=bi* LOC/DIR	Purpose of motion; temporal	SR clauses can take locative postpositions without a clear interpretational difference. Perhaps further locatives are possible in these constructions.
	=le LOC	temporal	
Miraña	*-tu* ABL	succession, comparative	
Yagua	*-jụ̀* DIR	purpose	
Kokama (potential)	*=ka* LOC/DIR	when (=puka)	
	=kuara INESS	reason (=ikua)	
Urarina	*hana* INESS	when	
	ahinia before	before (temp)	

Table 6. (*continued*)

Language	Case marker(s)	Extensions	Comments
Aguaruna	*-nī* LOC	simultaneity, condition, concession	DS clauses for second person are marked with a morpheme that is cognate with the locative case marker.
Shipibo (potential)	*-ain* DIR	simultaneity	The case marker seems to form part of one of the DS markers, but alternative connections to the ergative *-n* are also possible (R. Zariquiey, p.c.).
Huallaga	*-kama/-yaq* LIM	until	
	-pita ABL	reason, succession, comparative, neg. Purpose,	
	-chaw LOC	simultaneity	
Yaminahua	*-ax* ABL	succession	
Yanesha'	*-ot* LOC	reason	Possibly wider functionality.
	-o'mar LOC	reason	
Jarawara	*kaa* PERL	reason	Status as case marker not entirely certain.
Rikbaktsa	*-ere(ka)* INESS	temporal, conditional	
Ese Ejja	*=jo* LOC	reason, condition, precedence, when	Both are part of the SR paradigm.
	=xe PERL	reason, condition, precedence, when	
Mekens	*(e)se* LOC	temporal, conditional	
Cavineña	*=ju* LOC	temporal	The marker = *ju* is also DS marker.
Cuzco Q	*-manta* ABL	reason	
	-kama LIM	until	
Moseten	*-ya'* ADESS	when, conditional	
Leko	*-ra* LOC	temporal, conditional, reason	
Yurakaré	*=jsha* ABL	concession, succession	
Aymara (potential)	*-ta* ABL	temporal	

By far the most common extension is from spatial to temporal, which is not surprising given the close connectedness between space and time in human languages (see e.g. Haspelmath 1997). Stative locatives tend to extend towards simultaneity or when clauses (22), whereas ablative markers naturally extend to succession clauses (23).

(22) Embera [CHOCOAN] Mortensen 1999: 114
[mãwã b-ɪɪ]$_{NMLZ}$ ***-de*** *s'e-s^{h}i-a ũmãk^{h}ĩrã s'a b-ɪɪ-ta*
like.this be-PRES-**LOC** come-PAST-DECL man here be-PRES-ABS.FOC
'(…) while this was happening, there came a man who looked just like him.'

(23) Páez [ISOLATE] Jung 2008: 171
[lu:ts^{j}-k wewe?we-n^{j}i]$_{NMLZ}$ ***-xũ*** *jat-te ka:-pija-?x-ja?*
child-DIM speak-PERF.PTC-**ABL** house-LOC CAUS-learn-TR-INF
tak^{h}-e-?-t^{j}
begin-IPFV-HAB-ASSERT.3PL
'When the child knows how to speak, they start to teach it in the house.'

Since for many languages temporal and conditional clauses are marked in similar ways, the case markers that encode temporal relations can also code conditional relations in those languages.

(24) Leko [ISOLATE] Van de Kerke 2009: 316
[iya kelecha he-ir-a]$_{NMLZ}$ ***-ra*** *lamkas-tan*
you money have-NEG-PFV-**LOC** work-OBLG
'If you don't have money, you should work.'

Another common type of extension of spatial markers is towards reason clauses.

(25) Huallaga [QUECHUAN] Weber 1989: 195
[qella ka-y]$_{NMLZ}$ ***-pita*** *[osyoosu ka-y]*$_{NMLZ}$ ***-pita*** *chay-lla-ta miku-n*
lazy be-INF-**ABL** lazy be-INF-**ABL** that-just-ACC eat-3
'Because they are lazy, they just eat that.'

And those languages with a limitative marker usually extend it to mark temporal relations as well.

(26) Awa Pit [BARBACOAN] Curnow 1997: 263
[Demetrio kayl-na]$_{NMLZ}$***=kima*** *kal ki-nɨ-s*
Demetrio return-INF=**until** work(1) work(2)-FUT-LOCUT
'I will work until Demetrio returns.'

Summarizing, there is a very strong connection between spatial case and temporal interclausal relation when it comes to the use of case markers. In some languages,

the use of spatial case markers extends further, to conditionals. Other common extensions are reason and purpose.

3.6 Oblique cases

The oblique cases form a somewhat disparate group, and therefore the patterns yielded by this group of case markers is expected to show effects that are correspondingly diverse. Maps 16 and 17 show the distribution of oblique case markers and their extensions to subordinate clauses, respectively. Table 77 summarizes the information per language.

Map 17. Presence of oblique case

Languages with oblique case

Apurinã, Yanesha', Jarawara, Tsafiki, Miraña, Mocoví, Timbira, Hup, Mapudungun, Kakataibo, Yagua, Cubeo, Karitiana, Kokama, Kamaiurá, Puinave, Kanoê, Movima, Yurakaré

Map 18. Oblique case in subordinate clauses

Languages with oblique case in subordinate clauses

Cubeo (P), Hup, Miraña, Yagua, Timbira (P), Jarawara, Apurinã, Karitiana, Kanoe, Movima, Yurakare, Mapudungun

As can be seen in Table 8, most extensions of oblique markers are towards temporal clauses. This is further evidence of the close connection between location and time, since all oblique markers in the sample can have spatial interpretations.

In some cases, the multi-functionality of the oblique marker translates directly into multi-functionality as a subordinator. This is for instance the case in Jarawara (27) and Movima (28):

(27) Jarawara [ARAWAN] Dixon 2004: 489, 496

a. *[awa ini tati jaa bahi ite]*$_{NMLZ}$ ***jaa*** *otaa ka-ma*
tree branch head OBL sun sit **OBL** 1EXC.S in.motion-back
'When the sun sits on the topmost branches of the trees, we go back.'

b. *[ee hijari]*$_{NMLZ}$ ***jaa*** *ee hijara na-ba ee-ke*
1INC.S talk.NLZ **OBL** 1INC.S talk AUX-FUT 1INC-DECL
'If we want to talk , we should talk now.'

c. *[jobe wata-re]*$_{NMLZ}$ ***jaa*** *hinaka jobe-bona otaa hiri ne*
house exist-NEG **OBL** 3SG.POSS house-INTN 1EXC.A make AUX
'Since there wasn't a house (for Alan), we made a house for him.'

Table 7. Extensions of oblique case markers to interclausal relation markers

Name	Marker	Extension	Comments
Cubeo (potential)	*-re* DAT; LOC	temporal	Classification as case marker problematic.
Hup	*-Vt* LOC; INST; COM	simultaneity	
Miraña	*-ri* INST; LOC	succession; reason	
	-ma COM, INST, BEN	succession	
Yagua	*(i)ma* INST; LOC	simultaneity	
Timbira (potential)	*kãm* LOC; COM	general subordinator	
Jarawara	*jaa* OBL	temporal, conditional, reason, location	
Apurinã	*-ã* LOC; INST	conditional	
Karitiana	*-ty* OBL	desire, cognition, perception	
Kanoe	*-ni* OBL	temporal	Perhaps marginally.
Movima	*n-* OBL	temporal, purpose	
Yurakare	*=la* PERL, INST	reason, cognition	
Mapudungun	*-mew* OBL	reason, location	Glossed as an instrument marker, has a wide range of interpretations.

(28) Movima [ISOLATE] Haude 2006: 310

a. *jayna* ***n****-os* *[ena'* *łani-wa]*$_{NMLZ}$ *ń* *to'baycho*
DSC **OBL**-ART.NEUT.PAST DUR.STD bathe-NLZ 1INTR remember.MST
'Then, as I was bathing, I remembered.'

b. *ji<wa:->wa--'ne* ***n****-os* *sa-al-wa='ne* *us*
come<MID->--F **OBL**-ART.NEUT.PAST DR-search-NLZ=F ART.M
pa:pa='ne
father.of=F
'She came to look for her father.'

Perhaps somewhat counterintuitive, in some languages, oblique markers are used to mark complement clauses.

(29) Karitiana [TUPÍ-ARIKEM] Storto 2011: 229

a. *y-py-sondyp-yn* *yn [Inacio 'ep* *opiĩ]*$_{NMLZ}$ ***-ty***
1-ASSERT-know-NONFUT I Inácio tree cut-**OBL**
'I know that Inácio cut the tree´.

b. *y-py-so'oot-yn yn [Inacio 'ep opiĩ]$_{NMLZ}$ -**ty***
1-ASSERT-see-NONFUT I Inácio tree cut-**OBL**
'I saw that Inácio cut the tree'

(30) Yurakaré [ISOLATE] van Gijn 2006: 319
a. *nij wëshë-të-y [chitta mala-y-ti=**la** ti-petche]$_{NMLZ}$*
NEG notice-MID-1SG throw.SG go.SG-1SG=SUB=**OBL** 1SG-fish
'I did not remember that I left my fish.'
b. *ka-yle-ø-ya na ta-ppë tiri [ama-shku-ta*
3SG-know-3-REP DEM 1PL-grandfather Tiri WH-ADV-MID
*imbëtë-shta-ø-ti]$_{NMLZ}$=**la***
behave-FUT-3-SUB = **OBL**
'Our grandfather Tiri knew how he would behave.'

In summary, the extensions of the oblique markers seem to follow the extensions mentioned above in this paper, in particular towards temporal and reason clauses depending on the range of their semantics. Perhaps surprisingly, some languages use oblique markers to mark clausal complements.

3.7 Other cases

A number of other cases, which do not fall into any of the above categories, are also found in subordinate clauses. The patterns are too diverse and small to make sense in a genealogical or areal way, so I will not show any maps. Instead, Table 8 summarizes the relevant information.

Table 8. Extensions of other case markers to interclausal relation markers

Language	Case marker	Subordinator use
Kokama (potential)	*=ra* 'for the purpose of'	condition (-ra/-ri), purpose (-mira, -tara)
Warao	*ebe/kuare* 'because of'	reason
Páez	*-pa?ka* 'because of '	reason
Cubeo	*-pe* similative	comparative, purpose (P) *kijepe*
Hixkaryana	*horɨ* 'for the purpose of'	manner, purpose
Miraña	-d? similative	comparative
Urarina	*bana* 'at the time of'	when
	baja 'after'	sequential
	netohweĩ until	until
Shipibo-Konibo	*-tian* 'at the time of'	temporal (present in SR paradigms)
Jarawara	*tabijo* 'lack of'	reason

Table 8. (*continued*)

Language	Case marker	Subordinator use
Apurinã	*-sawaku* 'at the time of'	when
	-xika 'because of'	reason
Kakataibo	*=sa* similative	cognition
Huallaga Q	*-naw* similative	comparative, 'be about to'
	-rayku cause	reason
Cuzco Q	*-rayku* 'because of'	reason
Leko	*-bacha* 'because of'	reason

Many of the markers in this group have semantics that can be readily used either with respect to referential expressions or event-expressions. In fact, it is not clear to what extent these are extensions at all, and if they are, what their direction of diachronic development is: from nominal to verbal use or vice versa.

One type of situation is a similative nominal marker that can also be used as a similative verbal marker (the Quechuan languages, Cubeo, Miraña)

(31) Cubeo [TUCANOAN] Morse & Maxwell 1999: 101; 182

a. *ʧiai-ri-**pe*** *ãrõxã=abẽ* *ĩ*
cicada-CLS: 3D-**SIM** be.similar=N/H.3SG.M 3SG.M
'He (i.e. the ant-eater) is similar to the cicada.'

b. *pĩõ-jĩ-re* *xẽ-xa-kɨ* *'jo-pe* *xi*
blow-CLS:funnel-OBJ grab-IMP-M.SG this-SIM 1SG.POSS
*xẽ-ij-e-**pe***
grab-STV-INAN.PL.NLZ-**SIM**
'Hold the blowgun like I am holding it.'

Some languages have purposive case markers that can be used with nouns or verbs (Kokama, Hixkaryana).

(32) Hixkaryana [CARIBAN] Derbyshire 1985: 21; 39

a. *tono* *omsamtxemo tuna* ***horɨ***
she.went young.girl water **PRP**
'The young girl has gone for water.'

b. *kuraha* *wanɨmo* *ɨhoko* *ryesnirɨ* ***horɨ***
bow-wood I.picked.it.up occupied.with-it my.being **PRP**
'I picked up the bow-wood with a view to working on it.'

Similarly, temporal case markers that can be used either with nouns or verbs are found (Urarina, Shipibo, Apurinã). In Apurinã, the temporal marker *-sawaku* only attaches to nouns that express some kind of time concept (Facundes 2000: 388).

(33) Apurinã [ARAWAK] Facundes 2000: 388; 611
a. *õtu-**sawaku** n-apo-pe*
day-**TEMP** 1SG-arrive-PFV
'I arrived during the day'.
b. *a-makatxaka txa-ru komeru-pe u-payaka-**sawaku***
1PL-take AUX-3M.OBJ manioc-pulp 3M-be.soft-**TEMP**
'When it is soft, we take the manioc pulp.'

Finally, reason case markers are sometimes also found on verbs to mark reason clauses (Warao, Leko, Cuzco and Huallaga Quechua, Páez)

(34) Cuzco Q. [QUECHUAN] Lefebvre & Muysken 1988: 19; 23
a. *qan-ri ima-**rayku**-n mana saluda-wa-rqa-nki-chu*
you-EMPH what-**REAS**-AFF not greet-1OBJ-PAST-2-NEG
'You, why (because of what) did you not greet me?'
b. *[qaynunchaw pidru wiqchu-ku-sqa-n]$_{NMLZ}$-**rayku***
yesterday pedro slip-REFL-NLZ.REA-3-**REAS**
nana-chi-ku-sha-n
hurt-CAUS-REFL-PROG-3
'Because Pedro slipped yesterday he feels pain.'

4. Discussion

At the start of this paper, I indicated three major forces that may determine the distribution of any linguistic feature, in this case the use of particular case markers as subordinators: genealogical retention, areal contact-induced diffusion, and general communicative-cognitive principles. Starting with the latter, it was mentioned above that nominalization of subordinate clauses can be argued to follow functional principles in that an event-denoting unit is used in a syntactically atypical way, namely as an argument, modifier, or adjunct. In grammaticalization theory, the development from noun phrase to subordinate clause (e.g. Heine & Kuteva 2007) is regarded as a functionally motivated, common path. Therefore in a broad sense, general functional motivations play account for the patterns found in e.g. Van Gijn (2014) where nominalized subordinate clauses in South America are found to be very common.

We can approach the issue of 'naturalness' also from a semantic angle. Figure 1 gives a schematic representation of the associations between case markers and subordinators in the sample languages (the languages where the associations are uncertain have been left out), where the thickness of the line correlates with the frequency of the connections.

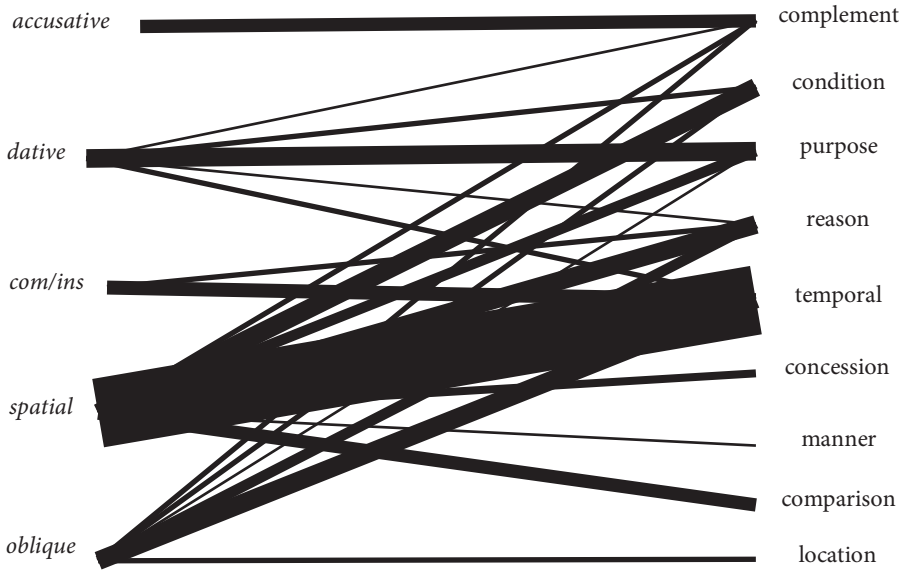


Figure 1. Association map of case markers and subordinators

If we look at Figure 1 from the perspective of the case marker, the following connections are most common:

(35)	Spatial	→	Temporal
	Dative/Benefactive	→	Purpose
	Oblique	→	Temporal
	Accusative	→	Complementation
	Comitative/Instrumental	→	Temporal

From the perspective of the subordinators, the following are the most frequent connections:

(36)	Complementation	←	Accusative
	Condition	←	Spatial
	Purpose	←	Dative/Benefactive
	Reason	←	Spatial
	Temporal	←	Spatial
	Concession	←	Spatial
	Manner	←	Spatial
	Comparison	←	Spatial
	Location	←	Oblique

Genetti (1986), discussing case-marker-based subordinators in the Bodic branch of the Tibeto-Burman family, develops a localist theory of grammaticalization of case markers into subordinators, based on work by Diehl (1975), who argues for

four general spaces: LOCATION, which locates physical objects in space, SOCIAL, which is location with respect to (human) beings, TEMPORAL, which locates events in time, and LOGICAL, which concerns the relations between propositions. On the basis of this abstract schema, Genetti (1986: 394) establishes the natural connections between cases and subordinators indicated in Table 9.[12]

Table 9. Natural extensions of locative case markers (Genetti 1986)

	LOCATION	SOURCE	GOAL
LOCATION	locative	ablative	allative
SOCIAL	comitative	ergative/instrument	dative
TEMPORAL	when/while	since, after	until
LOGICAL	if	because	purpose

In a general sense, the centrality of spatial cases in the grammaticalization path towards subordinators in South American languages is corroborated by the fact that, of all the cases, as can be seen in Figure 1, they are the prime source of subordinators across areas and language families, and have several different extensions, as shown in (36). Location → Subordinator is a path also observed more generally, especially towards temporal and reason clauses (e.g. Heine & Kuteva 2002: 205–6). In fact, as Heine and Kuteva (2002: 206) say: "It is hard to find languages where some expressions for locative concepts are not extended to also refer to temporal concepts."

In a more detailed way, the connections are more diffuse, but some naturalness effects can still be discerned. Table 10 shows a more refined classification of the data, giving the frequencies (in number of languages) of the occurrence of extensions of the case markers indicated in the top row to the subordinators indicates in the first column, following Genetti's proposal. The gray cells are the 'natural' connections.[13]

12. The terminology is slightly altered to better fit the set-up of this paper. Comitative is termed 'associative' in Genetti (1986) and ergative and instrument are taken together.

13. I have counted categories such as 'inessive' as locatives; limitative markers were grouped with allative, perlative markers are disregarded in Table 11. General temporal subordinators have been counted both as overlap and as sequence markers.

Table 10. Extensions in the sample and naturalness.

	LOC	ABL	ALL	COM	INS	DAT
Overlap	13	4	3	4	1	1
Condition	4					2
Sequence	8	7	4			1
Reason	2	2		1	1	1
Until			3			1
Purpose	1	1	3			7

The naturalness effects predicted by Genetti (1986) are certainly not all found in the data, though a few can be observed: a preference for locative markers to extend towards overlap markers, as well as between dative and purpose. Furthermore, ablative markers most often extend towards sequence subordinators, and comitative markers most often to overlap subordinators.

In summary, general principles do seem to play a role in the distribution of case-derived subordinators in South America, on at least three levels:

i. Nominalizations express the syntactically nominal status of subordinate clauses
ii. There is a clear space–time connection
iii. Certain specific 'natural' semantics extensions (locative-overlap, dative-purpose, and to a lesser extent ablative-sequence and comitative-overlap) can be observed.

A second potential influence on the distributional patterns of case-based subordinators is areality. Above I mentioned four major linguistic areas. I will briefly survey each of these areas with respect to case marking and extensions to subordinators.

The *Andean linguistic area* is generally associated with structural case and accusative alignment (Torero 2002; Adelaar 2008). Although object markers are certainly also found in the Amazonian macro area, they do seem to be subject to more restrictions than the accusative markers of the highlands. Extensions of accusative case markers to subordinate clauses is also predominantly restricted to Andean languages, though this seems to partly be a genealogical effect, as it is mainly restricted to Quechuan languages. The fact that Aguaruna (at the periphery of the Andean area) also has extensions to complement marking of the accusative marker may be the result of language contact. Another potentially areal phenomenon at the periphery of the Andean linguistic area is the extension of dative/benefactive markers to purpose clauses found in the foothill languages Mosetén and Leko.

The *Vaupés region*, in terms of case marking, is characterized by a nominative-accusative profile, differential object marking (with information structure

interfering), and a multi-purpose spatial marker (Aikhenvald 2002). It is unclear whether any of the extension patterns have areal motivations. Spatial case markers have been extended to temporal clauses in Tariana, Desano, Hup, and Dâw, but the extensions seem to be rather different, except for the extension of multi-purpose spatial markers to at least temporal overlap clauses in Hup and Tariana.

The *Guaporé-Mamoré linguistic area* is characterized by head-marking patterns, and at most peripheral case. Nominalized subordinate clauses are also mentioned as an areal trait (Crevels & van der Voort 2008: 171). In terms of case extensions, a number of languages on the Bolivian side of the area show extensions of oblique markers to subordinators, though with rather different semantics. There may also be a negative areal effect in the lack of extensions of spatial case markers to subordinators in the (north)eastern part of the Guaporé-Mamoré (and adjacent areas). As mentioned above, the western fringe of the area may show some contact effects with the Andean linguistic area in terms of the extension of the dative/benefactive.

The *Chaco* is mainly characterized by the absence of case (see also Comrie et al. 2010: 91).[14] This is in itself an interesting fact, since case markers seem to be generally present in South American indigenous languages. More particularly, Chaco languages seem to have a preference for expressing relations between an event and its arguments and/or adjuncts on the verb by means of applicatives and person markers, or – in the case of Tapiete – by person-marked adpositions.

Summarizing, there are some potential areal effects, both within and between linguistic areas, although areal accounts, in the absence of actual forms being borrowed, remains speculative.

A third potential factor in shaping distributions of case markers as subordinators is genealogy. Looking at the representatives of the larger families Arawak, Carib, Macro-Ge, Panoan, Quechuan, and Tupian may yield certain consistent genealogical trends, or further evidence for contact-induced change in the case of inconsistent patterns.

The members of the *Arawak* family in the sample are characterized by relatively small and semantically versatile case inventories. Although nominalization is a frequent subordination strategy for relative clause formation (Aikhenvald 1999: 100), complement and adverbial clauses are often formed by more verbal strategies. Nevertheless, some case extensions are found in the sample. Tariana and Yanesha' have extended their spatial case markers to temporal and reason clauses, respectively, and Apurinã uses its oblique case marker (locative/instrument) to mark temporal clauses. The latter language also has further markers (because of, at the time of) that can be used both with nouns and subordinate verbs.

14. Vilela (Lule-Vilela), not part of the sample of this paper, has peripheral case markers.

Nominalization is the dominant strategy in *Carib* subordination (Gildea 1998; Derbyshire 1999: 56–7; Gildea 2012: 481). Carib languages furthermore often have a host of postpositions, which can be inflected for person. These postpositions cannot always be counted as case markers according to the definition given at the beginning of this paper, but some of them can because they either do not take person inflection or because their inflection is in complementary distribution with an overt complement. Of the three Cariban sample languages, Hixkaryana is the only one that uses case markers as subordinators (instrument to reason, allative to purpose, as well as the marker *horɨ* 'for the purpose of' to purpose clauses). Some of the adpositions that cannot be counted as case markers seem to play a role in subordinate clause formation in all three Cariban sample languages, whether directly or historically (e.g. the locative adposition *tao* seems to be present in Tiriyó *ahtao* 'temporal/conditional' (Meira 1999), and Payne & Payne (2013: 421) mention several adpositions as potential sources for subordinating suffixes in Panare.

Not very much is known about general strategies that *Macro-Ge* languages use for subordinate clauses. Rodrigues (1999: 197) mentions the presence of switch-reference systems in some languages to mark coordinate clauses. The three Macro-Gean sample languages Bororo, Timbira, and Rikbaktsa show rather divergent patterns in their case marking systems as such (e.g. Timbira has an ergative case marker, Bororo and Rikbaktsa accusative), though the languages generally share the presence of a large inventory of peripheral case markers. Extensions to subordinate clauses are not abundant in the data that I have looked at, and are restricted to the extension of the accusative case marker to marking complements, as well as of the comitative marker to simultaneity clauses in Bororo, uncertain extensions of the ergative/genitive case marker to reason and the dative to (DS) perception complements in Timbira, and of the inessive *-ere(ka)* to temporal/conditional clauses.

The *Panoan* languages are relatively homogeneous in a number of aspects, such as the presence of an ergative marker (generally containing some nasal element) that has a number of other functions as well. Most Panoan languages also have a complex switch-reference system in common, which is used to encode many different adverbial clause types (especially temporal and purpose). The switch-reference markers may show potential connections with case markers, although in a number of instances this is due to the fact that both case markers and switch reference markers are part of the same bigger transitivity concord system. Apart from those correspondences, there does not seem to be a lot of extensions of case markers to subordinators (extension of the ablative to temporal clauses with a time lapse between the two events in Yaminahua, and an extension of the similative to cognition complements in Kakataibo).

The central and southern *Quechuan* languages share many properties, also in their case systems, and also in their extensions to subordinators: extensions of

the accusative marker to marking several complement types, of the dative to purpose clauses, the ablative to reason clauses, the limitative to 'until' clauses, and the 'because of' marker to reason clauses are typical Quechuan extensions. Although the northern Quechuan varieties underwent some changes in their case systems compared to the southern ones (such as a merger of the benefactive and genitive in Imbabura Quechua), the 'Quechuan' extensions still exist.

Like Cariban languages, many *Tupian* languages have adpositions that can inflect for person, although in a number of languages these adpositions have developed into inflectional case markers (Rodrigues & Cabral 2012: 517). With the exception of Kamaiurá which, in Seki's analysis, has a nuclear case marker, Tupian languages do not have structural case markers. Dative/benefactive, comitative (and to a lesser extent instrument), and spatial cases are common. Again, extensions to subordinate clauses are relatively marginal: Emerillon uses the comitative marker *-ehe* to mark reason clauses, and the spatial markers *-koti* 'allative' to cognition complements, and *-upi* 'perlative' to comparative constructions. The Karitiana oblique *-ty*, like *-koti*, is used to mark complement clauses.

The language families, then, show rather diverse effects, possibly related to the time-depth and/or geographical spread of the families, with Quechuan and Panoan languages being relatively homogeneous, the others less so. In conclusion it seems that genealogical effects can influence distributional patterns, but at the same time case systems (and their extensions to subordinators) seem to be diachronically rather unstable, so that older, or more dispersed language families show more disparate patterns.

5. Conclusion

Nominalized subordinate clauses are very common cross-linguistically, but perhaps especially in South American languages. It is therefore not unexpected to find many examples of case markers that have been extended to mark subordinate clauses, following proposed diachronic channels for subordinate clauses arising from or being equated with nominal structures (Heine & Kuteva 2007).

The semantics of the extensions show some unity across languages, which seems to be attributable to a combination of naturalness effects, areal effects, and genealogical effects, which can be summarized as follows.

General, functionally/cognitively motivated principles:

i. Nominalizations express the syntactically nominal status of subordinate clauses.
ii. There is a clear space–time connection.

iii. Certain specific 'natural' semantics extensions (locative-overlap, dative-purpose, and to a lesser extent ablative-sequence and comitative-overlap) can be observed.

Potentially areal effects:

i. Accusative to complement clause markers in northern Andes and adjacent areas.
ii. Dative to purpose clauses in southern-central Andes and adjacent foothills.
iii. General spatial markers to complement clauses in Tariana and Hup (Vaupés).
iv. Lack of extensions of case markers in the eastern Guaporé-Mamoré region and adjacent areas.
v. Lack of case in general in the Chaco.

Genealogical effects:

i. Relatively few extensions of case markers to subordinators in Arawak, Macro-Ge, and Tupian.
ii. Some potential diachronic but uncertain connections between case markers and subordinators in Carib and Panoan.
iii. Perhaps a (minor) pattern of extensions of peripheral case markers to complement-marking elements in Tupian.
iv. A relatively homogenous set of extensions of case markers to subordinators in Quechuan languages.

Acknowledgement

I thank Roberto Zariquiey, David Fleck and an anonymous reviewer for very useful comments on an earlier draft of this paper. Remaining errors are mine.

Abbreviations

1	first person	IRR	irrealis
2	second person	LIM	limitative
3	third person	LK	linker
4	fourth person (we inclusive)	LOC	locative
A	transitive subject	LOCUT	locutor
ABL	ablative	M	masculine
ABS	absolutive	MID	middle voice
ACC	accusative	MOT	motion
ACT	action	MST	mental state

ADH	adhesive
ADV	adverbial
AFF	affirmative
APPL	applicative
ART	article
ASSERT	assertive
ASUF	adjective suffix
AUX	auxiliary
BEN	benefactive
C	coreferent
CAUS	causative
CL	classifier
COM	comitative
COMP	complementizer
COP	copula
DAT	dative
DECL	declarative
DEM	demonstrative
DIM	diminutive
DIR	direction
DR	direct (voice)
DSC	discontinuative
DUR	durative
EMPH	emphasis
ERG	ergative
EXC	exclusive
F	feminine
FM	formative
FOC	focus
FUT	future
GER	gerund
HAB	habitual
HIAF	high affectedness
HSAY	hearsay
IMPFV	imperfective
INAN	inanimate
INC	inclusive
INESS	inessive
INF	infinitive
INST	instrument
INTN	intention
INTR	intransitive
INTS	intensive
MST	mental state
N/H	non-human
NEG	negation
NEUT	neuter
NLZ	nominalizer
NONA/S	non-subject
NONFUT	non-future
NPREF	noun prefix
NUC	nuclear case
OBJ	object
OBL	oblique
OBLG	obligatory, obligative
PAST	past
PERF	perfect
PERL	perlative
PFV	perfective
PL	plural
POSS	possessive
PREP	preposition
PRES	present
PROG	progressive
PRP	purpose
PTC	participle
REA	realis
REAS	reason
REC	reciprocal
RED	reduplication
REF	reference
REFL	reflexive
REL	relativizer
REP	reportative
S	intransitive subject
SG	singular
SIM	simulative
STD	standing
STV	stativizer
SUB	subordinator
TEMP	temporal
THM	theme
TR	transitive
VSM	verbal stem marker
WH	question word

References

Adelaar, Willem. 2008. Towards a typological profile of the Andean languages. In *Evidence and Counter-Evidence. Essays in Honour of Frederik Kortlandt*, Vol. 2: *General Linguistics*, Alexander Lubotsky, Jos Schaeken & Jeroen Wiedenhof (eds), 23–33. Amsterdam: Rodopi.

Adelaar, Willem with Pieter Muysken. 2004. *The Languages of the Andes*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511486852

Aikhenvald, Alexandra Y. 1999. The Arawak language family. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 65–106.

Aikhenvald, Alexandra Y. 2002. *Language Contact in Amazonia*. Oxford: OUP.

Aikhenvald, Alexandra Y. 2003. *A Grammar of Tariana, from Northwest Amazonia*. Cambridge: CUP. https://doi.org/10.1017/CBO9781107050952

Alves, Flavia. 2004. O Timbira falado pelos canela apãniekrá: uma contribuição aos estudos da morfossintaxe de uma língua Jê. PhD dissertation Universidade Estadual de Campinas, São Paolo.

Andrade-Martins, Silvana. 2004. Fonologia e gramática Dâw. PhD dissertation, Vrije Universiteit Amsterdam. Utrecht: LOT Publications.

Antunes, Gabriel. 2004. A grammar of Sabanê: A Nambikwaran language. PhD dissertation, Vrije Universiteit Amsterdam. Utrecht: LOT Publications.

Bacelar, Laércio. 2004. Gramática da língua Kanoê. PhD dissertation, Radboud University Nijmegen.

Baerman, Matthew & Brown, Dunstan. 2005. Case syncretism. In *The World Atlas of Language Structures*, Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds), 118–121. Oxford: OUP.

Bickel, Balthasar & Nichols, Johanna. 2007. Inflectional morphology. In *Language Typology and Syntactic Description*, Vol. III: *Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 169–240. Cambridge: CUP. https://doi.org/10.1017/CBO9780511618437.003

Bickel, Balthasar & Nichols, Johanna. 2008. The geography of case. In *The Oxford Handbook of Case*, Andrej Malchukov & Andrew Spencer (eds), 479–493. Oxford: OUP.

Blake, Barry J. 2001. *Case*. Cambridge: CUP. https://doi.org/10.1017/CBO9781139164894

Cole, Peter. 1982. *Imbabura Quechua*. Amsterdam: North Holland.

Comrie, Bernard, Golluscio, Lucía A., González, Hebe & Vidal, Alejandra. 2010. El Chaco como área lingüística. In *Estudios de lenguas amerindias*, 2: *Contribuciones al estudio de las lenguas originarias de América*, Zarina Estrada-Fernández & Ramón Arzápalo Marín (eds), 85–132. Hermosillo, Sonora: Editorial Universidad de Sonora.

Corbett, Greville G. 2006. *Agreement*. Cambridge: CUP.

Crevels, Mily. 2012. Itonama. In *Las lenguas de Bolivia*, Tomo II: *Amazonía*, Mily Crevels & Pieter Muysken (eds), 233–394. La Paz: Plural.

Crevels, Mily & van der Voort, Hein. 2008. The Guaporé-Mamoré region as a linguistic area. In *From Linguistic Areas to Areal Linguistics* [Studies in Language Companion Series 90], Pieter Muysken (ed.) 151–180. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.90.04cre

Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.

Croft, William. 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago IL: University of Chicago Press.

Crowell, Thomas H. 1979. A grammar of Bororo. PhD dissertation, Cornell University.

Curnow, Timothy. 1997. A grammar of Awa Pit (Cuaiquier): An indigenous language of south-western Colombia. PhD dissertation, Australian National University, Canberra.
Danielsen, Swintha 2007. *Baure: An Arawak Language of Bolivia*. Leiden: CNWS Publications.
Derbyshire, Desmond D. 1985. *Hixkaryana and Linguistic Typology*. Arlington TX: SIL and The University of Texas at Arlington.
Derbyshire, Desmond D. 1999. Carib. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 23–64.
Dickinson, Connie. 2002. Complex predicates in Tsafiki. PhD disssertation, University of Oregon.
Diehl, Lon 1975. Space case: Some principles and their implications concerning linear ordering in natural language. SIL working papers, University of Dakota.
Dixon, Robert M. W. 2004. *The Jarawara Language of Southern Amazonia*. Oxford: OUP.
Dixon, Robert M. W. & Aikhenvald, Alexandra Y. 1999. Introduction. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 1–22.
Duff-Trip, Martha. 1997. *Gramática del idioma yanesha' (amuesha)*. Lima: Ministerio de Educación; Instituto Lingüístico de Verano.
Eberhard, David. 2009. *Mamaindê Grammar: A Northern Nambikwara Language and its Cultural Context*. Utrecht: LOT Publications.
Epps, Patience. 2008. *A Grammar of Hup*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110199079
Everett, Caleb. 2006. Gestural, perceptual, and conceptual patterns in Karitiana. PhD dissertation, Rice University.
Everett, Dan & Barbara Kern. 1997. *Wari': The Pacaas Novos Language of Western Brazil*. London: Routledge.
Facundes, Sidney. 2000. The language of the Apurinã people of Brazil (Maipure/Arawak). PhD dissertation, State University of New York at Buffalo.
Faust, Norma & Loos, Eugene. 2002. *Gramática del idioma yaminahua*. Lima: SIL.
Fernández-Garay, Ana. 1998. *El Tehuelche: Una lengua en vías de extinción*. Valdivia: Andros Lta.
Fernandez-Garay, Ana. 2004. *Diccionario Tehuelche Español/Índice Español-Tehuelche*. Leiden: CNWS.
Fleck, David William. 2003. A Grammar of Matsés. PhD dissertation, Rice University.
Frank, Paul. 1985. A grammar of Ika. PhD dissertation, University of Pennsylvania.
Gabas Jr., Nilson. 1999. A grammar of Karo, Tupi (Brazil). PhD dissertation, University of California at Santa Barbara.
Galucio, Ana Vilacy. 2001. The morphosyntax of Mekens. PhD dissertation, University of Chicago.
Genetti, Carol. 1986. The development of subordinators from postpositions in Bodic languages. In *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, Vassiliki Nikiforidou, Mary VanClay, Mary Niepokuj & Deborah Feder (eds), 387–400. Berkeley CA: Berkeley Linguistic Society, University of California.
van Gijn, Rik. 2006. A grammar of Yurakaré. PhD dissertation, Radboud University Nijmegen.
van Gijn, Rik. 2014. Subordination strategies in South America: Nominalization. In *The Native Languages of South America: Origins, Development, Typology*, Loretta O'Connor & Pieter Muysken (eds), 274–297. Cambridge: CUP. https://doi.org/10.1017/CBO9781107360105.015
van Gijn, Rik. 2015. Morphological synthesis in the Guaporé-Mamoré. A contact feature? *Linguistic Discovery* 13(2): 96–122. https://doi.org/10.1349/PS1.1537-0852.A.464

van Gijn, Rik, Katharina Haude & Pieter Muysken 2011. Subordination strategies in South America: An overview. In *Subordination in Native South American Languages* [Typological Studies in Language 97], Rik van Gijn, Katharina Haude & Pieter Muysken (eds), 1–24. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.97.01van

Gildea, Spike. 1998. *On Reconstructing Grammar: Comparative Cariban Morphosyntax*. Oxford: OUP.

Gildea, Spike. 2012. Linguistic studies in the Cariban family. In *The Indigenous Languages of South America: A Comprehensive Guide*, Lyle Campbell & Verónica Grondona (eds), 441–494. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110258035.441

Girón, Jesús Mario. 2008. Una gramática del Wânsöjöt (Puinave). PhD dissertation, Vrije Universiteit Amsterdam. Utrecht: LOT Publications.

González, Hebe 2005. A grammar of Tapiete (Tupi-Guaraní). PhD dissertation, University of Pittsburgh.

Goodwin-Gómez, Gale. 1990. The Shirina dialect of Yanam (northern Brazil). PhD dissertation, Columbia University.

Grondona, Verónica. 1998. A grammar of Mocoví. PhD dissertation, University of Pittsburgh.

Guillaume, Antoine. 2008. *A Grammar of Cavineña*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110211771

Guillaume, Antoine & Rose, Françoise. 2010. Sociative causative markers in South American languages: A possible areal feature. In *Essais de typologie et de linguistique générale: Mélanges offerts à Denis Creissels*, Franck Floricic (ed.), 383–402. Charenton-le-Pont: ENS Éditions

Guillaume, Antoine & Rose, Françoise. 2011. Introduction to argument-encoding systems in Bolivian Amazonian languages. *International Journal of American Linguistics* 77(4): 459–468 https://doi.org/10.1086/662152

Guirardello, Raquel. 1999. A reference grammar of Trumai. PhD dissertation, Rice University.

Hardman, Martha J. 2001. *Aymara*. Munich: Lincom.

Haspelmath, Martin. 1997. *From Space to Time: Temporal Adverbials in the Languages of the World*. Munich: Lincom.

Haude, Katharina. 2006. A grammar of Movima. PhD dissertation, Radboud University Nijmegen.

Heine, Bernd. 2008. Grammaticalization of cases. In *The Oxford Handbook of Case*, Andrej Malchukov & Andrew Spencer (eds), 458–469. Oxford: OUP.

Heine, Bernd & Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511613463

Heine, Bernd & Kuteva, Tania. 2007. *The Genesis of Grammar: A Reconstruction*. Oxford: OUP.

Hopper, Paul & Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language* 56: 251–299. https://doi.org/10.1353/lan.1980.0017

Johanson, Lars. 2008. Case and contact linguistics. In *The Oxford Handbook of Case*, Andrej Malchukov & Andrew Spencer (eds), 494–504. Oxford: OUP.

Jung, Ingrid. 2008. *Gramática del páez o nasa yuwe: Descripción de una lengua indígena de Colombia*. Munich: Lincom.

van de Kerke, Simon. 2009. Leko. In *Las lenguas de Bolivia, Tomo I: Ámbito andino*, Mily Crevels & Pieter Muysken (eds), 287–332. La Paz: Plural.

Kulikov, Leonid. 2006. Case systems in a diachronic perspective: A typological sketch. In *Case, Valency and Transitivity* [Studies in Language Companion Series 77], Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds), 23–47. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.77.04kul

Kulikov, Leonid. 2008. The evolution of case systems. In *The Oxford Handbook of Case*, Andrej Malchukov & Andrew Spencer (eds), 439–457. Oxford: OUP.

Langacker, Ronald W. 1987a. *Foundations of Cognitive Grammar*, Vol. I: *Theoretical Prerequisites*. Stanford CA: Stanford University Press.

Langacker, Ronald W. 1987b. Nouns and verbs. *Language* 63: 53–94. https://doi.org/10.2307/415384

Lefebvre, Claire & Muysken, Pieter. 1988. *Mixed Categories: Nominalizations in Quechua*. Dordrecht: Kluwer.

Loos, Eugene. 1999. Pano. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 227–250.

Malchukov, Andrej. 2006. Constraining nominalization: Form/function competition. *Linguistics* 44 (5): 973–1009. https://doi.org/10.1515/LING.2006.032

Matras, Yaron & Sakel, Jeanette. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31(4): 829–865. https://doi.org/10.1075/sl.31.4.05mat

Maxwell, Michael & Morse, Nancy L. 1999. *Cubeo Grammar*. Dallas TX: Summer Institute of Linguistics and the University of Texas at Arlington.

McGregor, William B. 2009. Typology of ergativity. *Language and Linguistics Compass* 3(1): 480–508 https://doi.org/10.1111/j.1749-818X.2008.00118.x

Meira, Sérgio. 1999. A grammar of Tiriyó. PhD dissertation, Rice University.

Miller, Marion. 1999. *Desano Grammar*. Dallas TX: Summer Institute of Linguistics.

Mortensen, Charles. 1999. *A Reference Grammar of the Northern Embera Languages*. Arlington TX: Summer Institute of Linguistics.

Nichols, Johanna. 2003. Diversity and stability in language. In *The Handbook of Historical Linguistics*, Brian Joseph & Richard Janda (eds). Malden MA: Wiley-Blackwell. https://doi.org/10.1002/9780470756393.ch5

Nonato, Rafael. 2008. Ainore Boe egore: Um estudo descritivo da língua bororo. MA thesis, Universidade Estadual de Campinas.

Olawsky, Knut. 2006. *A Grammar of Urarina*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110892932

Overall, Simon. 2007. A grammar of Aguaruna. PhD dissertation, LaTrobe University.

Payne, Doris. 1985. Aspects of the Grammar of Yagua: A Typological Perspective. PhD dissertation, University of California, Los Angeles.

Payne, Thomas & Payne, Doris. 2013. *A Typological Grammar of Panare: A Cariban Language of Venezuela*. Leiden: Brill.

Popjes, Jack & Popjes, Josephine. 1986. Canela-Krahô. In *Handbook of Amazonian Languages* Vol. 1, Desmond C. Derbyshire & Geoffrey K. Pullum (eds), 128–199. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110850819.128

Rodrigues, Aryon D. 1999. Macro-Jê. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 165–206.

Rodrigues, Aryon D. & Cabral, Ana Suelly Arruda Câmara. 2012. Tupían. In *The Indigenous Languages of South America: A Comprehensive Guide*, Lyle Campbell & Verónica Grondona (eds), 495–574. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110258035.495

Romero-Figeroa, Andrés. 1997. *A Reference Grammar of Warao*. Munich: Lincom.

Rose, Françoise. 2011. *Grammaire de l'Émérillon Teko: Une langue Tupí-Guaraní de Guyane Française*. Leuven: Peeters.

Sakel, Jeanette. 2004. *A Grammar of Mosetén*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110915280

Seifart, Frank. 2005. The structure and use of shape-based noun classes in Miraña. PhD dissertation, Radboud University Nijmegen.

Seki, Lucy. 2000. *Gramática do Kamiurá: Língua Tupi-Guarani do Alto Xingu*. São Paolo: UNICAMP.

Silva, Leia. 2011. Morphosyntaxe du Rikbaktsa. PhD dissertation, Université Denis Diderot, Paris 7.

Smeets, Ineke. 2008. *A Grammar of Mapuche*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110211795

Stolz, Thomas, Stroh, Cornelia & Urdze, Aina. 2005. Comitatives and instrumentals. In *The World Atlas of Language Structures*, Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds), 214–217. Oxford: OUP.

Stolz, Thomas, Stroh, Cornelia & Urdze, Aina. 2006. *On Comitatives and Related Categories: A Typological Study with Special Focus on the Languages of Europe*. Berlin: Walter de Gruyter.

Storto, Luciana R. 1999. *Aspects of a Karitiana grammar*. PhD dissertation, MIT.

Storto, Luciana. 2011. Subordination in Karitiana. *Amerindia* 35: 219–236.

Terraza, Jimena. 2009. Gramática del Wichí: Fonología y morfosintaxis. PhD dissertation, Université du Québec à Montréal.

Thiesen, Wesley & Weber, David. 2012. *A Grammar of Bora with Special Attention to Tone*. Dallas TX: SIL International.

Thomason, Sarah & Kaufmann, Terrence. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley CA: University of California Press.

Torero, Alfredo. 2002. *Idiomas de los Andes: Lingüística e historia*. Lima: Editorial Horizonte/ IFEA.

Valenzuela, Pilar. 2003. Transitivity in Shipibo-Konibo grammar. PhD dissertation, University of Oregon.

Vallejos, Rosa. 2010. A grammar of Kokama-Kokamilla. PhD dissertation, University of Oregon

Vidal, Alejandra. 2001. Pilagá grammar (Guaykuruan family, Argentina). PhD dissertation, University of Oregon.

van der Voort, Hein. 2004. *A Grammar of Kwaza*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197280

Vuillermet, Marine. 2012. A grammar of Ese Ejja, a Takanan language of the Bolvian Amazon. PhD dissertation, Université Lumière Lyon 2.

Weber, David. 1989. *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley CA: University of California Press.

Weinreich, Uriel. 1953. *Languages in Contact: Findings and Problems*. New York NY: Linguistic Circle of New York.

Wise, Mary Ruth. 2002. Applicative affixes in Peruvian Amazonian languages. In *Current Studies on South American Languages*, Milly Crevels, Simon van de Kerke, Sergio Meira & Hein van der Voort (eds), 329–344. Leiden: CNWS.

Zariquiey, Roberto. 2011. A grammar of Kashibo-Kakataibo. PhD dissertation, LaTrobe University.

CHAPTER 5

Nominalized constructions with argument functions in the languages of the Chaco

A contribution to the typology of indigenous South American languages

Lucía A. Golluscio, Felipe Hasler and Willem de Reuse
Universidad de Buenos Aires and Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina / Universidad de Chile / University of North Texas and The Language Conservancy

The complementation strategies attested in two well-known areas, the Andean area and the Amazonian area are compared with those existing in Chaco languages, which possibly constitute a linguistic area as well. The Chaco languages examined belong to the Guaycuruan, Lule-Vilela, Mataguayan, Tupi-Guarani and Zamucoan families. While nominalization is marked grammatically in Andean and Amazonian languages, a more general dependency marker is typically used in the Chaco languages. Also, unlike nominalizers in Andean and Amazonian languages, Chaco language subordinators are not involved in lexical nominalization. Finally, parataxis is documented as an extended complementation strategy among the Chacoan languages.

Keywords: nominalization, complementation, parataxis, dependency markers, Chaco languages

1. Introduction

Nominalization has traditionally been defined as the process of transforming a syntactic element, be it a verb, a verbal phrase, or a clause, into a noun – see for example Comrie & Thompson (1985), Noonan (1985), Givón (2001), Malchukov (2006), among others. These authors agree that nominalization is, on the one hand, functionally motivated, that is, it generally takes place when the element in question fulfills a prototypical nominal syntactic function (subject, object, nominal predicate, etc.). On the other hand, they agree that this functional change is

https://doi.org/10.1075/tsl.124.05gol

correlated with the gradual loss of verbal categories and the acquisition of nominal properties along a continuum. Shibatani (2009) agrees with them in that nominalization is related to the element's nominal functional properties and, therefore, he defines it as a functional notion linked to the creation of a referential expression in contrast to a predicative expression. Shibatani (this volume) considers that this process does not necessarily correlate with the loss of verbal properties or the acquisition of nominal properties – what matters is that the nominalized structure behaves semantically as an entity-denoting expression and syntactically like a noun. In this paper, we will follow Shibatani's view.

The literature on the languages of the Andean highlands and on Amazonian languages has paid special attention to nominalization in recent years.[1] Adding Chaco languages to this literature, we contribute to the typology of nominalization in South America while exploring the relationship between nominalization and subordination. We focus on the description of Chaco nominalization strategies in argument functions – that is, those linked to complementation. We will also compare Chaco strategies with those found in Andean and Amazonian languages.

To that end, we have chosen two parameters used to describe syntactic dependency: (a) the desentencialization of the dependent clause, i.e. the process defined by the loss of dependent clause operators and its relational elements (Lehmann, 1988); and (b) flagging, i.e. the explicit marking vs. lack of marking of dependency (van Gijn, 2014). Specifically, regarding (a) we consider the resources used for the expression of subject in the dependent element, and regarding (b) the nature of the dependency marker. We focus on these features as they offer the most reliable data, in addition to enabling us to compare nominalization strategies of Chaco languages with those of Andean and Amazonian languages.

The Gran Chaco (henceforth the Chaco) is a geographical area of south central South America, encompassing a southeastern portion of Bolivia, northern Argentina east of the Andes, and the area west of the Paraguay river in Paraguay (Adelaar 2004: 488). It is a cultural area in which the permanent interaction between various human groups, dating from before the Spanish conquest, has generated shared features in its sociopolitical organization, its subsistence practices, rituals, music and cosmology (Braunstein 2005). Its status as a linguistic area is currently under discussion (Comrie et al. 2010; van Gijn et al. 2013; Bertinetto & Ciucci 2012; Campbell & Grondona 2012; Golluscio et al. 2014; Golluscio et al. 2015; González 2015; and Campbell 2017, among others). The Chaco languages

1. See Torero (2002) and Adelaar (2004) for an overview of Andean languages; Aikhenvald & Dixon (1999) for an overview of Amazonian languages; van Gijn et al. (2011) for an overview of subordination in South America, and van Gijn (2014) for a comparison of the Andes and Amazonia regarding nominalization.

considered in this paper belong to five families: Wichi and Nivacle (Mataguayan), Toba or Qom and Pilaga (Guaycuruan), Ayoreo and Chamacoco (Zamucoan), Tapiete and Chacoan Guarani or Ava-Guarani (Tupi-Guarani), Vilela and Lule (Lule-Vilela; genetic affiliation under discussion). To build our sample, we have considered the following criteria: (a) we have selected languages that are spoken or have been spoken in the Chaco region, (b) we have selected two languages per family for the sake of concision, and (c) we have only considered languages for which descriptions of nominalization are available. See Table 1.[2]

Table 1. Chaco languages considered in this study

Family	Language	Geographical location	Sources
Mataguayan	Wichi	Argentinian Provinces of Salta and Chaco, and the northern bank of the Pilcomayo River in Bolivia	Nercesian (2011, 2014)
	Nivacle	From the Pilcomayo River, on the border between Argentina and Paraguay, to the center of the Paraguayan Chaco (part of the Departments of Boquerón and Presidente Hayes)	Fabre (2014)
Guaycuruan	Pilaga	Argentinian Departments of Patiño and Bermejo	Vidal (2002)
	Toba	Argentinian Provinces of Formosa, Chaco, Salta, Santa Fe and Buenos Aires. It is also spoken in Paraguay (Cerrito) and Bolivia (Villa Montes)	Censabella (2002), Carpio (2009), Messineo & Cúneo (2009–10).
Zamucoan	Ayoreo	From the Grande River to the Paraguay River and from the east of Santa Cruz de la Sierra (Bolivia) to the north of Paraguay	Bertinetto (2009), Bertinetto & Ciucci (2012), Durante (2014).
	Chamacoco	Over the Paraguayan Chaco along the Paraguay River in the Department of Alto Paraguay (Paraguay)	Bertinetto & Ciucci (2012), Ciucci (2013)

(*continued*)

2. We did not include the Enlhet-Enenlhet family of the Chaco, also called the Mascoyan family. The available written documentation on these languages (e.g.,Unruh et al. 2003), although impressive, consists of pedagogical materials and is not easy to use because the examples are not glossed. Furthermore, linguistic papers on this family (e.g., Kalisch 2009) do not include suitable examples on the topic of this paper.

Table 1. (*continued*)

Family	Language	Geographical location	Sources
Tupi-Guarani	Tapiete	Argentinian, Bolivian, and Paraguayan Chaco	González 2005, Ciccone (2014)
	Chacoan Guarani	From the Bolivian region of Santa Cruz de la Sierra to the northwest of the Argentinian province of Salta and the east of Jujuy	Dietrich (2009–10)
Lule-Vilela (affiliation under discussion)	Vilela	Probably spoken before Spanish contact in Western Chaco or in the Andean foothills in northwestern Argentina. Since the end of the 19th century, documented in Chaco and Corrientes. Severely endangered. There is no speech community and only two speakers have been located.	Llamas (1910), Lozano (1970, 2006), Domínguez et al. 2006, Golluscio & González (2008), Golluscio (2009–10, 2015)
	Lule	Spoken by seminomadic peoples that lived between the Pilcomayo River and the northwestern foothills of the Argentinian Andes. Extinct since before the beginning of the 19th century.	Machoni (1732), de Reuse (2015), Zamponi & de Reuse (2012)

Since the languages in our sample differ in their degree of documentation and description, as well as in their degree of vitality, there are limitations upon our knowledge of nominalized constructions with argument functions. As a result, we have selected those complement constructions for which we could collect comparable information.

This article is organized as follows. After this introduction (§ 1), § 2 summarizes the relevant issues, § 3 describes nominalized expressions with argument functions in Chaco languages. In § 4 we systematize the nominalization strategies shared by the Chaco languages and compare them with Andean and Amazonian languages. In addition, we point to exceptions to expected patterns in two Chaco languages, i.e., Tapiete and Pilaga.

2. Preliminaries

Subordinate clauses in Amazonian languages "typically involve nominalized verbs, with the type of subordination being marked on the verb" (Aikhenvald & Dixon 1999: 9). Similarly, genealogically distinct Andean languages typically show nominalizer suffixes on the dependent clause, as in Quechuan and Aymaran languages (Cerrón-Palomino, 2008), Mapudungun (Smeets, 2008; Hasler, 2017) and Chipaya (Cerrón-Palomino, 2006).

Typically, Andean languages differ from Amazonian ones in the resources predominantly used for subject marking. While possessive morphology prevails in the former, morphology related to the person and number verbal inflection prevails in the latter. As van Gijn (2014: 289) points out, "the subjects as possessors are particularly common in the Andes and adjacent areas, probably by the Quechuan and Aymaran influence, but they also occur in non-contiguous regions in Amazonia". Also, both Amazonian and Andean languages share the same type of flagging: the dependent clause in its argument functions is marked by nominalizers. In general, these nominalizers are characterized by two features: (a) They can be involved in lexical nominalization (Comrie & Thompson, 1985; Shibatani, 2009, this volume), in addition to being able to generate grammatical nominalizations; and (b) when these nominalizations are adverbialized (resulting in adverbial subordination), they are marked by case affixes or adpositions (van Gijn 2014: 293).

In contrast, we claim that nominalized expressions with argument functions in Chaco languages are expressed by finite forms with clause-initial subordinators, or sometimes also by paratactic constructions. Thus, the subject of the dependent clause in these languages, as occurs in independent clauses, is marked by the person and number verbal inflection. We further claim that Chaco languages possess subordinators that do not exclusively mark nominalization. Rather, they function as general dependency markers that allow the dependent element to perform a nominal function. It is the fulfillment of this function, and not the presence of a special subordinator, that indicates nominalization.

Example (1) shows a nominalized expression with argument functions in Sateré-Mawé (Tupi, Amazonian), (2) presents a similar expression in Ayacucho Quechua (Quechuan, Andean) and (3) in Wichi (Mataguayan, Chaco).[3]

Sateré-Mawé (Tupian)

(1) *[kat pote it ere-to iu-wiwo* ***hap****]*$_{NMLZ}$ *ati-kuap teran*
what for NEG 2SG-go 1SG-with NMZR 1SG-know want
'I want to know why you won't go with me.' (Aikhenvald & Dixon 1999: 121)

Ayacucho Quechua (Quechuan)

(2) *[llapa ima-ruwa-****na****-n-ta]*$_{NMLZ}$ *ni-n*
all thing-do-NMZR-3.POSS-ACC say-IND.3SG
'He/she informs him/her of everything s/he has to do.'
(Soto Ruiz 1976, our glosses)

3. We translated the original Spanish examples into English. We have normalized the glosses following the Leipzig Glossing Rules (www.eva.mpg.de/lingua/resources/glossing-rules.php).

Wichi (Mataguayan)

(3) *n'-nuway [**toj** n'-yik]*$_{NMLZ}$
1SBJ-be.afraid SUB 1SBJ-leave
'I am afraid to leave.' (Nercesian 2011: 454)

In examples (1), (4) and (5), Sateré-Mawé uses the nominalizer *-hap* in grammatical nominalization, lexical nominalization, and adverbial subordination (Franceschini, 1999; Silva, 2010). In the case of adverbial subordination, the nominalizer must be accompanied by another suffix indicating the particular adverbial relation, for example, the *-mo* purpose suffix in (5). Similarly, the Ayacucho Quechua example (2) shows the nominalizer *-na* that can also be used in lexical nominalization (6) and in adverbial subordination (7). In adverbial subordination, the nominalizer must be accompanied by a case marker (here dative case *-paq*). In contrast, the Wichi example (3) shows the use of the *toj* subordinator, which is never involved in lexical nominalization and functions as an adverbial dependent marker without an adposition or case marker (8). Note that Wichi lexical nominalization uses other resources, such as the suffix *-yaj* (9).

Sateré-Mawé (Tupian)

(4) *hentup **hap***
pray NMZR
'place to pray' (Da Silva 2010: 164)

Sateré-Mawé (Tupian)

(5) *[te'eru-we-'yhop-sat-**ha(p)-mo**]*$_{NMLZ}$ *ra'yn P-tu-wat maués Kape*
3PL.A-REFL-leaf-hit-NMLZ-PURP ASP 3A-MID-go.PL Maués POSS
'They went to Maues in order to take back their money.'
(Franceschini 1999: 125)

Ayacucho Quechua (Quechuan)

(6) *qillqa-**na***
write-NMZR
'pencil' (Zariquiey & Córdova 2008: 222)

Ayacucho Quechua (Quechuan)

(7) *qam ceviche-ta yanu-nki [miku-**na**-y-**paq**]*$_{NMLZ}$
2SG ceviche-ACC cook-IND.2SG eat-NMZR-1SG.POSS-DAT
'You cook ceviche for me to eat.' (Zariquiey & Córdova 2008: 244)

Wichi (Mataguayan)

(8) *wet n'-yomey mayek [toj pajche toj kamaj neweche siwele]*
CONJ 1SBJ-talk stuff SUB before SUB still there.are.NEG white.men
'I talk about the past when no white men existed.' (Nercesian 2011: 455)

Wichi (Mataguayan)

(9) *la-chefwen-**yaj***
3.POSS-teaching-NMZR
'his/her teaching' (Nercesian 2011: 211)

3. Nominalization in Chaco languages

In this section, we describe nominalization with argument functions in Chaco languages. First, we consider data from the Mataguayan languages: Wichi (§ 3.1) and Nivacle (§ 3.2). Second, we examine data from the Guaycuruan languages: Toba or Qom (§ 3.3) and Pilaga (§ 3.4). Third, we turn to data from the Zamucoan languages: Ayoreo (§ 3.5) and Chamacoco (§ 3.6). Fourth, we focus on the Tapiete (§ 3.7) and Chacoan Guarani or Ava-Guarani (§ 3.8) languages. Finally, we look at Vilela (§ 3.9) and Lule (§ 3.10) which belong to a putative Lule-Vilela family. To test our claim that the nominalized expressions with argument functions in the Chaco languages are expressed by finite forms with clause-initial subordinators, and also sometimes by mere parataxis (§ 2), we will look, for each of the ten languages under discussion, at (a) the characteristics of subject marking in the dependent element and at (b) the nature of the dependency link.

3.1 Wichi (Mataguayan)

According to Nercesian (2011, 2014), nominalized constructions in Wichi express their subject through verbal person inflection, as in independent clauses. These constructions are introduced by the subordinator *toj* to express realis complement constructions (10), and *che*, to express irrealis complement constructions (11).

(10) *n'-nuway* *[**toj** n'-yik]*$_{NMLZ}$*
1SBJ-be.afraid SUB 1SBJ-leave
'I am afraid to leave'. (Nercesian 2011: 454)

(11) *iwusey* *[**che** nom]*$_{NMLZ}$*
3SBJ:can SUB.IRR 3SBJ:come
'S/he might come'. (Nercesian 2011: 456)

Desiderative verbs, such as *t'eynlo* 'want' and *iwatlo* 'need', require the irrealis suffix *-k* to receive a verbal clause as an object (see 12). If the argument is a noun phrase, this suffix is not used. The suffix *-k* can be replaced by the subordinator *che* and the two can co-occur. We take this behavior to be evidence that *-k* functions like *che* and *toj*: all three allow for complementation (Nercesian 2009).

(12) *n'-watlo-**k** [n'-yiset=hen tulu t'isan]*$_{NMLZ}$
1-need-IRR 1-cut=PL cow meat
'I need to cut the beef.' (Nercesian 2009: 190)

In addition, *toj* and *che* can appear in nominalized constructions with other functions, such as relativization. They can also occur in other constructions, such as dependent clauses with non-future semantics (in the case of *toj*, see (8)) or with future and conditional semantics (in the case of *che*, see (13)–(14)). Hence, we claim that *toj* and *che* are subordinators that express a general dependency relationship.[4] However, it is important to note that these subordinators cannot occur in lexical nominalization.

(13) *atsinha yenhilhna to-kuset **[che** wak'alh]*
woman 3SBJ:make:FUT POSS.INDF-trouser SUB.IRR 3SBJ:finish
hiw'en-hu= ma lhos
3SBJ:have-APPL= EMPH 3POSS:son
'The woman is going to make trousers and when she finishes it she is going to give it to her son.' (Nercesian 2011: 456)

(14) *[**che** iwumcho' chefwal] ha-n'-tiyoj-hi-che tayhi*
SUB.IRR rain tomorrow NEG.R-1SBJ-go-NEG.R-DIR brush
'If it rains tomorrow, I won't go to the brush.' (Nercesian 2011: 456)

3.2 Nivacle (Mataguayan)

According to Fabre (2014), when a nominalized construction with argument functions occurs, a dependency marker is added to the main verb: either the instrumental applicative (15), or the punctual applicative (16). These markers allow the embedding of the dependent clause. The dependent clause is co-referential with the independent clause, that is, it keeps the verbal person inflection of the independent clauses. In addition, the dependent clause is introduced by the subordinator *ti*, in the case of a realis predicate (17) or *ca*, in the case of an irrealis predicate (15)–(16).

(15) *ts'-ei.jatsjan-e-sh **[ca** ja-'vijatan-a pa-va matas*
(3A).1R-teach-3-INS SUB 1SBJ-respect-PUNCTDET-PL things
ta-'vuju'y-e-sh-a]$_{NMLZ}$
3SBJ-belong.other.people-3-INS-PUNCT
'S/he taught me to respect the things belonging to other people.'
(Fabre 2014: 272)

4. For other functions of *toj* and *che*, see Nercesian (2011, 2014).

(16) *ta-nachan-e-sh-a* *[**ca** n-vaf yaju]*$_{NMLZ}$
3-have.intention-3-APPL.INS-APPL.SUB SUB 3.IRR-die PROSP
'S/he wants/wanted to die.' (Fabre 2014: 160)

(17) *yi-'van* *[**ti** Ø-tsiisha-an Ø-lhacom'a]*$_{NMLZ}$
3A(>3P)-see SUB 3SBJ-be.naked-Q 3SBJ-be.all
'S/he saw that they were all naked.' (Fabre 2014: 273)

Like the Wichi subordinators, both *ca* and *ti* can occur in nominalized constructions with other functions, such as relativization, as well as in constructions that do not involve nominalization. For example, *ca* can introduce dependent clauses with future temporal (18) and conditional (19) meanings without the need for case markers or adpositions. In addition, *ti* is frequently used as "a simple link" (Fabre 2014: 288) (see example 20). Again, we claim that both *ca* and *ti* are subordinators that mark a more general dependency relationship than strict nominalization. As in Wichi, these subordinators cannot occur in lexical nominalization.

(18) *[**ca** ni-t'ôôi-taj lh-pa vat'-ônjanche] ca-pi Ø-vaf-sha'ne*
SUB 3SBJ.IRR-sound-CON F-DET POSS.INDF-trumpet DET-PL 3SBJ-die-3PL
pa neetshamch'e jayu
and 3SBJ-rise.up-UP-LONG PROSP
'When the trumpet sounds, the dead people will come back to life.'
(Fabre 2014: 275)

(19) *[**ca** a-fajulh-taj] na-va catiis pa a-tsjuulh*
if 2-be.able-CON.IRR DET-PL stars and 2A(>3P).IRR-count
'If you are able (to do it with) stars, count them!' (Fabre 2014: 275)

(20) *pa tatsha **ti** Ø-tat-vach'aclaj*
and quick SUB 3A-REFL-change
'And s/he quickly regretted it.' (Fabre 2014: 275)

3.3 Toba (Guaycuruan)

As in the Mataguayan languages, Toba (Censabella 2002; Carpio 2009) marks nominalized constructions with argument functions with the subordinator *da* (*ra* in the variety that Carpio (2009) describes). This subordinator is isomorphic with the demonstrative *da* indicating vertical position. Furthermore, the dependent clause marks its subject on the verbal person inflection, as in independent clauses. For example:

(21) *y-awatton [**ra** qa-n-qata-lek]*$_{NMLZ}$
3A-know SUB UA-3MID-chase-LOC
'(The rhea[5]) knows that it is being chased.' (Carpio 2009: 10)

This subordinator is not involved in lexical nominalization. In addition, it can occur in adverbial constructions without the need for case markers or adpositions (Messineo & Cúneo, 2009–10). Again, we claim that *da* is a subordinator indicating a more general dependency relationship than nominalization (see § 2).

(22) *Felipe n-pottak [**da** i-kewiyaougi']*$_{NMLZ}$
Felipe 3:MID-go.slowly SUB 3-walk
'Felipe walks slowly/Felipe is slow when he walks.'
(Messineo & Cúneo 2009–10: 241)

Another pattern, not attested in either Andean or Amazonian languages, and apparently not in Mataguayan languages either, consists of nominalization being expressed by a paratactic construction (i.e. juxtaposition of the matrix and dependent clause without any marking), as in the following example with a desiderative predicate.

(23) *wataGanaq da mashe qa-n-qo'ona zi Espinillo qa-y-awotayke*
soldiers SUB already IMPR.3MID-arrive CL Espinillo IMPS-3A-want
[qa-y-oda-wek itaGa zi qom]
IMPS-3A-expel-DIR:out again CL toba
'The soldiers who were arriving wanted to expel the Tobas once again.'
(Messineo & Cúneo 2009: 243)

We assume that parataxis is a typical Chaco pattern as well.

3.4 Pilaga (Guaycuruan)

Unlike the Mataguayan languages described above, and unlike Toba, Pilaga has two syntactic strategies to express nominalized constructions with argument functions: (a) a nominalizer *-naʂak* cooccurring with the subordinator *da*' and (b) the combination of finite forms in parataxis or with the subordinator *da*' (Vidal 2002).

We first consider nominalization with the suffix *-naʂak*. It should be noted that this suffix is used in both lexical and grammatical nominalization. (24) is an example of lexical nominalization with the suffix *-naʂak*.

(24) *l-onta-**naʂak***
3SG.POSS-work-NMZR
'his/her work' (Vidal 2002: 102)

5. A rhea is a flightless South American bird similar to an ostrich.

In grammatical nominalizations, *-naʂak* produces nominalized constructions with argument functions when the matrix and the complement subjects are co-referential. This strategy is used with desiderative matrix verbs such as 'want' and phasal matrix verbs such as 'finish'. In addition, the nominalized complement is introduced by the subordinator *da'*. The subject of the nominalized construction is marked with the appropriate possessive prefix.

(25) *se-take* *[**da'** y-onta-**naʂak**]*$_{NMLZ}$*
setA.1-want SUB POSS.l-work-NMZR
'I want to work.' (Vidal 2002: 356)

(26) *na'nolo' s-emat* *[**da'** y-onta-**naʂak**]*$_{NMLZ}$*
today setA.l-finish SUB POSS.l-work-NMZR
'I finish my work today.' (Vidal 2002: 356)

We now turn to the case of juxtaposition of the main and subordinate clauses either without subordinators (i.e. parataxis) (28) or with the subordinator *da'* (29). In these constructions, dependent verbs of non-nominalized complements can manifest the same verbal categories as matrix verbs, that is, they mark their subject via verbal inflection, as in independent clauses.

(27) *soʂote n-oma [soʂote* ***da'*** *Ø-ek]*$_{NMLZ}$
before setB.3-know before SUB setA.3-go
'He(i) knew that he(ii) had already gone.' (Vidal 2002: 359)

(28) *hayem d-aqta-n-em ga'* *[t-ae-yi]*$_{NMLZ}$
1SG setA.3-tell-ASP-DAT CLF setA.3-go-DIR
'He(i) told me that he(ii) has gone.' (Vidal 2002: 361)

The subordinator *da'* is isomorphic with the classifier *da'* which categorizes humans and animals as inherently vertical, or in a "standing" position (Vidal 2002), as in Toba. As in the two Mataguayan languages and Toba, this subordinator cannot be involved in lexical nominalization and can occur in adverbial subordination, such as temporal dependent clauses, without any case markers or adpositions (29).

(29) *[**da'** qo-y-aw'o* *ne-wosek* *qo-y-wetake* *lapat taʂa*
SUB IMPR.SBJ-setA.3-make INDF.POSS-stew IMPR.SBJ-setA.3-need meat and
l-alege
POSS.3-ingredients
'When one makes the stew, meat and ingredients are needed.'
(Vidal 2002: 370)

To summarize at this point, one of the constructions existing in Pilaga is not what we would expect for a Chaco language. Indeed, Pilaga shows a suffixal nominalizer which has the same characteristics as the Andean and Amazonian nominalizers.

However, this construction must be introduced by the subordinator *da'* which behaves like Chaco subordinators: it is not involved in lexical nominalization and it introduces adverbial constructions without case markers or adpositions.

Furthermore, Pilaga also has a dependent finite construction in juxtaposition with the matrix element or introduced by the subordinator *da'*. This second construction does fit in the general Chaco pattern.

3.5 Ayoreo (Zamucoan)

Like Toba, Ayoreo has two kinds of nominalized constructions with argument functions: a construction with subordinators and a paratactic construction.

In the construction with subordinators, the dependent clause can be introduced by the subordinator *uje* in realis verbal complements or by *ujetiga* in irrealis verbal complements. The dependent clause marks its subject as in matrix clauses. (Bertinetto & Ciucci 2012; Durante 2014).

(30) *disi-ode tangai* ***[uje*** *tai-a ch-uje pijan-e]*$_{NMLZ}$
child-M.PL IRREG.hear SUB truck-F.SG 3SG-honk horn-F.SG
'The children heard that the truck sounded its horn.' (Durante, p.c.)

(31) *y-apade pota* ***[ujetiga*** *abia di-ji*
1SG.POSS father SUB 3SG.POSS.daughter IRREG.3SG.come-in.towards
nguinguija-i]$_{NMLZ}$
home-M.SG
'My father wants his daughter to arrive his house.' (Durante, p.c.)

Both *uje* and *ujetiga* can occur in nominalized constructions with other functions, such as relativization. Additionally, they introduce dependent adverbial clauses which are unrelated to nominalization. For example, Durante (2014) points out that *uje* can introduce causal and temporal adverbial constructions and *ujetiga* can introduce hypothetical and final constructions.[6] For these reasons, we claim that both *uje* and *ujetiga* are subordinators that indicate a more general dependency relationship than nominalization, according to the parameters mentioned in (§ 2). As in Mataguayan and Guaycuruan languages, the subordinators are never involved in lexical nominalization.

(32) *y-abia e ch-isa* ***[uje*** *e ch-ise sique-e]*
1SG.POSS-daughter already 3SG-marry SUB already 3SG-reach year-F.SG
'My daughter was married when she had her birthday.' (Durante 2014: 18)

6. Because of these meaning differences, Bertinetto and Ciucci (2012) claim that these markers are related to the realis and irrealis moods, respectively.

(33) *[**jetiga** Juan di]* *Maria jo*
SUB Juan come.3SG María go.out.3SG
'If Juan comes, María goes out.' (Durante 2014: 18)

In the paratactic construction, the two clauses involved mark their subject through person and number verbal inflection (Durante 2014):

(34) *que y-iaja* *oaji-rique* *a*
no 1SG-know come.from-INDET MOD
'I do not know where s/he comes from.' (Durante 2014: 18)

3.6 Chamacoco (Zamucoan)

Chamacoco uses the subordinator *uje* in nominalizations with argument functions. The dependent clause introduced by this subordinator marks its subject through person verbal inflection, as in independent clauses. For example:

(35) *ch-unt [**uje** sehe o-d-uhu ilorz]*$_{NMLZ}$
3-hear SUB DES P-3.IR-kill this.way.IR
'He heard that they want to kill (him) in this way.' (Ciucci 2013)

As in Ayoreo, *uje* can occur in other nominalizing constructions, such as relativization (Bertinetto & Ciucci 2012), and can introduce other dependent clauses which do not involve nominalization, such as temporal, conditional and purpose clauses (Bertinetto & Ciucci 2012). Again, we claim that *uje* is a subordinator that indicates a more general dependency relationship than nominalization (see § 2). For example:

(36) *[**uje** ye t-uu.leeych], ich ese aahn-t s-erz yoo*
SUBNEG 1SG-fight CONJ DEM.M.SG evil.spirit-M.SG 3-win 1SG
'When/if I don't fight, that evil spirit will defeat me.'
(Bertinetto & Ciucci 2012: 99)

As in the other Chaco languages, the subordinatopr *uje* cannot be involved in lexical nominalizations.

3.7 Tapiete (Tupi-Guarani)

Tapiete constructions with desiderative predicates such as 'want', some cognition predicates such as 'believe', and phasal predicates such as 'start'/ 'finish' lack complementizers. Both clauses mark their subjects through person-number verbal inflection, as in matrix clauses (González 2005). For example:

(37) *ai-pota [pi-kwa'a shi-ñe'e]*$_{NMLZ}$
1SG-want 2PL-know 1SG.IN-language
'I want you to know my language.' (Ciccone 2014)

Constructions with other matrix verbs such as 'know', 'tell', 'see', and 'listen' are marked with the non-future nominalizer *-wa* (González 2005; Ciccone 2014). In these constructions, the subjects of the dependent clauses marked like the subjects of the matrix clause. For example:

(38) *ai-kwa'a he [arikai-pi koa ténta-ipi yururure-ha-**wa**]*$_{NMLZ}$
1SG.A-know 3.POST ago-LOC DEM town-MO.LOC beg-1PL.EXCL-NMZR
'I know about that time that we used to beg in town.' (Ciccone 2014)

This *-wa* nominalizer is used in lexical nominalization (39), following the pattern common in Amazonian languages. In addition, it can be involved in other nominalizing constructions, such as relativization (40), but it cannot be used in adverbial subordination. So it appears that that *-wa* is a nominalizer *stricto sensu*, which follows typical Amazonian strategies of nominalization.

(39) *mbe'u-**wa***
tell-NMZR
'story' (Ciccone 2014)

(40) *wimba'e [piwéra-pe karu-**wa**]*$_{NMLZ}$ *ha'e shi-kiwi*
man yard-LOC 3:eat-NMZR COP 1.POSS-brother
'The man who is eating in the yard is my brother.' (Ciccone 2014)

3.8 Chacoan Guarani (Tupi-Guarani)

According to Dietrich (2009–10), Chacoan Guarani[7] marks nominalized constructions with argument functions exclusively through parataxis, where linked verbs mark their subjects by verbal inflection, as in the following examples:

(41) *[re-ju re-wäe]*$_{NMLZ}$ *ro-echa*
2SG-come 2SG-arrive 1PL.EXCL-see
'You come in, I see you.' / 'I see that you have arrived.' (Dietrich 2009: 346)

(42) *ai-potä-a [kwa-pe re-ju]*$_{NMLZ}$
1SG-want-NEG DEM-LOC 2SG-come
'I do not want you to come here.' (Dietrich 2009: 346)

7. This language is also known as Ava-Guarani or Chiriguano. The latter designation is considered a derogatory ethnonym by speakers of the language (Dietrich 2009).

Interestingly, Dietrich (2009–10) points out that Chacoan Guarani does not have any suffixing nominalized constructions, a pattern quite unlike what is expected of a Tupi-Guarani language (Dietrich 2009: 346).

3.9 Vilela

In Vilela, nominalization with argument functions is carried out through parataxis, as in (43) and (44). For a more detailed discussion of Vilela, see Golluscio (2015).

(43) *[ko-e-rop]*$_{NMLZ}$ *ʔame-ki*
want-3-NEG see-1SG
'I see that s/he does not want (to do that).' (Lozano 2006: 99)

(44) *[sao-e]*$_{NMLZ}$ *ko-at*
come.in-3 want-3
'S/he wants to come in.' (Lozano 2006: 99)

3.10 Lule

In Lule, as in Vilela, nominalization with argument functions is carried out through parataxis, as in (45):

(45) *kristiano takmoks-in-t [lomoe-p alapea Dios si-ton]*$_{NMLZ}$
Christian believe-FUT-3SG be.alone-3SG one God exist-NMZR
'The Christian will believe in the existence of one God only.' (Lit.: The Christian will believe (that) only one God (is the) exister/the one that exists).
(Zamponi & de Reuse 2012)

Lule also has a nominalizer *-to* (sometimes *-ton*). An example is (46):

(46) *[etsi-**ton**]*$_{NMLZ}$ *tso=ta Dios metika-n-t*
be.good-NMZR heaven=LOC God take-FUT-3SG
'God will take the one who is good to heaven.' (Zamponi & de Reuse 2012)

Example (46) is a case where the nominalizer *-to(n)* is not clearly agentive. In general, this nominalizer is clearly agentive and is particularly common in relative clause formation, as can be seen in *si-ton* in example (45) above.[8] No examples of the nominalizer *-to(n)* nominalizing clauses with argument function have been found in our Lule corpus.

8. An account of Lule relative clause formation can be found in de Reuse (2015).

3.11 Summary

Table 2 summarizes the strategies of the Chaco languages for (a) expressing the subject of the dependent clauses and (b) marking the dependency relationship.

Table 2. Synthesis

Family	Language	Expression of the subject of the dependent clause	Parataxis	Subordinator	Nominalizer
Mataguayan	Wichi	Verbal inflection	–	+	–
	Nivacle	Verbal inflection	–	+	–
Guaycuruan	Toba	Verbal inflection	+	+	–
	Pilaga	Verbal inflection/ Possessive marker	+	+	+
Zamucoan	Ayoreo	Verbal inflection	+	+	–
	Chamacoco	Verbal inflection	–	+	–
Tupi-Guarani	Tapiete	Verbal inflection	+	–	+
	Chacoan Guarani	Verbal inflection	+	–	–
Lule-Vilela (affiliation under discussion)	Vilela	Verbal inflection	+	–	–
	Lule	Verbal inflection	+	–	–

4. Conclusion

The nominalized constructions with argument functions in the Chaco languages considered here are characterized by: (a) expression of the subject via verbal inflection, that is, through the same strategies used in independent clauses and (b) the use of a subordinator (not a nominalizer *stricto sensu* or parataxis). We thus recognize a difference between a subordinator and a nominalizer. A subordinator is not involved in lexical nominalization and introduces other dependency relationships that are unrelated to lexical nominalization, such as adverbial constructions, without case markers or adpositions. A nominalizer is involved in lexical nominalization and cannot introduce any other dependency relationship.

Feature (a) groups the Amazonian and the Chaco languages and separates them from the Andean languages which mark the subject of the dependent elements through nominals (possessive) inflection (van Gijn 2014). Feature (b) groups the Chaco languages together and distinguishes them from both Andean and Amazonian languages, as Table 3 shows.

Table 3. Comparison between Andean, Amazonian and Chaco languages

Area	(a) Marking of the subject in the dependent clause via:	(b) Means to indicate the dependency relationship:
Andes	Nominal inflection	Nominalizer
Amazon	Verbal inflection	Nominalizer
Chaco	Verbal inflection	Subordinator, parataxis, or both

Thus, while nominalization is explicitly marked in Andean and Amazonian languages, a more general dependency marker is typically used in the Chaco languages. Chaco nominalization is accomplished without a specific formal marker. Rather, the nominalization effect is seen from the function that the dependent element fulfills in the sentence. In sum, we distinguish Chaco language subordinators from Andean and Amazonian language nominalizers based on two features: (a) unlike nominalizers in Andean and Amazonian languages, Chaco language subordinators are not involved in lexical nominalization; and (b) Chaco subordinators can give rise to adverbial subordination like grammatical nominalizations in other languages, but without the need for a case marker or an adposition.

Finally, we should note the exceptions to the pattern mentioned in two Chaco languages: Tapiete and Pilaga. Tapiete follows the Amazonian pattern, possibly because the speakers of this variety of the language only recently settled in the Chaco. Finally, Pilaga, a Guaycuruan language which has one nominalizer of the type expected of Andean or Amazonian languages next to typically Chaco constructions, presents an interesting challenge for contact linguistics.

Abbreviations

1	first person	IN	inactive
2	second person	INS	instrumental
3	third person	IRR	irealis
A	agent	IRREG	irregular
AC	active	LOC	locative
ACC	accusative;	M	masculine
APPL	applicative	MID	middle voice
ASP	aspect	MO	motion
CLF	classifier	NEG	negative
CON	conative	NMLZ	nominalizer
CTR	control	OBJ	object
CVB	converb	PL	plural
CONJ	conjunction	POSS	possessive

COP	copula
DAT	dative
DEM	demonstrative
DES	desiderative
DET	determiner
DIR	directional
EMPH	emphasis
EXCL	exclusive
F	feminine
FUT	future
IMP	imperative
IMPR	impersonal
IND	indicative
INDF	indefinite
PROSP	prospective
PUNCT	punctual
PURP	purposive
Q	question marker
R	realis
REFL	reflexive
REP	reportative
setA	verbal inflection of the set A (Pilaga)
setB	verbal inflection of the set B (Pilaga)
SG	singular
SBJ	subject
SUB	subordinator
UA	unknown agent.

References

Adelaar, W. with P. Muysken. 2004. *The Languages of the Andes*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511486852

Aikhenvald, Alexandra W. & Dixon, Robert M. W. 1999. *The Amazonian Languages*. Cambridge: CUP.

Bertinetto, Pier Marco. 2009. Ayoreo (Zamuco). A grammatical sketch. *Quaderni del Laboratorio di Linguistica* 8:1–59.

Bertinetto, Pier Marco & Ciucci, Luca. 2012. Parataxis, hypotaxis and para-hypotaxis in the Zamucoan languages. *Linguistic Discovery* 10(1): 89–111. https://doi.org/10.1349/PS1.1537-0852.A.404

Braunstein, José. 2005. Los pueblos indígenas del Gran Chaco. *Mundo de Antes* 4: 127–137.

Campbell, Lyle. 2017. Why is it so hard to define a linguistic area? In *The Cambridge Handbook of Areal Linguistics*, Raymond Hickey (ed.), 19–39. Cambridge: CUP. https://doi.org/10.1017/9781107279872.003

Campbell, Lyle & Grondona, Veronica. 2012. Languages of the Chaco and Southern Cone. In *The Indigenous Languages of South America. A Comprehensive Guide*, Lyle Campbell & Grondona, Veronica (eds), 625–667. Berlin: De Gruyter. https://doi.org/10.1515/9783110258035.625

Carpio, Maria Belén. 2009. Accesibilidad referencial en toba (flia. guaycurú, Argentina). *Revista Estudios en Ciencias Humanas. Estudios y Monografías de los Postgrados de la Facultad de Humanidades*, 1–11.

Censabella, Marisa. 2002. Descripción funcional de un corpus en lengua toba (familia Guaycurú, Argentina). Sistema fonológico, clases sintácticas y derivación. Aspectos de sincronía dinámica. PhD dissertation, Universidad Nacional de Córdoba.

Cerrón-Palomino, Rodolfo. 2006. *El chipaya o la lengua de los hombres del agua*. Lima: Fondo Editorial PUCP.

Cerrón-Palomino, Rodolfo. 2008. *Quechumara: Estructuras paralelas de las lenguas quechua y aymara*. La Paz: PROEIB, Andes.

Ciccone, Florencia. 2014. Grammatical nominalization and complementation strategies in Tapiete (Tupi-Guarani). International Workshop "Nominalization in the Languages of the Americas", Humanities Department, Pontificia Universidad Católica del Perú (Lima, Peru), 28–30 August.

Ciucci, Luca. 2013. Inflectional Morphology in the Zamucoan Languages. PhD dissertation, Scuola Normale Superiore, Pisa.

Comrie, Bernard, Golluscio, Lucía Á., González, Hebe & Vidal, Alejandra. 2010. El Chaco como área lingüística. In *Estudios de lenguas amerindias*, Zarina Estrada & Ramón Arzápalo (eds), 84–130. Hermosillo: Editorial Unison.

Comrie, Bernard & Thompson, Sandra A. 1985. Lexical nominalization. In *Language Typology and Syntactic Description, Vol. III: Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 334–381. Cambridge: CUP.

de Reuse, Willem. 2015. The philological interpretation of small corpora in extinct languages, illustrated by relative clause data from Lule (Argentine Chaco). In *Memorias del VII Congreso de Idiomas Indígenas de Latinoamérica*, 29–31 October, 2015, University of Texas at Austin.

Dietrich, Wolf. 2009–10. Sintaxis del guaraní chaqueño (chiriguano, tupí-guaraní). In *Les langues du Chaco. Structure de la phrase simple et de la phrase complexe*, Lucía Á. Golluscio & Alejandra Vidal (eds). *Amerindia* 33–34: 333–363.

Domínguez, Marcelo, Golluscio, Lucía Á. & Gutiérrez, Analía. 2006. Los vilelas del Chaco: desestructuración cultural, invisibilización y estrategias identitarias. In Historias fragmentadas, identidades y lenguas: Los Pueblos Indígenas del Chaco Argentino, dossier, Lucía Á. Golluscio & Sylvia Hirsch (eds). *Indiana* 23: 199–226.

Dryer, Matthew S. & Haspelmath, Martin (eds). 2013. *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Durante, Santiago Gabriel. 2014. El ayoreo (familia zamuco). *Language Design: Journal of Theoretical and Experimental Linguistics* 16: 5–22.

Fabre, Alain. 2014. *Estudio gramatical de la lengua nivacle*. Ms, Kangasala, Finland.

Franceschini, Dulce. 1999. La langue Sateré-Mawé: Description et analyse morphosyntaxique. PhD dissertation, University of Paris 7.

Givón, Talmy. 2001. *Syntax: An Introduction*. Amsterdam: John Benjamins.

Golluscio, Lucía Á. 2009–10. Los converbos en -(e)l y la combinación de cláusulas en vilela (lule-vilela). In *Les langues du Chaco. Structure de la phrase simple et de la phrase complexe*, Lucía Á. Golluscio & Alejandra Vidal (eds). *Amerindia* 33–34: 250–285.

Golluscio, Lucía Á. 2015. Huellas de trayectorias y contactos en el sistema lingüístico: El caso vilela (Chaco). In *Language Contact and Documentation/Contacto lingüístico y documentación*, Bernard Comrie & Lucía Á. Golluscio (eds), 77–116. Berlin: De Gruyter.

Golluscio, Lucía Á. & González, H. 2008. Contact, attrition and shift in two Chaco languages: The cases of Tapiete and Vilela. In *Lessons from Documented Endangered Languages* [Typological Studies in Language 78], David K. Harrison, David S. Rood & Arienne Dwyer (eds), 195–242. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.78.08gol

Golluscio, Lucía Á., Hasler, Felipe & de Reuse, Willem. 2014. Adverbial subordination at the peripheries of the Andean and Chaco linguistic areas. In SSILA/Linguistic Society of America Annual Meeting, Minneapolis, January.

Golluscio, Lucía Á., Hasler, Felipe & de Reuse, Willem. 2015. Relative clause syntax in the Gran Chaco and the Andes: Overlapping typologies. In conference "Diversity Linguistics: Retrospect and Prospect", Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, 2 May.

González, Hebe. 2005. A grammar of Tapiete (Tupi-Guarani). PhD dissertation, University of Pittsburgh.

González, Hebe. 2015. El Chaco como área lingüística: Una evaluación de los rasgos fonológicos. In *Language Contact and Documentation/Contacto lingüístico y documentación*, Bernard Comrie & Lucía Á. Golluscio (eds), 201–247. Berlin: De Gruyter.

Hasler, Felipe. 2017. Relaciones interclausales en mapudungun desde una perspectiva tipológico-funcional. PhD dissertation, Universidad de Buenos Aires.

Kalisch, Hannes. 2009. Los constituyentes de la cláusula enlhet (enlhet-enenlhet). Esbozo de una cláusula omnipredicativa. In *Les langues du Chaco. Structure de la phrase simple et de la phrase complexe*, Lucía Á. Gollusci & Alejandra Vidal (eds). *Amerindia* 33–34:109–150.

Lehmann, Christian. 1988. Towards a typology of clause linkage. In *Clause Combining in Grammar and Discourse* [Typological Studies in Language 18], John Haiman & Sandra A. Thompson (eds), 181–225. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.18.09leh

de Llamas, Alfredo. 1910. *Uakambabelté o Vilela. Lenguas indígenas aborígenes. Mi contribución primera al estudio de la historia antigua*. Corrientes: Teodoro Heinecke.

Lozano, Elena. 1970. *Textos vilelas*. La Plata: CEILP.

Lozano, Elena. 2006. *Textos vilelas*. Second augmented edition and prologue by Lucía Á. Golluscio. Col. Nuestra América. Archivo de Lenguas Indoamericanas. Buenos Aires: Insituto de Lingüística, Facultad de Filosofia y Letras, Universidad de Buenos Aires.

Machoni, Antonio. 1732. *Arte y vocabulario de la lengua lule y tonocoté*. Madrid: Herederos de Juan García Infanzon.

Malchukov, Andrej. 2006. Constraining nominalization: Function/form competition. *Linguistics* 44(5):973–1009. https://doi.org/10.1515/LING.2006.032

Messineo, Cristina & Cúneo, Paola. 2009–10. *Construcciones seriales en toba. In Les langues du Chaco. Structure de la phrase simple et de la phrase complexe*, Lucía Á. Golluscio & Alejandra Vidal (eds). *Amerindia* 33–34:217–248.

Nercesian, Verónica. 2009–10. Construcciones de verbos seriales en wichi (mataguayo). Características sintácticas y semánticas. In *Les langues du Chaco. Structure de la phrase simple et de la phrase complexe*, Lucía Á. Golluscio & Alejandra Vidal (eds). *Amerindia* 33–34:187–216.

Nercesian, Verónica. 2011. Gramática del wichí, una lengua chaqueña. Interacción fonología-morfología-sintaxis en el léxico. PhD dissertation, Universidad de Buenos Aires.

Nercesian, Verónica. 2014. *Wichi lhomtes. Estudio de la gramática y la interacción fonología-morfología-sintaxis-semántica* [Lincom Studies in Native American Linguistics 74]. Munich: Lincom.

Noonan, Michael. 1985. Complementation. In *Language Typology and Syntactic Description, Vol. 2: Complex Constructions*, Timothy Shopen (ed.), 52–150. Cambridge: CUP.

Shibatani, M. 2009. Elements of complex structures, where recursion isn't. In *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution* [Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 163–198. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.07ele

Silva, Raynice. 2010. Estudo morfossintático da lingua Sateré-Mawé. PhD dissertation, Universidade Estadual de Campinas.

Smeets, Ineke. 2008. *A Grammar of Mapuche*. Berlin: De Gruyter. https://doi.org/10.1515/9783110211795

Soto Ruiz, Clodoaldo. 1976. *Gramática Quechua Ayacucho-Chanca*. Lima: Ministerio de Educación del Perú.

Torero, Alfredo. 2002. *Idiomas de los Andes: Lingüística e historia*. Lima: IFEA, Instituto Francés de Estudios Andinos.

Unruh, Ernesto, Kalisch, Hannes & Romero, Manolo. 2003. *Enenlhet Apaivoma. Nentengiai'a nengiangvaiakmoho neliateikamaha enenlhet apaivoma. Guía para el aprendizaje del idioma materno toba*. Biblioteca Paraguaya de Antropología, Ya'alve- Saanga: Nengvaanemkeskama Nempayvaam Enlhet.

van Gijn, Rik. 2014. Subordination strategies in South America: Nominalization. In *The Native Languages of South America: Origins, Development, Typology*, Loretta O'Connor & Pieter Muysken (eds), 275–296. Cambridge: CUP. https://doi.org/10.1017/CBO9781107360105.015

van Gijn, Rik, Golluscio, Lucía Á., González, Hebe & Vidal, Alejandra. 2013. Adverbial subordination strategies in the Chaco and beyond. In SSILA/Linguistic Society of America Annual Meeting, Boston.

van Gijn, Rik, Haude, Katharina & Muysken, Pieter (eds). 2011. *Subordination in South America* [Typological Studies in Language 97]. Amsterdam: John Benjamins.

Vidal, Alejandra. 2002. Pilagá grammar (Guaykuruan family, Argentina). PhD dissertation, University of Oregon.

Zamponi, Raoul & de Reuse, Willem. 2012. Lule Corpus. Work in progress.

Zariquiey, Roberto & Córdova, Gavina. 2008. *Qayna, kunan, pagarin. Introducción práctica al quechua chanca*. Lima: Pontificia Universidad Católica del Perú.

PART III

Case studies on nominalization in individual languages

CHAPTER 6

Nominalization in Central Alaskan Yup'ik

Yuki-Shige Tamura
Osaka University

This article analyzes related lexical and grammatical structures in Central Alaskan Yup'ik (CAY) in terms of the theory of nominalization (Shibatani 2018, this volume) focusing on the relationships between noun formation and grammatical structures paralleling so-called relative clauses in other languages. We first examine the characteristics of nominalizations lexicalized as nouns, showing that various types of nominalizers are employed in the formation of nouns, and then observe that essentially the same formal structure is utilized in relative-clause counterparts in CAY, with the use of some of the nominalizers seen in lexicalized forms (Jacobson 1995). We demonstrate that the differences between derived nouns and relative-clause counterparts are attributable to the different instantiations of the single process of nominalization–lexical nominalization and grammatical nominalization.

1. Introduction

A series of studies by Shibatani (2009, 2014, 2018, this volume) attempts to provide a unified theory for a broad range of grammatical phenomena that may have been described separately or in diverse arrangements in descriptive grammars and linguistic theories. One major point of Shibatani's theory is that a single process of nominalization underlies both (i) lexical noun formation including those that may be traditionally called event nouns such as *employ-ment*, participant nouns such as *employ-er*, as well as modifying nominals (i.e. genitive) such as *my/mine* and *employer-'s*, and (ii) grammatical structures including those that are traditionally labeled as relative and complement clause. While these grammatical phenomena may be structurally realized in various ways in languages of the world, the general theory of nominalization suggests that their formal variations should be regarded as being derived from a cognitive-linguistic function of nominalization, namely in terms of how languages represent the underlying unity of various types of nominalization, on the one hand, and, on the other, how they respond to the differences in the types and uses of nominalization structures, which all share the inherent

https://doi.org/10.1075/tsl.124.06tam

properties of ordinary nouns–entity-denotation and the referential and modification functions.

Shibatani (2018) argues that unless a fundamental functional notion is maintained in handling grammatical variations, grammar descriptions would be distorted by the characteristics of structural coding which happen to be observed in some specific languages (Shibatani 2018: 347–351; see also Croft 2003: Chapter 8). A danger is more acute when those specific characteristics are represented in dominant languages such as English and some other European languages. For example, although the terms such as "relative clause" and "relative pronoun" are widely used in describing a variety of languages across the globe, Comrie (1997: 77, 2006: 136) points out that the pronoun strategy in the formation of relative clauses is limited almost entirely to European languages. DeLancey (2002), observing the characteristics of relative-clause and complement-clause counterparts in Tibeto-Burman languages, comments as follows: "Early investigators of Tibeto-Burman languages seem to have had a sense of the centrality of nominalization to their syntactic organization, but were never really able to come to grips with its resolutely non-Indo-European manifestations" (ibid., 55–56). These two observations indicate that an accepted view in linguistics, that of avoiding "fitting languages of the world into the mold of 'Standard Average European'" (Croft 2001: 54), is still difficult in practice.

This article is aimed at analyzing nominalization phenomena in Central Alaskan Yup'ik (Eskimo-Aleut, hereafter CAY; see Jacobson (1995) and Miyaoka (2012) for comprehensive grammar descriptions), paying close attention to noun formation, the relative-clause counterparts, and the relationships between them. We suggest that Shibatani's reanalysis of so-called relative clauses in terms of his theory of nominalization provides a simple answer for a difficult question in CAY grammar: why is it that the relative-clause counterparts in CAY are marked by "nominalizers" in the first place, if relative clauses are assumed to be clausal verbal structures distinct from nominalizations (Dixon 2010: 313). Let us have a brief look at the following pair of examples to clarify the point:[1]

1. All the examples in this paper, except those for which the source is given, are attested by Caan Toopetlook (a speaker of CAY from Nunapicuaq, a village along the Kuskokwim River), and the semantic definitions of lexical words and derivational morphemes rely on Jacobson (1984). In addition, for the CAY data, the standard Yup'ik orthography (Jacobson 1984; Miyaoka 2012) is employed with minor adjustments for morphological boundaries if necessary. In particular, when the phonological changes in morphological boundaries blurs the morphemes we focus on, the forms after the phonological change (i.e. the authorized spellings) are put in parentheses. The orthographic symbols that differ significantly from those of the IPA are as follows: vv = [f]; ll = [ɬ]; ss = [s]; g = [ɣ]; gg = [x]; r = [ʁ]; rr = [χ] c = [č]; ng = [ŋ]; y = [j]; e = [ə]; and ' = gemination (Woodbury 2002: 80).

(1) a. *Tegu-ste-put ner-'u-q nervig-mi yaani.*
take-NMZR-1PL.3SG.ABS eat-INTR.IND-3SG table-LOC there
'Our policeman is eating at the table over there.'

b. *[Iqlu-ste-put mikelnguq-ø]$_{NP}$ ner-'u-q nervig-mi yaani.*
lie-NMZR-1PL.3SG.ABS child-3SG.ABS eat-INTR.IND-3SG table-LOC there
'The child who lied to us is eating at the table over there.'

c. *[Tegu-ste-put mikelnguq-ø]$_{NP}$ ner-'u-q nervig-mi yaani.*
take-NMZR-1PL.3SG.ABS child-3SG.ABS eat-INTR.IND-3SG table-LOC there
*(i) 'The child who took us somewhere/took our hands is eating at the table over there.'
(ii) 'Our child police is eating at the table over there.'

In (1a) and (1b), *tegu-ste-put* and *iqlu-ste-put* show essentially the same structure, both having the derivational suffix *-ste* and the nominal inflectional morpheme *-put*. However, as shown in the contrast between their English translations, *tegu-ste(-put)* 'take-NMZR-(1PL.3SG.ABS)' in (1a) is construed as being "lexicalized" (Jacobson 1995: 38), namely a noun that indicates a sense of "our policeman", rather than "someone who is taking us somewhere/taking our hands." On the other hand, *iqlu-ste-put* 'lie-NMZR-1PL.3SG.ABS' in (1b) may be regarded as a relative-clause counterpart. This is because it may be construed as indicating a single event (or a couple of events) of having lied to the people including the speaker rather than a habitual sense of liar, and it also serves as a modifier for the noun, *mikelnguq*. Example (1c) indicates that simple juxtaposition of *tegu-ste-put* with *mikelnguq* would not lead to a relative-clause interpretation as shown in the contrast between the two English translations. A major question pursued in this article is: What grammatical factors motivate such differences in interpretation between (1a), (1b) and (1c), all of which involve the nominalizer *-ste*? While cases like the above, involving a lexical process of deriving nouns and a grammatical structure modifying a noun, have long been analyzed as two distinct phenomena, typically being described separately in two different chapters in descriptive grammars, such a treatment misses an important fact that similar morphology is involved in both. The general theory of nominalization advanced by Shibatani (2009, 2014, 2018, this volume) allows us to treat the similarity and difference between (1a) and (1b) as two variants of the single process of nominalization; namely, lexical nominalization and grammatical nominalization.

This paper is organized as follows: we first outline the general characteristics of CAY lexical nominalization in Section 2, clarifying the way CAY nominalizers

contribute to noun formation. Then, in Section 3, after pointing out that the CAY nominalizers may be grouped into six major types with regard to the type of nominal entities denoted, we will observe two remarkable characteristics of CAY nominalized forms: (i) ambiguity between the productive and non-productive use of nominalized forms (Jacobson 1984, 1995: 38; Fortescue et al. 2010), and (ii) the incorporation of temporality. With these observations, we finally examine, in Section 4, modification structures construable as relative clauses, and describe the structural continuity between derived nouns and relative-clause counterparts including so-called participles. We suggest that the interrelationship between them is due to the functions of nominalization, which extend to so-called relative clauses in other languages, contrary to the view expressed by Dixon (2010) that: "In some languages, a verbal affix marking a relative clause is homophonous with a nominalizer. This should not be taken to mean that a relative clause is a type of nominalization" (p. 316).

Before moving on to the next section, however, let us take a brief look at CAY grammar with a focus on a particular characteristic of its inflectional morphology, which displays the property of double-marking (Nichols 1986; Woodbury 2002):

(2) *Angute-m nunur-a-a qimugta-ø.*
man-3SG.ERG scold-TR.IND-3SG.3SG dog-3SG.ABS
'The man is scolding the dog.'

(3) *Qimugta-ø qilug-tu-q.*
dog-3SG.ABS bark-INTR.IND-3SG
'The dog is barking.'

As shown in the contrast between (2) and (3) above, not only does CAY exhibit an ergative-absolute system on nominal inflection, but it also shows the morphological characteristic of the double-marking pattern (Nichols 1986: 85). In addition to case marking on the nouns, the verb inflections *-a-a* and *-tu-q* code the grammatical information on verb valency as to whether the verb is transitive or intransitive, and on the person and number of its arguments.

This double-marking property is also realized in the possessive construction shown in (4) below, in which the pronominal morpheme of *-a* in *naca-a* indicates that both the possessed thing and the possessor are third person singular.

(4) *angute-m naca-a*
man-3SG.ERG hat-3SG.3SG.ABS
'the man's hat'

Together with this double-marking pattern, the reader may have noticed that in CAY, the transitive construction as in (2) and the possessive construction as in (4)

show a parallel structure in terms of morphological marking: the dependent markings for the transitive subject and that for the genitive not only share morphology (*-m* here), but the head marking for the transitive verb and that for the possessed noun also use essentially the same pronominal suffixes (*-a* here). In sum, one important characteristic of CAY grammar is that the nominal structure and clause structure share the same structural template, employing the same case markers for the transitive subject and the genitive, and essentially the same pronominal suffixes for transitive verbs and possessed nouns, both showing the property of double morphological marking.[2] Other grammatical characteristics of CAY are introduced when necessary below.[3]

2. Lexical nominalizations

While the inflectional morphology in CAY may have a portmanteau characteristic as observed in the examples above, the derivational morphology exhibits a highly agglutinative nature (see Woodbury 2002: 80 for more information). The morphological segmentation of most words is obvious from the surface form. For example, *teacher* in English is expressed as *elitnauriste-*. This word, though lexicalized, can be divided into four morphemes: *elite-* 'to learn' *-naur-* 'habitually,' *-i-* (ANTIP) *-(s)te* (NMZR.ARG), which may be literally translated as '(a) person who habitually causes (someone) to learn.'[4] In CAY word formation, a root morpheme (called the *base* in Eskimo linguistics (Reed et al. 1977; Fortescue et al. 2010)) like *elite-* occurs first, followed by various types of derivational suffixes like *-naur-*, *-i-*, and *-(s)te* here (called *postbase*), and the word ends with the inflectional morphemes

2. When we suggest that CAY shows a parallel structure between nominal structure and verbal/clausal structure, one may wonder about the valency part of the verbal inflection, which seems to have no counterpart in the nominal structure. In fact, as shown in Bergsland (1997: 353) and Johns (1992: 82) and as confirmed in a comparative dictionary between Eskimo languages (Fortescue et al. 2010: 438), the verb valency morphemes (the verbal *mood* markers in Eskimo grammar) have developed from nominalizers. In other words, a transitive clause such as (2) is likely to have developed from a nominalization of the form such as "[the man's scolding] [a dog]]" (cf. Thalbitzer (1991: 1054)).

3. See Woodbury (2002) for a typological characterization of CAY word structure and grammar. And see also de Reuse (2009) for a proposal on the fundamental principles of CAY morphology.

4. While the morpheme *-i-* in *elit-naur-i-ste-* is regarded as an antipassive marker in general, a more precise literal translation would be "a person who makes habitual learning (to someone's disadvantage)", because the morpheme *-i-* is etymologically derived from *-aq* (a nominalizer, see (15c) below) and *-li* 'make,' and it further implies that what is made causes suffering for its recipient. See Jacobson (1984: 453) for the etymological account of this morpheme.

(called *ending*) that indicate person and number in addition to case, as outlined in the previous section. Of course, many derived nouns have no derivational suffixes discernable between the root and inflectional suffixes, which may be exemplified in *pani-ka* 'my daughter (daughter-1SG.3SG.ABS),' *ii-gka* 'my eyes (eye-1SG.3DU.ABS), *nuna-put* 'our land/village (land/village-1PL.3SG(/3PL).ABS), *napa-t* 'trees' (tree-3PL.ABS)'.

One remarkable characteristic of CAY nouns, however, is that even very common ones such as 'man', 'child', and 'dog' have been formed with nominalizing suffixes. Consider the examples in (5), taken from diverse semantic fields, in which nominalizers are attached to the root to create nouns. (The semantic properties after NMZR (nominalizer) will be accounted for in Section 3 below, and we use the noun form of absolutive, 3rd-person-singular unpossessed form of the noun, which utilizes the zero-coded form for the inflection.)

(5)	a.	*angu-n*	catch-NMZR.INST	'man'
	b.	*mikel-nguq*	be.small-NMZR.ARG	'child'
	c.	*una-n*	work.on (without.tools)-NMZR.INST	'hand'
	d.	*qimug-ta*	pull-NMZR.ARG	'dog'
	e.	*yaqu-lek*	wing-NMZR.ASSOC	'bird'
	f.	*aker-ta*	shine-NMZR.ARG	'sun'
	g.	*ella-lluk*	outside-NMZR.QUAL('bad')	'rain'
	h.	*aqum-lleq*	sit-NMZR.ARG	'chair'
	i.	*qerar-yaraq*	cross.over-NMZR.MAN	'bridge'

Of around 440 nouns in the CAY vocabulary list in Jacobson (1995: 485–506), some 285 nouns show the nominalizing morphology observed in (5) (around 65 percent of the nouns in the list). That is, we can tell even from a synchronic point of view that nominalization plays an important role in CAY word formation. From this word count, we exclude the following cases: first, the examples in which the roots can be realized by themselves together with a suitable inflection either as a noun or as a verb such as *mer-* 'water/to drink' and *nuteg-* 'gun/to shoot'; second, the nouns in which the attachment of nominalizers hardly changes the meaning expressed by root morphemes such as *nuli-aq* 'wife: wife-NMZR, or those in which the meanings of the roots are identified together with the attachment of the nominalizers such as *eneq* 'bone,' though we may speculate that these nouns bear traces of nominalization.[5] Finally, loan words are excluded such as *caay-uq* 'tea' (from Russian *chay*), *estuuul-uq* 'table' (from Russian *stol*), *John-aaq*/*Caan-aaq* 'John,' *caaska-q* 'cup' from Russian *cháshka,* though the sounds used for phonological

5. See Jacobson (1984: 664–667) for the etymological classifications of root morphemes.

adjustment are shared with those of certain nominalizers.[6] If we include these in our count, accepting such factors as zero-derivation, fossilized/fused nominalizers and phonological adjustments, the ratio would be much greater than 65 percent; more than 85 percent of the nouns in the vocabulary list are regarded as having been created with the aid of nominalizers.

Along with this characteristic is the fact that CAY has no compounding word formation that employs two (or more) independent root morphemes, a strategy which is observed in *sewing machine* or *bedroom* in English (Jacobson 1995: 33; Woodbury 2002: 80). CAY speakers express the former concept as *mingqe-ssuun* 'sew-NMZR.INST (lit. the instrument of sewing)' and the latter as *qavar-vik* 'sleep-NMZR.LOC (lit. somewhere to sleep in),' utilizing an instrumental nominalizer, *-ssuun* and a locative nominalizer, *-vik*, respectively.

These two characteristics observed here show that CAY speakers have coined nouns utilizing a relatively small number of root morphemes together with a relatively large number of nominalizing suffixes.[7] The following examples may help enhance our understanding of how crucial nominalizers are to CAY noun formation:

(6) *kenir-* 'to make a fire':

a.	*kenir-vik*	making.a.fire-NMZR.LOC	'kitchen'
b.	*kenir-gaq*	making.a.fire-NMZR.ARG	'cooked food'
c.	*kenir-cuun*	making.a.fire-NMZR.INST	'stove'
d.	*kenir-ta*	making.a.fire-NMZR.ARG	'cook'
e.	*keni(r)-lleq*	making.a.fire-NMZR.ARG	'cooker'
	(made from metal can used outside in summer)		

(7) *agayu-* 'to pray'

a.	*agayu-vik*	pray-NMZR.LOC	'church'
b.	*agayu-n*	pray-NMZR.INST	'God'
c.	*agayu-neq*	pray-NMZR.EVENT	'Sunday'
d.	*agayu-lir-ta*	pray-give-NMZR.ARG	'priest'
e.	*agayu-ma-lria*	pray-ASP-NMZR.ARG	'Christian'
f.	*agayu-ssuun*	pray-NMZR.INST	'hymn'

In (6), the nouns that roughly correspond to English *kitchen*, *cooked food*, *stove*, *cook*, and *cooker* are created as a form derived from the root concept of making a fire. Likewise, in (7), the concept of praying expressed by the root serves as a reference point to form nouns with respect to religious activity.

---

6. See Jacobson (1984: 678–680) for CAY loan words.

7. As pointed out by Woodbury (2002), CAY grammar shows virtually no morphological process but suffixation; "there is no compounding, no prefixing, infixing, ablaut, consonant mutation, reduplication, or morphological use of accentual change…" (p. 80).

Shibatani (2018: 345–346, this volume) shows that a cognitive process of metonymy underlies the process of nominalization, circumscribing the possible derivations of new forms via nominalization. Examples in (6) and (7) clearly show that the denotation evoked by each nominalized form has been built on the root concept through a metonymic process, as evidenced by the fact that all the denotations involved are those having inherent relationships or are typically associated with the root concept. We can suggest that the diverse types of nominalizer serve to narrow down the potential denotation/referent of a nominalized form. The same account can be provided for examples (5), in which each sense is created from the root concepts with a metonymic construal, being further specified by the type of the nominalizer employed. While this mechanism of noun formation itself is not uncommon crosslinguistically, what makes CAY noun formation stand out is that nominalization morphology is applied prolifically, even to the formation of basic level words that are coded as single roots in other languages such as 'man', 'dog', and 'rain', as in (5) to (7).

3. Types and characteristics of CAY nominalizers

Recognizing that nominalization plays an important role in CAY noun formation both in terms of quantity and quality, we now move on to the issues pertaining to the types of CAY nominalizers and to those remarkable characteristics of the derived forms in the language. Jacobson (1995) provides a list of derivational suffixes in his CAY grammar (pp. 567–569), in which 43 are regarded as nominalizers in the list of 173 derivational suffixes. Judging from the types of denotation indicated by the nominalizations on the basis of the root/stem concepts, we can classify these 43 nominalizers into the following six major types, though some nominalizers are cross-categorized.[8]

A. Instrumental Nominalizers (INST): e.g. *-n*, *-cuun*(/*ssuun*) 'device for V-ing'

(8) a. *ega-n* cook-NMZR.INST 'pot'
b. *cali-ssuun* work-NMZR.INST 'tool'
(see also (5a), (6c), (7b) and (7f) above)

B. Place Nominalizer (LOC): e.g. *-vik* 'place to V'

(9) *maqi-vik* take.a.steambath-NMZR.LOC 'steam bath house'
(see also (6a) and (7a) above, and also see (1); ner-vik eat-LOC 'table')

8. The order from (A) to (F) is just for the sake of account, the exemplified nominalizers should not be taken as an exhaustive list, the definitions are from Jacobson's (1984) CAY dictionary with some modifications because of the space limitation, and the examples are all in the absolutive, third-person singular, unpossessed form.

C. Manner Nominalizers (MAN): e.g. *-yaraq*, 'way of V-ing,'

(10) *qaner-yaraq* speak-NMZR.MAN 'language, manner of speaking, word'
(see also (5i) above)

D. Noun-Based Nominalizers:

(D-1) Association (ASSOC): e.g. *-lek* 'one with N;' *-miu* 'inhabitant of N;' among others.

(11) a. *umyua-lek* mind/thought-NMZR.ASSOC 'smart person'
b. *cénar-miu* shore-NMZR.ASSOC 'shore dweller'
(see also (5e) above)

(D-2) Size (SIZE): e.g. *-rpak* 'big N;' -cuaq 'small N;' *-yagaq* baby N;' among others.

(12) a. *nuna-rpak* land-NMZR.SIZE 'city'
b. *nuna-pi(k)-cuaq* land-NMZR.QUAL-NMZR.SIZE 'small village'
c. *tuntu-yagaq* caribou-NMZR.SIZE 'baby caribou'

(D-3) Quality (QUAL): e.g. *-pik* 'real, authentic;' *-lluk* 'bad N;' among others.

(13) a. *nuna-pik* land-NMZR.QUAL 'tundra'
b. *ii-lluk* eye-NMZR.QUAL 'bad eye' (see also (5 g) above)

(D-4) Time (TIME): *-kaq*$_1$ 'future N;'*-lleq*$_1$ 'former N;' among others.

(14) a. *neq-kaq* food-NMZR.TIME 'food (ready for consumption)'
b. *angya-lleq* boat-NMZR.TIME 'something which was a boat, former boat'

E. Argument Nominalizers (ARG): e.g. *-(s)ta* 'a V-er;' *-lria* 'the one who is V-ing;' *-(g)aq* 'that which has been V-ed;' *-nguq* 'one that is V-ing;' *-neq*$_1$ 'that which results from V-ing;' *-(ar)kaq*$_2$ 'that which will V;' *-lleq*$_2$ 'one that V-ed or wasV-ed,' among others.

(15) a. *cali-sta* work-NMZR.ARG 'worker'
b. *qava-lria* sleep-NMZR.ARG 'the one sleeping'
c. *ega-(g)aq* cook-NMZR.ARG 'boiled fish/food'
d. *uki-neq* make.a.hole-NMZR.ARG 'hole'
e. *tupig-arkaq* weave-NMZR.ARG 'something to be woven (e.g. grass)'
(see also (5b), (5d), (5f), (5h), (6b), (6d), (6e), (7d), and (7e) above)

F. Event Nominalizers (EVENT): e.g. *-neq*$_2$ 'process of V-ing;' *-ciq* 'condition of possessor with respect to V;' *-lleq*$_2$ 'state of V-ing,' among others.

(16) a. *quq-neq* split-NMZR.EVENT 'process of splitting'
b. *assi-ucia* good-NMZR.EVENT(*ciq*.3S.3S.ABS) 'whether it is good'
c. *aya-lleq* leave-NMZR.EVENT/ARG 'the act of leaving/ the one who left'
(see also (7c) above)

As for (D), note that CAY has no formal category of adjective (Jacobson 1995: 253), and property-concept roots are followed with nominal inflection, just like those of activity/event concepts. And for (E), all the nouns created with the nominalizers are formed on the absolutive basis, such that the derived forms stand for S (an intransitive subject 'one who is V_{intr}-ing') and P (a patient of a transitive event 'that which has been V_{tr}-ed'). Consider the way the words *teacher* and *student* are formed with nominalizers: (i) *elite-naur-aq* 'student; learn-habitually-NMZR.ARG' vs. (ii) *elite-naur-i-sta* 'teacher; learn-habitually-ANTI-NMZR.ARG.' As shown by the contrast, a voice-alternation marker, *-i-* here, is required in the verbal stem when the referent is agentive to the verbal root concept.

While the CAY nominalizers may be grouped into these six major types, this classification should be understood together with the following two interrelated properties that cross the semantic distinctions of (A) to (F): (i) the ambiguity between the productive and non-productive use of nominalizers and nominalized forms (Jacobson 1984, 1995: 38; Fortescue et al. 2010) and (ii) the incorporation of temporality.

First, let us observe that many, if not all, nominalizers not only can be recognized in "lexemes," but they can also be employed in an "innovative" way, while both remain nouns morphologically. For instance, compare the difference between *angu-n*, hunt-NMZR.INST 'man' (= 5a) and *ega-n*, cook-NMZR.INST 'pot' (= 8a). While the same instrumental nominalizer from (A) is employed for (5a) and (8a), the former is used exclusively to refer to a man (male person); *angu-n* never refers to a hunting woman or a hunting animal, or hunting instruments such as a snowmobile or rifle even if it can lead to successful hunting and it matches the meaning of the root 'to hunt' and the nominalizer 'an instrument' (John Toopetlook, p.c.; we use the term *lexicalized* for this type of nominalization having a conventionalized, specific referent). On the other hand, while the latter, *ega-n*, refers to a pot in a lexicalized sense, it can also be utilized by a speaker to create a noun for the nonce, referring to any container such as a kettle (*caanik*) and an empty can (*paankaq*) as long as it serves for cooking on the spot (John Toopetlook, p.c.; we use the term *non-lexicalized* for this type of nominalization having a wider range of referent or showing an innovative use of the nominalized forms). Furthermore, *yaku-lek*, wing-NMZR.ASSOC 'bird,' (5e) is so lexicalized that the nominalized noun could be only used to refer to a bird. On the other hand, CAY speakers can coin a noun for the nonce using this associative nominalizer, in order to refer to, for instance, a younger sibling who cannot stand to be away from his/her big sister: *alqa(q)-lek*, older.sister-NMZR.ASSOC 'sister's boy/girl' (John Toopetlook, p.c.). In fact, in Jacobson's dictionary (Jacobson 1984), *yaku-lek* 'bird' is cited as a lexicalized word in the entries, but *alqa(q)-lek* 'sister's boy/girl' is not listed at all, and *umyua-lek*, mind/thought-NMZR.ASSOC 'smart person' shown in (11a) is cited, but it is not

checked as a lexicalized word. What is interesting in CAY nouns is that, like *ega-n* above, many nominalized forms can show ambiguity between the lexicalized and non-lexicalized uses; for instance, *akut-aq* mix-NMZR.ARG 'Eskimo ice cream, anything that is mixed,' *kingu-qliq* be.rear-NMZR.ARG 'younger sibling, any rear part of something,' mikel-nguq be.small-NMZR.ARG 'child, any small thing,' *aka-lria* roll-NMZR.ARG 'car, wheel, anything that is rolling (see (19) and (20) below).[9]

Related to this property, certain nominalizers, as typically observed in nominalizers of (E) and (F) above, incorporate temporal/aspectual notions in themselves. For instance, when the verb, *tupig-* 'weave,' is nominalized with *-(g)aq* 'that which has been V-ed' in (E) as *tupig-a(q)-t* (plural), it indicates a woven thing like a grass mat, denoting a thing in the resultative state of weaving. On the other hand, when *-(ar)kaq*$_2$ 'that which will V' is suffixed to the verb, like *tupig-kaq*, the derived noun denotes grass or something that serves as the material used for weaving, i.e., a thing in the preparatory state of weaving. Let us now compare the difference between *cali-lria*; 'the one working; work-NMZR.ARG' and *cali-sta* '(a) worker; work-NMZR.ARG' (15a). The two nominalizers are both from (E) and both examples can mean '(a) worker,' but the former tends to be construed as referring to one in a temporal state of working; in contrast, the latter tends to be employed for an atemporal description, focusing on the generic state in which the referent is. If we replace *-lria* of *qava-lria* 'the one sleeping; sleep-NMZR.ARG' in (15b) with *-(s)ta*, the new word *qavar-ta* will refer to someone who likes to sleep a lot, highlighting a generic property of the referent.

One interesting consequence of these two general properties of the productivity and temporality observed in CAY nominalizers is that English relative clauses can be rendered in the form of simple nominalized nouns, as in the following pairs of examples:

(17) a. *Angute-m ner-a-a neqa-ø.*
man-3SG.ERG eat-TR.IND-3SG.3SG fish-3SG.ABS
'The man is eating the fish/food.'

b. *Angute-m ner-a-a ega-aq-ø* (=(*15c*)).
man-3SG.ERG eat-TR.IND-3SG.3SG cook-NMZR.ARG-3SG.ABS
'The man is eating the fish/food that has just been boiled/eating the boiled fish/food.'

(18) a. *Caan-aam qavar-vi(k)-i(a)-nek taig-u-q.*
John-3SG.ERG sleep-NMZR.LOC-3SG.3SG-ABL.POSS come-INTR.IND-3SG
'He came from John's bedroom.'

9. The discussions on the formal unity between lexicalized nominalization and nonce nominalization in Matses by Fleck (this volume) would be helpful and insightful to figure out the general characteristics of lexicalized and non-lexicalized nominalized forms.

b. *Caan-aam uita-vi(k)-i(a)-nek taig-u-q.*
John-3SG.ERG stay-NMZR.LOC-3SG.3SG-ABL.POSS come-INTR.IND-3SG
'He came from where John is staying/from John's staying place.'
(John Toopetlook, p.c.; based on Jacobson (1984: 586))

In (17), if one wants to say that the fish served to the man has just been boiled, the nominalized form *ega-aq* will be all that is needed, because the nominalizer *-aq* serves to indicate that the food is in a state of having been boiled, as shown in (17b). On the other hand, English needs some modifier (if not a relative clause) to indicate the temporal state of the referent, as exemplified in the translation of (17b). Likewise, when (18a) and (18b) are compared, one may notice that the two meanings of "from John's bedroom" and "from where John is staying" can be expressed in CAY with the same nominalized structure, i.e. Verb-vi(k)-i(a)-nek.

This property of the coexistence of lexicalized and non-lexicalized use in single nominalized forms may also engender interesting ambiguity as shown in (19) below:

(19) *Angute-m tangrr-a-a aka-lria-ø kanani.*
man-3SG.ERG see-TR.IND-3SG.3SG roll-NMZR.ARG-3SG.ABS down.there
a. 'The man is seeing the car down there.'
b. 'The man is seeing the one rolling/the log that is rolling down there.'

When the noun, *aka-lria-ø* 'roll-NMZR.ARG-3SG.ABS,' is intended in the lexicalized use, it refers to a car, as in interpretation (i); the car may be parked down there. On the other hand, when *aka-lria-ø* is intended in a non-lexicalized sense, it refers to something that is rolling. This ambiguity observed in *aka-lria-ø* is clarified with the definite interpretation of the referent. In CAY, as shown in (19), the object of the transitive verb is marked with the absolutive case, and the absolutive referent is assumed to be shared between the speaker and the hearer. Therefore, when *aka-lria-ø* is employed, it is normally clear as to what the actual referent is. However, without any contextual information, and when *aka-lria* is used in an indefinite way together with the ablative marking, as shown in (20) below, the form *aka-lria-mek* is truly ambiguous as to whether the lexicalized or non-lexicalized sense is intended (John Toopetlook, p.c.; we shall return to this in Section 4).

(20) *Angute-ø tangrr-u-q aka-lria-mek kanani.*
man-3SG.ABS see-INTR.IND-3SG roll-NMZR.ARG-3SG.ABL down.there
a. 'The man is seeing a car down there.'
b. 'The man is seeing something rolling/a log that is rolling down there.'

4. Grammatical nominalizations

Let us now turn to the issue of how relative-clause counterparts are formally expressed in CAY. We will argue here, following the theory of nominalization, that the CAY relative-clause counterparts are not much more than grammatical nominalizations in a modification-use, the principal function of which specifies a subset of the denotations of the head noun with its own denotations (see Shibatani (2018: 346–347, this volume)). Examining the relevant data, we show that in CAY a single construction of *nominal juxtaposition* serves as an umbrella construction for the semantic correspondents of, say, English adjective, participle, and relative clause (cf. Jacobson 1995: 250). While CAY shows no formal distinction between adjective, participle and relative clause, the language exhibits the same functional relationship between the head and its modifier within the juxtaposition construction. These observations direct us to seek a functional unity underlying the notions such as adjective, participle, infinitive, nominalization, and relative clause, contrary to the view expressed by Dixon (2010: 316), who divides forms performing the same function into several distinct structures on the basis of formal differences, noting that "it is important to distinguish between an RC – which has the structure of a clause, and can function as modifier within an NP – and a participle – which is an adjective derived from a verb, and can also function as modifier within an NP."

As mentioned in the previous section, CAY has no part of speech of adjective, and adjectival modification in other languages can be expressed by juxtaposing two nouns, the construction of which Jacobson (1995: 31, 78) calls *appositives*.[10] Let us first consider the following examples.[11]

(21) a. *Angute-m* *utaq-a-a* *arnaq-ø.*
man-3SG.ERG wait-TR.IND-3SG.3SG woman-3SG.ABS
'The man is waiting for the woman.'

b. *Angute-m* *utaq-a-a* *kass'aq-ø.*
man-3SG.ERG wait-TR.IND-3SG.3SG white.person-3SG.ABS
'The man is waiting for the white person.'

10. We avoid the term "appositive" in order to distinguish the relevant structures from the well-known appositive constructions such as *John, the butcher*, which involve juxtaposition of two definite descriptions. In ordinary relative clauses and modification involving noun modifiers, the modifying structures are considered to have denotations but not references, whether definite or not, according to Shibatani (this volume).

11. See also type (D) of nominalizers (noun-based nominalizers) in Section 3 for the creation of adjectival denotation.

c. *Angute-m* *utaq-a-a* *[kass'aq-ø*
man-3SG.ERG wait-TR.IND-3SG.3SG white.person-3SG.ABS
arnaq-ø]$_{NP}$
woman-3SG.ABS
(i) 'The man is waiting for the white woman.'
(ii) 'The man is waiting for the female white person.'

In (21a) and (21b), two nouns of *arnaq* 'woman' and *kass'aq* 'white person' are employed as an object of the verb of waiting, respectively, with the nominal inflection specifying the information on case, number and person (3SG.ABS here). In (21c), these two nouns are juxtaposed so that they can indicate the meaning of (a) white person (we shall return to the interpretations of (21c) soon below). As can be observed, there is no formal change in *arnaq* and *kass'aq* in (21c) from that of (21a) and (21b) respectively, keeping their nominal inflection of 3SG.ABS. The morphological agreement between *arnaq-ø* and *kass'aq-ø* guarantees that these two nouns form a constituent, whereby the denotation of one specifies a subset of that of the other, yielding restrictive interpretations (a woman who is a white person or a white person who is female), which gives the impression that the two nominals are referring to the same referent. If two nominals are not in a modification relation as in this example, the enclitic, *=llu* 'and' must be added to the rightmost noun in NP, and the verbal inflection of *-a* must be changed to *-k* to indicate that the two denotations do not intersect and are independent from each other, as exemplified in (22) below.[12]

(22) *Angute-m* *utaq-a-k* *[kass'aq-ø*
man-3SG.ERG wait-TR.IND-3SG.3DU white.person-3SG.ABS
arnaq-ø=llu]$_{NP}$
woman-3SG.ABS=and
'The man is waiting for the white person and the woman.'

As shown in the translations of (21c), the two nouns that are juxtaposed show ambiguity as to which noun serve as the head and which noun as the modifier. In (21), interpretation (i) may sound more natural than (ii) without any contextual information, but example (21) does have the interpretation (ii); for instance, this interpretation obtains when the speaker corrects the hearer's assumption that the man was waiting for a male person (John Toopetlook, p.c.; see also footnote 7 above).

Shibatani (2018: 307, this volume) emphasizes the fact that ordinary nouns show the three functions of denotation, modification and reference. Consider the simple noun *door* in English. It (i) has a certain denotation evoked by the sound

12. Word order is relatively free in CAY (e.g. Jacobson 1995: 119), though we need further investigation on this topic. In this paper, the order of nouns in bracketed NP in the examples goes unquestioned.

of [dɔ^{r}], and this denotation may be utilized in two ways: (ii) to "refer to" an entity called *door* in the real/discourse world by heading an NP, as in *Open the door for me* (Shibatani calls this *NP-use*), and (iii) to modify another denotation, as exemplified in noun compounds *door handle*, *door knocker*, *doorman*, and in noun phrases such as *door keeping robots* and *garage door opener* (Shibatani's *modification-use*). In other words, the theory of nominalization gives marked attention to the fact that nouns with a denotation play referential and modification functions. Given that in (21a) and (21b) the two items *arnaq* 'woman' and *kass'aq* 'white person' are confirmed as nouns, it is natural to assume that the juxtaposed construction forming a larger noun phrase in (21c) consists of two nouns that form the head-modifier relationship by virtue of one serving as an NP head and the other as a modifier specifying a subset of the denotation of the head noun. The fact that (21c) allows two interpretations, whereby either nominal can be construed as a head noun, corroborates this analysis. In what follows, we first show that this juxtaposed construction is uniformly utilized for the correspondents of adjectival, participial and relative clause modification constructions found in other languages. With this structural unity affirmed, we move on to the issues of (i) how the functional dependency between the head and modifier is to be described and (ii) how the ambiguity observed in the juxtaposed construction is resolved.

As the above illustration shows, the property concept roots that may be realized as adjective in other languages tend to be nominalized in CAY, and are instantiated in the juxtaposed construction:

(23) a. *Angute-m tamar-a-a [nutek-ø*
man-3SG.ERG lose-TR.IND-3SG.3SG gun-3SG.ABS
tungu-lria-ø]$_{NP}$
be.black-NMZR.ARG-3SG.ABS
'The man lost the black gun.'
Literally: 'The man lost the gun, a black one.'

b. *[Angute-m uquri-lria(lrii)-m]*$_{NP}$ *angqaq-ø*
man-3SG.ERG be.fat-NMZR.ARG-3SG.ERG ball-3SG.ABS
kitngig-a-a.
kick-TR.IND-3SG.3SG
'The fat man kicked the ball.'
Literally: 'The man, a fat one, kicked the ball.'

c. *Caan-(aaq)-ø an'-u-q [ak'a-llaq(llar)-mek*
John-3SG.ABS come.out-INTR.IND-3SG be.old-NMZR.ARG-3S.ABL
nem-'ek]$_{ABLP}$
house-3SG.ABL
'John came out of the old house.'
Literally: 'John came from an old one, from the house.'

d. *[Kuig-kun iqkite(l)-nguq(r)-kun]$_{VIA}$ yuu-t*
river-VIA be.narrow-NMZR.ARG-VIA person-3PL.ABS
cetu-u-t.
go.with.river-INTR.IND-3PL
'People are going down the narrow river.'
Literally: 'People are going through the river, through narrow one.'

In (23), the property concepts of *tungu-* 'be black', *uquri-* 'be fat', *ak'a-* 'be old,' and *iqkite-* 'be narrow' are each expressed together with a nominalizer in (23a), (23b), (23c), and (23d) respectively. Furthermore, the inflectional agreement between the nominalized forms and the nouns juxtaposed with them are observed, exhibiting its appropriate inflectional form that responds to its syntactic environments not only in absolutive and ergative, but in ablative and vialis case. In addition, the nominalized forms can stand as nouns heading an NP by themselves if contextual information sufficiently narrows down the intended denotation, as shown below:

(24) a. *Angute-m tamar-a-a [tungu-lria-ø]$_{NP}$*
man-3SG.ERG lose-TR.IND-3SG.3SG be.black-NMZR.ARG-3SG.ABS
'The man lost the black one.'

b. *[Uquri-lria(lrii)-m]$_{NP}$ angaq-ø kitngig-a-a.*
be.fat-NMZR.ARG-3SG.ERG ball-3SG.ABS kick-TR.IND-3SG.3SG
'The fat one kicked the ball.'

c. *Caan-(aaq)-ø an'-u-q [ak'a-llaq(llar)-mek]$_{ABLP}$.*
John-3SG.ABS come.out-INTR.IND-3SG be.old-NMZR.ARG-3S.ABL
'John came out of the old one.'

d. *[Iqkite(l)-nguq(r)-kun]$_{VIA}$ yuu-t cetu-u-q.*
be.narrow-NMZR.ARG-VIA person-3PL.ABS go.with.river-INTR.IND-3SG
'People are going down the narrow one.'

The same characteristics are also observed in "more verbal" concepts when they are nominalized.

(25) a. *Angute-m assikenrit-a-a [qimugta-ø qilu-lria-ø]$_{NP}$*
man-3S.ERG hate-TR.IND-3SG.3SG dog-3SG.ABS bark-NMZR.ARG-3SG.ABS
kiani.
there
'The man doesn't like the dog that is barking over there/the barking dog.'
Literally: 'The man doesn't like the dog, one that is barking.'

b. *[Mikelnguq-m qia-lleq(ller)-m]$_{NP}$ waniwa ner-a-a.*
child-3S.ERG cry-NMZR.ARG-3S.ERG right.now eat-TR.IND-3SG.3SG
'The child who was crying/the crying child is now eating.'
Literally: 'The child, one who was crying is now eating.'

c. *Qimugta-ø an'-u-q [[Caan-aam*
dog-ABS.3SG come.out-INTR.IND-3SG John-3SG.ERG
kipute-lleq-a(llra)-nek]$_{NMLZ\ (ABL)}$ ene-mek]$_{ABL}$
buy-NMZR.ARG-3SG.3SG-ABL.(POSS) house-3SG.ABL
'The dog came out of the house John bought.'
Literally: 'The dog came from a thing John bought, from the house.'

d. *Caan-(aaq)-ø suupa-li-u-q [egate-kun*
John-3SG.ABS soup-make-INTR.IND-3SG pot-VIA
kipute-lleq(lle)-m-kun]$_{VIA}$
buy-NMZR.ARG-1SG.3SG-VIA
'John made soup with the pot I bought.
Literally: 'John made soup through the pot, through a thing I bought.'

The nominalized forms, *qilu-lria-ø*, *qia-lleq*, and *kipute-lleq*, are juxtaposed with a noun in (25a), (25b), and (25c) (and also (25d)), respectively, showing the morphological agreement in nominal inflection. This property clearly indicates that the nominalized forms and the nouns together form the juxtaposed construction like (21c) above. Furthermore, recalling our account of examples (17) and (18) in the previous section, we recognize that verbal roots have been turned into nominal structures, as evidence by their ability to stand as NP heads, if contextual information sufficiently narrows down the intended denotation.

(26) a. *Angute-m assikenrit-a-a [qilu-lria-ø]$_{NP}$ kiani.*
man-3SG.ERG hate-TR.IND-3SG.3SG bark-NMZR.ARG-3SG.ABS there
'The man doesn't like the barking one over there.'

b. *[Qia-lleq(ller)-m]$_{NP}$ waniwa ner-a-a.*
cry-NMZR.ARG-3SG.ERG right.now eat-TR.IND-3SG
'The one who was crying is now eating.'

c. *Qimugta-ø an'-u-q [Caan-aam*
dog-3SG.ABS come.out-INTR.IND-3SG John-3SG.ERG
kipute-lleq-a(llra)-nek]$_{NMLZM\ (ABL)}$
buy-NMZR.ARG-3SG.3SG-ABL.POSS
'The dog came out of the one John bought.'

d. *Caan-(aaq)-ø suupa-li-u-q [kipute-lleq(lle)-m-kun]$_{VIA}$*
John-3SG.ABS soup-make-INTR.IND-3SG buy-NMZR.ARG-1SG.3SG-VIA
'John made soup with the one I bought.'

What examples (22) to (26) indicate may be summarized as follows. First, the nominalized forms in CAY can behave in the same way as simple nouns do. They can be used on their own as the head of an NP without accompanying any other nouns, as exemplified in (24) and (26), exhibiting appropriate case marking also like simple nouns. And they can also be utilized in the modification function by

forming a juxtaposed construction together with a head noun, corresponding to English adjectival, participial, and relative clause constructions. The CAY data clearly show that these traditional categories have underlying unity that the general theory of nominalization intends to capture. There is no obligatory tense marking in CAY and the relevant inflections are drawn from nominalizers deriving simple nouns. In addition, as discussed in Section 3, the nominalized forms themselves show the ambiguity between lexicalized and non-lexicalized uses, some of the nominalizers incorporating temporal/transient state into their noun formation.[13] Thus, CAY is a language which casts doubt on the distinction that Dixon (2010) wants to draw between nominalizations and relative clauses as quoted in Section 1, and also between adjectival participles and relative clause as quoted in the beginning of this section. Consider the following examples, in which more than two nouns are juxtaposed in a noun phrase:

(27) a. *Angute-m assikenrit-a-a [[qimugta-ø*
man-3SG.ERG hate-TR.IND-3SG.3SG dog-3SG.ABS
tungu-lria-ø] qilu-lria-ø]$_{NP}$ *kiani.*
be.black-NMZR.ARG-3SG.ABS bark-NMZR.ARG-3SG.ABS there
'The man doesn't like the black dog that is barking over there.'
Literally: 'The man doesn't like the dog, a black one, one that is barking/barking one.'

b. *Kia pik-a-u [augna*
who.3SG.ERG own-INTER-3SG.3SG that.3SG.ABS
[[ak'a-llaq-ø mikel-nguq-ø
be.old-NMZR.ARG-3SG.ABS be.small-NMZR.ARG-3SG.ABS
[snuukuuq-ø]] cuka-ite(l)-nguq-ø]$_{NP}$*?*
snowmachine-3SG.ABS fastness-NEG-NMZR.ARG-3SG.ABS
'Who owns that old slow small snow-machine?'
Literally: 'Who owns that one, an old one, a small one, a snow-machine, a thing that is not fast?'
(John Toopetlook, p.c.; based on Jacobson 1995: 252)

As shown in (27a), as said above, the property concept of being black and the verbal concept of barking are juxtaposed together as nominalized nouns showing the same nominal inflection. Likewise, in (27b), though three nominalized forms, *ak'a-llaq-ø*, *mikel-nguq-ø*, and *cuka-ite(l)-nguq-ø* are employed in the NP,

13. As the reader can see from examples (17b), (18b) and (26), our non-lexicalized use of the nominalized forms corresponds to so-called headless relative clauses when they are in NP-use. In other words, the non-lexicalized use of nominalized forms is all regarded as an instantiation of grammatical nominalization (see Fleck (this volume) on the characteristics of nonce nominalization in Matses).

the forms themselves could not tell which one shows clause-hood, while it is clear that they all evoke nominal denotations, and that they contribute to narrowing down the potential NP reference by specifying subsets of their head nominal.

What a language like a CAY suggests is that we need to reassess, as attempted by Shibatani (this volume), situations in other languages like English, whose relative clauses have traditionally been analyzed differently from nominalizations. Shibatani recognizes that the three functions of denotation, modification, and reference of ordinary nouns and nominalizations (both lexical and grammatical) can also be served by grammatical constructions that have not been traditionally analyzed as nominalizations. Shibatani demonstrates that an English-style relative clause such as [*who you love*] in *Marry a man* [*who you love*] and its equivalents in other languages have all these three properties and uses. Formal unity, as we see in CAY, and diversity between lexical nominalizations and grammatical nominalizations, as in English, show that languages respond differently to the underlying unity and the differences in type and use. Some languages, like CAY, capture the functional unity by the morphological uniformity, while in some others, like English, the diversity in type and use is formally represented at the expense of the underlying unity.[14]

Now let us move on to further characteristics of the juxtaposed construction. As for the relationship between the nouns in juxtaposition, Jacobson (1995: 66) provides us with the following observation: "…possession, that is ownership, is indicated only on the main noun of the pair." Consider examples (28) and (29) below:

(28) a. *kass'aq-ø* *ui-ka*
white.person-3SG.ABS husband-1SG.3SG.ABS
'my white husband'

b. **kass'a(q)-qa* *ui-ka*
white.person-1SG.3SG.ABS husband-1SG.3SG.ABS
'my white husband'

(29) a. *qimugte-ka* *tungu-lria-ø*
dog-1SG.3SG.ABS be.black-NMZR.ARG-3SG.ABS
'my black dog' literally: 'my dog, the one who is black'

b. **qimugte-ka* *tungu-lria-ka*
dog-1SG.3SG.ABS be.black-NMZR.ARG-1SG.3SG.ABS
'my black dog'

Following Jacobson's description, we can say that in the juxtaposed construction, the nominal form that can take possessor marking is considered as the head. Note that, though the portmanteau characteristic renders nominal inflections less

14. I owe a lot about this point to editors of this volume.

transparent, the two nouns still share the same inflectional information on the referent (3SG.ABS above).

With this property in mind, observe that when the juxtaposed construction consisting of two ordinary nouns produces some uncertainty in its interpretation, the speaker tends to employ a nominalized form to make clear the relationship between the head and the modifier.

(30) a. ?*[Arnaq-ø nulia-qa]*$_{NP}$ *ner-u-q nervig-mi yaani.*
woman-3SG.ABS wife-1SG.3SG.ABS eat-INTR.IND-3SG table-LOC there
Intended meaning: 'The woman who is my wife (not 'my wife who is a woman') is eating at the table over there.'

b. *[Arnaq-ø nulia-qe-ke-ka (nuliaqek'a)]*$_{NP}$ *ner-u-q*
woman-3SG.ABS wife-as.have-NMZR.ARG-1SG.3SG.ABS eat-INTR.IND-3SG
nervig-mi yaani.
table-LOC there
'The woman who is my wife is eating at the table over there.'
'Literally: The woman, one that I have as a wife, is eating at the table over there.'

In (30a), the hearer may make an inference different from what the speaker intends, thinking that because wives are usually female, the two people must be intended by the speaker like the woman and my wife, even though the verb is marked by the third person singular. In such a case, CAY speakers prefer to employ a nominalized form made for the nonce; in (30b) *nuliaq* 'wife' is first verbalized with *-qe* 'as.have' and then turns *nulia-qe* 'wife-as.have' into the nominalized form of *nulia-qe-ke* 'wife-as.have-NMZR.ARG (lit. one that X has as wife)' (John Toopetlook, p.c.).

In example (30b), one may notice that, semantically speaking, the modifier noun rather than the head noun is marked with the possessor marking contrary to the description above. However, in *nulia-qe-ke-ka* (wife-as.have-NMZR.ARG-1SG.3SG.ABS), the possessor marking is being used with regard to an ergative argument of *nulia-qe* 'wife-as.have' rather than to a possessive relationship (cf. Woodbury 1985: 76; Miyaoka 1986: 114–115). Some other examples of the same type are shown below:

(31) a. *[Calissuute-n atu-lleq-qa (atulqa)]*$_{NP}$
tool-2SG.3SG.ABS use-NMZR.ARG-1SG.3SG.ABS
nanva-uma-llini-u-q.
break-PERF-probably-INTR.IND-3SG
'Your tool that I used seems to have been broken.'

b. *[Calissuute-n atu-lleq-pet (atulpet)]*$_{NP}$
tool-2SG.3SG.ABS use-NMZR.ARG-2SG.3SG.ABS
nanva-uma-llini-u-q.
break-PERF-probably-INTR.IND-3SG
'Your tool that you used seems to have been broken.'

While the possessor marking of the nominalized form in the modification-use behaves differently from the head noun, we also have to point out that it does not straightforwardly follow the ergative marking of the verbal inflection either. Recall that we mentioned in Section 1 that in CAY the same case markers are employed for the transitive subject and the genitive with the double marking property:

(32) a. *Qimugte-m (qimugtiim) qilug-a-a mikelnguq-ø.*
dog-3SG.ERG bark-TR.IND-3SG.3SG child-3SG.ABS
'The dog is barking at the child.'

b. *mikelnguq-ø [qimugte-m (qimugtiim) qilug-lleq-a (qilullra)]*$_{NMLZ}$
child-3SG.ABS dog-3SG.ERG bark-NMZR.ARG-3SG.3SG.ABS
'the child that the dog is barking at'
Literally: 'the child, the dog's, one that be barked at'

c. *qimugta-ø [mikelngur-m (mikelnguum) qilug-ste-a (qilugtii)]*$_{NMLZ}$
dog-3SG.ABS child-3SG.ERG bark-NMZR.ARG-3SG.3SG.ABS
'the dog that is barking at the child'

d. *Qimugta-ø qilug-tu-q mikelngur-mun.*
dog-3SG.ABS bark-INTR.IND-3SG child-TERM
'The dog is barking at (to) the child.'

(John Toopetlook, p.c.; based on Jacobson 1995: 255)

For (32a), a simple transitive sentence, we may have two possible juxtapositions, as shown in (32b) and (32c), depending on which argument is placed as the head noun. The former takes *mikelnguq-ø* as its head noun, and the ergative case marked *qimugte-m* may be considered as the agent of barking, *qilug-*, forming the modification NP together with *qilug-lleq-a*, which shows exactly the same pattern as (30b) and (31) above. On the other hand, in the latter, (32c), *qimugta-ø* is chosen as the head noun, and in the nominalized modifier, the nominalizer *-ste* (*-sta*) is employed, which requires an intransitive verb for its nominalization as mentioned in the previous sections with the example of *elit-naur-i-sta* 'teacher' (We shall return to this soon below). When the verb of barking is employed in the intransitive clause, the target of barking should be marked with the terminalis case, *mikelngur-mun*, as shown in (32d) (John Toopetlook, p.c.). However, in (32c) it is marked with the ergative case as *mikelngur-m*. This fact implies that the possessive marking of the nominalized form in the modification use does not simply follow the argument-marking system of the clause structure.

As another signal indicating the presence of grammatical nominalization, let us return to example (19), repeated below, in which we observed that the same nominalization structures can show ambiguity between the lexical and non-lexical use.

(19) *Angute-m tangrr-a-a aka-lria-ø kanani.*
man-3SG.ERG see-TR.IND-3SG.3SG roll-NMZR.ARG-3SG.ABS down.there
a. 'The man is seeing the car down there.'
b. 'The man is seeing the one rolling/the log that is rolling down there.'

First, as mentioned above (in footnote 13), the nominalized forms used in the non-lexical sense can head an NP in CAY, as you can see from examples (17b), (18b), (19), and (26) above. With this point in mind, let us consider how possessor marking behaves in the NP-use of grammatical nominalization in the cases like (19).

As shown in the translations of (33a) below, possessor marking cannot be applied when the nominalized form is understood in the non-lexicalized sense. Even if the speaker and the hearer both know whose thing is rolling, the nominalized form referring to it cannot be "possessed", while the marking can be utilized for the same form when it is used in the lexicalized sense (John's car here). According to my consultant John Toopetlook, if one wants to express the ownership of the thing that is rolling, he/she would have to say it in an explicit way, as shown in (33b), in which a distinct head noun, e.g. *laaqaq* 'log', occurs together with a possessor marker, e.g. *laaqa-a* (*laaqii*) log-3SG.3SG.ABS.

(33) a. *Angute-m tangrr-a-a [Caan-aam*
man-3SG.ERG see-TR.IND-3SG.3SG John-3SG.ERG
aka-lria-a (akalrii)]$_{NP}$ kanani.
roll-NMZR.ARG-3SG.3SG.ABS down.there
i. 'The man is seeing John's car down there.'
*ii. 'The man is seeing John's thing rolling/John's log that is rolling down there.'

b. *Angute-m tangrr-a-a [[Caan-aam laaqa-nga (laaqii)]*
man-3SG.ERG see-TR.IND-3SG.3SG John-3SG.ERG log-3SG.3SG.ABS
aka-lria-ø]$_{NP}$ kanani.
roll-NMZR.ARG-3SG.ABS down.there
'The man is seeing John's log that is rolling/John's log rolling down there.'

In other words, the possibility of possessive marking may distinguish between lexical and grammatical nominalizations for those forms that allow lexicalized and literal interpretations, indicating that the former show a greater affinity to ordinary nouns, which can display possessor marking freely (see the related discussion by Fleck this volume).

Finally, CAY grammatical nominalization shares an important principle with its lexical counterpart, pointing to the unified nature of these processes. As recalled from the earlier discussion on the forms in (E) in Section 3, only the absolutive interpretations (namely the S and P readings) are possible with lexical nominalizations. We observed that a form such as *elitnauriste-* 'teacher' includes the antipassive morpheme *-i-* (before the nominalizer *-ste*), signalizing that the agentive (ergative) entity aligns with the derived S role. The same "ergative" principle of targeting S and P (as well as S derived via antipassivization) applies to grammatical nominalization.

For example, the verb *nere-* can be used as either a transitive or an intransitive verb, as below:

(34) a. *Angute-m ner-a-a neqa-ø nervig-mi.*
man-3SG.ERG eat-TR.IND-3SG.3SG fish-3SG.ABS table-LOC
'The man is eating the fish/food.'

b. *Angun-ø ner-'u-q nervig-mi.*
man-3SG.ABS eat-INTR.IND-3SG table-LOC
'The man is eating at the table.'

With this type of verb, the agentive denotations are possible only with the intransitive versions, countenancing an agentive S, as (35a) and (35b.i) below show. On the other hand, the transitive version, as in (35b), is possible when the intended denotation is in P role, as in (35b.ii), though this particular sentence is odd for a semantic reason. The same structure is perfectly grammatical when the semantic issue has been resolved, as in (35c).

(35) a. *[Angun-ø ner-lleq-ø (nerrleq)]*$_{NP}$ *an-'u-q.*
man-3SG.ABS eat-NMZR.ARG-3SG.ABS go.out-INTR.IND-3SG
'The man who was eating/ate went out.'

??b. *[Arna-m [ner-lleq-a (nerellra)]]*$_{NP}$ *an-'u-q.*
woman-3SG.ERG eat-NMZR.ARG-3SG.3SG.ABS go.out-INTR.IND-3SG

*i. intended meaning: 'The woman who was eating/ate something went out.

??ii. 'The one that the woman was eating went out.'

c. *[Arna-m [ner-lleq-a (nerellra)]]*$_{NP}$ *tuntuvag-mek*
woman-3SG.ERG eat-NMZR.ARG-3SG.3SG.ABS moose-3SG.ABL
suupa-u-llru-u-q.
soup-be-past-INTR.IND-3SG
'What the woman was eating/ate was moose soup.'

When the verb has only a transitive use in its basic form, as in the case of *tamar-* 'lose' in (36a) below, the agent role cannot be nominalized directly. In order to

nominalize such an A role, it must be first turned into a derived S via antipassivization, as in (36b). Such an S can then be nominalized, as in (36c) and (36d).

(36) a. *Angute-m tamar-llru-a-a (tamallrua) nacaq-ø.*
man-3SG.ERG lose-PST-TR.IND-3SG.3SG cap-3SG.ABS
'The man lost the cap.'

b. *Angun-ø tamar-i-llru-u-q (tamarillruuq) naca-mek.*
man-3SG.ABS lose-ANTIP-INTR.IND-3SG cap-3SG.ABL
'The man lost a cap.'

c. *[[Naca-m tamar-i-lleq-ø (tamarilleq)]$_{NMLZ}$]$_{NP}$*
cap-3SG.ERG lose-ANTIP-NMZR.ARG-3SG.ABS
iter-tu-q. (NP-use)
come.back-INTR.IND-3SG
'The one who lost a cap returned.'

d. *[[Naca-m tamar-i-lleq-ø (tamarilleq)]$_{NMLZ}$ angun-ø]$_{NP}$*
cap-3SG.ERG lose-ANTIP-NMZR.ARG-3SG.ABS man-3SG.ABS
iter-tu-q. (Modification-use)
come.back-INTR.IND-3SG
'The man who lost a cap returned.'

The unity between lexical and grammatical nominalizations in terms of sharing of the nominalizing morphology and the ergative patterning in argument nominalizations is likely due a historical development, whereby grammatical nominalizations give rise to conventionalized interpretations of those engaged in frequently occurring habitual activities (e.g., from "one who hunts" to "a hunter"). We have seen many forms above that are in transition, allowing both literal ("one who hunts') and conventionalized ("a hunter") readings.

5. Concluding remarks

This article analyzes nominalization phenomena in Central Alaskan Yup'ik in terms of the generalized theory of nominalization developed by Shibatani (2018 and this volume, inter alia), with a focus on its noun formation, relative-clause counterparts and the relationships between the two. Observing the general characteristics of CAY lexical nominalizations in Section 2, we recognized the crucial role that nominalizers play in forming nouns in both quantity and quality. In Section 3, with the classification of nominalizers, we pointed out that the CAY nominalizations may be used either productively in referring to a variety of entities satisfying the denotational properties of the forms in question or in a restricted manner denoting only specific objects. We also noted that some of the

nominalizers incorporate temporal/transient state into their noun formation, which may have to be expressed independently from nominalizations in other languages. In Section 4, with the examples that can be considered as relative-clause counterparts, we observed a structural continuity between lexical nominalization and the relative-clause counterparts, both functioning either as an NP-head or as a modifier of a head noun. In conclusion, CAY noun modification involves a nominal (an ordinary noun, a lexical nominalization, or a grammatical nominalization) that specifies a subset of the denotation of the head noun. That is, what corresponds to a relative clause in other languages is clearly a grammatical nominalization in CAY in both morphology and it uses, demanding an answer to why so-called relative clauses function in a similar way to grammatical nominalizations of the CAY type if they are not nominalizations, as argued by Shibatani (this volume).

Acknowledgements

I wish to express my profound gratitude and sincere thanks to Masayoshi Shibatani for his invaluable comments and suggestions on earlier drafts. I hugely appreciate John W. Toopetlook for providing me with irreplaceable information on Central Alaskan Yup'ik. Furthermore, this article benefited from helpful comments, suggestions and encouragements at various stages from Oscar Alexie, Sophie Alexie, Matthew Carmody, Sung-Yeo Chung, David Fleck, Katsunobu Izutsu, Takeshi Koguma, Haruhiko Murao, Makoto Sumiyoshi, Eri Tanaka, and Roberto Zariquiey.

Morpheme gloss abbreviations

1	first person	INST	instrumental
2	second person	INTER	interrogative
3	third person	INTR	intransitive
ABL	ablative	LOC	locative/place
ABS	absolutive	MAN	manner
AG	agentive	NEG	negation
ANTIP	antipassive	NOM	nominative
ARG	argument	NMZR	nominalizer
ASP	aspect	PL	plural
ASSOC	associative	REFL	reflexive
COM	comitative	SG	singular
DU	dual	SUBR	subordinator
ERG	ergative marking	TERM	terminalis
EVID	evidential	TR	transitive
IND	indicative	VIA	vialis

References

Bergsland, Knut. 1997. *Aleut Grammar*. Fairbanks AK: Alaska Native Language Center.

Comrie, Bernard. 1997. Attributive clauses in Asian languages: Towards an areal typology. In *Sprache in Raum und Zeit: In Memoriam Johannes Bechert, Band 2: Beiträge zur empirischen Sprachwissenschaft*, Winfried Boeder, Christoph Schroeder, Karl Heinz Wagner & Wolfgang Wildgen (eds), 51–60. Tübingen: Gunter Narr.

Comrie, Bernard. 2006. Syntactic typology: Just how exotic are European-type relative clauses. In *Linguistic Universals*, Ricardo Mairal & Juana Gil (eds), 130–154. Cambridge: CUP. https://doi.org/10.1017/CBO9780511618215.007

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

Croft, William. 2003. *Typology and Universals*. Cambridge: CUP.

DeLancey, Scott. 2002. Relativization and nominalization in Bodic. In *Proceedings of the twenty-eighth Annual Meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics, in honor of Prof. James A. Matisoff*, Patrick Chew (ed.), 55–72. Berkeley CA: BLS.

Dixon, Robert M.W. 2010. *Basic Linguistic Theory*, Vol 2. Oxford: OUP.

Fortescue, Michael, Jacobson, Steven A. & Kaplan, Lawrence. 2010. *Comparative Eskimo Dictionary with Aleut Cognates*. Fairbanks AK: Alaska Native Language Center.

Jacobson, Steven A. 1984. *Yup'ik Eskimo Dictionary*. Fairbanks AK: Alaska Native Language Center.

Jacobson, Steven A. 1995 *A Practical Grammar of the Central Alaskan Yup'ik Eskimo Language*. Fairbanks AK: Alaska Native Language Center.

Johns, Alana. 1992. Deriving ergativity. *Linguistic Inquiry* 23(1): 57–87.

Miyaoka, Osahito 1986. Eskimogo no nookakusei (Ergativity in Alaskan Yupik Eskimo). *Gengo Kenkyu* 90:97–118.https://doi.org/10.11435/gengo1939.1986.90_97

Miyaoaka, Osahito. 2012. *A Grammar of Central Alaskan Yupik (CAY)*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110278576

Nichols, Johanna. 1986. Head-Marking and Dependent-Marking Grammar. *Language* 62(1): 56–119. https://doi.org/10.1353/lan.1986.0014

Reed, Irene, Miyaoka, Osahito, Jacobson, Steven A., Afcan, Paschal & Krauss, Michael. 1977. *Yup'ik Eskimo Grammar*. Fairbanks AK: Alaska Native Language Center.

de Reuse, Willem J. 2007. Polysynthesis as a typological feature: An attempt at a characterization from Eskimo and Athabascan perspective. In *Variations on Polysynthesis: The Eskaleut languages* [Typological Studies in Language 86], Marc-Antoine Mahieu & Nicole Tersis (eds), 19–34. Amsterdam: John Benjamins.

Shibatani, Masayoshi. 2009. Elements of complex structures, where recursion isn't: The case of relativization. In *Syntactic Complexity Diachrony, Acquisition, Neuro-cognition, Evolution* [Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 163–198. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.07ele

Shibatani, Masayoshi. 2014. Rethinking relative clauses. Handout distributed at International Workshop on Nominalization in the languages of the Americas. Lima: Pontificia Universidad Católica del Perú, 28 August.

Shibatani, Masayoshi. 2018. Nominalization in crosslinguistic perspective. In *Handbook of Japanese Contrastive Linguistics*, Prashant Pardeshi & Taro Kageyama (eds), 345–410. Berlin: Walter de Gruyter. https://doi.org/10.1515/9781614514077-013

Thalbitzer, William. 1911. Eskimo. In *Handbook of American Indian Languages* [Bureau of American Ethnology Bulletin 40.1], Franz Boas (ed.), 967–1069. Washington DC: Government Printing Office.

Woodbury, Anthony C. 1985. Noun phrase, nominal sentence, and clause in Central Alaskan Yupik Eskimo. In *Grammar Inside and Outside the Clause: Some Approaches to Theory from the Field*, Johanna Nichols & Anthony C. Woodbury (eds), 61–88. Cambridge: CUP.

Woodbury, Anthony C. 2002. The word in Cup'ik. In *Word: A Cross-Linguistic Typology*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 79–99. Cambridge: CUP.

CHAPTER 7

The 'relative' illusion and the origin of non-subject nominalizers in Cahita (Uto-Aztecan)

Albert Álvarez González
Universidad de Sonora

This paper proposes to analyze non-subject relative clauses in Cahita languages (†Tehueco, Yaqui, and Mayo, from the Uto-Aztecan family) as non-subject grammatical nominalizations used as appositives in restrictive appositions. This analysis is based on a diachronic perspective which considers the origin of the markers involved in these constructions as well as the evolutionary paths that lead to the relativization function. The study shows that the different origins proposed for non-subject nominalizers in Cahita (a possessed nominal classifier, a subject nominalizer, a directional postposition, and a locative postposition) are all different types of nominal markers indicating that the base clause/verb is treated as a nominal constituent; that is, that the construction has been nominalized. As for the evolution to the relativization function, the first nominalization uses of these markers are always associated with a referential function and the modifying function corresponds to a secondary development made via apposition, indicating that non-subject relativization in Cahita corresponds in fact to a specialized function of the non-subject nominalization: the modifying function of a possessed non-subject nominalization used as an appositive in a restrictive apposition.

1. Introduction

Based on the approach developed in Shibatani (2009) and in Shibatani & Awhad (2009) concerning the connection between relativization and nominalization in different languages around the globe, in two previous works (Álvarez 2012, 2016), I have defended the idea that relative clauses in Yaqui (a language from the Uto-Aztecan family, spoken in the northwest of Mexico) and, more generally in Cahita group of languages (which includes the †Tehueco, Yaqui, and Mayo languages), are better analyzed as argument grammatical nominalizations used in restrictive appositions (Kubo 2009). In this paper, I will focus more specifically on non-subject

https://doi.org/10.1075/tsl.124.07gon

nominalizers (object and oblique/locative nominalizers) and I will explore the origin of these markers in Cahita. In doing so, I will propose that what have been commonly analyzed as object relative clauses in Yaqui by different authors (Lindenfeld 1973; Martínez & Langendoen 1996; Dedrick & Casad 1999; Guerrero 2012), are in fact object grammatical nominalizations.[1] I will argue that these grammatical nominalizations have two major functions (referential and NP-modifying functions) depending on the syntactic uses of the nominalized expression (respectively, as NP-head and as NP-dependent in apposition) and that relativization in Cahita should be considered merely one specialized function of nominalization: the modifying function of an appositive grammatical nominalization.

Regarding the diachrony of Cahita non-subject nominalizations, in Álvarez (2016) I have mentioned that the markers used in these nominalizations are historically associated with oblique case markings (namely, instrumental, directional and locative postpositional markings). Here, I will go deeper into the historical reconstruction of the non-subject nominalizations in Cahita and I will propose that the origin of the Old Cahita object nominalizer (the suffix *-ye*) is, in fact, not an instrumental postposition but an old inalienable inanimate possessive classifier. I will also posit that the recruitment of a directional postposition in the development of Modern Cahita object nominalizers can be explained through the reanalysis and evolution undergone by this old nominal classifier, but also by the possible influences of the locative nominalization marking and the evolution of an old subject nominalizer.

The organization of the information is as follows: In Section 2, I will present the nominalization approach to relativization recently exposed in Álvarez (2016), as well as the terminological consequences that this approach has in the domain of the relative clauses (RC). The nominalization/relativization syncretism will be also discussed from a synchronic and a diachronic perspective. After a brief introduction to the Cahita languages in Section 3, I will illustrate in Section 4 the non-subject nominalizations from Old Cahita and Modern Cahita (Yaqui and Mayo) and I will briefly comment on the synchronic evidences supporting the nominalization status of these constructions. Section 5 will deal with the origin of the markers involved in these nominalizations as well as in the evolution undergone by these nominalized expressions. This historical analysis will show that the three different

1. Some other Cahita scholars seem to follow the same idea, although they did not propose an analysis of the constructions under study here. For instance, in the first available description of Cahita, these constructions are named verbal nouns (Buelna 1890), and although Johnson (1962: 43–44) and Lionnet (1978: 31, 43, 78) label the suffixes involved in these constructions as relative markers, they also recognize that the constructions suffixed by these markers function as nouns.

sources (possessed classifier, subject nominalizer, and spatial postpositions) are all noun phrase markers and that the referential use is chronologically the first use of these constructions. This diachronic study will also propose that the object nominalizations in Cahita are achieved through contextual inferences that have caused a chain of metonymic shifts from possessed nominalizations. In Section 6, some final remarks will be provided in regard to how these origins and evolution support the nominalization approach to relativization assumed in this paper.

2. The nominalization approach to relativization

This approach, which has been recently presented in Álvarez (2016), is closely related to Shibatani's recent work on nominalization (2009, 2010, this volume). I summarize here the information provided in Álvarez (2016).

2.1 The notion of nominalization

Nominalization is understood here as a process that creates nominal constituents, which independently of its internal structure, can be associated with two different functions: (i) a referential function when used as NP-head, this use corresponds to the prototypical uses of nominal constituents and (ii) a noun-modifying function when used as NP-head modifier.

Depending on the lexical status of the nominalized expression, two different types of nominalization can be distinguished: **lexical nominalization** that creates nominalized forms with lexical status, and **grammatical nominalization** that creates nominalized forms without lexical status. Both kinds of nominalization are exemplified in 1 with English data:

(1) a. [*The work-er*] *is drunk.* agentive lexical nominalization
b. [*The one who worked*] *is drunk.* agentive grammatical nominalization

As shown in 1, English uses a different strategy for each kind of nominalization: a synthetic strategy for agentive lexical nominalization (Example (1a) with the suffix *-er*) and an analytic one for agentive grammatical nominalization (Example (1b)). Other languages use the same strategy for both processes, as can be observed in Examples 2 from Yaqui with the nominalizer *-me*.

(2) a. *u tekipanoa-**me*** agentive lexical nominalization
DET work-NMZR_SBJ
'the worker'

b. *u tekipanoa-ka-**me*** agentive grammatical nominalization
DET work-PERF-NMZR_SBJ
'the one who worked'

Depending on the denotational properties of nominalized forms, it is possible to distinguish two other types of nominalizations: **action/state nominalizations** and **argument nominalizations** (Comrie & Thompson 1985; Shibatani 2010).

Action/state nominalizations denote a state of affairs characterized by the action or state denoted by the verbal base. The resulting nominalized expression serves to name the action or state designated by the base verb. So, action nominalizations are created out of action verbs, and state nominalizations are created out of state verbs.

Argument nominalizations denote an entity in terms of its involvement in the state of affairs denoted by the verbal base. The resulting nominalized expression serves to represent one of the arguments of the base verb. Depending on the syntactic/semantic role assigned to this argument/entity with regard to the verbal base, different argument nominalizations are possible: for instance, subject, object, and oblique nominalizations if we consider syntactic roles, or agent, patient, location, and comitative nominalizations if we consider semantic roles. Examples in (3) illustrate these kinds of grammatical nominalization in English:[2]

(3) a. Action nominalization:
[*The teacher's examining*] *is continuous*.
b. State nominalization:
I don't enjoy [*my being in love with you*].
c. Argument nominalization:
i. Subject/agent: *(the one) who broke the window*
ii. Object/patient:[3] *what I broke / the one that I broke*
iii. Oblique/locative: *where he works*
iv. Oblique/comitative: *(the one) with whom I have laughed*

2. The distinction between event and argument nominalizations also operates for lexical nominalization. For instance, from the English action verb *examine,* we have the following derived lexical nouns: *examin-ation* (action nominalization), *examin-er* (subject/agent nominalization), and *examin-ee* (object/patient nominalization).

3. Argument nominalizations can be syntactically-oriented (selecting the syntactic function independently of the semantic role, only subjects for instance), semantically-oriented (selecting the semantic role independently of the syntactic function, only patients for instance), or both syntactically and semantically-oriented (selecting a syntactic function and a semantic role, only patientive objects for instance).

2.2 The notion of relativization

Traditionally, relativization is seen as a type of clause linkage strategy. It can be viewed as a process creating Relative Clauses (RCs), which are considered to be part of a multi-verb construction, representing one of the different construction types of interclausal connectivity, along with serial verbs, complement clauses, adverbial clauses, clause-chains and coordination (Payne 1997: 306).

Syntactically, RCs would be dependent on the matrix clause, exhibiting a type of subordination. Andrews (2007: 206) points out this relation of dependency and embedding between two clauses when he defines a RC as "a *subordinate clause,* which delimits the reference of a noun phrase by specifying the role of the referent of that noun phrase in the situation described by the RC." Semantically, this definition recognizes the noun-specifying function associated to RCs. Comrie and Kuteva (2005) insist on the same function when they define a RC as "a clause narrowing the potential reference of a referring expression by *restricting the reference* to those referents of which a particular proposition is true." (See the critical assessments of these definitions of RCs by Shibatani, this volume.)

Thus, a RC would function as a nominal modifier (Keenan 1985) by restricting the semantic domain covered by a syntactic constituent (typically a noun termed the "head noun" or the "relativized noun"). Based on these definitions, it is clear that the terms "headless/free" and "non-restrictive" RCs are problematic since none of these constructions fulfills the defining function of RCs (see Section 1.5).

So far, it is easy to identify the major differences between nominalization and relativization: nominalization is associated with nominal constituents conveying prototypically a referential function, whereas relativization is associated with clausal constituents conveying a noun-modifying function.

However, a nominalized expression may also have a noun-modifying function if used with another nominal constituent in an appositive construction, opening thus a possibility for syncretism in some languages.

2.3 The notion of apposition

An apposition is a construction that combines two nominals: the anchor and the appositive (Kubo 2009). The anchor precedes the appositive and serves as the appositive's referent, whereas the nominal constituent following the anchor represents the appositive. An apposition implies thus a co-referential relationship between two nominal constituents.

Depending on the function of the appositive, two types of appositive expressions are distinguishable (Kubo 2009: 27):

– **Non-restrictive apposition:** The appositive functions to provide the preceding noun with additional information (an alternative description) but is not needed to identify the reference of the anchor. In this type, there are two heads and the appositive expression has a referential function. The prosodic properties correspond to a detached appositive, i.e. the anchor and the appositive are in different information units, which are indicated in speech by their inclusion in separate tone units and in writing by commas.

	NP		**NP**		
	[anchor]		[appositive]		
English: (4)	*My*	*daughter,*		*Mary,*	*will come.*
Yaqui: (5)	*U*	*yoeme,*	[*(U)*	*tekipanoa-ka-me],*	*naamukia.*
	DET	man	DET	work-PERF-NMZR_SBJ	drunk

'The man, the one who worked, is drunk".

– **Restrictive apposition**: here the function of the appositive is to identify the reference of the preceding noun by restricting the denotation of the anchor. In this type, the anchor is the head of the construction and the appositive has a modifying function. In prosodic terms, the nominal constituent following the anchor is an integrated appositive: the restrictive appositive and the anchor cannot be separated by a comma or pause, and the appositive must be immediately adjacent to the anchor.

	NP		**Modifier**	
	[anchor]		[appositive]	
English (6)	*My*	*daughter*	*Mary*	*will come.*
Yaqui: (7)	*U*	*yoeme*	[*tekipanoa-ka-me*]	*naamukia.*
	DET	man	work-PERF-NMZR_SBJ	drunk

'The man who worked is drunk".

Interestingly, Yaqui examples of a non-restrictive apposition in (5) and of a restrictive apposition in (7) show that, besides the prosodic properties of the appositive, the possible presence of the determiner *U* in (5) represents another evidence of the NP-use for the appositive nominalization in Yaqui, because this determiner is not possible in (7).

2.4 The nominalization/relativization syncretism

The syncretism between grammatical nominalization and relativization arises when argument grammatical nominalizations are used in a noun-modifying function as a restrictive appositive. In this case, the anchor is the head noun and the restrictive appositive is an argument grammatical nominalization functioning as a

head-noun modifier. The syncretism between relativization and nominalization is then originated by the modification use of an argument grammatical nominalization in apposition.

Even though the relation between relativization and nominalization is cross-linguistically solid, it is important not to consider both processes as the same. Indeed, nominalization may be used in some languages for relativization purposes but it is more generally used for another purpose, distinct from relativization: to refer to an entity, i. e. the defining function of nominal constituents.

In sum, grammatical nominalization turns finite clauses into nominalized expressions that may be associated with two distinct functions, according to the syntactic context in which they appear:

– a **referential function** when the nominalized expression is used as the head of an independent NP (8) or as the head of an appositive-NP within a non-restrictive apposition (5).

(8) [*U tekipanoa-ka-me*] *naamukia*. 'The one who worked is drunk'.

– a **noun-modifying function** when the nominalized expression is used as appositive within a restrictive apposition (7). In this case, the nominalized appositive expression functions as a RC, modifying the head noun (the anchor).

2.5 Terminological problems in the domain of RCs

As already observed by some scholars (Shibatani & Awadh 2009; Queixalos 2012; Shibatani this volume), this nominalization approach strongly questions the well-foundedness of the notions of "headless" and "non-restrictive" RCs.

When a nominalized expression is used alone (i. e. without an anchor) as in (8), it has a referential function. This situation logically leads to the reinterpretation of the so-called "headless relative clauses" as simply nominalized expressions. In fact, these constructions do not fit the definitions of relative clause constructions given above, for two good reasons: first, they do not have a modifying function but a referential one (there is no overt head to modify, so they are "relative" to nothing), and second, they are not clauses, but nominal expressions. Thus there is no reason to call these constructions "headless relative clauses" (also known as "free relatives"), and viewing them from the perspective of RCs is misleading. They have to be considered as grammatically nominalized expressions standing alone.

The difference between restrictive and non-restrictive appositions matches up with the distinction between what has been called "restrictive" and "non-restrictive" RCs. However, in search of terminological consistency, the term "non-restrictive RC" has to be avoided because this kind of construction does not have the

defining function of a RC. A dependent clause with a non-restrictive function cannot be a RC, since the defining function of a RC is precisely a restrictive function.

Several authors consider indeed that non-restrictive RCs are not true RCs since they merely make a comment about a referential entity without delimiting its reference (Keenan 1985: 168; Carlson 1977; Lehmann 1984).

Both designations – non-restrictive and headless/free RCs – represent thus an inconsistent use of terminology, since none of the so-called constructions fits the defining function of a RC. They may be structurally similar to RCs, but they do not have the same function. In languages with nominalization/relativization syncretism it is easy to solve this naming problem, because these constructions are in fact all nominalized constructions, which can have two different functions depending on the context:

- RCs are grammatical nominalizations used as appositives in restrictive apposition contexts (modifying function).
- Non-restrictive RCs are grammatical nominalizations used as appositives in non-restrictive apposition contexts (referential function).
- Headless/free RCs are grammatical nominalizations used as standing alone independent NPs in non-apposition contexts (referential function).

The nominalization approach of relativization proposed here is not just a matter of labels, it is correlated to a proper understanding of these linguistic structures. Treating these constructions as argument grammatical nominalizations instead of RCs makes it possible to cover the three types of construction without creating inconsistency by using the term RC for constructions that do not have the function of a RC. Additionally, by using this nominalization approach a clear distinction is made between form/structure (grammatical nominalization) and use/function (referential or modifying function).

2.6 The origin of the nominalization/relativization syncretism

This confusion between structure and function in the domain of RCs is also present when the source of RCs is addressed. Considering the diachrony of RCs, Givón (2009, 2012) proposes that one way to the RC formation is through the Paratactic channel illustrated in (9):

(9) **The Paratactic Channel to RC**
Diachronic pathway of the nominalized RCs
Headless RC > Non-restrictive RC > Restrictive RC

Givón (2012: 11) explains this pathway saying that: "The early source of modifying RC is often a headless RC used first as a paratactic, non-restrictive RC. All it takes

to move from non-restrictive (paratactic) to restrictive (syntactic/embedded) RC is the merger of intonation contours." (Givón 2012: 11).

This diachronic explanation insists on the fact that an initial structure (named by Givón a headless RC) is used first with a non-restrictive function in a parataxis, and second with a restrictive function in a hipotaxis. The use of the term 'RC' at the three stages implies that the same structure (RC) is used in three different contexts (headless, paratactic, and syntactic) with different functions (non-restrictive, restrictive). Besides the fact that no function is mentioned to be associated with the first stage (headless RC), this evolution shows that the real function of relativization is achieved only at the third stage. So, if Stages 1 and 2 are two stages previous to relativization, there is no good reason to name the structures involved at Stages 1 and 2 as RCs. Therefore, instead of using the term RC for the first two stages structures, it is then preferable to use the name of the first structure in order to name the others two.

As explained above, 'headless' RCs are in fact nominalized expressions used as independent NPs. Givón seems to admit this nominalized origin when he claims that one strong evidence from Ute (Uto-Aztecan) in favor of this evolution lies on the fact that "the non-restrictive RC still functions synchronically as nominalized clauses" (Givón, 2012: 14), as illustrated in the following example given by Givón:

(10) *puchuchugwa-y* [*'uway yoghovuchi 'ua-qa-tu*]
know-IMM-3S that/O coyote/O trap-ANT-NMZR
'(I) know *the one who trapped the coyote*'.

In fact, the construction in (10) is not a non-restrictive RC but a 'headless' RC; that is, the first stage within the Paratactic channel. However, contrary to the Givón's comment that seems to indicate that the notion of nominalized clause is restricted to 'headless' RCs, we consider that the three types of "RC" are in fact three types of argument grammatical nominalization used in different syntactic contexts and with two possible functions (referential and modifying function). In consequence, the Paratactic channel presented in (9) can be re-expressed as:

(11) **Diachronic pathway of the restrictive appositive grammatical nominalization:**

Syntactic context:	Non-appositive	> Non-restrictive appositive	> Restrictive appositive
STRUCTURE:	GRAMMATICAL NMZL	GRAMMATICAL NMZL	GRAMMATICAL NMZL
Function:	Referential Function	> Referential Function	> Modifying Function

The evolution in (11) represents the nominalization approach of the evolution presented in (9). This approach shows that relativization has to be considered as merely one specialized function of nominalization – namely, the modifying function of an appositive restrictive grammatical nominalization. As mentioned above,

grammatical nominalizations, like any nominal constituents, can have two different functions (referential and noun-modifying functions), depending on the syntactic status of the nominalized expression, and diachronically, the referential function is prior to the modification function. In consequence, the noun-modifying function is one of the nominalization functions in its own right and hence nominalizations are not derivatives of relative clauses. Therefore, it is preferable to view relativization from the perspective of nominalization, not the other way round.

In the following, I will present how the origin and the evolution of non-subject nominalizers in Cahita support this nominalization approach to relativization. But before that, I will briefly introduce the Cahita language(s).

3. The Cahita language(s)

Cahita is represented by Yaqui, Mayo, and Tehueco, which belong to the Taracahitan branch of the Sonoran group within the Southern Uto-Aztecan languages. The different languages included in this Southern division of the Uto-Aztecan family are presented in Table 1.

Table 1. The Southern Uto-Aztecan languages (adapted from Miller 1984)

Sonoran
- a. Tepiman: Upper Piman (Tohono O'odham, Akimel O'odham, †Nevome), Lower Piman (Pima Bajo, Northern Tepehuan, Southern Tepehuan, †Tepecano)
- b. Taracahitan:
  - Tarahumaran: Rarámuri (Tarahumara), Guarijío
  - Opatan: †Opata, †Eudeve, (†Jova?)
  - *Cahita: Yaqui, Mayo,†Tehueco*
  - Tubar: †Tubar

Corachol-Aztecan
- a. Corachol: Cora, Huichol
- b. Aztecan:
  - †Pochutec
  - General Aztec (or Nahuatl): Pipil, Aztec (many varieties)

Although the term "languages" is commonly used when referring to Yaqui, Mayo, and Tehueco, since they are structurally very similar, it is possible to recognize them as three varieties of the same language, that is to say the Cahita language, as named in the first available documentation of this linguistic group.

The *Arte de la lengua Cahita escrita por un Padre de la Compañía de Jesús* that represents the earliest known description of the Cahita language, was printed in Mexico City in 1737 by Francisco Xavier Sanchez although it was probably

compiled prior to 1650 (Dedrick & Casad 1999: 3; Álvarez 2018). This first version was later edited and published in 1890 by Eustaquio Buelna, who recognized in his introduction (Buelna 1890: X) that the Cahita language is represented by three dialect variants: Yaqui, Mayo, and Tehueco. In his preface, the author of the *Arte* comes to the same conclusion, considering that, in spite of their differences, Yaqui, Mayo, and Tehueco can be viewed as the same language (Buelna 1890: 5).

According to Buelna (1890: X), Tehueco was spoken by three different indigenous tribes, the Sinaloas, Tehuecos, and Zuaques, who were settled on the bank of the Fuerte River in the north of the present-day state of Sinaloa in northwestern Mexico. In the days of the *Arte,* Yaqui was spoken by indigenous people located along the Yaqui River, a little closer to the north, in the south of the state of Sonora, while Mayo was spoken by a community settled on the bank of the Mayo River, between the Yaqui territory to the north and the Tehueco territory to the south. Nowadays the Mayo language has extended its territory, occupying its original location (Mayo of Sonora) and the former location of Tehueco (today, Mayo of Sinaloa). It has almost 40,000 speakers, whereas Yaqui has approximately 15,000 speakers, located in their same original homeland along the Yaqui River.[4]

The linguistic forms documented in the *Arte* come from Tehueco, but the original author was careful to point out throughout the *Arte* the existing differences between Tehueco and the other two Cahita variants. Buelna (1890: XI) recognizes that these differences are very few in number, and thus it is possible, as the title suggests, to view the linguistic information provided in the *Arte* as data from the Cahita language (including Tehueco, Yaqui, and Mayo) of the first half of the seventeenth century. I will refer to it here as Old Cahita.

The Tehueco variant has since disappeared, probably assimilated by Mayo during the eighteenth century (Álvarez 2018). Although present-day speakers of Mayo and Yaqui can still communicate with each other quite readily (Escalante 1990: 16; Moctezuma & López 1990, 1994; Álvarez & Felix 2011), both communities believe that they do not share the same language, mostly for sociocultural and historical reasons (Moctezuma & López 1991). In this paper, I will use the term "language" when referring to Mayo or Yaqui, in spite of the fact that they can be considered variants of the same language (Cahita) on strictly linguistic (structural) grounds. Both languages are examples of Modern Cahita.

4. There is in the USA another Yaqui community of approximately 5,000 speakers. This community moved to the bordering state of Arizona from its original homeland in the south of the state of Sonora in Mexico at the beginning of the twentieth century, fleeing persecution by the Mexican dictator Porfirio Diaz.

Given that this work is concerned by the diachrony of non-subject nominalizations in Cahita, in the next sections I will compare data from Old Cahita with data from Modern Cahita.

Cahita, represented by Yaqui, Mayo, and Tehueco, is a nominative-accusative language with explicit marking in nominals (only for accusative: suffix *-ta*) and with different sets of paradigm in the pronoun system. The suffix *-ta* also serves as the genitive case marker with full nouns. It is an agglutinative language with a very predominant use of suffixes and postpositions. The unmarked order of constituents is SOV, except with pronominal subject arguments where the order tends to change to OSV. Determiners are optional, especially in object position.

4. Non-subject nominalizations in Cahita

The subsections of the present section present the different types of non-subject nominalizations in the Cahita languages, focusing first on Old Cahita (from the data in the *Arte)* and then the Modern Cahita languages (Yaqui and Mayo). At the end of this section, a diachronic perspective will be adopted, and a comparison will be provided in order to highlight the evolution undergone by these nominalization processes and also the changes in the domain of relativization, due to the syncretism between nominalization and relativization in Cahita. In Section 5, the origin of the different non-subject nominalizers involved will be addressed.

4.1 Old Cahita

From the information provided in the *Arte de la lengua Cahita,* we know that non-subject nominalizations were rendered by two means in Old Cahita: the suffix *-ye* was used for object nominalization, and this same suffix was combined with the locative suffix *-po* for locative nominalization.

Examples of object nominalizations in Old Cahita marked by the suffix *-ye* are given in (12).[5] This type of construction could function either as a referential nominalized expression (12a) or as a modifying nominalized expression in an appositive construction (12b); that is, with a pre-posed nominal head. In both

5. Interestingly, the anonymous author of the *Arte* considered the constructions suffixed by *-ye* as verbal nouns "*nombres verbales en -ye*" (Buelna 1890: 86). Additionally, it is worth noting that, contrary to the other subordinate clauses, no chapter is devoted to relative clauses in this first description of Cahita. Most examples of Old Cahita RCs (in fact, restrictive appositive grammatical nominalizations) provided in this paper are extracted from the Catechism also included in Buelna (1890).

cases, the agent of the action denoted by the verbal base is marked by a possessive pronoun preceding the nominalized verb. The examples in (12) show that these grammatical nominalizations could receive any marking of finiteness (such as temporal/aspectual markers).

(12) **The suffix *-ye***

– Object NMLZ (Headless Object-RC) **Referential use**

a. *in-eria-**ye*** *in-eria-ca-**ye***
1SG.POS-love-NMZR_OBJ 1SG.POS-love-PERF-NMZR _OBJ
'the one that I love' 'the one that I loved'
*in-eria-naque-**ye***
1SG.POS-love-FUT-NMZR_OBJ
'the one that I will love'

– Object NMLZ Appositive (Object-RC) **Modifying use**

b. *ioreme* [*em-veb-tevo-**ye***]…
person 2SG.POS-beat-order-NMZR_OBJ
'the person that you order to beat…'

For locative nominalizations in Old Cahita, the object nominalization suffix *-ye* was combined with the locative suffix *-po*. Here too we can observe the two possible functions of a nominalized expression: a referential function when it is used alone (13a), and a modifying function when a noun (functioning as an anchor) precedes it (13b). The possessive marking corresponds again to the participant who carries out the action denoted by the base verb; apparently no temporal/aspectual restriction is present.

(13) **The suffix *-ye* + the locative suffix *-po***

– Locative NMLZ (Headless RC) **Referential use**

a. *in-hibua-**ye-po***
1SG.POS-eat-NMZR-LOC
'(the place) where I eat'

– Locative NMLZ Appositive (Locative oblique-RC) **Modifying use**

b. *tapeti* [*in-voie-**ye-po***]
bed 1SG.POS-lie_down-NMZR-LOC
'the bed where I lie down…'

4.2 Modern Cahita

4.2.1 *Yaqui*

In Álvarez (2012), I have shown that Yaqui has two nominalizing suffixes that can be added to verbal bases in order to create a non-subject nominalization: the suffix *-'u* for object nominalization, and the suffix *-'Vpo* for locative nominalization.

These nominalized expressions have a referential function if used alone without a preceding noun (Examples (14a), and (15a)) and a modifying function if used as an appositive, that is to say, following a head noun (Examples (14b), and (15b)).

(14) **The suffix *-'u***
- Object NMLZ (Headless Object-RC) **Referential use**

a. (*u*) *in tea-ka-**'u***
DET 1SG.POS find-PERF-NMZR_OBJ
'what I found'

- Object NMLZ Appositive (Object-RC) **Modifying use**

b. *U chu'u* [*in tea-ka-**'u***] *chukuli.*
DET dog 1SG.POS find-PERF-NMZR_OBJ black
'The dog that I found is black.'

(15) **The suffix *-'Vpo***
- Locative NMLZ (Headless RC) **Referential use**

a. (*u*) *nim bo'o-pea-**'apo***
DET 1SG.POS sleep-DES-NMZR_LOC
'(the place) where I want to sleep'

- Locative NMLZ Appositive (Locative oblique-RC) **Modifying use**

b. *Wa kari* [*nim bo'o-pea-**'apo***] *ujyooli.*
DEM house 1SG.POS sleep-DES-NMZR_LOC pretty
'That house where I want to sleep is pretty.'

4.2.2 *Mayo*

Like Yaqui, Mayo has two nominalizing suffixes used for non-subject nominalizations (De Wolf 1997, Peña 2012): the suffix *-'Vwi* for object nominalization (Examples (16a)–(16b)) and the suffix *-'Vpo* for locative nominalization (Examples (17a)–(b)). The referential uses of the nominalized expressions with the corresponding suffixes are illustrated in (16a) and (17a), and the modifying uses in (16b) and (17b).

(16) **The suffix *-'Vwi***
- Object NMLZ (Headless Object-RC) **Referential use**

a. *em neeréwwa-ka-**'awi***
2SG.POS lend-PERF-NMZR_OBJ
'what you lent me'

- Object NMLZ Appositive (Object-RC) **Modifying use**

b. *Machéeta-m* [*em neeréwwa-ka-**'awi***] *ne-híppure.*
machete-PL 2SG.POS lend-PERF-NMZR_OBJ 1SG.NOM-have
'I have the machete you lent me.'

(17) **The suffix *-'Vpo***

– Locative NMLZ (Headless RC) **Referential use**

a. [*Póhporo-m em é'ekria-**'apo***] *ka-ne ta'aya.*
match-PL 2SG.POS keep-NMZR_LOC NEG-1SG.NOM know
'I don't know (the place) where you keep the matches.'

– Locative NMLZ Appositive (Locative oblique-RC) **Modifying use**

b. *Inapo hinu-k u-ka ka:-ta* [*in tomte-ka-**'apo***]
1SG.NOM buy-PERF DET-ACC house-ACC 1SG.POS born-PERF-NMZR_LOC
'I bought the house where I was born.'

All these constructions from Old and Modern Cahita exhibit several synchronic features clearly indicating that they are nominalizations. For instance, they receive the same syntactic treatment as other nominal constituents in the following aspects:

i. Like any possessed noun, they always contain a possessive pronoun, signaling that the subject/agent of the base verb is metonymically interpreted as the possessor of the action denoted by the verb. It is a clear indication that the clause has been nominalized.
ii. They can have the same syntactic functions as other nominals, like subject, object, or noun-modifier.
iii. They combine with the same determiners (articles, demonstratives) with the same optionality.

Additionally, most uses of nominalizations are referential rather than modifying. As can be observed in Table 2, the study on the frequency of these nominalized structures in Yaqui narrative and dialogic texts shows that the most frequent, and therefore the primary, function is overwhelmingly referential; that is, the prototypical function of nominal constituents (91.18% for the suffix *-(')u*, 80.95% for the suffix *-'Vpo*).

Table 2. Frequency of referential and modifying uses for non-subject nominalizations in Yaqui discourse

Text type	Suffix *-'u*		Suffix *-'Vpo*	
	Modifying function	Referential function	Modifying function	Referential function
Narrative: 367 clauses	1 (4.76%)	21 (95.24%)	3 (18.75%)	13 (81.25%)
Dialogic: 348 clauses	2 (16.67%)	10 (83.33%)	1 (20%)	4 (80%)
All texts	3 (**8.82%**)	31 (**91.18%**)	4 (**19.05%**)	17 (**80.95%**)

Total modifying use in all texts for both suffixes combined: 7 (12.72%)
Total referential use in all texts for both suffixes combined: 48 (87.28%)

Diachronically, the comparison of the non-subject nominalizations marking in Tehueco, Yaqui and Mayo reveals that the markers used for this kind of nominalization have changed over time. This evolution is summarized in (18):

(18) **Evolution in Cahita non-subject nominalization marking:**

	Old Cahita	**Modern Cahita**
Object Nominalizer:	*-ye*	> *-'u* (Yaqui), *-'Vwi* (Mayo)
Locative Nominalizer:	*-ye* + *-po*	> *-'Vpo* (Yaqui, Mayo)

The nominalizer *-ye* that was used in Old Cahita for object and locative nominalizations has been lost and substituted in the Modern Cahita languages with the suffixes *-'u* (Yaqui) / *-'Vwi* (Mayo) for object nominalizations. In the case of locative nominalizations, the combination of the suffix *-ye* and the locative suffix *-po* has been reduced to *-'Vpo* (Yaqui/Mayo). Based on this last change, it seems easy to propose that the echo-vowel[6] present in the locative nominalizer of Modern Cahita is the reflex of the object nominalizer from Old Cahita (suffix *-ye*). As the same echo-vowel is present in the current object nominalizer in Mayo (suffix *-'Vwi*), it is possible that this Mayo marker is the result of a similar combination. So the evolution in (18) can be restated as:

(19) **Evolution in Cahita non-subject nominalization marking:**

	Old	**Cahita**	**Modern Cahita**
Object Nominalizer:	*-ye*	> **-yeu*	> *-'u* (Yaqui)
		> **-yewi*	> *-'Vwi* (Mayo)
Locative Nominalizer:	*-ye* + *-po*		> *-'Vpo* (Yaqui, Mayo)

Except for the case of the locative suffix *-po*, the evolution in (19) raises the question of the source of the other markers involved in these non-subject nominalizations: the suffixes *-ye, -wi* and *-u*. The next section investigates the origin of these markers and their evolution to the object nominalization function.

5. The origin of the Cahita non-subject nominalization markers

At first sight, when we look for the origin of the Cahita non-subject nominalization markers, it is easy to recognize that this origin seems to be related to postpositions (Álvarez 2016). Indeed, we find in Old and Modern Cahita several postpositions that are formally similar to the different non-subject nominalizers that have been

6. Dedrick & Casad (1999: 28–29) use the term "echo-vowel" for referring to the epenthesis of a glottal stop with a copy of the stem final vowel when a word ending in a vowel is followed by an affix.

presented in the previous section. The object nominalizer *-ye* is similar to the Old Cahita instrumental postposition (20), the locative nominalizer *-'Vpo* is related to the locative postposition *-po* (21), the Yaqui object nominalizer *-'u* and the Mayo object nominalizer *-'Vwi* seem to be associated with a directional postposition, suffix *-u* in Yaqui and suffix *-wi* in Mayo (Examples 22a and 22b, respectively).

(20) *Cuta-**ye***[7] *nee* *mea-c*. **Old Cahita**
wood-**INST** 1SG.ACC hit-PERF
'He hit me with a stick'.

(21) a. *baa-**po*** **Old Cahita**
water-**LOC**
'in the water'

b. *ba'a-**po*** **Modern Cahita (Yaqui/Mayo)**
water-**LOC**
'in the water'

(22) a. *Navojoa-**u*** *ne-siika-k*. **Yaqui**
Navojoa-**DIR** 1SG.NOM-go-PERF
'I went to Navojoa'.

b. *Navojoa-**wi*** *ne-siika-k*. **Mayo**
Navojoa-**DIR** 1SG.NOM-go-PERF
'I went to Navojoa'.

If the presence of the locative postposition in the locative nominalization suffix is quite clear, given the transparent semantic relation between the postpositional meaning and the nominalization meaning, the postpositional source of the other non-subject nominalization markers is more problematic and needs to be explored in more detail. Some issues remain opaque. For instance, what is the relationship between the instrumental and directional meanings and the object/patient nominalization? or what is the reason for combining the instrumental and the locative postpositions in order to form the locative nominalization marker? These questions raise some doubts about the postpositional source of the Cahita object nominalizers. In the following, I will focus on the origin of Cahita non-subject nominalizations, beginning with the suffix *-ye,* and I will propose that this object nominalizer from Old Cahita does not originate from the instrumental postposition but instead both object nominalization and postpositional uses are parallel developments from another origin.

7. The current instrumental postposition is *-e* (plural: *-mea*) in Yaqui (Dedrick & Casad 1999: 187) and *-(y)i* in Mayo (De Wolf 1997: 71).

5.1 The origin of the suffix *-ye*

Cristofaro (this volume) insists on the importance of contexts for explaining the evolutionary paths implied in many nominalization processes, specially for understanding the extension from one context to another. So, if the Old Cahita object nominalizer came from the instrumental postposition, we have to ask what kinds of contexts could explain the extension of the nominalization from the notion of instrument to the notion of patient. Such bridging contexts are hard to find in this case since instruments can be easily reanalyzed as agents, but not so much as patients (Heine & Kuteva 2004).

In fact, it seems that the nominalizing use of this suffix would originate within the possession domain. More specifically, I propose that the nominalizing suffix *-ye* comes from an old possessive classifier that has been reconstructed for inalienable possession in Proto-Uto-Aztecan (PUA).

Dakin (1991) has reconstructed for PUA a system of possessive nominal classification based on the alienability (AL)/inalienability (INAL) distinction (mediated and direct possession in Dakin's terms) and marked by **-*wa*[8] for AL and **-*yɨ* for INAL, as illustrated in (23).

(23) **Proto-Uto-Aztecan** (Dakin 1991: 319)

a. **ni-náka-wa*
1SG-meat-CLF.POS.AL
'my meat'

b. **ni-káma-yɨ*
1SG-mouth-CLF.POS.INAL
'my mouth'

In Proto-Nahuatl, this classification system was reorganized incorporating the animacy distinction as a basic category. According to Dakin (1991), in Proto-Nahuatl, all animate nouns, including kinship terms, were marked by the suffix **-wa:*, whereas the suffix **-ye* was reduced to the marking of INAL inanimate nouns. Concerning this suffix **-ye*, she states:

> Inanimate nouns in Proto-Nahuatl usually directly possessed included body parts and other part-to-whole constructions, **abstract nouns, instrumentals**, probably * *k*^{w}*e-* 'skirt', and perhaps most **patient participial nouns derived from verbs**. When these nouns were possessed in a direct relation, they must all have carried the **-ye* suffix, although most lost it. (Dakin: 1991: 313)
>
> (Bolds are mine)

8. Following the convention of Campbell & Langacker (1978) also used in Dakin (1991), * indicates reconstructions for a single branch of the family (Proto-Nahuatl, Proto-Tepiman, etc.) and ** indicates Proto-Uto-Aztecan (PUA).

Dakin (1991) recognized traces of the suffix *-*ye* in the following cases pertaining to the uses of the suffix *-yo:/-ya:*[9] in Classical Nahuatl:

- with some possessed body-parts: *to-pa:k-ya* 'our kidneys'
- for distinguishing alienable and inalienable possession for those body and plant parts which were also used for food or other purposes: *no-naka-yo* 'my flesh', *no-nak* 'my meat'
- for forming abstract nouns as the following: *te:n-yo:-λ* 'fame', *to:na-ka:-yo:-λ* 'sustenance'[10]
- for forming possessed instrumentals: *in teki-lo:* 'knife', *no-λa-tek-ya* 'my knife'

Although Dakin does not propose an example of a patient participial noun, there are several examples in which the suffix *-yo/-ya* seem to be combined with verbal bases. For instance, the abstract nouns seem to be derived from verbs (*te:neua-* 'to make famous' and *to:na-* 'to thrive') as well as the instrumental (from the verb *teki* 'to cut'). These cases would be clear examples of nominalizations, as shown by the abstract nouns examples given by Dakin that can be seen as action/result nominalizations. And what Dakin (1991: 313) names 'patient participial nouns derived from verbs', clearly corresponds to what I call "object nominalizations".

So, I think there are good reasons to consider that the non-subject nominalizer *-ye* of Old Cahita comes from the PUA inalienable possessive classifier **-yi*, which has been in a first stage associated with inanimate possessed nouns (as *-*ye* in Proto-Nahuatl), opening thus the possibility to combine with possessed deverbal nouns; that is, possessed (action/result/patient) nominalizations.

Several arguments can be proposed for supporting this origin:

- The inanimate possessive noun classifier is a noun phrase marker and, as Yap et al. (2011: 22) have pointed out, "noun phrase markers are frequently used to signal nominalization constructions, and some of these noun phrase markers may be reanalyzed as nominalizers themselves." This noun phrase marker can thus serve as a "substantivization strategy" whereby a clause can be more readily recognized as a nominalization construction. (See Shibatani this volume)
- Cross-linguistically, the change from a possessive morpheme to a nominalizer and from classifier to nominalizer are two common grammaticalization paths (Yap et al. 2011; Gerner 2012) (See Shibatani this volume)

9. Dakin (1991: 313–314) considers probable that the *-*yo:* form comes from the **-yi* plus the augmentative *-*wi* and that the *-*ya*: form derives from *-*ye* plus the objective case marker *-*a*.

10. Launey (1981: 286) proposes the translation 'our harvest' for *to-t:ona-ca:-yo.*

- The object nominalization is always accompanied by the possessive pronoun that refers to the subject of the nominalized base verb. This possessed object nominalization receives thus the same treatment as any possessed inalienable inanimate nouns in the proto-language; that is, it is suffixed by a reflex of *-*ye*.
- The suffix -*ye* disappears in the locative nominalization marking when the passive suffix -*wa* is involved, since the -*wapo* nominalization does not require the possessor pronominal expression because the passivization implies the agent defocalization (Shibatani 1985). An example from Old Cahita is given in (24).

(24) *hin-ua-po* **Old Cahita** (Buelna 1890: 173)
buy-PASS-LOC
'market/where is bought'

This example represents a clear evidence supporting the fact that the suffix -*ye* is connected to the possession domain, since this marker is absent when the agent/possessor is unexpressed, but present when expressed overtly (compare with examples in (13)).

- Inanimacy is strongly associated with patients/objects, since prototypical patients/objects are inanimate entities (Hopper & Thompson 1980; Givón 2001)
- As seen in (20), the suffix -*ye* is also the instrumental marker in Old Cahita and instruments represent another inanimate semantic role. The object nominalizer and the instrumental marker would thus come from the same inanimate nominal classifier, illustrating two parallel developments from a same origin.
- At least, two other Uto-Aztecan languages show a syncretism between the object-RC marker and a case marker associated with inanimacy, and the marker used for this syncretism in the languages present a formal similarity with **-*yi*. This is the case in Shoshone (Object-RC marker: -*i*, locative marker: -*i*, Miller 1986) and in Luiseño (Object-RC marker: -*i*, accusative marker: -*I*, Hyde 1971).[11]

Considering this origin, a very plausible evolution from the classification function to the object nominalization would be as follows: the suffix -*ye* was firstly used as an inalienable inanimate possessed noun classifier with nominal bases (**Stage 1**) as in Proto-Nahuatl (Dakin 1991). It extended its classification use to verbal bases (**Stage 2**), which have been probably nominalized by conversion (zero derivation). These possessed nominalizations would be associated at this stage with an abstract

11. In these Uto-Aztecan languages, my guess would be that relativization is made through restrictive appositive nominalizations, and that the Object-RC marker is in fact an object nominalizer, like in Cahita.

meaning, such as action/state nominalizations (depending on the base verb type). These abstract nominalizations are then metonymically interpreted as more concrete nominalizations, such as result and object nominalizations (**Stage 3**).[12] Obviously, the object nominalization interpretation only appears with transitive base verbs.

In this reconstruction, the non-subject nominalization marking emerges from a possessive deverbal NP, in which the inanimate possessed classifier used in this possessed deverbal NP is reanalyzed as a nominalizer. Originally, the nominalization is made via zero derivation (conversion); that is, the verb is nominalized by being used in a nominal context (a possessive NP). The classifier suffix originally signals that the possessed element is inalienable and inanimate. The speakers can easily recuperate this information from their encyclopedic knowledge. So, this classification function is not essential to the speakers and they can assign another function to this marker. The reanalysis from inalienable inanimate possessed classifier to nominalizer is then possible, since nominalization is initially realized without any marking (zero derivation) and languages tend to overtly mark this transpositional operation. This evolution would then imply a change from no-marking to overt marking in the domain of nominalization. This evolution is schematized in (25):

12. This concretization of the action nominalization is very common in Indo-European languages (Luján & Ruiz Abad 2014: 243–244). It is the case for instance in English with the possessed lexical nominalization *my creation* that commonly refers more to a result or an object (*what I have created*) than to an action. Result and object are in fact very close, since the result of an action can be interpreted as the object of the verb denoting this action (*I have created a statue = my creation is a statue*). Result nominalization can be thus associated with object nominalization. Comrie and Thompson (1985) recognized the same, naming as objective nominalization, the formation of "nouns designating the result, or the typical or 'cognate' object of an action" (1985: 335).

(25) **Evolution from the classification function to the object nominalization**

Stage 1. Inalienable inanimate possessed noun classifier.		
Nominal base:	my mouth-INAL.INAM.POS.CLF	'my mouth'
Stage 2. Inalienable inanimate possessed deverbal noun classifier		
NMZR=ZERO DERIVATION		
Stative base verb:	my be-INAL.INAM.POS.CLF	'my being'
Intransitive action base verb:	my act-INAL.INAM.POS.CLF	'my acting', 'my act'
Transitive action base verb:	my love-INAL.INAM.POS.CLF	'my loving', 'my love'
Stage 3. State/Action/Result/Object nominalization marker		
REANALYSIS:	CLASSIFIER > NMZR	
Stative base verb:	my be-NMZR	'my being'
Intransitive action base verb:	my act-NMZR	'my acting', 'my act'
Transitive action base verb:	my love-NMZR	'my loving', 'my love', 'the one/ what I love'

Stage 3 corresponds to the Headless RC stage of the Paratactic Channel proposed by Givón (2012); that is, when the nominalized expression is used as an independent NP standing alone. The evolution to the relativization function is achieved through two more stages. Firstly, the object nominalization interpretation associated with transitive base verbs is reinforced when the nominalized expression is used with a pre-posed noun in a non-restrictive apposition (**Stage 4** in (26)), since the relationship between the pre-posed noun (the anchor) and the nominalized expression (the appositive) is inferred as being coreferential. The syncretism between object nominalization and object RC appears when the anchor and the appositive object nominalization fused into the same intonation contour (**Stage 5** in (26)). This evolution is schematized in (26):

(26) **Evolution from object nominalization to object relativization**

Stage 4. Object NMLZ, with a pre-posed noun in a non-restrictive apposition		
book, my love-NMZR_OBJ	'the book, my love'	'the book, the one I love'
woman, my love-NMZR_OBJ	'the woman, my love'	'the woman, the one I love'
Stage 5. Object NMLZ, with a pre-posed noun in a restrictive apposition		
woman/book my love-NMZR_OBJ	'the woman/book my love'	'the woman/book I love'

When Stage 5 is achieved and consolidated, the use of the nominalizer can no longer be associated with the action/result interpretation, like it did in Stage 3. The change to object nominalization is fully achieved and henceforth this nominalized expression is limited to the object interpretation illustrated in (27).

(27) **Final stage**

Stage 6.	Object NMLZ standing alone (Headless Object-RC)
my love-NMZR_OBJ	'the one/what I love'

In this evolution, we can observe that Stage 2 is the bridging context to the nominalization function; that is, the context that allows the reanalysis of the possessive classifier as a nominalizer.

Functionally, it is clear that the first nominalization function is referential (**Stage 3**) and that the noun-modifying function appears later via a restrictive apposition (**Stage 5**). Diachronically, the noun-modifying function is the result of the evolutionary process within the nominalization domain, a development from the referential function that is the primary function of nominalization.

Regarding how this change is developed, this evolution in the domain of possessed nominalizations is made through metonymic shifts, first from Action to Result and second from Result to Object nominalization. The object interpretation is obtained with transitive base verbs and reinforced via apposition because of the inferred co-referentiality between the pre-posed noun (the anchor) and the appositive nominalization. This evolution also shows the specialization of the suffix *-ye* that has reduced its use from all kinds of base verbs to transitive base verbs only, but also its generalization since it has generalized its use from inalienable inanimate entities to all kind of object entities.

5.2 The origin of the suffixes *-'u/-'Vwi* in Yaqui and in Mayo

Based on the evolution of the object nominalizers presented in (19) and on the formal similarities between the object nominalizers and the directional postpositions in Modern Cahita (see examples in (22)), it could be proposed that the suffixes *-'u/-'Vwi,* in Yaqui and in Mayo respectively, come from the combination between the old possessed classifier reanalyzed as an object nominalizer, the suffix *-ye,* and the corresponding directional postposition *-u/-wi*, copying thus the strategy at the origin of the locative nominalizer formation. In order to support this hypothesis, since object clauses do not specifically involve the notion of direction, it is important to explore whether the use of the nominalization in these clauses may have developed in contexts involving this notion. Contrary to what happens between instrumental and object/patient meanings, although no direct connection can be established between the notion of object/patient and the original directional meaning of the nominalizer, some contexts are clearly compatible with both. For instance, the object nominalization use of the directional postposition could have developed from the directional use through processes of inference

in contexts such as '*he looked towards what I was doing*' / '*he saw what I was doing*', which are compatible both with a directional interpretation and an object nominalization one.

Another possible bridging context for the extension from the notion of direction to that of object/patient is illustrated in (28) and (29). In this case, we can see that the forgotten entity was introduced in Old Cahita by a directional postposition (28), whereas in current Yaqui the directional marker is no longer required (29).

(28) *Emo-u ne copte.* **Old Cahita**
2SG-DIR 1SG.NOM forget.INTR
'I am forgetting (to) you'.

(29) a. [*In yaa-bae-'u*] *ne kopta-k.* **Yaqui**
1SG.POS DO-DES-NMZR_OBJ 1SG.NOM forget.TR-PERF
'I forgot what I was going to do'.
b. *Tomi-ta ne kopta-k.*
money-ACC 1SG.NOM forget.TR-PERF
'I forgot the money'.

In such contexts, the nominalizing use of the suffix *-'u* could plausibly have derived from the directional meaning through processes of context-driven inference, as described in grammaticalization studies and studies of language change in general (Heine 2003; Traugott & Dasher 2005, among many others). This grammaticalization from directional (allative) to object/patient is well-known in Spanish, where the human/definite object marker *a* comes from a directional preposition, and it has been also documented for instance in Imonda (Seiler 1985: 165) and in Lezgian (Haspelmath 1993: 89),[13] both cases cited by Heine & Kuteva (2004: 38).

The change from the directional postposition to the object nominalizer would imply that the construction is reanalyzed as transitive, as confirmed in (29). Interestingly, we can observe that the nominalized clause corresponds to the object of the transitive construction in (29a), but it does not receive the accusative marking, contrary to the nominal object in (29b). The absence of the accusative marking in the object nominalization could thus be a consequence of the postpositional source construction, a result of the original structure. Besides the fact that the construction has changed from adjunct to argument, this evolution also shows that the first nominalization use would have been referential.

13. Interestingly, in Lezgian, the directional marker, suffix *-z*, has been grammaticalized as an object marker but only with perception verbs like 'see'; that is, contexts in which both directional and stimulus/object interpretation are compatible.

Another strong argument in favor of the directional origin of the object nominalizer in Modern Cahita is the fact that the directional meaning is still possible when the main verb is a motion verb, as illustrated in the following examples from Dedrick & Casad (1999: 383) for Yaqui, and De Wolf (1997: 223) for Mayo. Interestingly, contrary to the object nominalization uses of suffixes *-'u* and *-'Vwi,* which are always combined with transitive verbs, in these examples, the object of the postposition is an intransitive clause and this clause denotes the place to where the subject participant is going.

(30) [*Ba'á-ta kó'om-sika-'u*] *née wée-bae*. **Yaqui**
water-POS down-go.SG-NMZR_DIR I go.SG-DES[14]
'I am going to Where the Water Goes Down'.

(31) *'Áapo kom-siíka* [*bá'a-m 'ayuka-'awi*]. **Mayo**
3SG.NOM down-go water-PL exist-NMZR_DIR
'He/she went downward, to where there is water'.

In these constructions, the use of suffixes *-'u* and *-'Vwi* seems to convey both directional and nominalizing functions; that is, these markers appears to function as directional nominalizers. But in fact, according to the reconstruction proposed in (19), the nominalizing function should be rendered by the glottal stop and the echo-vowel present respectively in *-'u* and *-'Vwi,* which are the reduced forms of the object nominalizer *-ye* from Old Cahita. However, probably facilitated by the loss in phonetic substance of the suffix *-ye* and by the influence of the locative nominalizer in Modern Cahita, the delimitation between these two different markers is blurred and the sequence is reinterpreted as a whole. Additionally, this directional nominalization use may change to an object nominalization in some new contexts. Indeed, when the directional meaning is bleached as a consequence of using this type of complement with transitive main verbs that are not motion verbs, the combination of the reduced form of *-ye* and the directional postpositional as a whole can be reanalyzed (only when the verb base is transitive) as an object nominalizer, a function initially achieved only via the suffix *-ye*.

The evolutionary path proposed for the suffixes *-'u* and *-'Vwi* would be thus as follows: instead of a nominal object (**Stage 1**), the directional postposition can take a possessed clausal nominalization marked by *-ye* as its postpositional object (**Stage 2**). Probably due to the erosion of *-ye* and the influence of the locative nominalizer *-yepo* reduced as *-'Vpo*, the sequence **-yeu* /**-yewi* reduced as *-'u* / *-'Vwi* is reanalyzed as a directional nominalizer (**Stage 3**). When this nominalization suffixed by *-'u* / *-'Vwi* is used with a transitive main verb that is not a motion

14. Glosses are adapted from Dedrick & Casad (1999: 383).

verb (**Stage 4**), the directional meaning is bleached and the reanalysis to the object nominalization can operate, but only if the base verb is transitive. This evolutionary scenario is schematized in (32).

(32) **Evolutionary path from directional postposition to object nominalization**

Stage 1. Directional postposition			
Nominal base:	the mountain-DIR		'to the mountain'
Stage 2. Directional postposition with a possessed clausal NMLZ and a motion main verb			
Verbal base:	GO	my Verbing-*ye*-DIR	'go to where I Verbing'
Stage 3. Directional NMZR with a possessed clausal NMLZ and a motion main verb			
Verbal base:	GO	my Verbing-*'u*/*-'Vwi*	'go to where I Verbing'
		my Verbing-NMZR_DIR	
Stage 4: Object NMZR used in an independent NP standing alone and with transitive main verbs of no motion			
Transitive verb base:	LOVE	my Verbing-NMZR_OBJ	'love what I Verbing'

From this point, the evolution of *-'u*/*-'Vwi* to object relativization follows the evolution of *-ye* to the same nominalization use schematized in (26). Like in the evolution of the suffix *-ye*, the object nominalization that is only possible with transitive verbs, is reinforced when a noun is pre-posed in a non-restrictive apposition (**Stage 5**). Again, the merger of the intonation contours between the pre-posed noun (anchor) and the object nominalization (appositive) converts the construction into a restrictive apposition (**Stage 6**); that is, a construction with a relativization-type function. This evolution is summarized in (33):

(33) **Evolution from object nominalization to object relativization**

Stage 5: Object NMZR in a non-restrictive appositive NMLZ (Object nominalization)		
Transitive verb base:	N, my Verbing-NMZR_OBJ	'N, the one I Verbing'
Stage 6: Object NMZR used in a restrictive appositive NMLZ (Object relativization)		
Transitive verb base:	N my Verbing-NMZR_OBJ	'N that I Verbing'

This grammaticalization from directional postposition to object nominalizer would thus involve the usual four interrelated mechanisms (Heine & Kuteva 2004: 2, 2007, Section 1.2): phonetic erosion (*-yeu* > *-'u, -yewi* > *-'Vwi*), semantic bleaching, extension of contexts and decategorialization (from postpositions to nominalizers). This would be in fact an instance of a well-known grammaticalization process whereby spatial expressions develop new, more abstract meanings through processes of context-induced inference (Heine, Claudi & Hünnemeyer 1991, among several others).

However, there is another type of construction that could also represent a good source for the object nominalization use of the suffix *-'u* in Yaqui, and may be both hypothetical source constructions could have interacted in the same evolutionary direction. Indeed, the historical data found in the *Arte* indicate that the Yaqui object nominalizer could also come from another nominalizing suffix existing in Old Cahita. The suffix *-u* was indeed a marker used in different types of subject nominalizations (Álvarez 2016).

5.3 An alternative source for the suffix *-'u* in Yaqui

5.3.1 *The suffix* -u *in Old Cahita*

5.3.1.1 *Former agent nominalizer.* Agent nominalization was associated in Old Cahita with two suffixes: the suffix *-me* was used either as a marker of referential nominalized expression (34a) or as a marker of a modifying nominalized expression in an appositive construction (34b) and it could appear in past perfective, present, and future situations, whereas the suffix *-u* was restricted to past terminative situations, and it is only documented in the *Arte* as a referential nominalizer (35).

(34) **The suffix *-me***

– Subject NMLZ (Headless Subject-RC) **Referential use**

a. *eria-**me*** *eria-ca-**me*** *eria-naque-**me***
love-NMZR_SBJ love-PERF-NMZR_SBJ love-FUT-NMZR_SBJ
'the one who loves' 'the one who loved' 'the one who will love'

– Subject NMLZ Appositive (Subject-RC) **Modifying use**

b. *Itom Atzai* [*teueca-po cateca-**me***]…
1PL.POS father sky-LOC be_seated-NMZR_SBJ
'Our Lord, who art in heaven, …' (Lit. Our Father who is seated in the sky…)

(35) **The suffix *-u***

– Former subject NMLZ (Headless Subject-RC) **Referential use**
*eria-**u***
love-NMZR_SBJ
'the one who was loving (not anymore)'
*eria-ca-**u***
love-PERF-NMZR_SBJ
'the one who had loved (not anymore)'

These examples show that the main difference between the uses of these two subject nominalization suffixes was aspectual: the suffix *-u* was mainly used to refer

to animate entities in terms of the action they were performing in the past. The terminative meaning associated with this suffix (Buelna 1890: 84) indicates that the action denoted by the verbal base was usually performed by the referent in the past, but not any longer at the time of reference, as can be seen in (35).[15] As this temporal/aspectual meaning may be rendered in English by the temporal adjective *former,* I have proposed to call this kind of nominalization a *former agent* nominalization (Álvarez 2016).

It is interesting to note that the function of the *-u* nominalization was apparently restricted to a referential function, since there is no modifying use of this kind of nominalization documented either in the *Arte* or in the Catechism in Buelna (1890). Apparently, this is another difference between these two agent nominalizers from Old Cahita.

5.3.1.2 *Former attribuand nominalizer.* Besides the use as a former agent nominalizer, the author of the *Arte* points out another nominalizing use of the suffix *-u* when associated with the perfective suffix *-ca*: attached to a nominal base within a possessive construction, it indicates that the person being referred to is no longer what he/she used to be (because of a long absence or a death, for example) (Buelna 1890: 85–86).[16] An example of this use is provided in (36).

(36) *in* *atzae-ca-u*
1SG.POS father-PERF-NMZR_SBJ
'the one who was my father (now deceased)'

When this construction had no possessive marking like the possessive pronoun *in* 'my' in (36), the suffix *-tu* 'to be, to become' was added to the nominal base before the combination with *-cau* (Buelna 1890: 86), as in (37a)–(b).[17] Note that the nominal base could be an inanimate entity, as in (37b).

(37) a. *iorem-tu-ca-u*
person-VBZ-PERF-NMZR_SBJ
'the former people / the ones who were people (not any more)'

---

15. This terminative meaning appears in the explanation given by the author of the *Arte,* when he points out that this nominalized expression is used to mean that the verb action has been accomplished and finished. The original texts states: "De estos participios en *u* se usa para significar haberse acabado ó consumido la acción, ó término del verbo" (Buelna 1890: 85).

16. The original text states: "También con estos participios significan las ausencias largas de aquellos, que parece que ya no los han de volver á ver" (Buelna 1890: 85–86).

17. The original text says: "… añaden al nombre si trae nota de posesión, ó el semipronombre posesivo, esta partícula *cau,* … pero si no trae nota de posesión ó semipronombre posesivo, se añade al nombre en el recto esta partícula *tucau*" (Buelna 1890: 86).

b. *bachi-tu-ca-u*
corn-VBZ-PERF-NMZR_SBJ
'the one that was corn (not any more because it has been eaten)'

In terms of the nominalization process, the examples in (37) show that the verbal base to which the suffix *-u* is attached is a past stative verb 'to be X in the past', where X is the nominal base referring to an entity (e.g. *iorem* 'person' in (37a) and *bachi* 'corn' in (37b)). Thus, this nominal base has been verbalized by the suffix *-tu*. The presence of the past perfective suffix *-ca* locates this X state in the past, and finally the suffix *-u* nominalizes this past stative expression, yielding the referred subject argument with the terminative meaning of 'the one who was in the state denoted by X, but is so no longer at the reference time'. Therefore, the difference between the nominalization uses in (36) and (37) and the former agent nominalization use in (35) is the type of verbal base – stative in (36) and (37) and eventive in (35). The stative verb is derived by means of a zero copula in (36) and by means of the suffix *-tu* in (37).

As this kind of argument nominalization creates a nominal constituent referring to an entity that is characterized in terms of a former attribute (a state, a quality), I have proposed to label this kind of nominalization a *former attribuand* nominalization (Álvarez 2016).[18]

Former agent nominalization and former attribuand nominalization are two different types of subject nominalization. In both cases, the nominalized expression denotes an entity functioning syntactically as the subject of the nominalized base verb. The base verb is an action verb in the former agent nominalization, and a state verb in the former attribuand nominalization. The denoted entity is then characterized in terms of its involvement in a former action in the former agent nominalization, and in terms of a former state in the former attribuand nominalization.

5.3.1.3 *Action/result nominalizer.* Former agent nominalization and former attribuand nominalization are the only uses documented in the *Arte* for the nominalizing suffix *-u*. However, if we consult the Catechism following the *Arte* (Buelna 1890), we can observe a few examples where the suffix *-u* is used in a different way that seems to mix the structures of Examples (35) and (36). Indeed, Example 38 shows that, like (35), the sequence *-ca-u* is suffixed to an eventive verbal base;

18. The term "Attribuand" has been proposed to refer to the case role that covers the subject of quality predicates; that is, the participant who carries an attribute (Halliday 1968: 190; Omamor 1978: 271). In sentences such as *Mary is beautiful* or *Mary is a teacher, Mary* fills the case role Attribuand, since the only role of *Mary* in these cases is as the carrier of an attribute (*be beautiful, be a teacher).*

however, unlike (35) but like (36), the construction in (38) is a possessed nominalized expression.

(38) Action/result nominalization with possessive pronoun:
Itom iautzia Jesu-Cristo a-lulutiri-naque
1PL.POS father Jesus-Christ 3SG.ACC-forgive-FUT
[***in-ca-ala-ane-ca-u***].
1SG.POS-NEG-well-act-PERF-NMZR
'Our Lord, Jesus Christ, will forgive **my sins** (lit. what I had not acted well).'
(Buelna 1890: 245)

This example illustrates an important change compared to the use exemplified in (35): this kind of nominalization created by the suffix *– u* is no longer associated with a former agent meaning, but with an action/result meaning, and the agent of the verb *ane* 'to act' is expressed in (38) by the pre-posed possessive pronoun *in* 'my'.

Unlike former agent nominalization and former attribuand nominalization, this action/result nominalization is no longer associated with subject nominalization. The nominalized expression denotes here the activity designated by the base verb/clause, not the subject/agent of this verb/clause. The possessive pronoun pre-posed to this action nominalization is interpreted as the subject/agent. As in the case of the evolution of the suffix *-ye*, this possessed action nominalization can also be metonymically interpreted as a result nominalization. This metonymic shift from action to result represents an important step, since it allows the change from subject nominalization to object nominalization (cf. footnote 12).

Interestingly, the nominalized verb in (38) is an intransitive verb (*ane* 'to act' (Buelna 1890: 204)), which implies that the nominalized expression cannot refer here to an object/patient, but to an action/result; *my bad actions* or *my sins* respectively. As observed in the evolution of the suffix *-ye* from possessed classifier to object nominalization, it seems that the possessed action/result nominalization represents thus a very plausible intermediate step to the object nominalization in Cahita.

In (39), I illustrate another example of an action/result nominalization taken from the Catechism, but in this case the nominalized base verb is the transitive verb *aua* 'ask for, request, call' (Buelna 1890: 204). Again, we can observe the same presence of the possessive pronoun corresponding to the subject/agent of the nominalized clausal base and the suffix *-u* combined to the past suffix *-ca*. The function assigned to this expression is again referential, and it is possible to have both action and result interpretations, but also an object/patient nominalization (*what we had requested for him*), since the base verb is here transitive.

(39) *itom a au-ria-ca-u*
1PL.POS 3SG.ACC ask_for-APL-PERF-NMZR
'our request for him' (Buelna 1890: 239)

It is also interesting to note again that, like in the former agent and former attribuand nominalizations, the function of this nominalized expression was apparently restricted to a referential function, since there is no modifying use of this kind of nominalization documented either in the *Arte* or in the Catechism.

5.3.2 *The suffix -*'u *in Yaqui*

Unlike what we have seen in Old Cahita (Example 35), the suffix *-u* is no longer involved in former agent nominalization in Yaqui. Instead, this suffix has developed an object nominalizer use, as seen in Examples (14a)–(b). This new use is clearly related to the use as an action/result nominalizer documented in Old Cahita (Examples (38) and (39)) where the subject/agent of the nominalized verb was already expressed by a possessive pronoun. As in the case of the suffix *-ye* (Stages 3 and 4 in (25) and (26), respectively), the metonymic change from action/result nominalizer to object nominalizer is allowed by the fact that the nominalized base verb is transitive and is reinforced by the extension of use from independent NPs (as in (39)) to appositive NPs (as in (40)). Compare the following Yaqui example with (39) from Old Cahita:

(40) *u karo,* [*itom a a'au-ria-ka-'u*]
DET car 1PL.POS 3SG.ACC ask_for-APL-PERF-NMZR_OBJ
'the car, the one that we requested for him'

In the non-restrictive apposition exemplified in (40), the inferred co-referentiality between the anchor/head noun *karo* and the appositive nominalized expression *itom a a'auriaka'u* forces the activation of a patientive interpretation for the nominalized expression, unlike (39), where the use of the nominalization as an independent NP was still associated with the action/result meaning, although the object/patient interpretation was also possible.

As mentioned above for the suffix *-ye* (Stage 5 in (26)), the modifying function of the possessed nominalization, and therefore the overlap with relativization, appear when the construction changes to a restrictive apposition via the merger of the intonation contours between the anchor and the appositive, like in (41).

(41) *u karo* [*itom a a'au-ria-ka-'u*]
DET car 1PL.POS 3SG.ACC ask_for-APL-PERF-NMZR_OBJ
'the car that we requested for him'

The suffix *-(')u* associated in Old Cahita with former agent nominalization, former attribuand nominalization, and action/result nominalization, has thus evolved into the current object nominalization marker in Yaqui and the old uses of the suffix *-u* as a former agent nominalizer and an action/result nominalizer have disappeared.[19]

Although the source is probably distinct, we can observe that the evolutionary path of the suffix *-u* is similar to the evolution of the suffix *-ye,* since the two paths go to the object nominalization use through the possessed action/result nominalization stage, and to relativization through apposition.

The evolutionary path proposed for the suffix *-(')u* would thus be as follows: From former agent/attribuand nominalization (as in (35) and (36) from Old Cahita) to possessed action/result/object nominalization when combined with an eventive verbal base and a possessive pronoun in independent NPs (as in (38) and (39) from Old Cahita). From this point, the evolution of *-(')u* to object relativization follows the evolution of *-ye* to the same nominalization use schematized in (26) and (27). Like in the evolution of the suffix *-ye*, the object nominalization that is only possible with transitive verbs, is the only available interpretation when a noun is pre-posed in a non-restrictive apposition, as in (40). Again, the merger of the intonation contours between the pre-posed noun (anchor) and the object nominalization (appositive) converts the construction into a restrictive apposition; that is, a construction with a relativization-type function (as in (41)) Once this object interpretation is established via apposition, the action/result interpretations are no more available. This evolutionary scenario is schematized in (42):

19. In Modern Cahita, the corresponding suffix *-'u* (Yaqui) / *-'Vwi* (Mayo) is only used as an object nominalizer, except for its relic use as a former attribuand nominalizer in the sequence *-tuka'u* (Yaqui) / *-tukaywi* (Mayo). See Álvarez (2016) for a discussion on the current uses of these markers in Modern Cahita.

(42) **Evolution from subject nominalization to object nominalization and to object relativization**

Stage 1. Former agent/attribuand NMZR in a non-appositive NMLZ used as NP (Subject-nominalization)

love-*ca-u*	'the one who had loved (not any more)'
act- *ca-u*	'the one who had acted (not any more)'
my father-*ca-u*	'the one who was my father (not any more)'

Stage 2. Action/result/object NMZR in a non-appositive NMLZ used as possessive NP (Non-subject nominalization)

my love-*ca-u*	'my former loving/love',	'the one/what I had loved'
my act- *ca-u*	'my former acting/action',	'what I had acted'

Stage 3. Object NMZR in a non-restrictive appositive NMLZ (Object nominalization)

woman/book, my love-*ca-u*	'the woman/book, my former love'
	'the woman/book, the one I had loved'

Stage 4. Object NMZR in a restrictive appositive NMLZ (Object relativization)

woman/book my love-*ca-u*	'the woman/book my former love'
	'the woman/book I had loved'

Stage 5. Object NMZR in a non-appositive grammatical NMLZ used as NP (Headless Object-RC)

my love-*ca-u*	'the one I had loved'

Again, we can observe that:

- Although the sources could be different (a possessive classifier for suffix *-ye* and a subject nominalizer for suffix *-(')u*), the evolution of the suffix *-(')u* from Subject/Agent nominalization to Action/Result nominalization and to Object/Patient nominalization is similar to the evolution of the suffix *-ye* since it is made through metonymic changes caused by the extension of use contexts (first with possessive pronoun, and second with transitive verbs and a pre-posed noun) and via apposition (because of the inferred co-referentiality between the pre-posed noun and the appositive grammatical nominalization).
- The suffixes *-ye* and *-(')u* have followed the same channel to the object nominalization. Stage 2 in (25) for suffix *-ye* corresponds to Stage 3 in (42) for *-(')u*. Henceforth, the evolution is the same.

- The metonymic changes within grammatical nominalizations clearly indicate that the modifying function of the object nominalization comes later than the referential function, chronologically speaking. The first nominalization function is referential (**Stage 1**), and the noun-modifying function only appears after via restrictive apposition (**Stage 4**).
- Like for the evolution of *-ye*, the evolution of *-(')u* implies the generalization of use contexts (from restricted past terminative uses to generalized uses with regard to the TAM marking) as well as its specialization (from all types of base verbs to transitive base verbs only).

In sum, as mentioned in Section 4.2, it appears that Yaqui and Mayo could have formed their object nominalization markers by analogy with the locative nominalization marking, which originally combined the object nominalization suffix *-ye* with a spatial postposition *-po*. As shown in (19), the suffix *-ye* has been reduced to an echo-vowel in the marking of locative nominalizations in Modern Cahita. Following the pattern of the locative nominalization formation, it is clear that Mayo has recruited another spatial postposition, the directional marker *-wi* exemplified in (22b), for the formation of the object nominalizer. This pattern copy is confirmed by the use of the same echo-vowel present in the current locative nominalizer from Mayo. It is very likely that this same pattern (except for the presence of a glottal stop instead of the echo-vowel) has also been in action for Yaqui, since the object nominalizer and the directional marker present the same formal similarity. However, another source can be proposed for Yaqui, namely an old subject nominalizer with tempo-aspectual restrictions, which could also have participated in the evolution to the current object nominalizer in Yaqui, showing two plausible sources that could have been working in concert.

It is even possible that the choice of the directional postposition recruitment for the object nominalization marking in Mayo could have been also influenced by this Old Cahita former agent/attribuand nominalizer. Indeed, the analogy between the locative nominalization marking (suffix *-'Vpo*) and the Mayo object nominalization marking (suffix *-'Vwi*) could have been induced by another analogy associated with the evolution of the suffix *-u* from being an old subject nominalizer with tempo-aspectual restrictions to being the current object nominalizer in Yaqui. As the directional postposition from Old Cahita suffix *-wi* has been reduced in Yaqui as *-u* (see Example (22a)) but maintained in Mayo (Example (22b)), the reduction observed in Yaqui could have led to a formal similarity between the object nominalizer and the directional postposition in Yaqui. Then, this homonymy could have caused a confusion and an assimilation between the directional marker and the object nominalization marker *-u* in Yaqui, since both markers are similar. This assimilation in Yaqui could thus lead to the recruitment of the corresponding

directional marker *-wi* in Mayo for the object nominalization marking, with the parallel influence of the locative nominalizer formation (*-yepo*). A plausible argument for this possibility is found in (43), in which we can observe that the former attribuand nominalization is still present in Modern Cahita as a relic of the old nominalizing use of the suffix *-u*. But while Yaqui uses the suffix *-'u*, Mayo uses for this kind of nominalization a new marker the suffix *-ywi*, which seems to be another erosioned result of the possible combination **-ye-wi*; that is, the same strategy at the origin of the object nominalizer in Mayo (*-'Vwi*).

(43) a. *in achai-tu-ka-'u* **Yaqui**
1SG.POS father-VBZ-PERF-NMZR
'the one who was my father (now deceased)'

b. *in compare-tu-ka-ywi* **Mayo**
1SG.POS buddy-VBZ-PERF-NMZR
'the one who was my buddy (now deceased)'

6. Final remarks

The nominalization approach to relativization that has been proposed here for non-subject RCs in Cahita is supported by the origin of the markers involved in these constructions as well as by the evolutionary paths that lead to the relativization function.

The different origins proposed for the marking of non-subject nominalizations in Cahita (a possessed nominal classifier for the suffix *-ye* in Old Cahita, a subject nominalizer for the suffix *-(')u* in Yaqui, a directional postposition for the suffixes *-(')u/-'Vwi* in Yaqui and in Mayo, and a locative postposition for the marker *-yepo / -'Vpo* in Old and Modern Cahita) are indeed different types of nominal markers indicating that the base clause/verb is treated as a nominal constituent; that is, that the construction has been nominalized. Thus, all these markers serve as a "sustantitivization" strategy (Yap et al. 2011) and they signal that the resulting construction is no longer a clause but a nominal constituent.

As for the evolution to the relativization function, it is clear that the first nominalization uses of these markers are always associated with a referential function and that the modifying function corresponds to a secondary development made via apposition, showing that object relativization in Cahita corresponds in fact to a specialized function of the object nominalization: the modifying function of a possessed object nominalization used as an appositive in a restrictive apposition.

More generally, this study about the origin of non-subject nominalizers in Cahita will hopefully contribute to a better understanding of nominalization

phenomena, since it also gives information about the processes that trigger the extension of nominalized clauses from one context to another. In this regard, the diachrony of object nominalizations in Cahita has shown the importance of metonymical changes in this kind of evolution. For instance, the subject of the nominalized base verb/clause is first interpreted as the possessor of an action nominalization, then this action nominalization changes to a result nominalization and finally to an object nominalization with transitive verbs and via apposition. This evolution that seems to have been in action twice in Cahita (for the suffix *-ye* in its evolution from the possessed nominal classifier function, and for the suffix *-u* in its evolution from the subject nominalizer function) is thus made through contextual inferences causing a chain of metonymic shifts from the possessed action nominalization to the object nominalization. In this evolution, the object nominalization interpretation is triggered by the use of a possessed action/result nominalization with a transitive base verb and is reinforced by the presence of a pre-posed noun in an apposition. In the case of the directional postposition source, the first nominalization use is a possessed directional nominalization, and the change to object nominalization is also caused by the metonymic uses of this possessed nominalization in new contexts (with no-motion main verbs and with transitive base verbs) and is reinforced by the same kind of apposition.

The change to object relativization in Cahita appears when the juxtaposition between the anchor noun and the possessed nominalization functions like an apposition, with a correferential relationship between its two members. This inferred correferentiality implies that the possessed nominalization is no longer associated with an event (action nominalization) but with an entity (argument/participant nominalization). The possessive pronoun in the possessed nominalization thus refers to the subject of the event denoted by the nominalized base verb, whereas the entire possessed nominalization refers to the object of the event denoted by the nominalized base verb, which is the same entity as the one referred by the pre-posed anchor noun. The final step to object relativization is then achieved via the merger of the intonation contours between the anchor noun and the possessed nominalization, illustrating a clear example of the Paratactic channel to RCs (Givón 2009, 2012) as well as an instance of clause expansion; that is, of a strategy whereby clausal participants are treated like nominal participants (Heine 2009: 26).

Abbreviations

ACC	accusative
AL	alienable
ANT	anterior
APL	applicative
CLF	classifier
DEM	demonstrative
DES	desiderative
DET	determiner
DIR	directional
FUT	future
INAL	inalienable
INAM	inanimate
INSTR	instrumental
INTR	intransitive
IMM	immediate
LOC	locative
NEG	negation
NMLZ	nominalization
NMZR	nominalizer
NOM	nominative
OBJ	object
PASS	passive
PERF	perfective
PL	plural
POS	possessive
SG	singular
SBJ	subject
TR	transitive
VBZ	verbalizer

References

Álvarez González, Albert. 2016. The evolution of grammatical nominalizations in Cahita languages. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada-Fernández (eds), 107–140. Amsterdam: John Benjamins.

Álvarez González, Albert. 2018. Apuntes sobre el *Manual para administrar a los Indios del idioma Cahita los Santos Sacramentos*. In *Las lenguas en México: Diálogos historiográficos*, Barbara Cifuentes & Rodrigo Martinez Baracs (eds), 211–232. Mexico: IIB UNAM/Somehil.

Álvarez González, Albert. 2012. Relative clauses and nominalizations in Yaqui. In *Relative Clauses in Languages of the Americas. A Typological Overview* [Typological Studies in Language 102], Bernard Comrie & Zarina Estrada-Fernández (eds), 67–96. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.102.04gon

Álvarez González, Albert & Felix Quesada, Rafael Jesús. 2011. Morfología derivativa y variación lingüística en cahita. In *Estudios morfológicos, sintácticos, semánticos y de lingüística aplicada* [Colección Estudios Lingüísticos 3], Maria del Carmen Morúa Leyva & Gerardo López Cruz (eds), 47–76. Hermosillo: University of Sonora.

Andrews, Avery D. 2007. Relative clauses. In *Language Typology and Syntactic Description*, Vol. II: *Complex Constructions*, 2nd edn, Timothy Shopen (ed.), 206–236. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.004

Bascom, Burton William. 1965. Proto-Tepiman. PhD dissertation, University of Washington, Seattle.

Buelna, Estaquio. 1989[1890]. *Arte de la lengua cahita por un padre de la Compañía de Jésus*. Mexico: Editorial Siglo XXI.

Campbell, Lyle & Langacker, Ronald. 1978. Proto-Aztecan vowels: Parts I-III. *International Journal of American Linguistics* 44: 85–102, 197–210, 262–279. https://doi.org/10.1086/465526

Carlson, Greg N. 1977. Amount relatives. *Language* 53: 520–542. https://doi.org/10.2307/413175

Comrie, Bernard & Kuteva, Tania. 2005. Relativization strategy. In *World Atlas of Language Structures (WALS)*, Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds), 494–497. Oxford: OUP.

Comrie, Bernard & Thompson, Sandra A. 1985. Lexical nominalization. In *Language Typology and Syntactic Description*, Vol. III: *Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 349–398. Cambridge: CUP.

Cristofaro, Sonia 2019. Nominalization in diachronic cross-linguistic perspective. In *Nominalization in Languages of the Americas* [Typological Studies in Language 124], Roberto Zariquiey, Masayoshi Shibatani & David W. Fleck (eds). Amsterdam: John Benjamins. (this volume)

Dakin, Karen. 1991. Nahuatl direct and mediated possession: A historical explanation for irregularities. *International Journal of American Linguistics* 57(3): 298–329. https://doi.org/10.1086/ijal.57.3.3519722

Dedrick, John M. & Casad, Eugene H. 1999. *Sonora Yaqui Language Structures*. Tucson AZ: University of Arizona Press.

Escalante, Fernando. 1990. Voice and Argument Structure in Yaqui. PhD dissertation, University of Arizona.

Gerner, Matthias. 2012. The typology of nominalization. *Language and Linguistics* 13(4): 803–844.

Givón, T. 2012. Toward a diachronic typology of relative clause. In *Relative Clauses in Languages of the Americas. A Typological Overview* [Typological Studies in Language 102], Bernard Comrie & Zarina Estrada-Fernández (eds), 3–25. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.102.01giv

Givón, T. 2009, *The Genesis of Syntactic Complexity*. Amsterdam: John Benjamins.

Givón, T. 2001. *Syntax: An Introduction*, Vol. II. Amsterdam: John Benjamins.

Guerrero, Lilian. 2012. On relative clauses and related constructions in Yaqui. In *Relative Clauses in Languages of the Americas. A Typological Overview* [Typological Studies in Language 102], Bernard Comrie & Zarina Estrada-Fernández (eds), 97–126. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.102.05gue

Halliday, Michael A. K. 1968. Notes on transitivity and theme in English: Part 3. *Journal of Linguistics* 4(2):179–215. https://doi.org/10.1017/S0022226700001882

Haspelmath, Martin. 1993. *A Grammar of Lezgian*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110884210

Heine, Bernd. 2003. Grammaticalization. In *The Handbook of Historical Linguistics*, Brian D. Joseph & Richard D. Janda (eds), 576–601. Oxford: Blackwell. https://doi.org/10.1002/9780470756393.ch18

Heine, Bernd. 2009. From nominal to clausal morphosyntax: Complexity via expansion. In *Syntactic Complexity. Diachrony, Acquisition, Neuro-cognition, Evolution* [Typological Studies in Language 102], T. Givón & Masoyoshi Shibatani (eds), 23–51. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.02fro

Heine, Bernd, Claudi, Ulrike & Hünnemeyer, Frederike. 1991. *Grammaticalization*. Chicago IL: University of Chicago Press.

Heine, Bernd & Kuteva, Tania. 2004. *World Lexicon of Grammaticalization*. Cambridge: CUP.

Heine, Bernd & Kuteva, Tania. 2007. *The Genesis of Grammar: A Reconstruction*. Oxford: OUP.
Hopper, Paul J. & Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language* 56(2): 251–299. https://doi.org/10.1353/lan.1980.0017
Hyde, Villiana. 1971. *An Introduction to the Luiseño Language*. Banning CA: Malki Museum Press.
Johnson, Jean B. 1962. *El idioma yaqui*. México: INAH.
Keenan, Edward L. 1985. Relative clauses. In *Language Typology and Syntactic Description*, Vol. II: *Complex Constructions*, Timothy Shopen (ed.), 141–170. Cambridge: CUP.
Kubo, Yoshihiro. 2009. A note on two types of apposition. *The Bulletin of Central Research Institute Fukuoka University, Series A: Humanities* 9(7): 27–35.
Launey, Michel. 1981. *Introduction à la langue et à la littérature aztèques*, Vol. 1: *Grammaire*. Paris: L'Harmattan.
Lehmann, Christian. 1984. *Der Relativsatz*. Tübingen: Günter Narr.
Lindenfeld, Jacqueline. 1973. *Yaqui Syntax*. Berkeley CA: University of California Press.
Lionnet, Andrés. 1977. *Los elementos de la lengua cahita*. Mexico: UNAM.
Luján, Eugenio R. & Ruiz Abad, César. 2014. Semantic roles and word formation: Instrument and location in Ancient Greek. In *Perspectives on Semantic roles*. [Typological Studies in Language 106], Silvia Luraghi & Heiko Narrog (eds), 241–269. Amsterdam: John Benjamins.
Martínez Fabián, Constantino & Langendoen, D. Terrence. 1996. Sobre las llamadas cláusulas relativas en yaqui. In *Memorias del tercer encuentro de lingüística en el Noroeste, Tomo 1: Lenguas indígenas*, Vol. 2, Zarina Estrada-Fernández, Max Figueroa Esteva & Gerardo López Cruz (eds), 443–464. Hermosillo: Editorial Unison.
Miller, Wick R. 1986. Sketch of Shoshone, a Uto-Aztecan language. In *Handbook of American Indians*, Vol. 17: *Languages*, Ives Goddard (ed.), 693–720. Washington DC: Smithsonian Institute.
Miller, Wick R. 1984. The classification of the Uto-Aztecan languages based on lexical evidence. *International Journal of American Linguistics* 50(1): 1–24. https://doi.org/10.1086/465813
Moctezuma, José Luis & López, Gerardo. 1990. Variación dialectal yaqui-mayo. *Noroeste de México* 9: 84–105.
Moctezuma, José Luis & López, Gerardo. 1991. El yaqui y el mayo como lenguas históricas. *Noroeste de México* 10: 79–84.
Moctezuma, José Luis & López, Gerardo. 1994. En torno a la dialectología cahita. In *Estudios de lingüística y sociolingüística*, José Luis Moctezuma & Gerardo López (eds), 221–274. Hermosillo: University of Sonora/INAH.
Omamor, Augusta P. 1978. Case grammar and viability. In *Valence, Semantic Case and Grammatical Relations* [Studies in Language Companion 1], Werner Abraham (ed.), 261–293. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.1.17oma
Payne, Thomas E. 1997. *Describing Morphosyntax. A Guide for Field Linguists*. Cambridge: CUP.
Peña Valenzuela, Ana Patricia. 2012. *Cláusulas relativas en el mayo de Sonora*. MA thesis, University of Sonora.
Queixalós, Francesc. 2012. Notes sur la nominalisation. *Amerindia* 35:105–111.
Seiler, Walter. 1985. *Imonda, a Papuan Language*. Canberra: Department of Linguistics Research School of Pacific Studies, Australian National University.
Shibatani, Masayoshi. 1985. Passives and related constructions: A prototype analysis. *Language* 61: 821–848. https://doi.org/10.2307/414491
Shibatani, Masayoshi. 2009. Elements of complex structures, where recursion isn't it. The case of relativization. In *Syntactic Complexity. Diachrony, Acquisition, Neuro-Cognition, Evolution*

[Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 163–198. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.07ele

Shibatani, Masayoshi. 2010. Nominalization. Workshop presented during the Seminario de Complejidad Sintáctica, 15–16 November. Hermosillo: University of Sonora.

Shibatani, Masayoshi. 2019. What is nominalization? Towards the theoretical foundations of nominalization. In *Nominalization in Languages of the Americas* [Typological Studies in Language 124], Roberto Zariquiey, Masayoshi Shibatani & David W. Fleck (eds). Amsterdam: John Benjamins. (this volume)

Shibatani, Masayoshi & Awadh Bin Makhashen, Khaled. 2009. Nominalization in Soqotri, a South Arabian language of Yemen. In *Endangered Languages: Contributions to Morphology and Morpho-syntax*, Leo W. Wetzels (ed.), 9–31. Leiden: Brill.

Traugott, Elizabeth C. & Dasher, Richard B. 2005. *Regularity in Semantic Change*. Cambridge: CUP.

de Wolf, Paul P. 1997. *Esbozo del Mayo Sonorense* 2 Vols. Hermosillo: University of Sonora.

Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick. 2011. *Nominalization in Asian Languages: Diachronic and Typological Perspectives* [Typological Studies in Language 96]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.96

CHAPTER 8

On habitual periphrasis in Cuzco Quechua

Rammie Cahlon
The Hebrew University of Jerusalem

In Cuzco Quechua, a periphrastic construction composed of a lexical subject nominalization in conjunction with the copula is used to express habitual events anchored in the past, regardless of formal tense marking. The aim of this paper is to analyze the construction and evaluate its diachronic development with respect to established grammaticalization clines of past habitual. The variation appears to be the direct result of an on-going process which replicates, in part, an established grammaticalization pathway of past-habitual grams (Bybee et al. 1994). This process of language change sheds light both on the pathway mentioned and reasserts the claim that nominalizations serve source for main clause morphology (Gildea 2008).

1. Introduction

In Cuzco Quechua, a periphrastic construction composed of a lexical subject nominalization in conjunction with the copula is used to express habitual events anchored in the past (1–2). The construction follows the pattern *V*-NMLZ (AUX-), where *V* is any verbal stem, to which a subject nominalizer, realized as the affix *q*, is attached, and an auxiliary that optionally supplements the construction. The construction is well attested in Quechuanist literature and is normally dubbed "*pasado habitual*" (Cusihuamán 1976; Soto-Ruíz 1976, among others). It has been noted that there exists a certain temporal variance within the construction: the auxiliary, when it appears, may be marked for either the past tense or the unmarked tense. The so-called "unmarked tense" is better thought of a general imperfective, since its value in terms of temporal deixis is somewhat low, and is also used to mark the present habitual.

(1) *Ñuqa-qa tiro-pi-qa* *[iskay chunka punto-ta-puni-n*
I-TOP shooting-LOC-TOP 20 point-ACC-indeed-EV.1
rura-q]$_{NMLZ}$ *ka-ra-ni.*
do-NMZR AUX-PST-1.B
'As for me, in shooting I used to score twenty points.' (GKM.42)

https://doi.org/10.1075/tsl.124.08cah

(2) *Sapa tuta-manta-n las seis [alojado-kuna-q mikhuna puchu-n-ta*
every night-ABL-EV.1 at six guest-PL-GEN food rest-3.B-ACC
hurqu-mu-q] $_{NMLZ}$ *ka-ø-ni.*
gather-CIS-NMZR AUX-UNMRKD-1
'Every morning at six I would gather the food leftovers of the guests.'
(GKM.78)

Although structurally different, the two events in (1) and (2) are identical in terms of their event structure. In Reichenbachian terms (Reichenbach 1947), both sentences refer to an event (E) set prior to the speech time (S) and the reference time (R) is concurrent with (E). These events are protracted over an indefinite time span during which they are held to be true, though not necessarily at every single point in time. Although (2) does not contain any past marking, it is made clear by extralinguistic context that the event in question no longer holds true at speech time.

The aim of this paper is to analyze the construction mentioned above and evaluate its diachronic development with respect to established grammaticalization clines of past habitual (PAST-HAB). To my knowledge, no such description exists for Cuzco Quechua, nor do I know of any systematic description of the Cuzco Quechua TAM system. Some literature regarding the TAM systems of other varieties of Quechua exist; for example, Hintz (2007) for his comprehensive work on the aspectual system of South Conchucos Quechua, Adelaar (1977) for Tarma Quechua, and Weber (1989) for Huallaga Quechua.

The alternation in formal tense marking illustrated above appears to be the direct result of an on-going grammaticalization process which replicates similar processes of language change that have taken place in Quechuan and whose direct results are now evident in the TAM system of Cuzco Quechua. The aspectual functions of the pattern were found to exceed most definitions of habitual aspect. The periphrasis serves as yet another example of a well-established grammaticalization pathway where nominal elements grammaticalize to mark an imperfective aspect (cf. Bybee, Perkins & Pagliuca 1994; Gildea 2008).

This paper reports the results of an analysis of the testimonials of Don Gregorio Mamani (1908?-1979) and his wife Asunta Quispe Huamán (?-1983) as they were recorded by Ricardo Valderrama Fernández and Carmen Escalante Gutiérrez in the mid-1970's. (Mamani, Valderrama & Escalante Gutiérrez 1983) Due to the personal nature of the texts, a control corpus was also selected in an attempt to minimize the effect of the genre on the results.

The control corpus comprises two bodies of texts. The first comes from texts collected between 2001–2003 and published as annexes by the ethnologist Xavier Ricard Lanata for his work on *Les voleurs d'ombre – L'univers religieux des bergers de l'Ausangate*, published in 2010. These texts contain transcriptions of short

conversations with people from the vicinity of Mount Ausangate, in the department of Cuzco in Peru. The second is the *Antologia Quechua del Cuzco*, which was published in 2012 by the municipality of the Cuzco district and the Centro Guaman Poma de Ayala. The anthology consists of texts of various genera, including oral narratives, sermons and speeches, traditional stories, poetry and theatrical work. Texts which were dated before 1900 were excluded unless specifically noted.

The paper is structured as follows. First, in Section 2 I address the term *habitual* and what it brings under its scope for the purpose of this paper. In Section 3 I will then describe the morphosyntactic behavior of the construction as well as its functions from a synchronic point of view. Drawing a parallel with the English 'used to' construction, in Section 4, will aid in making the construction's evolution more apparent. I then trace the development of the construction and its on-going verbalization. Finally, conclusions and a brief summary follow in Section 5.

2. The aspectual selva oscura of habitual constructions

The most commonly noted function of the construction in question is to mark an aspectual distinction. Although there are multiple approaches to characterizing habituals (Comrie (1976: 26), Sasse (2002: 210), Boneh & Doron (2013)), The most prevalent view is that the habitual aspect is a sub-type of the imperfective (Comried 1976: 26; Sasse (2002: 2109), and as such it falls under the domain of what Sasse terms ASPECT_1. In essence, ASPECT_1 refers to aspect proper whereas ASPECT_2 refers to Aktionsarten. Kleiber (1987) writes that the term 'habitual' has been applied to several different linguistic phenomena in the literature and gives (3a)–(3c) as examples of the three most common phenomena that fall under the scope of the term.

(3) a. *Paul va à l'école à pied.*
'Paul goes to school on foot.'
b. *Les chats sont intelligents.*
'Cats are intelligent.'
c. *Jean hait les baleines.*
'Jean hates whales.'

According to Comrie (1976: 26), the habitual aspect describes "a situation which is characteristic of an extended period of time, so extended in fact that the situation referred to is viewed not as an incidental property of the moment but, precisely, as a characteristic feature of the whole period." This definition was largely adopted here, but it was further restricted, since it covers both Kleiber's (a) and (c), whereas the definition adopted in this paper covers only (a).

Here, habitual events are regarded as macro-events, which are composed of an unspecified number of instances of micro-events which then are generalized and protracted over a period of time anchored by the verbal tense. Those micro-events are then viewed whollyand generalized to be held true over the entire period of time denoted therebycharacterizing the referent. It follows then that habituals, semantically, require iteration for generalization to take place, which is the first restriction on Comrie's definition.

The prerequisite of iteration leads to the second. Due to the characterizing quality of habituals, it has been posited that they share certain features with stative verbs. The similarity, however, is limited to the habitual event in its entirety and not to the micro-events of which it is composed, which are by definition episodic in nature. This relates to the composition of states. Unlike episodic verbs, stative verbs are held to be true during the entire given time frame: their durativity is the element that conditions their stativity. This, in turn, has led scholars to note that stative verbs constitute a different phenomenon and that they are inherently incompatible with habitual aspect. This further deviates from Comrie's definition, as his take on the matter is somewhat different: for him, the fact that stative verbs combine with constructions like the English 'used to' suggests that habituality does not per se require iteration (for more see Comrie 1976: 27 and Brinton 1987: 203).

In summary, the habitual aspect is understood to fall under the domain of Sasse's ASPECT1 and to refer to a semantically complex expression which necessitates iteration of micro-events. Therefore it is not compatible with verbs denoting a permanent stative property.

3. Past Habitual Periphrasis in Cuzco Quechua – Form and function(s)

The Cuzco Quechua Past Habitual Periphrasis (PHP), as mentioned previously, is constructed from an agentive nominalized form in combination with the copula, which serves as the auxiliary. The inflected forms are given in Table 1. Although this analysis does not hold from a historical point of view, it is synchronically useful. As with other copular sentences, the paradigm is asymmetrical, and the copula does not appear in the third person.

This function of the copula is well documented with other periphrastic constructions in Quechuan, and since the verb *ka-y* is the main auxiliary used with compound tenses, parallels may be drawn. Among these periphrastic constructions is the past perfect, which is composed of the nominalizer *-sqa-* and the auxiliary.

Table 1. Inflection of the '*past habitual*'

	Sg.		Pl.	
1	mikhu-q	ka-ni	mikhu-q	ka-y-ku (Exc.)
	eat-NMZR	AUX-1	eat-NMZR	AUX-1-PL1
			mikhu-q	ka-nchis (Inc.)
			eat-NMZR	AUX-1INC
2	mikhu-q	ka-nki	mikhu-q	ka-nki-chis
	eat-NMZR	AUX-2	eat-NMZR	AUX-2-PL2
3	mikhu-q		mikhu-q-ku	
	eat-NMZR		eat-NMZR-PL1	

(4) *Llipi-y-ku ñuqayku manchari-sqa ka-yu-sha-ø-y-ku.*
all-1.A-PL.1 we.EXC fear-PP AUX-AUG-DUR-UNMRKD-1.B-PL1
'All of us had been scared.' (GKM.20)

(5) *Chay pacha ñawi-y, sik-y totalmente punki-sqa ka-ra-n.*
DEM.2 time eye-1.A butt-1.A totally swell-PP AUX-PST-3.B
'That time, my eyes and butt had swollen completely.' (GKM.20)

The copula's function here is first and foremost to index the subject and in some cases carry other TAM markers. Much like in the pattern investigated in this article, the perfect construction need not be overtly marked with the past tense suffix, and may carry different derivational affixes. Somewhat surprisingly, and unlike the examples given for the past perfect, it appears that the aspectual marking of the past habitual's auxiliary is disfavored. In a personal communication, Willem Adelaar notes that this is not the case for Tarma Quechua, where aspect markers are frequently used in combination with the PHP. There were only a couple examples where the auxiliary had been marked with the durative marker *-sha-*. These few examples were only attested in subordinate clauses, which by themselves are quite rare with the past habitual. In fact, only two examples were found in such an environment, both of which were subordinated with the adverbial subordinator *-qti-*. The similarity between the two, together with their low frequency, suggests that perhaps this is not a case of the construction in question, but perhaps just a subordinated copular clause where the nominal is a lexicalized nominalization.

(6) (One day the order came to move camp to a place near Arequipa,)
Ña [wiksa-y phata-naya-sha-q] $_{NMLZ}$ *ka-sha-qti-n-ña.*
'already' belly-1.A burst-DESID-DUR-NMZR AUX-DUR-DS-3.A-'already'
'When my belly was already about to burst.' (AQW.102)

(7) (Maybe if he had taken care of me,)
sapa [unqu-q] $_{NMLZ}$ *ka-sha-qti-y chay-qa,*
every sick-NMZR COP-DUR-DS-1.A DEM.2-TOP
'whenever I got sick/pregnant,
(my children wouldn't have died.)' (AQW.107)

Quechuan languages have a large inventory of derivational (i.e., non-inflectional) affixes. Some of this affixes may verbalize nouns however here I mean non-class changing derivation. In the case of PAST-HAB, both the nominalized verb and the auxiliary display a strong tendency to appear with very few affixes if any. If the derivational affix is applied, it normally appears within the nominalization. Affixes that do appear tend to be related to the valence properties of the verb (the causative *-chi-*, the reflexive *-ku-*, the benefactive *pu-*), rather than to aspect.

In terms of tense compatibility, the construction was left unmarked in the vast majority of cases. Where explicit marking takes place, the following markers were found: past (*-ra-*), past perfect (*-sqa-* (*ka-*)), and the speaker-non-experienced past tense (*-sqa*). No future tense markers were attested, although this might very well be due to the nature of the corpora. With respect to frequency, the most common marker by far is in fact the lack of one, i.e., the unmarked tense. In my corpus, there were but two attestations of the past tense marker. This, however, has very little consequence for the semantics of the construction as a whole, since, as noted by Cusihuamán (1976) and Soto-Ruíz (1976), the event expressed by the construction is always set prior to speech time.

The nominalization of the past perfect tense is particularly interesting. As noted above, the past perfect is a periphrastic construction which is quite similar to the past habitual construction interms of 'building blocks'. In (8), two habitual sentences connected by *otaq* 'or' areequated. The first contains the nominalized past perfect which appears in the form *V-sqaka-q*(*ku*), while the second contains the simpler form *muna-q* 'want-*q*'. That is, when thepast perfect is marked for the habitual aspect, its auxiliary is nominalized to act as thelexical component of the entire phrase. The scheme has the following hierarchal order: [[[[V-sqa] ka-]q] [ka-]]. This reflects the basic structure [[[LEX-][q]][AUX]].

(8) (There were always lots of drunken hooligans picking fights. Who knows why they'd go to that house,)
[mana-chus warmi-n-ku ka-q] $_{NMLZ}$ *icha warmi-n-ku-wan-chu*
NEG-EV.4 wife-3-PL1 be-*NMZR* or wife-3-PL1-COM-NEG
[maqa-na-sqa ka-q-ku-pas] $_{NMLZ}$, *o-taq-chus [mana warmi-n-ku*
fight-RECP-PP be-NMZR-PL1-ADD or-CONT-EV.4 NEG wife-3.A-PL1
qu-y-ta muna-q] $_{NMLZ}$.
give-INF-ACC want-NMZR

> 'perhaps they didn't have wives or they had a fight with their wives or perhaps their wives did not want to.' (GKM.83)

The construction is negated much like other verbal clauses, using the discontinuous *mana … -chu*. The enclitic usually appears on the nominalized form (9–10). In a minority of cases, the enclitic was missing (as in the case of *muna-q* in (8)).

(9) (When the year was good and there was a lot of harvest, the Arariwa[1] would take for himself one furrow from every field).
Ni-taq pipas [ima-ta-pas ni-q-chu] $_{NMLZ}$.
NEG$_1$-CONT someone thing-ACC-ADD say-NMZR-NEG2
'And no one would say anything.' (GKM.39)

(10) *Mana ñuqa-qa [deja-ku-q-chu]* $_{NMLZ}$ *ka-ni.*
NEG$_1$ I-TOP leave-REFL-NMZR-NEG$_2$ AUX-1.B
'I would not be left alone.' (GKM.21)

Having described the morphosyntactic behavior of the PHP, I now turn to its functions. Most relevant here is the fact that the past-tense meaning is coded regardless of whether the auxiliary, if present, is formally marked for past tense. The main function of the construction is to denote a habitual event. This characterization appears, amongst others, in Cusihuamán (1976) and Soto-Ruíz (1976) where the pattern is dubbed *pasado habitual*. Cusihuamán's definition, reproduced here in my own translation, reads:

> This past form describes: (a) those repeated or everyday activities that correspond to the speaker's past experience or to that of his associates, and (b) the way of life, customs or traditions of different people or group, which were observed for the first time by the speaker. (p. 163)

This definition largely covers the main function of the periphrasis exemplified in (11a), although it should be noted that the difference between (a) and (b) in Cusihuamán's description is morphological in nature where Cusihuamán's (b) is marked with reportative *-sqa*, used to mark non-experienced past events. That is to say, it is not the periphrasis on its own that marks the remoteness of the speaker to the expressed clause, but rather Quechuan morphology, specifically, the marker *-sqa*. In (11a), the speaker Gregorio tells of a time when he and his wife, Josefa, tried having children but to no avail. He then addressed his friends and asked them, "What's good for having children?" They in return mocked him and called him names.

1. A ritualistic guardian of crops and animals (Bolin 2010: 51).

(11) a. *[Uspha aransach'a ni-wa-q-ku]* $_{NMLZ}$.
ash aransach'a say-1OBJ-NMZR-PL1
'They would call me ash-aransach'a (a spindly barren looking tree).'
(GKM.65)

b. *Hina ni-wa-ra-n-ku, chay Don Jacinto Mamani suti-yuq-mi.*
so say-1OBJ-PST-3.B-PL1 DEM.2 Don Jacinto Mamani name-HAVE-EV.1
'So they told me his name was Don Jacinto Mamani.' (GKM.18)

By contrasting the PHP construction in (11a, and (11b), where the verb is explicitly marked solely for PAST, the particular aspectual variety becomes more apparent. The scenario coded in (11a) consists of multiple instances of the same event, which consists of Gregorio's friends calling him "*aransach'a*". All of those micro-events are bundled into one phrase, which is used to characterize Gregorio's friends to whom he came seeking advice. Temporally, the whole event is unbounded, and so are the micro-events composing it. Moreover, the regularity of their calling him names is unspecified and appears to be irrelevant. However, what is relevant is that it happened enough times for him to designate them as such that call him *aransach'a*. This contrasts with (11b), in which he was told by Don Jacinto's workers the name of their master. This event took place once and was completed successfully; the two sentences' viewpoints are in opposition here, in terms of the way Gregorio looks at these two events.

Besides showing the construction's compatibility with compounded tenses, Example (8) shows that the auxiliary may also be nominalized and marked for the past habitual aspect, and so serves as an interesting case. The fact that the verb *ka-y* is fully compatible with this habitual device may seem somewhat problematic, given the intrinsic stativity of copular verbs. However, it should be noted that some instances of *being* can be habitual. A habitual reading is mostly acceptable when the verb is used as a copula and the proposition is delimited adverbially; it does not refer to an ever-lasting state, the core meaning being *being in the capacity of X* or *being while having the property of X*. These copular in-between cases may have acted as bridging contexts that facilitated the compatibility with the lexical *ka-y,* EXIST, and then in turn enabled the expansion to other stative verbs.

(12) *Fabrica llank'a-sha-qti-n-taq-mi, ñuqa-qa [wasi-lla-pi ka-q]* $_{NMLZ}$
Factory work-DUR-DS-3.A-CONT-EV.1 I-TOP home-DEL-LOC be-NMZR
ka-ni wayk'u-spa.
AUX-1.B cook-SS
'And while he was working in the factory, I would be home alone cooking.'
(AQW.111)

In (12), the link between the speaker, Asunta, and her being is delimited by the adverbial "home alone" which segments her existence or better yet narrows the

scope of the predication to a specific aspect of her existence. This in turn enables a habitual reading because it is not her *being* which is habitual, but rather her *being home alone*.

The construction occurs with other stative verbs such as the verb *tiya-y*, which also has two readings – a stative one, 'to dwell', and an episodic one, 'to sit'. In (13a), the speaker tells about a period in his life when he would stay on his own while his master would go on business trips, leaving him to fend for himself. The *staying in the house alone* took place several times during the span of this period. The viewpoint is contrasted in (13b) where the entire *staying alone* event is seen not as a succession of events but as a whole, one singular moment from the perspective of the speaker.

(13) a. *Hina ñuqa-lla sapa-lla-y uywa-kuna-ta michi-spa [wasi-pi*
So I-DEL alone-DEL-1.A animal-PL-ACC pasture-SS house-LOC
tiya-q] $_{NMLZ}$ *ka-ni.*
sit-NMZR AUX-1.B
'So I alone would stay in the house, herding the animals.' (GKM.24)

b. *Chay-pi sapa-lla-y tiya-ra-ni kinsa alqu-lla-ntin.*
DEM.2-LOC alone-DEL-1.A sit-PST-1 .B 3 dog-DEL-COM
'That time, I stayed there alone together with three dogs.' (GKM.29)

However, as noted above, not all examples of stative verbs inflected for the construction could be resolved as being delimited, and therefore pass as habituals. A case in point is *ka-y*'s other two closely related functions denoting existence and possession, where delimitation is not possible.

(14) (Back in those days, I could not go to Cuzco on my own that same day because the Rumiqulqa passage was infamous:)
[Asaltador-kuna ka-q] $_{NMLZ}$
bandit-PL be-NMZR
'There used to be bandits.' (GKM.61)

(15) *Chay estancia-pi [mana unu ka-q-chu]* $_{NMLZ}$
DEM.2 farm-LOC NEG1 water be-NMZR-NEG2
'There used to be no water on that farm.' (GKM.29)

(16) *Pero [mana pago ka-q-chu]* $_{NMLZ}$
but NEG1 rent be-NMZR-NEG2
'But there was no rent.' (GKM.64)

(17) (Next to the butcher's house)
huk canchón hatun rumi qulqa-yuq-ku[sic] ka-ra-n.
a fenced.lot big stone granary-HAVE-PL1 COP-PAST-3.B
'There was a fenced lot with large piles of rocks.' (GKM.24)

Much like (17), which is explicitly marked for the past tense using the affix *ra-*, examples (14–16) refer to past states. Unlike previous minimal pairs, here the difference between (14–16) and (17) appears to be not the composition of the "event" in itself, but rather the significance attributed to its span, which, in turn, characterizes the period: (14) is understood to be a period in the past when bandits lurked in the mountain passes leading to Cuzco; it serves as the reason why "back in those days" the speaker could not go on his own. This suggests a change in the state expressed is possible or perhaps even already took place. Example (17), on the other hand, depicts a singular time frame, which has no duration. We do not know what happened to that pile of rocks and probably neither does the speaker. It may or may not still be there – the point is that it existed for the time frame when the speaker attested it – the rest is deemed irrelevant.

The use of the copula *ka-y* in denoting possession closely relates to it being a verb of existence. Similarities between HAVE and BE have been well studied, notably by Benveniste (1970), Isačenko (1974) and more recently Stassen (2009). Isačenko even suggests that HAVE could be analyzed, in essence, as BE + transitivity for Whorf's Standard Average European. Clancy (2010) dedicates an entire chapter showing how BE and HAVE cohere. If one applies Isačenko's terminology, Quechua is a typical BE-Language where the concept of possession is expressed by the same verb denoting EXIST. The connection between BE and HAVE in Quechua appears to be very tight, and other verbs, which have acquired the meaning COPULA and EXIST, also function in HAVE constructions in some varieties such as those spoken in Bolivia (Stark 1971).

The semantics of such a construction are tied in with the binary property of EXIST – namely, something either exists or it does not. Literally, the construction denotes *Pm BE for Pr*, where *Pm* is the possessum and *Pr* is the possessor. Because of this, it too cannot be delimited either. HAVE's stativity is essentially the same as that denoted in (14–16). In (18), it is not implied that the speaker's head could grasp the use of the alphabet intermittently. We actually know for a fact that he was illiterate from Valderrama and Escalante's writings. The whole situation is confined to when they tried teaching him to read and write.

(18) (In the army the taught me the alphabet)
Pero mana-chu hina [uma-y ka-q] $_{NMLZ}$ *abedecedario-paq.*
but NEG1-NEG2 so head-1.A be-NMZR alphabet-DAT
'But I didn't have the head for the alphabet.' (GKM.45)

In (19), the speaker tells of her employer's husband who tried to rape her three times and describes him as being a devil. A habitual reading where the husband would turn into a devil from time to time is also possible, though highly unlikely due to the context. A habitual reading seems incompatible if we understand the

speaker to tie his devilishness to him trying to rape her three times. An event that takes place three times is not habitual. The speaker seems to have confined the devilish existence of her tormentor to a certain period in her past, as if the tormentor is no longer relevant to her present state. Whether the husband is still alive or not is unknown, but the use of *ka-q* signals that the situation is no longer relevant for the speaker's present state. This reading is reiterated in (20) and perhaps more strongly in (21).

(19) *Qosa-n-pas [huk diablo-lla-taq ka-q]* $_{NMLZ}$.
husband-3.A-ADD a devil-DEL-CONT be-NMZR
'And her husband was just a devil too.' (AQW.96)

(20) a. *Allin mikhuna ka-ra-n.*
good food COP-PST-3.A
b. *Hinaspa-pas siempre ñuqayku-paq [[ima-lla-pas*
SO-ADD always we-DAT something-DEL-ADD
ka-q] $_{NMLZ}$ *-puni-n].*
be-NMZR-indeed-EV.1
'There was good food, and so there was always some small thing for us to eat.' (GKM.78)

(21) (During service)
Ñuqa [huq pobre ka-q] $_{NMLZ}$ *ka-ni, mana ima-y-pas ni hayk'a-y-pas*
I a poor be-NMZR AUX-1.B NEG 1 thing-1.A-ADD nor thing-1.A-ADD
ka-n-chu. Kunan ñuqa-pas allin.
COP-3.B-NEG2 … now I-ADD good
'I used to be poor, I had not a thing nor anything,. … Now I am well.'
(RXL2–010801)

Much like the irrelevancy of the state denoted by *kaq* to the speaker's current state in (20b) and (21), in HAVE sentences, there seem to be an implicature that the possessum is no longer possessed. This is especially evident with the obligative. The obligative is built upon the HAVE construction, except the possessum is actually a nominalization that makes use of the affix *-na-*.[2] In (22), there is no doubt that the necessity for working in that house no longer holds for the speaker, because she had moved out. It is a description of the entire period as a whole since a period of time when she did not have to work while staying in that house is not implied.

2. [sic] is not unlike English *I have to work*. For more on the possession-to-obligation cline see Łęcki (2010) and Bhatt (1997).

(22) *Chay wasi-pe-qa allin-ta-n trata-wa-ra-n-ku, pero*
DEM.2 house-LOC-TOP GOOD-ACC-EV.1 treat-1OBJ-PST-3.B-PL1 but
[llank'a-na-y ka-q] $_{NMLZ}$.
work-OBL-1.A be-NMZR
'They treated me well in that house, but I had to work.' (AQW.97)

The use of *kaq* and other stative verbs suggests that the scope of the so-called *pasado habitual* exceeds most definitions of the habitual aspect since it allows for verbs that are semantically incompatible with habitual devices. This larger domain subsumes the habitual, thereby enabling two different albeit related readings.

In a small number of tokens, the inherent pluractionality associated with the habitual aspect colored a verbal argument – the actor. This in turn leads to distributive pluractionality and specifically what Dressler 1968 calls "subject distributive". Typically, these token were achievement verbs and the event depicted took place only once. In (23), the past event is not protracted over a time expanse. The plurality is event-external and the result of an action of multiple actors rather than a singular one. The 'bundling' in this case is not of one generalized macro-event, but rather of multiple separate events, each with a different actor, all sharing in the same activity of eating. In (24), the lexeme in itself seems to be incompatible with habitual reading.

(23) *Asno-kuna-qa [mikhu-q]* $_{NMLZ}$ *huk paisano mankiru Sicuani*
donkey-PL-TOP eat-NMZR a compatriot pot-trader Sicuani
lado-manta caballo-kuna-wan llama-kuna-wan kuska.
side-ABL horses-PL-COM llama-PL-COM together
'The donkeys ate next to a pot trader from Sicuani, together with the horses and llamas.' (GKM.27)

(24) *Khayna-n [gente wañu-q]* $_{NMLZ}$.
Like.this-EV.1 people die-NMZR
'Like that, people were dying.' (AQW.92)

The connection between habituality and distributive pluractionality is not rare typologically. Yu (2003), following Cusic (1981), lists a distributive parameter under verbal plurality and notes that the pluractional in Chechen can be used to signal habituality and distributivity. Similarly, the so-called past iterative or habitual Lithuanian affix *-dav-* normally implies a habitual meaning (Mathiassen 1996: 9). However, as noted by Lea Sawicki and Efrat Miller in a personal communication, in some cases, the affix allows for a preferred distributive reading. In (25) for example, the interpretation of the plurality of the actor is supported by the use of the generic *žmogus* ('man'), which cannot be interpreted as having the habit of facing the monster simply because upon doing so, the actor associated with

that particular event-instance dies. This means that a better way to translate the Lithuanian sentence would be with "any person".

(25) *Susidūręs su tos baisybės žvilgsniu, žmogus nutirpdavo*
meet.PART with this monster's look man
ir krisdavo kaip perkūno trenktas
freeze.FREQ.PART and fall.FREQ.PART like thunder struck
'Upon meeting the gaze of the monster, one would freeze and fall as if stricken by thunder.'

In this section, I have described the functions of a construction that is said to mark the past habitual. Although it does indeed do so in the majority of the cases, the compatibility of periphrasis with stative verbs seems to indicate that its scope exceeds most current definitions for the habitual aspect. It was shown that in some instances habitual readings can be acceptable with the verb *ka-y*, which has several functions. When used in copular sentences, some readings were acceptable because the scope of the predication could be limited to a certain aspect of being (12). However, in some cases it appears undelimited, which gives rise to a non-habitual reading (19a). The copula *ka-y* was also shown to be compatible with the periphrasis when used to denote EXIST (14–16) or HAVE (18), binary functions for which the inherent stativity of the verb is essential. The stativity, in turn, blocks habitual readings that require episodic events in order for a characterizing generalization to take place. In such cases, the habitual aspect is neutralized and the construction marks an unbounded event which is set prior to the speaker's utterance. In essence, the nominalizer *-q* was found to mark a more general past imperfective. This construction also conveys that the state is irrelevant to the speaker's current state or no longer holds true (20–21). In some cases, the pluractionality associated with the macro-event was displaced and colored a different verbal argument – the actor. This led to a distributive reading (23–24).

4. The emergence and expansion of the construction

The past habitual periphrasis in Quechua most likely developed from a simple copular clause with the proposition *X BE V-er*. This copular clause then underwent reanalysis, and instead of having the constituency portrayed in (26a), the one in (26b) took precedence, which gave rise to the periphrastic construction evident today.

(26) a. Nominal sentence: [V-q] [(ka-)]
b. Auxiliary: [V][-q (ka-)]

In the copular stage, the copula inflected freely and took tense marking, similarly to how it is marked today. In the stage suggested in (26b), the copula *ka-y* still retained the same functions as it had when it was used in proper copular clauses; it was still inflected for present and past. This enabled the PHP to encode both present and past habituals.[3] The compatibility with the present tense is evident in (27), which is taken from the 16th century Huarochiri manuscript. Here, Wayna Qhapaq, after having conversed with Quniraya, wishes to send shamans and magicians to Ura Tiksi. Several shamans reply and proclaim their abilities and some state that they *fly* as swallows do.

(27) *Ñuqa-m [[wayanay pisqu pawa-q]* $_{NMLZ}$*]*$_{NP}$ *ka-ni.*
I-EV.1 swallow bird fly-NMZR COP-1.B
'I fly as a swallow.' (Huarochiri.9)

In more ways than one, the case of English 'used to', bears resemblance to the past habitual in Quechua. 'Used to' can be traced back to around 1400 when the verb *user* entered the English language from Old French in the sense "to follow a usage or custom". As a main verb, it inflected freely, much like the Quechuan copula of Stage II. Because of the perfective nature of the past tense and the imperfective nature of the present tense in English, the use of the habitual gram in the present tense disappeared, which transformed the general habitual gram to a dedicated past-habitual gram (Bybee, Perkins & Pagliuca 1994: 155).

In Cuzco Quechua, as well as in Conchucos Quechua (Hintz 2007: 320), something else happened. The past tense meaning remained but its overt marking was no longer deemed obligatory. If we consider past marking in Cuzco Quechua to be unstable, at least to some extent, since the vast majority of examples are left unmarked (~present tense), one could hypothesize the following stages:

(28) a. Stage I – Nominal sentence: V-*NMZR* + [COPULA[present/past{past}]]
b. Stage II – Past/Present Habitual V-*NMZR* + [AUX[present/past{past}]]
c. Stage III – Past Habitual V-*NMZR* {past} + AUX[present/past]
d. Stage IV – Past habitual V-*NMZR* {past} + AUX[present]
e. Stage V – P. Imperfective V-*NMZR* {past} + AUX[present]

In Stage I, the construction was a simple copular clause, composed of a nominalized clause which served as a predicate. The copula here was marked for tense and person. The copula's function still did not change in Stage II; however, there was a shift in meaning – it acquired a new meaning, i.e., habituality. In Stage III, the construction grammaticalized further and marked PAST-HAB exclusively, even

3. In Cerrón-Palomino 2008: 144, the habitual paradigm is reconstructed in full.

though the auxiliary was still inflected for both the present and the past tense. I propose that because the construction in its most minimal form, the bare nominalization, still encodes past tense meaning, it is the nominalizer that encodes it. And in fact, no other reading is available nowadays (i.e., present tense interpretation). This is most clear when we consider that the copula is not required to anchor the situation temporally, as evident by the zero marking in the third person singular and plural. This is untrue of simple copular clauses where the copula is required to signal third person past in the form *ka-ra-n*, as in (30). Hintz (2007) suggests similarly that Conchucos Quechua had undergone similar development and "the past meaning was absorbed by the remaining phonological material." (p. 320).

(29) *[Mikhu-q]*$_{NMZR}$*-mi.*
eat-NMZR-EV.1
'He used to eat.'

(30) *Mana carro ka-ra-n-chu.*
NEG$_1$ car COP-PAST-B.3-NEG$_2$
'There were no cars.' (GKM.25)

Stage IV marks the fossilization of the present tense; since the temporal deixis falls under the domain of V-*q*, the AUX is basically left unmarked for tense (~present). In Conchucos Quechua too, it seems, no variation in tense occurs nowadays and the auxiliary never takes past tense marking.

Two extensions that occurred with English 'used to' most likely also occurred in Cuzco Quechua and gave rise to the emergence of the past imperfective (stage V). According to Bybee, Perkins and Pagliuca (1994) 'used to' appeared solely with animate subjects at first, and in the 1600's extended to inanimate subjects as well. Although a deeper investigation might be in order, it is not implausible that a periphrasis based on an agentive affix, which presupposes animacy, if we understand it as bearing the semantics *doer of X*, should share the same path.[4] In my corpus, the majority of examples were used with animate subjects, but inanimate subjects were also attested. In (31), the speaker describes a location in which she resided for some time as being inhospitable. Here, she uses the periphrastic construction with an impersonal utterance ('it was cold') and with the noun 'wind'.

(31) *Pero nisyu-ta [chiri-q]*$_{NMZR}$*. wayra-taq cuchillo hina-raq [phawa-q]*$_{NMZR}$
but very-ACC be.cold-NMZR wind-ADD knife like-'just' fly-NMZR
'But it was very cold and the wind would fly like a knife.' (AQW.102)

4. The agentive in Quechuan can be used to refer to inanimate nouns, but actancy appears to be connected on a semantic level with the agentive. One way of achieving a non-agentive noun is by introducing a valency-reducing affix such as the reflexive *-ku-*.

The second extension that took place relates to the compatibility of the habitual gram with stative verbs, which in turn facilitated the emergence of a past imperfective. This compatibility caused some scholars to deny habitual value to the habitual grams that show such compatibility. Bybee, Perkins and Pagliuca (1994) see very little difference between stative verbs with 'used to' and stative verbs marked for past (p. 156). Bertinetto (1994: 130, my translation) states that "the habitual value does not constitute the basic meaning of the construction, but rather a mere occasional and pragmatic extension." For him, the function of a 'habitual' gram that combines freely with purely durative verbs is better described as continuous aspect. Because the periphrasis also locates the event temporally in the past, he suggests the label "*confinamento nel passato*" (p. 40). Binnick (2005) claims that habitual grams such as 'used to', and for our purposes *q (ka-*), have more to do with relevance to the current state. He states that 'used to' "functions as a kind of present perfect, but instead of linking a past event with the present state (…), the *used to* construction does just the reverse, divorcing the past situation from the present era." However, he adds that "that a past habit is no longer the case is not part of the meaning of the expression itself but a conversational implicature." (p. 351) In keeping with this, Hantson (2005) suggests labeling the English construction as an anti-perfect, which might very well be an apt label for the the Cuzco Quechu PNP construction.

This past tense semantics bearing the conversational implicature that the past tense no longer holds was shown in Example (21).

The extension to stative verbs could have occurred in several ways. It may have occurred due to the similarities in the event composition of both statives and the habituals. Essentially, they are both non-eventives that are used to denote states. The habitual employs generalizations of several perfective events to denote their being held true, while statives are inherently held true throughout a given time frame.

Another means of extension could be the existence of a bridging context. In this case, the polyfunctional verb *ka-y,* which serves both as a copula and a lexical verb, probably became compatible with the habitual device only while serving as a copula (*being in the capacity of X* or *being while having the property of X*) and only then for its second function of a verb denoting EXIST. Once it was made compatible with existential *ka-y*, it was extendable to other stative verbs.

It has been noted that the Quechuan *Sprachraum* could be viewed as "an auspicious laboratory in which to test old and new hypotheses formulated on the basis of European languages".[5] The purpose of this section is to put to the test grammaticalization clines for the habitual aspect suggested by scholars, with specific

5. Cerrón-Palomino 2003: 41; here in Hintz's (2007) translation (p. 20).

attention to the cline suggested in Bybee, Perkins & Pagliuca (1994). The evolution of the construction in question largely replicates the cline suggested and supports the findings, but not in full.

In their study, they note that sources for habitual grams are related to the verbs *know* and *live* in various languages. The connection between *know* and *live* and the habitual appears to be straightforward – knowing how to do something implies doing it more than sporadically. 'Live' is perhaps less straightforward and Heine and Kuteva (2002) suggest it first passes through a continuous stage and only then grammaticalizes to habituals or progressives. This is quite similar in meaning to the core meaning of the agentive, which indicates that the actor partakes in the activity denoted by the lexical verb so much so that it is used to characterize him or her.

Bybee, Perkins and Pagliuca (1994) adopt, *grosso modo*, Comrie's definition of the habitual aspect but they also state that "the difference between habitual and stative is minimal" (p. 152). Their findings suggest that habitual grams are "highly affected by tense", for which they find evidence in the asymmetry that characterizes past habitual and present habitual markers (p. 151). The overt marking of the habitual aspect is much more common when the habitual event refers to the past than when it refers to the present. They note that only two languages had overt present-habitual marking in comparison to ten languages that had overt past-habitual marking. The tense-affected asymmetry is also attested in Cuzco Quechua where two different strategies are employed to mark the habitual: the unmarked tense for the present-habitual and the periphrastic construction which is the subject of this paper.

They also postulate that the reason for such asymmetry could be traced to an overlap between simple past meaning and perfectivity and a connection between present meaning and imperfectivity. The Quechua data corroborates this claim too, perhaps more strongly so, since the unmarked ('present') tense does not strictly bind an event temporally. It can be used to refer to all realized events, and because of that would be better considered as a general imperfective. The imperfective subsumes present meaning and therefore, unmarked utterances have several different readings, one being habitual-gnomic. When an utterance is marked for the past tense, on the other hand, "in order to explicitly talk about habitual or ongoing situations, (…), one needs to add extra elements into one's utterance" (p. 153).

The ten languages in the GRAMCATS sample that had a dedicated past-habitual gram had obligatory past tense marking. It is hypothesized that this is tied to the evolution of such grams:

> These habitual grams began as more general habitual locutions that were compatible with either past or present. However, since explicit reference to habitualness

> is necessary in the past, but not in the present, the habitual construction was used more frequently in the past than in the present. Hence it grammaticized in the past and was lost in the present. (Bybee, Perkins & Pagliuca, 1994: 154)

Past tense marking, which renders the periphrasis *V–q ka-ra-*, though attested in some dialects such as Huanca, Imbabura and Lambayeque (Hintz 2007: 320), is not obligatory in Cuzco Quechua. This suggests that for the grammaticalization of past habitual, it is not past tense *marking* which is obligatory, but rather it suffices that the construction as a whole comes to be associated with past tense *meaning*.

The suggested pathway of grammaticalization for the past habitual is largely replicated in Cuzco Quechua. Here too a nominal serves as a source for a flavor of the imperfective aspect. The typological asymmetry in habitual marking found in the GRAMCATS corpus was also found here, where the unmarked (~present) tense, being a general imperfective, is used for present-habitual, while an overt gram is used for the past-habitual.

Unlike in the languages studied by Bybee, Perkins and Pagliuca, the past-habitual gram in Cuzco Quechua does not co-occur with an obligatory past marker. Here, the past tense was associated with the nominalizer, which acquired a past tense meaning, thereby enabling the copula to not appear with it. The fact that the nominal now has past-tense meaning suggests that perhaps it is no longer a nominal and is perhaps undergoing verbalization. Another case in point here is the construction's pluralization in the third person. Unlike copular clasues where the copula appears and marked for the third person plural, with the "periphrasis", the verbal non-second-person pluralizer *-ku-* appears, attaching itself directly to the now evidently verbal stem.

(32) a. *Kay atuq-kuna-n mañoso ka-n-ku.*
DEM.1 fox-PL-EV.1 crafty COP-3.B-PL1
'These foxes are crafty.' (GKM.38)

b. *Lliw asno-lla-pi [puri-q-ku]$_{NMZR}$ (*kanku)*
all donkey-DEL-LOC walk-NMZR-PL1
'Everyone would travel on donkeys.' (GKM.25)

It would appear that the periphrasis is moving further up the cline and is being more and more grammaticalized. The compatibility with stative non-pluractional verbs suggests that the function of *q* expanded to a more general past imperfective. As noted above, there is a resemblance between English 'used to' and the construction described here, both in terms of evolution and current meaning. The imperfectivity attributed here to past habitual construction is also attributed its counterpart in English by Bertinetto and Lenci (2012), who suggest that the term *gnomic imperfective* is better suited to describe such constructions. Bybee, Perkins and Pagliuca (1994) state that imperfectives develop from progressive meaning,

which is generalized into continuous meaning (p. 142). Here, however, though requiring further study and more a diverse corpus, it seems that habitual meaning is being generalized into continuous, which suggests that it too could probably develop into a more generalized imperfective meaning.

The construction also appears to be undergoing verbalization, as evident from the past meaning and verbal pluralization it acquired. This lends further credence to the claim that nominalizations serve as sources for main clauses, as suggested by Gildea for the Carib languages (2000: vii). He ascribes such a process to languages where actor nominalizations occur instead of finite complement clauses. This enables the development of TAM markers, not unlike #-*q* (*ka*-)#, which in turn evolve into main clauses. Gildea notes that a mid-stage involving an auxiliary usually takes place "but sometimes not, especially in those languages that do not require a copula for all forms of nonverbal predication." (2008: 12).

5. Concluding remarks

The purpose of this paper was to describe the functions of the periphrastic construction typically dubbed *pasado habitual* in Quechuan linguistics Although it does signal this function, its aspectual domain was found to be much broader. It was found to exceed the scope of most definitions of the habitual aspect, a subflavor of the imperfective aspect. The construction's full compatibility with verbs of state means that event-generalizations, an inherent property of the habitual devices, are contingent upon ASPECT_1. It follows then that signaling habituality is not the exclusive or perhaps even dominant function of the periphrastic construction but rather a common potential expression thereof. This is not unlike English 'used to', which some scholars agree does not signal habituality (Binnick 2005; Boneh & Doron 2008; Bertinetto & Lenci 2012 among others). The terms applied to 'used to' are 'confinement in the past' (Bertinetto & Lenci 2012) or even 'present perfect' (Binnick 2005).

Much like the English periphrasis, the Quechua PNP construction was found to signal that an event took place prior to speech time. The event is usually protracted over a time expanse which the speaker deems to be substantial, or else it could not be used to characterize the time frame. There is also an implicature that the event's endpoint, though not specified, is located prior to the speech time – that is, that the event no longer holds. Whether this is merely a conversational implicature, as suggested by Binnick for the English construction, requires further investigation.

A different behavior of the construction, one that was found only with achievement verbs, is event-external subject distributivity. In these marginally attested

cases, the pluractionality of the habitual pluralized, not the event itself, but a different argument – the actant. That is, instead of having a singular event pluralized and repeated, a number of sufficiently similar events, with potentially different actants, are bundled up together and presented as a set. One possible explanation for such a behavior could be the perfectivity associated with telicity in combination with punctuality: If perfectivity is coerced on an imperfective pluractional construction, maybe distribution over participants could ensue. This, of course, requires more examples and further study.

The evolution of the construction is valuable for typologists and those interested in the domain of diachronic typology, specifically the grammaticalization of TAM grams. As it was shown, contrary to what Bybee, Perkins and Pagliuca (1994) found in the GRAMCATS corpus, it is not past *marking* which is required in the grammaticalization of PAST-HAB but rather past *meaning*. Although the distinction might appear to be inconsequential, it exemplifies the evolution, not of PAST-HAB, but of verbal expressions and main clauses.

Acknowledgments

I gratefully acknowledge the numerous helpful comments, suggestions and insights on earlier drafts of this article made by Jan Szeminski, Eitan Grossman, Willem Adelaar, Nora Boneh and Tally Shitrit. I would also like to thank the editors of this volume, Masayoshi Shibatani and Roberto Daniel Zariquiey Biondi, for their patience and sapience.

Abbreviations

1	First person	EV.1	Direct evidential
2	Second person	EV.2	Reportative evidential
3	Third person	GEN	Genitive
1OBJ	First person object	HAVE	Possession
A	Verbal index set	INF	Infinitive
ACC	Accusative	LOC	Locative
ADD	Additive	NEG	Negator
AUG	Augmentative	NEG1 …NEG2	Predicative negation
AUX	Auxiliary	NMZR	Nominalizer
B	Nominal index set	PL	Nominal plural marker
CIS	Cislocative	PL.1	Non-second person pluralizer
COM	Comitative	PL.2	Second person pluralizer
CONT	Continuative	PP	Past participle
COP	Copula	PST	Past

DAT	Dative	RECP	Reciprocal
DEL	Delimitative	REFL	Reflexive
DEM	Demonstrative	SS	Same subject
DESID	Desiderative	TOP	Topic
DS	Different subject	UNMRKD	Unmarked tense
DUR	Durative		

References

Adelaar, Willem F. H. 1977. *Tarma Quechua: Grammar, Texts, Dictionary*. Lisse: The Peter de Ridder Press.

Adelaar, Willem F. H. & Muysken, Pieter. 2004. *The languages of the Andes*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511486852

Alan, C. L. 2003. Pluractionality in Chechen. *Natural Language Semantics* 11(3): 289–321. https://doi.org/10.1023/A:1024997718758

AQW = See Mamani et al.

Benveniste, Émile. 1970. The linguistic functions of to be and to have. In *Problems in General Linguistics*, Mary Elizabeth Meek (trans), 163–179. Coral Gables FL: University of Miami Press.

Bertinetto, Pier Marco. 1994. Le perifrasi abituali in italiano ed in inglese. *Quaderni del Laboratorio di Lingüística* 8: 32–41.

Bertinetto, Pier Marco & Lenci, Alessandro. 2012. Habituality, pluractionality, and imperfectivity. In *Oxford Handbook of Tense and Aspect*, Robert Binnick (ed.), 852–880. Oxford: OUP.

Bhatt, Rajesh. 1997. Obligation and possession. In *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect* [MIT Working Papers in Linguistics 32], Heidi Harley (ed.), 21–40. Cambridge MA: MITWPL.

Binnick, Robert I. 2005. The markers of habitual aspect in English. *Journal of English Linguistics* 33(4): 339–369. https://doi.org/10.1177/0075424205286006

Bolin, Inge. 2010. *Rituals of Respect: The Secret of Survival in the High Peruvian Andes*. Austin TX: University of Texas Press.

Boneh, Nora & Doron, Edit. 2008. Habituality and the habitual aspect. In *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect* [Linguistik Aktuell/Linguistics Today 110], Susan Deborah Rothstein (ed.), 321–349. Amsterdam: John Benjamins. https://doi.org/10.1075/la.110.14bon

Boneh, Nora & Doron, Edit. 2013. Hab and Gen in the expression of habituality. In *Genericity*, Alda Mari, Claire Beyssade & Fabio del Prete (eds), 176–191. Oxford: OUP.

Brinton, Laurel J. 1987. The aspectual nature of states and habits. *Folia Linguistica* 21(2–4): 195–214.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago IL: University of Chicago Press.

Cerrón-Palomino, Rodolfo. 2008. *Quechumara: Estructuras paralelas de las lenguas quechua y aymara*. La Paz: Centro de Investigación y Promoción del Campesinado.

Cerrón-Palomino, Rodolfo. 2003. *Lingüística Quechua*. Cuzco: Centro De Estudios Regionales Andinos "Bartolomé De Las Casas".

Clancy, Steven J. 2010. *The Chain of Being and Having in Slavic* [Studies in Language Companion Series 122]. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.122

Comrie, Bernard. 1976. Aspect: An Introduction to the Study of Verbal Aspect and Related Problems. Cambridge: CUP.

Cusihumán, Antonio G. 1976. *Gramática quechua: Cuzco-Collao*. Lima: Ministerio de Educación: Instituto De Estudios Peruanos.

Dressler, Wolfgang U. 1968. *Studien zur verbalen Pluralität*: Iterativum, Distributivum, Durativum, Intensivum in der allgemeinen Grammatik, im Lateinischen und Hethitischen [Österr. Akad. der Wiss., Phil.-Hist. Kl., 259.1]. Wien: Hermann Böhlaus.

Gildea, Spike. 2008. Explaining similarities between main clauses and nominalized phrases. *La Structure des Langues Amazoniennes* 32: 57–75.

GKM = See Mamani et al.

Hantson, André. 2005. The English perfect and the anti-perfect used to viewed from a comparative perspective. *English Studies* 86(3): 245–268. https://doi.org/10.1080/0013838042000339862

Heine, Bernd and Tania Kuteva 2002. *World lexicon of grammaticalization*. Cambridge: CUP.

Hintz, Daniel J. 2007. Aspect and Aspectual Interfaces in South Conchucos Quechua: The Emergence of Grammatical Systems. PhD dissertation, University of California, Santa Barbara.

Huarochiri = See Taylor.

Isačenko, Alexander V. 1974. On have and be languages: A typological sketch. *Slavic Forum: Essays in Linguistics and Literature*, 43–77. The Hague: Mouton.

Kleiber, Georges. 1987. *Du côté de la référence verbale: Les phrases habituelles*.Berne: Peter Lang.

Lanata, Xavier Ricard. 2010. Annexes. *Les voleurs d'ombre: l'Univers religieux des bergers de l'Ausangate, Andes centrales*. Société d'ethnologie. < http://www.mae.u-paris10.fr/socethno/ (4 June 2013).

Łęcki, Andrzej M. 2010. *Grammaticalisation Paths of Have in English*. Frankfurt: Peter Lang. https://doi.org/10.3726/978-3-653-00288-1

Mamani, Gregorio Condori, Valderrama Fernández, Ricardo & Escalante Gutiérrez, Carmen. 1983. *De nosotros, los Runas: Autobiografía*. Madrid: Alfaguara.

Mathiassen, Terje. 1996. *Tense, Mood and Aspect in Lithuanian and Latvian*. Oslo: Universitetet i Oslo, Blavisk-baltisk avdeling.

Reichenbach, Hans. 1947. The tenses of verb. In *Elements of Symbolic Logic*, 287–298. New York NY: Macmillan.

RXL = See Lanata.

Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state. *Linguistic Typology* 6(2): 199–271. https://doi.org/10.1515/lity.2002.007

Soto Ruíz, Clodoaldo. 1976. *Gramática quechua: Ayacucho-Chanca*. Lima: Ministerio de Educación.

Stark, Louisa Rowell. 1971. *Sucre Quechua: A Pedagogical Grammar*. Madison WI: University of Wisconsin.

Stassen, Leon. 2009. *Predicative Possession*. Oxford: OUP.

Taylor, Gérald. 1999. *Ritos y tradiciones de huarochirí*. Lima: Instituto Francés de Estudios Andinos.

Weber, David. 1989. *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley CA: University of California Press.

CHAPTER 9

Life of =*ti*: Use and grammaticalization of a clausal nominalizer in Yurakaré

Sonja Gipper[1,2] and Foong Ha Yap[3,4]
[1]University of Cologne / [2]Australian National University / [3]Hong Kong Polytechnic University / [4]Chinese University of Hong Kong, Shenzhen

In this paper, we investigate the uses of the clausal nominalizer =*ti* in Yurakaré, a linguistic isolate spoken in Bolivia. Clauses nominalized with =*ti* can serve a variety of functions: filling an argument position, relativization, forming the complement of a complement-taking verb, and expressing adverbial modification. On the basis of synchronic spoken corpus data, we propose a grammaticalization path for =*ti.* We argue that its most plausible source is the demonstrative *ati,* thus suggesting that =*ti* is a demonstrative-based nominalizer. Further, we show that =*ti* has developed a range of insubordinate uses, indicating 'intersubjective commitment'. We propose that from there, =*ti* is currently on its way toward becoming a stance marker, contrasting with other clause-final enclitics of Yurakaré.

1. Introduction

In South American languages, nominalization constructions are frequently used as subordination strategies (van Gijn 2014). This is also the case in Yurakaré, a language isolate spoken in central Bolivia. In this paper, we examine the different uses of Yurakaré subordination constructions with the nominalizing enclitic =*ti* based on evidence from synchronic corpus data (van Gijn, Hirtzel & Gipper 2011).

As a clausal nominalizer, =*ti* can form nominalized clauses with a variety of functions: filling an argument position, relativizing, forming complements of complement-taking verbs, and expressing adverbial modification with temporal, causal, concessive, conditional, and locative interpretations. We describe these uses in Section 2. Relativizing, complementizing, and adverbial functions are also reported for other nominalizers in this volume, for instance in Harakmbut (isolate; Van linden, this volume), Kakataibo (Panoan; Valle & Zariquiey, this volume) and Mochica (extinct, isolate; Eloranta, this volume).

https://doi.org/10.1075/tsl.124.09gip

Previous studies have shown that nominalization constructions are often marked by (sometimes erstwhile) referentiality-marking devices such as demonstratives (e.g. Tagalog *ang* and *yung*; see Nagaya 2011). We suggest in Section 3 of this paper that *=ti* is derived from, or is at least related to, the demonstrative *ati* 'that', a connection that has been suggested before (van Gijn 2006: 313; 2011: 182).

In Section 4, we further demonstrate that clauses marked with *=ti* show a variety of insubordinate uses (see Evans 2007), which we argue finally leads to the grammaticalization of *=ti* as a clausal enclitic indexing the speaker's stance. Such stand-alone uses of nominalizations are also reported for other languages described in this volume, for instance Kakataibo (Panoan; Valle & Zariquiey, this volume) and Xavante (Jê; Machado, this volume).

Where appropriate, the results from our corpus study will be compared to previous findings from other, mostly Asian, languages (e.g. Noonan 1997, 2008, 2011; Watters 2008; Simpson 2008; Yap & Grunow-Hårsta 2010; Yap, Grunow-Hårsta & Wrona 2011 and papers therein; Shibatani 2013) to help identify robust grammaticalization pathways across languages whereby nominalization constructions extend from referential to non-referential domains.

In the following subsections 1.1 and 1.2, we give a brief background on the Yurakaré language and introduce the data used in the corpus analysis, before we proceed to the analysis of the data in Section 2.

1.1 Background on Yurakaré

Yurakaré is a linguistic isolate spoken in the area of the Andean foothills in central Bolivia by about 2,000 speakers.[1] The language is considered endangered due to a break in intergenerational transmission. Most speakers nowadays are bilingual in Yurakaré and Spanish, and most children no longer acquire active competence in Yurakaré. Van Gijn (2006) is a full reference grammar on the language. An ethnographic account is provided by Hirtzel (2010).

In Yurakaré, core grammatical relations are not identified by case marking but rather by cross-referencing on the verb. Only non-core participants are marked with postpositional clitics (see van Gijn 2006: 106, 143). This means that subject and object arguments do not differ from each other in their formal marking; they are only identified by cross-referencing on the verb.

1. We consider the language an isolate since all attempts at genetic classification have failed; see van Gijn (2006: 7-8) for a discussion.

1.2 Data

The data for this study were recorded between 2006 and 2011 during a documentation project of the Yurakaré language, and during a field trip in 2013 by the first author. The 2013 data were recorded using the Family Problems Picture Task, an interactive problem-solving task where the speakers describe a range of pictures and form a story out of them (see San Roque et al. 2012). Most of the data cited in this paper are conversational in nature and include at least two participants. The data consist of video-recordings that were transcribed and translated by native speakers of Yurakaré. The data are part of the Yurakaré section (van Gijn, Hirtzel & Gipper 2011) of the DobeS archive for endangered languages at the MPI for Psycholinguistics, Nijmegen, The Netherlands.

We conducted a quantitative study on a subcorpus of 5.75 hours of data from various villages. The data come from 24 different speakers, of which 13 are female and 11 male. The youngest speaker is 15, and the oldest 73. The mean age of the sample is 49.62, and the median 49.5 years. We analyzed every instance of =*ti* in this subcorpus.

2. Subordinating uses of =*ti*

In this section, we examine the different uses of constructions with the nominalizer =*ti* in our Yurakaré corpus. The marker =*ti* is only used to nominalize finite clauses; lexical nominalizations of the type *sing-er* (see Shibatani, this volume) are not attested for =*ti.* As a versatile nominalizer, =*ti* is used to form subordinate structures with four different functions: argument function, relativizing function, complement function, and adverbial modification function. These four functions are exemplified in (1) to (4) respectively.

(1) *ana ku-i-mala-uma-y sëë ana* [*a-dyindye*
DEM 3SG.OBJ.COM-VPL-go.SG-DST-1SG.SBJ 1SG DEM IPFV-sad
tütü-ø=ti]$_{\text{NMLZ}}$
be-3SG.SBJ=NMZR
'This is what I think about this one who is sad.' (SocCog-YUZ107-2)[2]

2. The names of the files in the Dobes-Yurakaré Archive are given in brackets following the respective example.

(2) *yokkoshe bobo-m tiya-m=chi=laba ush naa sewwe*
really kill-2SG.SBJ eat-2SG.SBJ=FR=SUBJ before DEM boy
[*mi-n-kaya-tu=ti*]$_{\text{NMLZ}}$ *ku-ta-w*
2SG.OBJ-BEN-give-1PL.SBJ=NMZR 3SG.OBJ.COM-say-3PL.SBJ
'"Did you really kill and eat the boy that we brought you?" they said to her.'[3]
(YURGVDP08oct06–01)

(3) *nij ta-kusu-ø* [*ati dia=y sawata-ntu=ti*]$_{\text{NMLZ}}$
NEG 1PL.OBJ-want-3SG.SBJ DEM day(SP)=LOC work-INT:1PL.SBJ=NMZR
'We do not want to work on that day.' (loreto)

(4) [*ta-ka-n-toro=ti*]$_{\text{NMLZ}}$[4] *baytu tishi ta-sibbë=chi*
1PL.OBJ-3SG.OBJ-BEN-finish=NMZR go.1PL.EXH now 1PL.POSS-house=DIR
'When we finish it, let's go to our house immediately.'
(YURGVDP04oct06–02)

The four different functions of *=ti*-marked clauses can be split into two subgroups: participant nominalizations (argument and relativizing function) and event nominalizations (complement and adverbial function). These two subgroups differ regarding their morphological potential. Malchukov (2004, 2006) notes that in the process of deriving one category from another, there can be two processes at work, decategorization and recategorization:

> [A] verb used as a referring expression, apart from losing some of its verbal trappings, usually also acquires a number of nominal properties such as case, determiners, etc. Thus a traditional term like "nominalization" actually conflates two distinct operations; "deverbalization" and "substantivization".
> (Malchukov 2006: 974)

In all four syntactic environments, the verb of the *=ti*-marked nominalized structure is fully inflected for TAM (tense, aspect and mood) and person; i.e. structures nominalized with *=ti* are full finite clauses. For example, in (3) we can observe that the *=ti*-marked verb is marked for intentional modality as well as for first person plural subject. Moreover, *=ti*-clauses can have overtly expressed arguments as well as adjuncts. An example of a case where an adjunct is overtly expressed is (3). Since in Yurakaré, overt expression of arguments is not obligatory, *=ti*-clauses often occur without overt expression of arguments.

---

3. This example is taken from a traditional narrative; the children of the jaguar bring a boy to their mother for her to eat him. In this example, they ask her whether she really killed and ate the boy.

4. The verb *toro* 'finish' encodes the actor as a benefactive object; therefore, this verb is marked with the first person plural object marker and a benefactive applicative.

Our analysis of =*ti* as a nominalizer is based on the observation that =*ti*-clauses can carry certain types of nominal morphology. In terms of 'substantivization', the nominalizer =*ti* imparts to the nominalized clause some of the morphological potential of nouns, but not the full range. Moreover, there is a difference between participant and event nominalizations: While participant nominalizations with =*ti* can be accompanied by demonstratives as determiners, event nominalizations marked with =*ti* can be followed by certain postpositional enclitics. This is probably related to the fact that in Yurakaré, core grammatical relations are not marked for case but only cross-referenced on the verb. Participant nominalizations with =*ti* occur in argument functions that are cross-referenced on the verb; that is, they function as core arguments and are not marked for participant role, unlike non-core arguments that often serve adverbial functions (see Section 2.4). Relativization by means of a =*ti*-clause is also only found for semantic roles that can be cross-referenced on the verb (see van Gijn 2006: 317–318).

Possessive and nominal number marking are not possible for either participant or event nominalizations. In other words, =*ti*-nominalized clauses are not treated by the external syntax completely like nouns. Nevertheless, the fact that some nominal morphology can be attached to clauses marked with =*ti* justifies its analysis as a nominalizer, and the analysis of =*ti*-marked clauses as nominalized structures.

Clauses nominalized with =*ti* have two different syntactic functions that cross-cut their nominalization types: argument and modifying functions. In argument and complement function, =*ti*-marked clauses take the position of arguments of the main clause verb. In (1), the =*ti*-marked nominalized clause functions as an argument of the main clause and is cross-referenced on the verb with the applicative object prefix *ku-*. This type of object marking is obligatory with the matrix verb *imaluma* 'think/worry'. Example (3) shows a complement use, where the =*ti*-marked clause appears as the subject of the verb *kusu* 'want'. This verb encodes the experiencer as an object and the desired entity or event as its subject. When used in relativizing and adverbial functions, =*ti*-clauses have a modifying function. In (2), the =*ti*-marked nominalized clause with relativizing function modifies the noun *sewwe* 'boy', while in (4), the =*ti*-clause provides an adverbial temporal modification of the main clause.

To summarize, Table 1 gives an overview of the four functions of =*ti*-nominalized clauses and their properties as discussed so far.

Since the four types of =*ti*-nominalized constructions cannot stand on their own and are embedded within a matrix clause, we consider them subordinate structures. In its subordinating use, =*ti* contrasts with two other, non-nominalizing enclitics marking subordinate clauses, =*ja* 'realis' and =*ya* 'irrealis'. Mostly but not exclusively, =*ti* is used when the subjects of the main clause and the subordinate

Table 1. Summary of functions of clauses nominalized with *=ti*

Function	Nominalization type	Nominal morphological potential	Syntactic function
Argument	Participant	Demonstratives	Argument
Relativizing	Participant	Demonstratives	Modifier
Complement	Event	Postpositional enclitics	Argument
Adverbial	Event	Postpositional enclitics	Modifier

clause are different, as can be observed in Examples (1)–(4), while *=ja* and *=ya* are mostly used when the subjects are the same. Van Gijn (2011) offers an analysis of this phenomenon in terms of spatio-temporal integration. Moreover, clauses with *=ja* and *=ya* cannot receive nominal morphology; therefore, these two markers cannot be considered nominalizers. A similar situation is found in Iskonawa (Panoan), where same-subject switch reference markers are not related to nominalization, while nominalization constructions are used in different-subject and different-object contexts (Zariquiey, this volume).

In terms of frequency, the adverbial use is by far the most frequent, accounting for 80% of the subordinating uses of *=ti*. The frequencies are summarized in Table 2. In addition to the clear cases, there are 23 cases that show some kind of ambiguity, allowing more than one of the three possible readings: relative clause, complement clause or adverbial clause. This phenomenon is described in more detail in Section 2.5.

Table 2. Frequencies of subordinating uses of *=ti*

Function	Observed frequency	Proportion
Argument	20	3%
Relativizing	12	2%
Complement	65	11%
Adverbial	464	80%
Ambiguous cases	23	4%
Total	**584**	**100%**

In the following subsections, we examine the formal and functional characteristics of these *=ti* constructions, beginning with its argument function (2.1), then moving on to the relativizing function (2.2), complementation (2.3), and subsequently to adverbial modification uses (2.4). In Section 2.5, we analyze the connection between these uses of *=ti*.

2.1 =*ti*-marked clauses with argument function

Clauses nominalized with =*ti* can be used to fill argument positions of the matrix verb. An example is given in (5), which involves a =*ti*-marked nominalized clause that functions as the object argument of the main verb *dula* 'make' – namely, *ta-ñuma-shama ka-winani-ø=ti* 'the one our late mother used' (or, 'this which our late mother used').

(5) *naa ta-choo-shama a-dula-ø=ya ushta naa*
DEM 1PL.POSS-uncle-DCSD IPFV-make-3SG.SBJ=REP before DEM
[*ta-ñuma-shama ka-winani-ø=ti*]$_{\text{NMLZ}}$
1PL.POSS-deceased_parent-DCSD 3SG.OBJ-walk-3SG.SBJ=NMZR
'That late uncle of ours made the one our late mother used.'
(Conversation-NL)

In (5), we can observe that the object argument formed by the =*ti*-marked clause is not cross-referenced on the matrix verb *dula* 'make'. This is due to the occurrence of the imperfective marker *a-* which impedes the verb from taking objects other than third person singular (see van Gijn 2006: 186). In addition, object cross-referencing would be unmarked for third person singular objects, because the object prefix for the third person singular is *ø-*. This might also be the reason why only third person singular objects can co-occur with *a-:* the third person object cross-referencing prefix does not interfere with *a-* morphologically or phonologically.

We can also observe in (5) that the =*ti*-marked clause is accompanied by the demonstrative *naa.* Demonstrative marking is the only type of nominal morphology found with =*ti*-marked participant nominalizations in argument function.

The argument use of =*ti*-marked clauses is not particularly frequent in our corpus, with 20 out of 584 uses (see Table 2 above).

2.2 =*ti*-marked clauses with relativizing function

Nominalizer =*ti* can be used in a relativizing function to form clauses that modify a noun. All semantic roles that can be expressed by cross-referencing suffixes (subjects) and prefixes (objects, including applicative objects) can be relativized (see van Gijn 2006: 317–318). In natural data, this use is less frequent than the other uses. Out of 584 subordinating uses of =*ti* in the corpus, only 12 are unambiguously relativizing (see Table 2 above).

Example (6) shows an example of a =*ti*-clause with relativizing function. The noun *dibujo* 'picture' is modified by the nominalized clause *ta-ka-n-wiwi-w=ti* '(the one) they brought us'.

(6) *ana dibujo* [*ta-ka-n-wiwi-w=ti*]$_{\text{NMLZ}}$
DEM picture(SP) 1PL.OBJ-3SG.OBJ-BEN-arrive.PL-3PL.SBJ=NMZR
'this picture that they brought us' (SocCog-YUZ107-2)

In (6), it can also be observed that the whole noun phrase including the modified noun *dibujo* 'picture' and the *=ti*-marked clause are preceded by the demonstrative *ana*. This shows that participant nominalizations with *=ti* in relativizing function can fall under the scope of demonstratives in combination with the noun they modify.

2.3 *=ti*-marked clauses with complement function

As seen in Example (3) above, complements of complement-taking verbs function similarly to arguments, in the sense that they serve as grammatical arguments of the main verb in a clause (see Cristofaro 2003: 159). In general, complement clauses answer well to 'what'-type questions (e.g. Q: She likes *what*? A: She likes *watching Bollywood movies*), and they are sometimes accompanied by referentiality-marking devices such as case markers (see Comrie & Thompson 1985; Malchukov 2004, 2006; Morey 2011). In Yurakaré, there are three types of complement constructions with the nominalizer *=ti*. First, *=ti* can be used in combination with the intentional TAM-suffix *-ni*, phonologically reduced to *-n* in combination with the first person plural cross-reference marker *-tu* as in (3). Van Gijn (2011: 185) interprets the marker *-ni* in these constructions as subjunctive, noting that verbs expressing ability, desire and liking appear in these constructions with non-factual complements. Second, *=ti* can be used on its own to form a complement clause with a factual interpretation. Third, there is a complementizer *=ti=la,* a probably grammaticalized combination of nominalizer *=ti* and the instrumental postposition *=la*. Instrumental *=la* is the only postposition that occurs with *=ti*-clauses in complement function in our corpus.

In the corpus, *=ti*-clauses with complementizing function occur mostly with speech act verbs, perception verbs, and modal verbs indicating desire and ability, as well as with cognition verbs and phasal verbs. The most frequent complement-taking verbs taking unambiguous complement clauses are *kusu* 'want' and *puede* 'be able' (16 instances each), followed by complement-taking verbs *dyuju* 'tell' (8 instances) and *bëjta* 'see' (6 instances). In total, there are 65 unambiguous complementizing clauses marked with *=ti* (see Table 2 above) and 13 ambiguous cases (see Section 2.5 below).

An example of a verb that takes a complement marked with *=ti* in combination with intentional *-ni* to form a non-factual complement is the Spanish loan verb *puede* 'be able, can'. Example (7) demonstrates its use.

(7) *achaya puede-tu tuwa* [*nish dula-ntu=ti anuta*]$_{\text{NMLZ}}$ *=la*
so_that can(SP)-1PL.SBJ we NEG do-INT:1PL.SBJ=NMZR like_that=COMM
'So that we are able to not behave like that.' (SocCog-YUZ106-2)

In (7), both the matrix clause and the complement have a first person plural cross-referenced by *-tu* and *-ntu* as their subject. The complement is non-factual, which explains the use of the construction with nominalizer *=ti* and intentional *-ni* here, the latter realized as *-ntu* as a result of phonological fusion between intentional *-ni* and first person plural subject cross-referencing suffix *-tu*.

An example of a complement with the complementizer *=ti=la* is (8). The matrix verb phrase *ku-i-mala-uma*, meaning 'think/worry (< go) about it' takes a complement marked with *=ti=la*.

(8) *mapakka sufri li-tütü-y=ja*
big suffer VLOC-be-1SG.SBJ=REA
ku-i-mala-uma-y=la [*amashku*
3SG.OBJ.COM-VPL-go.SG-DST-1SG.SBJ=COMM how
li-tütü-p=ti]$_{\text{NMLZ}}$*=la*[5] *mu-ta-ø=ya*
VLOC-be-2PL.SBJ=NMZR=INS 3PL.OBJ.COM-say-3SG.SBJ=INTSUBJ
'"While I was there, suffering very much, I was thinking about how you [my family] are," he probably said to them.' (SocCog-YUZ109-2)

2.4 *=ti*-marked clauses with adverbial function

When used in adverbial function, *=ti*-marked clauses can express a variety of relations between the subordinate structure and the main clause. These include temporal, causal, concessive, conditional, purposive, and locative relations (see also van Gijn 2011). The adverbial use is the most frequent of the subordinating uses. 464 out of 584 subordinating uses of *=ti* are unambiguous adverbial uses. There are furthermore 20 ambiguous uses (see Section 2.5 below). The temporal reading is the most frequent; 287 out of the 464 clear adverbial uses have a temporal reading. The second most frequent reading is the concessive (51 uses), followed by the conditional reading (26 uses). An example of a temporal relation is given in (9).

5. We chose to set the brackets after *=ti* to emphasize that nominalized clauses with *=ti* can carry postpositional enclitics. In the future, it will be interesting to investigate the question of whether the combinations of *=ti* with these postpositions have in fact become grammaticalized to form a single unit.

(9) [*anu ma-n-wita-ø=ti*]$_{\text{NMLZ}}$ *a-bashti=ja*
like_that 3PL.OBJ-BEN-arrive.SG-3SG.SBJ=NMZR 3SG.POSS-wife=TOP
resibi-ø latiji ana=y
receive(SP)-3SG.SBJ then DEM=LOC
'When he arrived upon them like that, his wife received him then, here.'
(SocCog-YUZ108-2)

An example of a conditional reading of a clause marked with the nominalizer *=ti* is given in (10).

(10) [*nij wita-ø=ti*]$_{\text{NMLZ}}$ *nij bata-tu*
NEG arrive-3SG.SBJ=NMZR NEG go.FUT-1PL.SBJ
'If he doesn't arrive, we won't go.' (290906_convI)

The use of case markers and adpositions in combination with nominalized constructions is a very common clause-linking strategy in South American languages (van Gijn 2014: 293). In Yurakaré, the postpositional enclitics result in more specific interpretations of clauses marked with *=ti* that serve adverbial functions, as summarized in (11). Note that in Yurakaré, only peripheral arguments are marked by postpositional enclitics, while core arguments are only cross-referenced on the verb (see van Gijn 2006: 106, 143).

(11)

Postpositional enclitic	Extended use with *=ti* as subordinator
=la 'instrumental'	→ *=ti=la* 'causal'
=jsha 'ablative'	→ *=ti=jsha* 'sequential', 'concessive'
=chi 'direction'	→ *=ti=chi* 'direction', 'location'
=y 'locative'	→ *=ti=y* 'location'

Example (12) shows a use of *=ti* with the instrumental postpositional clitic *=la,* resulting in a causal interpretation.

(12) *kusuti* [*nij li-ujwa-tijti=ti*]$_{\text{NMLZ}}$*=la a-wëwë-ø=ya*
maybe NEG VLOC-look-HAB.1PL.SBJ=NMZR=INS IPFV-CRY-3SG.SBJ=INTSUBJ
'Maybe because we did not look at her she is crying.' (loreto)

In (13) we can observe a sequential interpretation of the use of *=ti* with the ablative enclitic *=jsha.*

(13) *kusu* [*mala-y=ti*]$_{\text{NMLZ}}$ *=jsha wiwi-w=ya*
maybe go.SG-1SG.SBJ=NMZR=ABL arrive.PL-3PL.SBJ=INTSUBJ
ma-bashti-w
3PL.POSS-wife-PL
'Maybe their wives arrived after I left.' (290906 convI)

A concessive use of =*ti* in combination with the ablative clitic =*jsha* is exemplified in (14).

(14) *nij ma-bëjta-y sëë* [*naa=chi wiwi-w=ja*[6]
NEG 3PL.OBJ-see-1SG.SBJ 1SG DEM=DIR arrive.PL-3PL.SBJ=REA
ma-tësë-w=ti]$_{\text{NMLZ}}$ *=jsha*
3PL.OBJ-stand-3PL.SBJ=NMZR=ABL
'I did not see them, even though they were standing there after they arrived.'
(290609_convI)

The interpretation of =*ti=chi* can be either directional or locative. Example (15) demonstrates a locative use with the directional marker =*chi*.

(15) *nij mashi kompadre* [*mi-kompadre*
NEG rain compadre(SP) 2SG.POSS-compadre(SP)
li-winani-ø=ti]$_{\text{NMLZ}}$*=chi*
VLOC-walk-3SG.SBJ=NMZR=DIR
'It did not rain, compadre, where your compadre was.' (160906_convI)

2.5 Connection between subordinating uses

As we have seen in Table 1, the uses differ in terms of their morphological potential and their syntactic function. Participant nominalizations can be modified with demonstratives, while event nominalizations can carry postpositional enclitics. Regarding their syntactic function, =*ti*-clauses with argument and complement functions fill an argument position of the matrix verb, while those with relativizing and adverbial functions occur in modifying position.

In fact, however, clauses nominalized with =*ti* are sometimes ambiguous regarding their surface structure and can give rise to various possible interpretations. An example from our spoken corpus is (16), which can be interpreted as relativizing, complement or adverbial, because all three clause types can have the same surface structure involving a =*ti*-nominalized clause. Note, however, that there are differences regarding the scope of =*ti*: while the noun *yee* 'woman' serves as the head noun in the relativizing interpretation in (16)(i), it becomes part of the nominalized clause in both the complement and in one of the possible adverbial readings, as in (16)(ii) and (iv), respectively.

6. We can observe here that the verb *wiwi* 'arrive.PL' is marked with the subordinator =*ja*, indicating a higher degree of integration of events (see van Gijn 2011). This verb forms a unit with the verb marked with =*ti=jsha*. The nominalizer =*ti* takes scope over both verbs here.

(16) *ana buyta=ja ujwa-ø [yee [ma-n-dyuju-ø=ti]*$_{\text{NMLZ(i), (iii)}}$*]*$_{\text{NMLZ(ii), (iv)}}$
DEM chief=TOP watch-3SG.SBJ woman 3PL.OBJ-BEN-tell-3SG.SBJ=NMZR
(i) 'This chief is watching the woman [who is telling it to them].'
– relativizing
(ii) 'This chief is watching [the woman telling it to them].'
– complement
(iii) 'This chief is watching the woman [while she is telling it to them].'
– adverbial, narrow scope of *=ti*
(iv) 'This chief is watching [while the woman is telling it to them].'
– adverbial, broad scope of *=ti*
(SocCog-YUZ103-2)

In the relativizing interpretation, as seen in (16) (i) above, the *=ti*-marked clause is used to modify another noun (*yee* 'woman'), thus forming part of the argument of the main clause. In the complement clause reading in (16)(ii), the noun *yee* forms part of the nominalized clause marked with *=ti*. In the narrow scope adverbial interpretation, as seen in (16) (iii), the *=ti*-marked nominalized clause provides specifying information about the woman not only in terms of what she is doing, but also in terms of when the action that defines her is unfolding. In the broad scope adverbial reading in (iv), the noun *yee* 'woman' becomes part of the simultaneous event described by the adverbial, falling under the scope of *=ti*.

Ambiguous examples such as (16) can serve as "bridging contexts" (Heine 2002) for the extension from relativizing to complement and adverbial functions. A bridging context in the sense of Heine (2002) is a stage of language change where a linguistic item can have two possible interpretations: its original meaning and a meaning different from its original meaning that arises through pragmatic inference; this inferred meaning is still cancellable (Heine 2002: 84). This is exactly the situation we find in (16): The temporal meaning of the adverbial reading arises through pragmatic inference and is cancellable. This identifies structurally ambiguous examples such as (16) as bridging contexts.

There are 23 ambiguous uses of *=ti* in our corpus. In 8 of these, the argument reading is possible, while in another 12, the relativizing reading is among the options. This demonstrates that, while ambiguous uses are not very frequent in the corpus, they often involve one of the participant nominalization functions of *=ti*.

We suggest that the function of *=ti* in adverbial contexts generally involves more abstract referents than in relativizing and complement contexts (see also Yap & Grunow-Hårsta 2010; Yap, Grunow-Hårsta & Wrona 2011 and papers therein, in particular Genetti 2011 and Sung 2011; Yap, Deng & Caboara, 2017). Since grammaticalization often involves an increase in abstractness (e.g. Traugott & Heine 1991: 4), it is likely that the adverbial use developed after *=ti* became grammaticalized to form nominalized clauses with argument, relativizing and complement

functions. One further reason is that it is plausible to assume that in order to be able to attach postpositional enclitics to a clause to form semantically more precise adverbials, there needed to a be structure that was already nominal in nature.

3. The demonstrative *ati* as a possible source for =*ti*

It is not completely clear how the element =*ti* became grammaticalized as a clausal nominalizer rendering subordinate structures. The most plausible scenario is that =*ti* is in some way related to the demonstrative *ati* (van Gijn 2006: 313; 2011: 182). There are two arguments for this. First, the forms of the two elements are similar enough to posit a relation between the two items. Second, as we show below, the two items share some of their morphological potential. We elaborate on these arguments below.

The demonstrative *ati* can be used as a demonstrative pronoun as well as a demonstrative determiner. In determiner function, it is prenominal. Examples (17) and (18) demonstrate these two uses of *ati,* respectively.

(17) *tëpshë=chi ati*
what=FR DEM
'What is that?' (al_ce_frogstory)

(18) *nijta sëë mu-ta-ø ati sewebonto*
NEG 1SG 3PL.OBJ.COM-say-3SG.SBJ DEM young man
'"Not I," said that young man.' (SocCog-YUZ106-2)

The exact relation between the nominalizer =*ti* and the demonstrative *ati* is not very clear. There are at least two possibilities. First, =*ti* could have been derived from *ati,* with the onset vowel of the demonstrative having become elided. Second, there could have been a nominal element *ti* that developed in two directions: on the one hand, becoming attached to a demonstrative element *a*[7] to form the demonstrative *ati,* and on the other hand becoming attached to clauses, developing into the clausal nominalizer =*ti.* Regarding the second account, we are not able to determine the exact etymology of *ti*.

We cannot decide here which account is more plausible, since we lack historical data for Yurakaré. Based on our synchronic data, we suggest that it is more likely that the first path is the correct one, where the demonstrative *ati* is taken to be the source of =*ti.* In the following, we discuss the synchronic evidence.

7. It is plausible to assume that *a* is a demonstrative element, because the forms of the other two demonstratives of the language are *ana* and *naa/na'a.*

To argue for *ati* as the source of *=ti,* we have to identify a context in which *ati* could have developed into *=ti.* One possibility is a clause-final use of the demonstrative pronoun *ati* 'that' in argument function, as seen in (19) where *ati* is used to refer to a resumptive third person singular subject.[8]

(19) *achu* *ma-n-dyërërë* *ma-tütü-ø* *ati*
like_that 3PL.OBJ-BEN-talk 3PL.OBJ-sit-3SG.SBJ DEM
'He/that one sits with them talking in that manner.' (SocCog-YUZ104-2)

In Example (19), the demonstrative *ati* is the subject of the clause used in clause-final position. The two verbs form a serial verb construction with the meaning 'sit with them talking'. Note that the verb *dyërërë* 'talk' encodes the actor as a benefactive applicative object. In sum, (19) is a normal clause involving a verb (in this case, a serial verb construction) with an overtly expressed subject. Yurakaré allows subject and object drop, and word order is relatively free, so the use of an overt subject in clause-final position is only one way of expressing the same content.

From an argument use such as (19), it is well possible that the finite predicate preceding *ati* was sometimes interpreted as having a modifying function, which would result in the interpretation 'the one who sits with them talking in that manner'. Such a construction could then be used to fill argument positions of verbs. This use of *ati* is reconstructed; we have no data where the demonstrative *ati* is used in this context. However, it is plausible to assume this, since as noted by Cristofaro (this volume), nominalization constructions often develop from contexts where the predicate was already used in modifying function, specifying some nominal element that later develops into the nominalizer.

Assuming a modifying function for the predicate preceding *ati* in some contexts, the elision of the first vowel of *ati* in fast speech could have led to the cliticization of *ati,* going hand in hand with a semantic bleaching leading to a functional change from a referential function to a nominalizing function. The result would look like (20):

(20) [*ma-n-dyërërë* *ma-tütü-ø=ti*]$_{\text{NMLZ}}$
3PL.OBJ-BEN-talk 3PL.OBJ-sit-3SG.SBJ=NMZR
'the one who sits with them talking'[9] (SocCog-YUZ104-2)

8. In fact, the demonstrative *ati* can also be used in locative function meaning 'there'. In this use, it carries the locative postpositional clitic *=y,* which does not change the overt form of the demonstratives because it ends in the same vowel. Because the forms are homophones, we cannot decide with full certainty which of the two interpretations is the intended one.

9. Actually, we are dealing with an insubordinate use in (20) (see Section 4); however, the interpretation in terms of a participant nominalization given here is structurally possible and is therefore used here to exemplify the interpretation we argue is the most reasonable to have

The open (or coda-less) syllables in Yurakaré verbs with covert third person cross-reference marking (indicated by ø) could also have played a role in the rise of =*ti* as a clausal nominalizer or marker of a complement. As seen in (21), a morphosyntactic configuration involving a final vowel in the clause preceding the pronoun *ati* would often result in vocalic reduction that gives rise to the encliticization of =*ti*, and its subsequent reanalysis as a clausal nominalizer.

(21) *ma-tütü-ø ati* → *ma-tütü-ø ti* → *ma-tütü-ø=ti*

From such third person argument constructions, =*ti* could have developed toward a more general subordinator, forming relativizing constructions with a nominal head, complements and adverbials, as proposed above in Section 2.5.

If it is correct that =*ti* is derived from *ati,* the development we observe is one from a referentiality-marking device toward a marker of textual cohesion. The demonstrative *ati* is usually used for referents previously mentioned in a discourse (see Gipper 2017). Since Yurakaré does not have third person personal pronouns, the demonstrative *ati* is often used in such contexts. The reinterpretation would thus go from a resumptive (demonstrative) pronoun to a nominalizer for participant nominalizations in argument function, and subsequently extending to marking relativization, complements, as well as adverbials to mark textual cohesion.

Another argument for the connection between =*ti* and *ati* is constituted by the morphological properties of the two elements. However, this argument does not help us to distinguish between the two accounts proposed above, since both accounts assume a demonstrative origin. As seen in Table 3, like =*ti, ati* can take postpositional enclitics which can result in similar meanings as the combinations of the case markers with =*ti.*

Table 3. Functions of post-positional enclitics with demonstrative *ati* and nominalizer =*ti* in adverbial use

Postposition	**Combination with *ati***	**Combination with =*ti***
=*la* 'instrumental'	'instrumental', 'path', 'causal'	'causal'
=*jsha* 'ablative'	'ablative', 'sequential'	'sequential', 'concessive'
=*chi* 'direction'	'direction', 'location'	'direction', 'location'
=*y* 'locative'	'location'	'location'
=*tina* 'comitative'	'comitative'	-

Like =*ti, ati* does not take possessive marking (van Gijn 2006: 129). An important difference in their morphological potential is that *ati* can be marked for number

developed first. We chose Examples (19) and (20) because they are near minimal pairs showing the use of *ati* and =*ti,* respectively.

(i.e. the singular/plural distinction) and take some derivational suffixes such as *-jti* 'limitative' (see van Gijn 2006: 129), while *=ti*-marked clauses cannot. This may suggest that *=ti* is the more grammaticalized form and hence is likely to be the derived form.

4. Insubordination of *=ti*

The data demonstrate that *=ti* is developing uses of insubordination. Insubordination is defined as "the conventionalised main-clause use of what, on prima facie grounds, appear to be formally subordinate clauses" (Evans 2007: 367). As we have seen in Section 2 above, structures with *=ti* are formally embedded or subordinate. Insubordinate uses develop out of main clause ellipsis constructions, but it is only when there is a conventionalized meaning that we consider them to be insubordinate constructions (see Evans 2007: 377). Thus, if *=ti*-marked clauses are used without a main clause to which they are subordinate, we are dealing either with an insubordinate use, or with an ellipsis of the main clause. If the reconstruction of a main clause is possible, it depends on the degree of conventionalization whether or not an analysis in terms of an insubordinate use is preferable. At any rate, from a grammaticalization perspective, main-clause ellipsis is a highly robust morphosyntactic mechanism for the development of insubordinate constructions. This is attested in a wide range of languages, including Japanese (e.g. Ohori 1995; Higashiizumi 2006; Shinzato 2015; Shibatani, this volume) and Korean (Rhee 2014; Ahn & Yap 2014, 2015).

Some insubordinate uses of *=ti* in Yurakaré appear to be developing out of adverbials expressing a causal relation. Example (22) contains two uses of *=ti* that involve a main-clause ellipsis strategy in which the *=ti*-marked clause can be reconstructed to convey a causal relation. In line 001, Speaker A suggests that his interlocutor and he should talk about another topic, proposing two possible subjects. In line 002, Speaker B states that the researchers had told them to talk about the pictures of the Family Problems Picture Task, using a stand-alone construction with *=ti*. The interpretation here is a causal one, and the elided main clause (indicated in square brackets) is understood to be a rejection of A's proposal in line 001. Possibly, the main clause is elided in this context because a direct or explicit disagreement could be considered a breach in politeness. In line 005, Speaker B makes a similar statement, again using a stand-alone construction with *=ti*.

(22) 001 A: *ka-la-niri-ima-shta-tu tiri tata koycho-shama*
3SG.OBJ-MAL-tell-COL-FUT-1PL.SBJ PROPN father PROPN-DCSD
neyne ku-ta-ø
whatever 3SG.OBJ.COM-say-3SG.SBJ
'We could tell about Tiri, about father Coycho, about whatever he is called.'

002 B: *achuta-ø lash ana-jti bëjma* [*ana dibujo-w*
be_like_that-3SG.SBJ then DEM-LIM look.IMP DEM picture(SP)-PL
ma-ujwa-cha-p tu-ta-w=ti]$_{NMLZ}$
3PL.OBJ-look-JUS-2PL.SBJ 1PL.OBJ.COM-say-3PL.SBJ=NMZR
'That's true, [but] you see, because they said, "You should only look at these pictures" [we should not consider the other ones].'

003 A: *otte*
'Yeah.'

004 B: *nish tu-yte-w na atta-shama*
NEG 1PL.OBJ.COM-ask-3PL.SBJ DEM other-DCSD
'They didn't ask about the others.'

005 [*solamente ana-jti ujwa-cha-p*
only(SP) DEM-LIM look-JUS-2PL.SBJ
tu-ta-w=ti]$_{NMLZ}$
1PL.OBJ.COM-say-3PL.SBJ=NMZR
'Because they said, "You should only look at this" [we won't bother with the others].' (SocCog-YUZ106-2)

Example (22) demonstrates one way in which the insubordination use of =*ti* could have grammaticalized. It is probable that it developed out of such an elliptical construction in which the =*ti*-marked clause stands in a causal relation to the elided main clause. Previous studies in other languages have shown that main-clause ellipsis often yields insubordinate constructions that express the speaker's subjective and intersubjective stance (Ohori 1995; Higashiizumi 2006; Evans 2007; Ahn & Yap 2014, 2015; Rhee 2014; Shinzato 2015; Shibatani, this volume). From a cognitive perspective, ellipsis has the effect of engaging the hearer more actively in the conversation by allowing the hearer himself/herself "to mentally fill in the information gap through scaffolded and shared knowledge" (Wai & Yap 2014: 1; see also Kurzon 1995 and Clifton & Frazier 2010). In this way, main-clause ellipsis contributes to the development of insubordinate constructions as interactional discourse strategies that express both the speaker's subjective and intersubjective stance.

However, in our corpus of spoken Yurakaré we mostly find tokens of =*ti*-marked insubordinate clauses with no trace of main-clause ellipsis. Out of 164 insubordinate uses of =*ti,* only 10 involve main-clause ellipsis. This suggests that there are other means of insubordination for =*ti*-clauses.

Let us examine one such example that does not involve main-clause ellipsis. In (23), Speaker A and Speaker B have finally finished arranging the pictures for the Family Problems Picture Task, and Speaker A notes in line 001 that the story goes according to the way they have arranged the pictures. In line 002, Speaker B asks for reconfirmation, using the enclitic *=ye*. This clitic expresses that some information is not yet integrated in the speaker's knowledge system. It can be used when the speaker is not quite sure about something, as in the case of fading memory. In (23), *=ye* indicates that the speaker is not quite sure whether the story is correct in the way the speakers arranged it. In line 003, Speaker A reconfirms with an insubordinate construction with *=ti*. Since no main clause can be easily reconstructed in this context, we argue that we are dealing with a fully grammaticalized insubordinate use of *=ti* here. That is, *=ti* has been reinterpreted as a marker of the speaker's stance, and the *=ti*-marked insubordinate clause can be used independently as a finite structure without any need for the addressee to mentally reconstruct a biclausal relationship between a *=ti*-marked adverbial and an elided main clause.

(23) 001 A: *achama* *buybu*
be_like_that story
'The story is like that.'
002 B: *achu* *mala-ø=ye*
like_that go.SG-3SG.SBJ=ADAP.F
'It goes like that?'
003 A: [*achu* *mala-ø=ti*]$_{NMLZ}$
like_that go.SG-3SG.SBJ=NMZR
'It does go like that.' (SocCog-YUZ104-2)

We propose that the pragmatic meaning conveyed by insubordinate uses of *=ti,* as in (23) is one of 'intersubjective commitment'. This overarching meaning captures best the different uses of insubordinate *=ti*-clauses. Basically, such clauses are used in four contexts, as well as in mixtures of those contexts. First, they are used when a speaker tries to convince or persuade another speaker, in effect to signal an intersubjective move. Second, they are used to indicate that the information is in some way common knowledge between the speakers, either through previous discourse, through being observable in the external context, or through general shared knowledge. In the third reading, *=ti*-clauses connect the speaker's statement to some other stretch of preceding discourse. In this reading, *=ti* is used to create discourse cohesion. We subsume the uses which do not fit these three contexts under a fourth category, under a fourth category.

What the first three contexts have in common is that they are contexts where speakers convey commitment, as in (23), where it is used in a response to reconfirm some information questioned by the addressee. In addition to that, stand-alone

=*ti*-clauses in these contexts seem to convey an intersubjective component, referring back to some preceding discourse or to some information present in the context. In the case of (23), Speaker A seems to refer back to the whole preceding discussion in which Speakers A and B have created their story with the pictures. When Speaker B expresses some uncertainty in line 002, Speaker A uses =*ti* to refer back to their preceding discussion to provide support for his commitment to the statement that the story goes like that.

The intersubjective component can be derived from at least two possible source constructions of stand-alone =*ti*-clauses. The first possible source is an adverbial expressing a causal relation. On this account, uses such as that in (23) would be extensions of main-clause ellipsis uses, such as that in (22), where the main clause is elided and the adverbial expressing a causal relation stands alone. For English, it has been shown that clauses with *because* can have various functions relating to discourse and speaker stance. In these cases, the *because*-clauses are no longer subordinate, but introduce a coordinate clause (e.g. Couper-Kuhlen 1996; Schleppegrell 1991). An important function of such *because*-clauses is to establish links with some preceding discourse (Schleppegrell 1991: 333–5). An example is given in (24):

(24) Michael: 001 Yeah it's kind of like when um,
002 like if um you play the champion at the game of war?
003 and someone bets on the champion but he loses?
004 you don't really know who's gonna win
005 <u>because</u> WAR you don't know what cards you get.
(Schleppegrell 1991: 334)

The *because*-clause in line 005 of (24), in addition to providing a causal link to the preceding utterance in 004, establishes a thematic link with the earlier utterance in line 002, reintroducing the topic of 'war' (Schleppegrell 1991: 334). The observation that *because*-clauses in English can establish links across larger stretches of discourse is reminiscent of the discourse cohesion use of =*ti*-clauses. Further, *because*-clauses can be used to support or challenge another speaker's statement (Schleppegrell 1991: 331). As proposed above, this reading of *because* is similar to the main-clause ellipsis uses of stand-alone clauses with =*ti.* Main-clause ellipsis with causal =*ti*-clauses can thus account for the convincing as well as the discourse cohesion uses of =*ti.*

Another possible path of development of insubordinate uses of =*ti*-clauses like (23) is the ellipsis of a main clause in a complement construction involving a vision verb, such as *bëjta* 'see' and *ujwa* 'look'. In Yurakaré, these two verbs can be used to form particles that can be used as a strategy to convince the other speaker of something by directing their attention to some piece of information they should be able to 'see'; that is, to be able to believe, much like English parenthetical *you see.* Some

possible forms for these particles are *bëjma* (imperative form of the verb 'see') and *ujchampu* (presentative second person singular form of the verb 'look'). We consider these particles to convey an intersubjective meaning, because they point to some information that the speaker considers accessible and plausible to the other speaker.

The verbs *bëjta* 'see' and *ujwa* 'look' can both take complements marked with *=ti* when they are used as regular complement-taking verbs rather than attention-directing particles. However, the particles based on these verbs, even though they may no longer be complement-taking verbs but rather parenthetical expressions, can also occur with *=ti*-marked clauses. In the corpus, 21 out of 164 insubordinate uses of *=ti* are accompanied by such particles. This use may have developed from the regular complement-taking uses of these verbs; speakers still use a complement marked with *=ti* even though the verbs are parentheticals rather than regular complement-taking verbs. From there, an elision of the intersubjective particle could lead to an 'intersubjective commitment' interpretation of the stand-alone *=ti*-marked clause. In fact, it is well possible that the grammaticalization of the intersubjective particles derived from vision verbs and of stand-alone *=ti*-clauses conveying intersubjective commitment went hand-in-hand, with the particles and the *=ti*-clauses both becoming stand-alone structures indicating intersubjectivity.

An example of a use of the particle *bëjma* 'see.IMP' is given in (25). The particle *bëjma* is used to convey that the information given in the second clause should be accessible to the other speaker (i.e. that she should be able to 'see' it), thus conveying an intersubjective stance. The second clause is marked with *=ti,* probably coming from a complement reading, but at the same time leading to the use of *=ti*-marked clauses as stand-alone structure, taking over the intersubjective reading of the particle.

(25) *bëjma* [*anu* *ma-la-aya* *iba-ø=ti*]$_{\text{NMLZ}}$
see.IMP like_that 3PL.OBJ-MAL-advise should-3SG.SBJ=NMZR
'Look, he should be giving them advice like that.' (SocCog-YUZ104-2)

From both source constructions (main-clause ellipsis with an adverbial and main-clause ellipsis with a complement), we can derive the interpretation of connecting to either the preceding discourse or something else in the interlocutors' common ground for stand-alone *=ti*-clauses without a reconstructable main clause. The speaker expresses commitment on the basis of something that should be obvious to the addressee as well, either because it has been said before or because it is visible or otherwise perceivable from the context. Since both uses occur in our synchronic corpus and since both uses result in similar intersubjective commitment readings for *=ti,* it is probable that both uses contributed to the insubordination of *=ti*-clauses in that the existence of one use boosted the insubordination use emerging from the other.

Let us return to the fourth context where insubordinate =*ti*-clauses are used. An example of such a use in combination with the speech act verb *ta* 'say' is (26).

(26) 001 *peta-cha-p=naja*
lie_down-JUS-2PL.SBJ=already
'"Go to sleep now!" [I said to my children]'.
002 *nijta kerosina*
NEG kerosene
'"There is no more kerosene," [I said to my children]'.
003 [*peta-ntu=naja* *ta-w=ti*]$_{NMLZ}$
lie_down-INT:1PL.SBJ say-3PL.SBJ=NMZR
'"Let's go to bed now," they said.' (290906_convI)

In Example (26), the speaker tells her sister about the events of a certain evening. In lines 001 and 002, she explains how she told her children to go to bed because there was no kerosene left to give them light. In line 003, she gives the children's response with a direct quotation framed by the verb *ta* 'say'. The clause is further marked with =*ti,* clearly representing an insubordinate use.

We suggest that =*ti* has become reinterpreted as a clause-final enclitic indicating 'intersubjective commitment' (i.e. speaker commitment based on information available to the addressee as well), either through the preceding discourse or through the context. Examples such as (26) represent this use of =*ti*. The step from an insubordination construction to its use as a clause-final enclitic is not very far, since we cannot distinguish between these two structures at the level of overt morphosyntax. Yurakaré has a set of clause-final particles related to aspect and speaker stance, as shown in (27) (a) and (b), respectively, and =*ti* fits very well into the speaker stance category.

(27) a. aspect:
=*naja* 'already'
=*bëla* 'still'
=*bë* 'momentaneous'
b. speaker stance:
=*la* 'commitment'
=*se* 'presupposition'
=*yu*/=*ri* 'resignative'
=*ye*/=*ra* 'adaptive'
=*ti* 'intersubjective commitment'

The difference between =*ti* and the commitment marker =*la* is that while =*ti* conveys an intersubjective component grounding the speaker's commitment to information that is available to the addressee as well, =*la* expresses the

speaker's commitment without making any reference as to the source of his or her commitment.

To summarize, we argue that the enclitic *=ti* in Yurakaré first developed from a demonstrative into a clausal nominalizer with the function of linking a subordinate structure to a matrix clause. From there, an insubordination construction developed out of a causal adverbial and/or a complement use with complement-taking vision verbs that refer to something obvious (e.g. *bëjma* 'see.IMP'), probably from contexts where the main clause is elided. As a final step, we propose that *=ti* is currently acquiring a function as a clause-final enclitic marking intersubjective commitment. The path we propose is visualized in Figure 1.

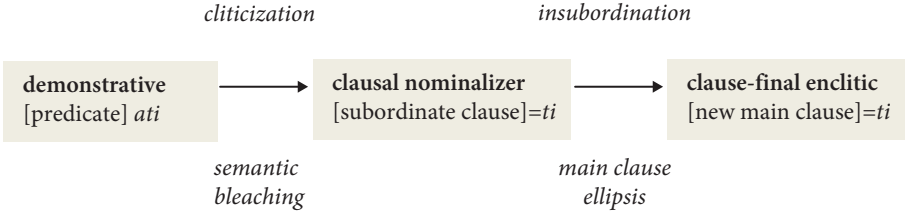


Figure 1. Grammaticalization path for the nominalizer *=ti*

5. Conclusion

In this paper we identify various uses of *=ti*-nominalized clauses in Yurakaré discourse: argument function, relativizing function, complement function, and adverbial function. While *=ti*-clauses in these four functions differ regarding their morphological potential and their syntactic environment, in language use they are often structurally very similar and can have the same surface form. Such syncretism between nominalization, relativization, complementation and adverbial subordination is not uncommon cross-linguistically (see, for example, Genetti 2011 for Tibeto-Burman languages and Sung 2011 for Formosan languages). The *=ti*-marked adverbials are often accompanied by postpositional enclitics that help to disambiguate the specific relationships (e.g. causal, sequential, concessive) between the subordinate clause and the main clause, which is common in South American languages (van Gijn 2014: 293; see also van Gijn, this volume). Event nominalizations with *=ti,* i.e. *=ti*-clauses with complement and adverbial functions, are also prone to insubordination, whereby the subordinate clause is reanalyzed as a finite, 'stand-alone' (i.e. independent) structure. These *=ti*-marked insubordinate structures appear to be 'syntactically incomplete', but they nevertheless are anchored in discourse by various other means, including non-syntactic strategies such as rich contextual information through prior interactions between the interlocutors. We

have seen, for example, how insubordinate uses of =*ti*-marked constructions can emerge in discourse contexts where 'unsaid-but-pragmatically-understood' information in elided main clauses can be easily retrieved from prior turns in dyadic or multi-party exchange structures. We further suggested that in some stand-alone contexts, nominalizer =*ti* has been reinterpreted as a clause-final enclitic marking the speaker's intersubjective commitment.

In this paper we also examined whether there is a link between the Yurakaré nominalizer =*ti* and the demonstrative *ati*. Our conclusion is still somewhat tentative; however, we propose that a development from the demonstrative *ati* to the nominalizer =*ti* could well have happened, starting from the use of *ati* with a preceding predicate in modifying function followed by elision of the first vowel of *ati* and semantic bleaching, leading to the development of the enclitic =*ti* with a nominalizing function. We also identified some common morphosyntactic features (e.g. the use of postpositional enclitics) between the two forms that suggest a possible extension from the demonstrative *ati* to a phonologically-reduced enclitic form =*ti* that serves as a clausal nominalizer.

The development of the demonstrative *ati* into the participant nominalizer =*ti,* with further extension to relativizing, complement and adverbial functions, appears to involve increasingly abstract referents, with the demonstrative *ati* and =*ti*-marked participant nominalizations used for more concrete referents, such as people, places and things (Lyon's (1977: 443) first-order ontological entities), while the nominalizer =*ti* in complement and adverbial contexts (i.e. event nominalizations) is used instead for more abstract 'referents' that take the form of situations and events (i.e. Lyon's (1977: 443) second-order ontological entities). Particularly in its adverbial use, =*ti* can be seen as a marker of textual cohesion, because it is used to express relations between events.[10] Further, the extension from clausal nominalizer =*ti* to a clausal enclitic indexing the speaker's intersubjective commitment involves an extension from the referential to the non-referential domain. Similar developments where phonological reduction (e.g. *ati* > =*ti*) accompanies scope expansion (e.g. referential > textual > pragmatic uses) have also been observed in other languages. This unidirectional extension from referential to non-referential uses is consistent with findings from diachronic and typological research (e.g. Yap, Choi & Cheung 2010; Schapper & San Roque 2011; Kratochvil 2011; Nagaya 2011, 2014; Tanangkingsing 2012; Yeh 2010, 2014; Adachi 2014).

10. Lyons (1977: 443) identifies three major types of nominal groups, which he refers to as 'ontological entities'. First-order ontological entities refer to concrete nominals such as people, things and places. Second-order and third-order ontological entities refer to more abstract nominals such as 'event-type' and 'proposition-type', respectively.

If, as we have hypothesized, the nominalizer *=ti* is derived from the demonstrative *ati,* then the uses of *=ti* construcions in Yurakaré show that it is not so much the type of source form (i.e. noun/light noun vs. demonstrative) that determines the outcome of the grammaticalization process, but rather the source construction, as proposed by Cristofaro (this volume). On this assumption, demonstratives do not necessarily behave differently from nouns in the development toward becoming a nominalizer.

Of particular relevance to our present study on the possible development of Yurakaré=*ti* from demonstrative to nominalizer, a number of grammaticalization studies have shown that demonstratives can develop into nominalizers and from there develop extended pragmatic uses. For example, Nagaya (2011, 2014) has shown that the distal demonstratives *ang* and *yung* in Tagalog (a Philippine-type Austronesian language) are frequently used to form participant and event nominalizations (e.g. ***ang***=*ka~kanta* 'the one who will sing' and ***yung***=*kapag nag-birthday=ako marami=ng regalo* 'the situation that when I have my birthday, there are a lot of gifts',) which then further develop extended uses with pragmatic nuances (e.g. *Ang=ganda ni=Kim!* 'How beautiful Kim is!') (Nagaya 2011: 590, 610, 616). Kratochvil (2011), based on his analysis of Abui (a Papuan language), further shows that not all demonstrative-based nominalizers necessarily yield participant nominalizations, but may instead be deployed in the formation of clausal nominalizations. In the future, it will be interesting to further explore and compare the grammaticalization pathways of demonstratives across a wider range of languages.

Acknowledgements

We wish to gratefully acknowledge the two Volkswagen Foundation DobeS grants (81821 and 83448) that provided funding for the collection of part of the data used in this study. The project was carried out by Rik van Gijn, Vincent Hirtzel and the first author. Commercial rights of the data collected during this project lie with CENY (Consejo Educativo de la Nación Yurakaré). We would also like to gratefully acknowledge the Research Grants Council of Hong Kong (General Research Fund 2010-2014; PolyU 5513/10H) awarded to the second author for the project entitled "Stance Marking in Asian Languages: Linguistic and Cultural Perspectives". The first author is further indebted to Nicholas Evans who funded the Family Problems project and a visit to the Australian National University, Canberra from his Anneliese Maier Research Award. Further, we thank all the speakers of Yurakaré who generously participated in the data collection and thus made this work possible. In particular, we thank Jeremías Ballivián, who transcribed most of the data used here, and who also conducted the interviews and facilitated contacts with speakers during the Family Problems data collection. Thanks also to Julia Colleen Miller for her invaluable support regarding the setup of the recording equipment for the 2013 field trip. We are also very grateful to editors David Fleck, Matt Shibatani and Roberto Zariquiey and an anonymous reviewer for many inspiring comments.

Abbreviations

1	first person
2	second person
3	third person
ABL	ablative
ADAP	adaptive
BEN	benefactive
COL	collective
COM	comitative
COMM	commitment
DCSD	deceased
DEM	demonstrative
DIR	direction
DST	distributive
EXH	exhortative
F	female speaker
FR	frustrative
FUT	future
HAB	habitual
IMP	imperative
INS	instrumental
INT	intentional
INTSUBJ	intersubjective
IPFV	imperfective
JUS	jussive
LIM	limitative
LOC	locative
MAL	malefactive
NEG	negation
NMLZ	nominalization
NMZR	nominalizer
OBJ	object
PL	plural
POSS	possessive
PROPN	proper noun
REA	realis
REP	reportative
SBJ	subject
SG	singular
(SP)	Spanish
SUBJ	subjective
TOP	topic
VLOC	verbal locative
VPL	verbal plural

References

Ahn, Mikyung & Yap, Foong Ha. 2014. On the diachronic development of Korean 'say' evidentials and their extended pragmatic functions. *Diachronica* 31(3): 299–336. https://doi.org/10.1075/dia.31.3.01ahn

Ahn, Mikyung & Yap, Foong Ha. 2015. Evidentiality in interaction: A pragmatic analysis of Korean hearsay evidentials. *Studies in Language* 39(1): 46–84. https://doi.org/10.1075/sl.39.1.03ahn

Adachi, Mayumi. 2014. Demonstratives as sentence final particles in Vietnamese. Paper presented at the Second American Pragmatics Conference, UCLA, 17 October.

Clifton, Charles Jr. & Frazier, Lyn. 2010. Imperfect ellipsis: Antecedents beyond syntax? *Syntax* 13(4): 279–297. https://doi.org/10.1111/j.1467-9612.2010.00142.x

Comrie, Bernard & Thompson, Sandra A. 1985. Lexical nominalization. In *Language Typology and Syntactic Description*, Vol. III: *Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 349–398. Cambridge: CUP.

Couper-Kuhlen, Elizabeth. 1996. Intonation and clause combining in discourse: The case of *because*. *Pragmatics* 6(3): 389–426. https://doi.org/10.1075/prag.6.3.04cou

Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.
Cristofaro, Sonia. 2019. Nominalizations in diachronic cross-linguistic perspective. In Zariquiey, Shibatani & Fleck (eds). (this volume)
Eloranta, Rita. 2019. Lexical and grammatical nominalization in Mochica. In Zariquiey, Shibatani & Fleck (eds). (this volume)
Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and Empirical Foundations*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.
Genetti, Carol. 2011. Nominalization in Tibeto-Burman languages of the Himalayan area. In Yap, Grunow-Hårsta & Wrona (eds), 163–193.
van Gijn, Rik. 2006. A grammar of Yurakaré. PhD dissertation, Radboud University Nijmegen dissertation.
van Gijn, Rik. 2011. Semantic and grammatical integration in Yurakaré subordination. In *Subordination in Native South American Languages* [Typological Studies in Language 97], Rik van Gijn, Katharina Haude & Pieter Muysken (eds), 169–192. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.97.07van
van Gijn, Rik. 2014. Subordination strategies in South America: Nominalization. In *The Native Languages of South America*, Loretta O'Connor & Pieter Muysken (eds), 274–296. Cambridge: CUP. https://doi.org/10.1017/CBO9781107360105.015
van Gijn, Rik. 2019. Case markers as subordinators in South American indigenous languages. In Zariquiey, Shibatani & Fleck (eds). (this volume)
van Gijn, Rik, Hirtzel, Vincent & Gipper, Sonja. 2011. *The Yurakaré Archive*. Online language documentation, DobeS Archive, MPI Nijmegen. < https://corpus1.mpi.nl/ >
Gipper, Sonja. 2017. Pre-semantic pragmatics encoded: A non-spatial account of Yurakaré demonstratives. *Journal of Pragmatics* 120: 122–143. https://doi.org/10.1016/j.pragma.2017.08.012
Heine, Bernd. 2002. On the role of context in grammaticalization. In *New Reflections on Grammaticalization*, [Typological Studies in Language 49], Ilse Wischer & Gabriele Diewald (eds), 83–101. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.08hei
Higashiizumi, Yuko. 2006. *From a Subordinate Clause to an Independent Clause: A History of English because-clause and Japanese kara-clause*. Tokyo: Hituzi Syobo.
Hirtzel, Vincent. 2010. Le maître à deux têtes: Enquête sur le rapport à soi d'une population d'Amazonie bolivienne, les Yuracaré. PhD dissertation, École des Hautes Études en Sciences Sociales, Paris.
Kratochvíl, Frantisek. 2011. Discourse-structuring functions of Abui demonstratives. In Yap, Grunow-Hårsta & Wrona (eds), 757–788.
Kurzon, Dennis. 1995. The right of silence: A socio-pragmatic model of interpretation. *Journal of Pragmatics* 23: 55–69. https://doi.org/10.1016/0378-2166(94)00036-E
Lyons, John. 1977. *Semantics*, Vol. 2. Cambridge: CUP.
Machado, Adriana. 2019. Nominalization and its pervasiveness in Xavante. In Zariquiey, Shibatani & Fleck (eds). (this volume)
Malchukov, Andrej. 2004. *Nominalization/Verbalization: Constraining a Typology of Transcategorial Operations* [Lincom Studies in Language Typology 8]. Munich: Lincom.
Malchukov, Andrej. 2006. Constraining nominalization: Function/form competition. *Linguistics* 44(5): 973–1009. https://doi.org/10.1515/LING.2006.032
Morey, Stephen. 2011. Nominalization in Numhpuk Singpho. In Yap, Grunow-Hårsta & Wrona (eds), 289–311.

Nagaya, Naonori. 2011. Rise and fall of referentiality: Articles of Philippine languages. In Yap, Grunow-Hårsta & Wrona (eds), 589–626.

Nagaya, Naonori. 2014. Notes on stand-alone *yung*-nominalizations in Tagalog. *Language and Linguistics* 35: 177–186.

Noonan, Michael. 1997. Versatile nominalizations. In *Essays on Language Function and Language Type: Dedicated to T. Givon*, Joan Bybee, John Haiman & Sandra A. Thompson (eds), 373–394. Amsterdam: John Benjamins. https://doi.org/10.1075/z.82.21noo

Noonan, Michael. 2008. Nominalization in Bodic languages. In *Rethinking Grammaticalization: New Perspectives* [Typological Studies in Language 76], Maria José Lopez-Couso & Elena Seoane (eds), in collaboration with Teresa Fanego, 219–237. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.76.11noo

Noonan, Michael. 2011. Aspects of the historical development of nominalizers in the Tamangic languages. In Yap, Grunow-Hårsta & Wrona (eds), 195–214.

Ohori, Toshio. 1995. Remarks on suspended clauses: A contribution to Japanese phraseology. In *Essays on Semantics and Pragmatics* [Pragmatics & Beyond New Series 32], Masayoshi Shibatani & Sandra A. Thompson (eds), 201–218. Amsterdam: John Benjamins.

Rhee, Seongha. 2014. Analogy-driven grammaticalization: A case of grammaticalization of sentence-final markers from concomitance-connectives. *Linguistic Research* 31(3): 591–614. https://doi.org/10.17250/khisli.31.3.201412.008

San Roque, Lila, Gawne, Lauren, Hoenigman, Darja, Miller, Julia Colleen, Rumsey, Alan, Spronck, Stef, Carroll, Alice & Evans, Nicholas. 2012. Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation & Conservation* 6: 135–174.

Schapper, Antoinette & San Roque, Lila. 2011. Demonstratives and non-embedded nominalisations in three Papuan languages of the Timor-Alor-Pantar family. *Studies in Language* 35(2): 380–408.

Schleppegrell, Mary J. 1991. Paratactic *because*. *Journal of Pragmatics* 16: 323–337. https://doi.org/10.1016/0378-2166(91)90085-C

Shibatani, Masayoshi. 2013. What can Japanese dialects tell us about the function and development of the nominalization particle *no*? In *Japanese/Korean Linguistics 20*, Bjarke Frellesvig & Peter Sells (eds), 429–451. Stanford CA: CSLI.

Shibatani, Masayoshi. 2019. What is nominalization? Towards the theoretical foundations of nominalization. In Zariquiey, Shibatani & Fleck (eds). (this volume)

Shinzato, Rumiko. 2015. Two types of conditionals and two different grammaticalization paths. In *Sentence Final Particles*, (ed.) Hencal, Sylvie, 157–180. Berlin: Mouton de Gruyter.

Simpson, Andrew. 2008. The grammaticalization of clausal nominalizers in Burmese. In *Rethinking Grammaticalization* [Typological Studies in Language 76], Maria José Lopez-Couso & Elena Seoane (eds), in collaboration with Teresa Fanego, 265–288. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.76.13sim

Sung, Li-May. 2011. Clausal nominalization in Budai Rukai. In Yap, Grunow-Hårsta & Wrona (eds), 523–559.

Tanangkingsing, Michael. (2012). Demonstratives as stance markers: An analysis of non-referential and pragmatic uses of deictic expressions in Cebuano. Ms, National Taipei University of Technology.

Traugott, Elizabeth C. & Heine, Bernd. 1991. Introduction. In *Approaches to Grammaticalization, II: Focus on Types of Grammatical Markers* [Typological Studies in Language 19], Elizabeth C. Traugott & Bernd Heine (eds), 1–14. Amsterdam: John Benjamins.

Valle, Daniel & Zariquiey, Roberto. 2019. Clausal nominalization in Kakataibo (Panoan). In Zariquiey, Shibatani & Fleck (eds). (this volume)

Van linden, An. 2019. Nominalization in Harakmbut. In Zariquiey, Shibatani & Fleck (eds). (this volume)

Wai, Brian & Yap, Foong Ha. 2014. *"Better left unsaid": The role of silence and ellipsis in political discourse*. Ms, Hong Kong Polytechnic University.

Watters, David. 2008. Nominalization in the Kiranti and Central Himalayish languages of Nepal. *Linguistics of the Tibeto-Burman Area* 31(2): 1–43.

Yap, Foong Ha, Choi, Pik-ling & Cheung, Kam Siu. 2010. De-lexicalizing *di3*: How a Chinese locative noun has evolved into an attitudinal nominalizer. In *Formal Evidence in Grammaticalization Research* [Typological Studies in Language 94], An Van linden, Jean-Christophe Verstraete & Kristin Davidse, in collaboration with Hubert Cuyckens (eds), 63–91. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.94.04yap

Yap, Foong Ha, Yi Deng & Marco Caboara. 2017. Attitudinal nominalizer(s) in Chinese: Evidence of recursive grammaticalization and pragmaticization. *Lingua* 200: 1–21. https://doi.org/10.1016/j.lingua.2017.07.011

Yap, Foong Ha & Grunow-Hårsta, Karen. 2010. Non-referential uses of nominalization constructions: Asian perspectives. *Language and Linguistics Compass* 4(12): 1154–1175. https://doi.org/10.1111/j.1749-818X.2010.00250.x

Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick (eds). 2011. *Nominalization in Asian Languages: Diachronic and Typological Perspectives* [Typological Studies in Language96]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.96

Yeh, Meili Marie. 2010. Grammaticalization of the demonstrative *'isa:a'*in Saisiyat. Paper presented at theWorkshop on Pragmatic Markers in Asian Languages: A pre-conference workshop of The 4th Conference on Language, Discourse and Cognition, National Taiwan University, 30 April.

Yeh, Meili Marie. 2014. *Demonstrative as a discourse connective device: A case study of 'isa:a' in Saisiyat*. Ms, National Hsinchu University of Education, Taiwan.

Zariquiey, Roberto. 2019. Nominalization and switch-reference in Iskonawa (Panoan, Peru). In Zariquiey, Shibatani & Fleck (eds). (this volume)

Zariquiey, Roberto, Shibatani, Masayoshi & Fleck, David W. (eds). 2019. *Nominalization in Languages of the Americas* [Typological Studies in Language 124]. Amsterdam: John Benjamins.

CHAPTER 10

The rise of the nominalizations

The case of the grammaticalization of clause types in Ecuadorian Siona

Martine Bruil
University of Leiden

The nominalization of verbs in Ecuadorian Siona is carried out by nominal classifiers. The major types of nominalization in the language are event nominalization and subject nominalization. Object nominalization needs additional morphology. When verbs are not nominalized they carry portmanteau morphology that marks the categories of subject, clause type (assertive, reportative, interrogative, and dependent clauses) and tense. A peculiarity of the system is that the reportative, interrogative and dependent clause subject paradigms show a remarkable resemblance with nominal classifiers in the language. This paper proposes that the reportative forms grammaticalized from a reported speech construction through clause union, the interrogative forms grammaticalized from a (pseudo-)cleft construction through insubordination and the dependent clause forms developed from nominalizations that were used adverbially.

1. Introduction

One of the striking features of Ecuadorian Siona is that it shows massive homophony in its verbal morphology. Identical subject agreement suffixes are used in different functions in this Western Tukanoan language. This can be observed in Example (1):

(1) a. *Kahka-**hi**.* (Assertive)
enter-3S.M.PRS.ASS
'He is entering.'
b. *Kahka-**ki**?* (Interrogative)
enter-2/3S.M.PRS.N.ASS
'Are you (M)/ is he entering?'

https://doi.org/10.1075/tsl.124.10bru

c. *Kahka-**kɨ**-jã.* (Reportative)
enter-2/3S.M.PRS.N.ASS-REP
'You (M) are/ he is entering, they say.'

d. *Kahka-**kɨ**-na jãã-wɨ.* (Dependent)
enter-S.M.PRS.DEP-DS see-N3S.PST.ASS
'While he was entering, I saw (him).'

e. *Kahka-**kɨ**-bi jãã-bi.* (Nominalization)
enter-CLS:ANIM.M-SBJ see-3S.M.PST.ASS
'The one who was entering saw (it).'

All clause types illustrated above employ a suffix *-kɨ* to mark similar functions (1b)–(1e), except for the assertive, which employs the suffix *-hi* in (1a). The homophonous *-kɨ* suffixes that are used in interrogative clauses such as in (1b), in reportative clauses such as in (1c), and in dependent clauses such as in (1d) all express some type of subject agreement. The nominal classifier *-kɨ* functions as an subject nominalizer in (1e).

Synchronically, the suffixes with the form *-kɨ* need to be analyzed as separate suffixes although they are formally similar. These suffixes have a distinct morphosyntactic behavior: non-assertive *-kɨ* marks second and third person singular subjects, dependent verb *-kɨ* marks all singular masculine subjects, and nominalizer *-kɨ* marks masculine agents and nominalizes the verb. One can recognize the function of *-kɨ* in a certain context by virtue of the position in the sentence, the intonation, and additional morphology, such as the reportative suffix *-jã [ɲa]* (1c), the different subject suffix *-na* (1d), the case suffix *-bi* (1e), or the lack of additional morphology in the interrogative (1b), among other indicators. However, the similarities in form and function are not likely to be coincidental; diachronically, the suffixes derive from one source. Nominalizations seem to have a central role in the development of the different clause type marking constructions.

One comment needs to be made with respect to the separation of functions: while the different subject agreement categories with verbs illustrated in examples (1b)–(1e) correspond to similar but separate suffixes, I treat the classification and nominalization uses as two functions of the same suffix, the classifying function obtaining when attached to nominals, and the nominalizing function when attached to verbs.

In this paper, I will discuss the grammaticalization paths by which these similar but distinct constructions probably developed. In order to discuss these pathways, I will first introduce the nominalization strategies in Ecuadorian Siona in Section 2. In Section 3, I will present the different subject agreement paradigms that are used to mark different clause types in the language. I will discuss the source of the subject agreement suffixes in Section 4 and I will reconstruct the

pathways of development of the reportative, interrogative, and dependent verbs in Section 5. Section 6 is a summary of the proposed reconstruction.

2. Nominalization in Ecuadorian Siona

Nominalization is a productive process in Ecuadorian Siona. The suffixes that are used to nominalize the verb are nominal classifiers. The relation between nominalization and nominal classification will be discussed in Subsection 2.1. It is possible to create different types of deverbal nouns. These will be discussed in Subsection 2.3. The uses of these nominalizations will be discussed in Subsection 2.3.

2.1 The role of nominal classifiers

One may have observed in the glosses of the examples above that most suffixes that can be used to nominalize verbs are nominal classifiers. The creation of deverbal nouns is only one function of the nominal classifiers. This suffix type is employed in Ecuadorian Siona to individuate an item, to mark that a nominal is of a specific class, or to express agreement within a noun phrase. Nominal classifiers are used with some nouns, demonstratives, numerals and adjectives (see Shibatani this volume). Examples of the use of classifiers with nouns are provided in (14) and (15):

(2) *baa-ko*
spouse-CLS:ANIM.F
'wife'

(3) *bɨ'ka'-kɨ*
parent-CLS:ANIM.M
'father'

The feminine classifier *-ko* in (2) and the masculine classifier *-kɨ* in (3) are applied here to indicate the gender of the referents. Most of the suffixes that operate as nominalizers with verbs are found with other word classes with distinct classifying functions, as illustrated above for the feminine and masculine classifiers.

The nominalizer *-se'e* does not seem to have a generalized use as nominal classifier. It is not found with a similar function in the nominal domain. Therefore, it is analyzed here as a plain nominalizer.

2.2 Types of nominalizations

One type of nominalization in Ecuadorian Siona is subject nominalization. I use the term subject nominalization instead of agentive nominalization, because the deverbal noun refers to the subject of the event. This subject is not always an agent. This type of nominalization is regularly used to refer back to the actors within stories. Examples (4)–(5) illustrate this use:

(4) *Jũ-'i-kɨ-bi* *kaa-ɨ-jã…*
sit-IMPF-CLS:ANIM.M-SBJ say-2/3S.M.PST.N.ASS-REP
'The one who was sitting said…'.

(5) [*Jɨ'-de* [*o-i-ko-bi*]$_{NP}$ *ihko nee-goja-o-na* *jɨ'ɨ*
[1S-OBJ take.pity-IMPF-CLS,ANIM.F-SBJ] here do-heal-S.F.PRS-DS 1S
ba-'i-jɨ.
live-IMPF-N3S.PRS.ASS
'The one who took pity upon me healed me here and that is why I am alive.'

The suffix *-kɨ* nominalizes the verb *jũ'i-* 'to sit' in (4) and the suffix *-ko* nominalizes the verb *oi-* 'to take pity on' in (5). Both nominalizations now refer to the referent that would be a subject in the non-nominalized variant of these verbs: 'the one who sits' and 'the one who takes pity on.' The difference between the suffixes *-kɨ* and *-ko* lies in the gender of the referent; *-kɨ* refers to a masculine referent and *-ko* mostly to a feminine referent, but in some cases to an inanimate one. If a nominalized verb is used to refer to a group of people, the nominalizing suffix *-k^{w}a'i* is used as illustrated in Example (6):

(6) [*Jure ah-k^{w}a'i-kato*]$_{NP}$ *goa gohcho-jɨ.*
now COP-CLS:ANIM.PL-TOP just lie-N3S.PRS.ASS
'The ones from the modern times just lie.'

The suffix *-k^{w}a'i* is a contraction of the feminine suffix *-ko* and the pluralizing suffix *-wa'i*. When *-kɨ*, *-ko*, and *-k^{w}a'i* are used without any other nominalizing morphology these suffixes create subject nominalizations.

However, these suffixes are not used only in subject nominalizations. They can also be employed in combination with other suffixes in order to refer to object nominalizations, as in (7), and place nominalizations, as in (8):

(7) [*Jɨ'ɨ wa-i-sih-k^{w}a'i-de*]$_{NP}$ *hẽã-goa-je* *bãã-hɨ̃'ɨ̃.*
1S kill-IMPF-CMPL-CLS:ANIM.PL-OBJ throw.away-just-CLS:GEN NEG.AUX-IMP
'Don't just throw away the ones (animals) that I have killed.'

(8) [*ĩmɨ-toa-jo-wɨ gahe-hã'-ko-de*]$_{NP}$ *ma'a*
high-fire-canoe-CLS:CONTAIN descend-PRP-CLS:ANIM.F-OBJ path
hɨo-ka-ɨ'ɨ.
clear-BEN-N3S.PST.ASS
'We cleared the path where the plane would land.'

The deverbal noun *waisihkʷa'ire* 'the one that I killed' in (7) refers to the object of the verb *wai-* 'to kill'; the speaker is referring to the killed animals and not to the killer. It refers to the object of the verb *hẽãgoaje* 'just throw away' in the main clause. The nominalization *ĩmɨtoajowɨ gahehã'kore* 'where the plane would land' in (8) refers to the location where the 'landing' (*gahe-*) was going to take place.

Object and place nominalizations can also be created by the use of different nominalizing suffixes. The suffix *-se'e* is used to refer to the object of an action, as illustrated below:

(9) *sɨjo-se'e*
smoke-NMZR.OBJ
'the smoked thing'

The deverbal noun *sɨjose'e* in (9) can be used to refer to any entity that is a possible object of the verb *sɨjo-* 'to smoke,' such as fish. The suffix *-se'e* is used when the action that the object underwent is already completed. That is, the object in (9) is already smoked. The nominalizing suffixes *-sih*, illustrated in (7), and *-se'e*, illustrated in (9) are probably historically related; the functions of the suffixes is very similar. The major difference is that *-sih* is followed by an additional classifier and *-se'e* is not. Since these additional classifiers are mostly human classifiers, its combination with *-sih* refers mostly to a human object.

It is possible to create further types of place nominalizations using classifiers that refer to places. This is illustrated in the examples below:

(10) *ba-'ɨ-we'jã*
be-IMPF-CLS:PLACE
'the place where one lives'

(11) *moo-ja*
fish-CLS:RIVER
'the river where one fishes'

The place classifier *-we'jã* is used in Example (10) to create the deverbal noun *ba'iwe'jã* 'the place where one lives.' The classifier *-ja* in Example (11) is used to create a more specific place, namely a river where one fishes.

An additional type of nominalization that is found in Ecuadorian Siona is the event nominalization. Event nominalizations are likewise formed with a nominal

classifier, in this case, the generic classifier -je. An example of this type of nominalization is presented below:

(12) *De'o-hi* [*ĩne* *mɨhto-je.*]$_{NP}$
be.good-3S.M.PRS.ASS palm.peach peel-CLS.GEN
'Peeling peach palm fruits is nice.'

The deverbal noun *mɨhtoje* is used to refer to the event of peeling palm peaches. It refers to an event that is unbounded with respect to time and place.

2.3 Functions of nominalizations

The types of nominalizations that were presented in the previous subsection can be used in various functions. One function is forming a nominal element that fills an argument slot in the matrix clause, as shown in Examples (4), (5), and (7). The nominalizations *jũ'ikɨbi* 'the one who is sitting' in (4), *oikobi* 'the one who takes pity' in (5), and *ĩne mɨhtoje* 'peeling peach palm fruits in (12) fill the subject slot in the matrix clauses. The nominalization *waisihk^{w}a'ire* in (7) fills the object slot of the matrix clause. These nominalizations occur without any additional noun or nominal element. These nominalizations and especially nominalizations in sentence initial position, are often employed for tail-head linking. This usage helps to disambiguate the subject of the following verb. This is illustrated in the example below:

(13) a. *Jãã-ko* *de'o-ko* *bohai-dajã-kɨ'-o*
see-S.F.PRS be.good-S.F.PRS white-hair-have-S.F.PRS
bah-ko-jã *ĩõ.*
be-2/3S.F.N.PST.N.ASS-REP she
'She was watching and was beautiful and had blond hair.'

b. ***Ba-'i-ko-bi*** *hãõ behto-hubɨ* *hãõ kaa-to*
be-IMPF-CLS.ANIM.F-SBJ this coconut-bunch she say-CLS.PLACE
hũĩ-a-o, *ohko-a-'ɨ* *kaa-o-jã.*
green-COP-3S.F.ASS water-COP-OTH.ASS say-2/3S.F.PST.N.ASS-REP
'The one who was (in possession of blond hair) said: "This bunch is still green, it is just water"'.

In (13a), the tail of the sentence is expressed by the verb *bahkojã* 'she was'. Sentence (13b) begins with a nominalization of the same verb: *ba'ikobi* 'the one who was.' The speaker only repeats the final verb and not the entire verb phrase *bohaidajãkɨ'o bahkojã* 'she was in the possession of blond hair.' This use of a nominalization is one of the reference tracking devices between sentences that are used in the language.

Nominalizations can also be used to modify a noun. An example of this relativizing function of nominalizations is presented in (8). The nominalization *ĩmɨtoajowɨ gahehã'kore* 'where the plane would land' modifies the noun *ma'a* 'path' in (8).

The third function of nominalizations is their use in periphrastic constructions. The expression of some functions in Ecuadorian Siona such as negation (14), some modal meanings (15), and future (16) is accomplished using periphrastic constructions that involve nominalizations:

Negation:

(14) *We'e-je bãã-hɨ̃'ɨ̃.*
carry-CLS:GEN NEG.AUX-IMP
'Don't carry (it)!'

Abilitative modality:

(15) *Kʷɨ-ma'-kɨ ba-ha'i jɨ' jĩhkʷ-ɨ.*
swim-NEG-CLS:ANIM.M be-3S.M.PST.ASS 1S grandparent-CLS:ANIM.M
'My granddad couldn't swim.' (Lit. My granddad was a non-swimmer)

Future tense:

(16) *Jɨ'ɨ saa-ni trabaha-hã'-kɨ-a-'ɨ.*
1S go-SS.PST work-PRP-CLS:ANIM.M-COP-N3S.PRS.ASS
'I am going to go to work.' (Lit. After I go, I am a prospective worker)

The periphrastic negation in which the negative auxiliary *bãã-* is used always contains an event nominalization that is marked by the suffix *-je*, as illustrated in Example (14). Most modal meanings are expressed periphrastically as well. Ability (or inability) is expressed by the combination of a subject nominalization, using the suffixes *-kɨ*, *-ko*, or *-kʷa'i* to nominalize the verb in combination with the existential verb *ba'i-*, as shown in Example (15), or the copula *a-*. The periphrastic construction that expresses future shown in (16) is similar to the expression of ability: it contains a subject nominalization and the copula *-a*. Additionally, it makes use of the purposive suffix *-hã'*. In summary, a nominalization can be used by itself to fill an argument slot, to modify a noun, and in periphrastic constructions that express negation or various TAM meanings.

3. Subject agreement morphology

When verbs are not nominalized, they generally carry some type of subject agreement morphology. Ecuadorian Siona possesses various different subject agreement

paradigms, because the marking is different for the distinct clause types and for the distinct tenses. I will discuss the paradigms for the different clause types in 3.1 and for (past) tense in 3.2.

3.1 Clause type paradigms

Ecuadorian Siona distinguishes four clause types in its verbal morphology. The term 'clause type' is used here to refer to the grammatically marked function of a sentence, following Portner (2009: 262-263). Other scholars have also referred to this category, as 'sentence type' (Sadock & Zwicky 1985; König & Siemund 2007). The three common cross-linguistic categories are the declarative, which makes statements, the interrogative, which expresses questions, and the imperative, which is used for orders or requests (Sadock & Zwicky 1985; König & Siemund 2007; Portner 2009). Ecuadorian Siona shows a similar pattern albeit with slight differences. The language formally distinguishes assertions from reports, categories that are cross-linguistically often considered to be part of the declarative (Aikhenvald 2004).

With respect to subject agreement marking, there are three major formal categories with respect to clause types in Ecuadorian Siona: assertions, non-assertions and dependent clauses. These three clause types all have different subject agreement paradigms. The category of non-assertions requires further explanation. It consists of interrogative clauses and reportative clauses. The conflation of questions and reports as one category may be surprising, since assertions and reports are often considered to pertain, at least semantically, to the realm of declarative clauses, as mentioned above. There is some semantic evidence that reports are similar to questions in Ecuadorian Siona. To be specific, both questions and reports do not assert the information in the proposition. The speaker enquires about the information in the proposition in questions. In Ecuadorian Siona reports, the speaker does not assert the information either:

(17) *Ligia duhta-ko-jã,* *kaa-de-na,* *duhta-je*
Ligia take.out-2/3S.M.PRS.N.ASS-REP say-PL.PST.DEP-DS take.out-CLS.GEN
bãã-ko.
NEG.AUX-3S.F.PRS.ASS
'Ligia is taking (it) out, they say, (but although) they said (that), she isn't taking (it) out.'

In the first part of the sentence in (17), the speaker is just presenting the information that Ligia is taking something out, but he/she is not vouching for its truth.

That is why the speaker can deny it is true in the following sentence without it being a contradiction.[1]

The non-assertive character of both questions and reports may suggest that these should be grouped together not only on the basis of their subject agreement morphology, but also on the basis of their semantics. For a more detailed semantic analysis of the Ecuadorian Siona clause-typing system, see Bruil (2014, 2015). The focus of this paper is on the similarities in form of the non-assertive clause types. A historical explanation for the assertive versus non-assertive division will be discussed in Section 4. I will discuss the way in which assertive paradigms are structured in Subsection 3.1.1. The organization of the non-assertive is described in Subsection 3.1.2 and that of dependent verbs in Subsection 3.1.3.

3.1.1 *Assertive clauses*

Assertive clauses are used when speakers have good reason to believe that the information that they are presenting is true; the speaker asserts that the information in the proposition is true. One can identify this clause type by its distinctive subject agreement morphology. All assertive clauses are marked for person, number and gender. The paradigm consists of three subject agreement suffixes. The mapping of person, number, and gender categories onto these suffixes is illustrated in the example below:

(18) a. *Kaa-ko.*
say-3S.F.PRS.ASS
'She says.'

b. *Kaa-hi.*
say-3S.M.PRS.ASS
'He says.'

c. *Kaa-ɉi.*
say-N3S.PRS.ASS
'I / you / we/ you all / they say.'

The subject agreement markers *-ko* and *-hi* mark very specific categories: third person singular feminine and third person singular masculine, respectively. The third form *-ɉi* is essentially a default category that expresses every person, number, gender combination that is not third person singular.

1. For more detailed explanations of this test, also referred to as the known truth/falsity test, see (Bruil 2014; Faller 2002; Matthewson et al. 2007; Peterson 2010; Waldie et al. 2009).

3.1.2 *Non-assertive clauses*

Non-assertive clauses (i.e. interrogative and reportative clauses) have a distinct subject agreement paradigm. It consists of three suffixes just as the assertive paradigm does, but the forms are different, except for one, and the categories of person, number, and gender are differently distributed over the three forms. The paradigm is presented for interrogative clauses in (19):

(19) a. *Kaa-ko?*
say-2/3S.F.PRS.N.ASS
'Do you (F)/Does she say (so)?'

b. *Kaa-kɨ?*
say-2/3S.M.PRS.N.ASS
'Do you (M)/Does he say (so)?'

c. *Kaa-je?*
say-N2/3S.PRS.N.ASS
'Do I / we / you (all) / they say (so)?'

Just as in the case of the assertive paradigm, the non-assertive paradigm has two suffixes with highly specific reference, namely *-ko* for second or third person singular feminine and *-kɨ* for second or third person masculine, and one suffix that comprises more person, gender, and number combinations, *-je,* for all non-second or third person singular categories.

Reportative clauses have exactly the same subject agreement paradigm. Example (20) shows the reportative forms of the verb:

(20) a. *Kaa-ko-jã.*
say-2/3S.F.PRS.N.ASS-REP
'You (F) say / she says.' (It is said.)

b. *Kaa-kɨ-jã.*
say-2/3S.M.PRS.N.ASS-REP
'You (M) say / he says.' (It is said.)

c. *Kaa-je-jã.*
say-N2/3S.PRS.N.ASS-REP
'I / we / you (all) / they say.' (It is said.)

If one compares Example (19) with (20), one can observe that questions and reports have identical subject agreement paradigms. There are two ways to distinguish questions from reports in Ecuadorian Siona. The two clause types differ with respect to the presence of the suffix *-jã*, which is only used in reportative clauses. Additionally, the intonation of the two clause types is different: questions have rising intonation at the end of a sentence and reports a falling intonation.

3.1.3 *Dependent clauses*

Dependent verbs are used in clause-chaining. It is very common to have many dependent verbs in one sentence, especially in narratives. Sentences in this genre will generally have various dependent verbs and only the final verb is an independent verb (either assertive or non-assertive). An example of this is provided below:

(21) *Waa-ni daa-ɨ-na ĩ dīhõ soe-ni te'te-ni*
kill-SS.PST bring-S.M.PST.DEP-DS his wife pluck-SS.PST cut.into.pieces-SS.PST
k^{w}a'ko-ni mama-hɨ-de ãõ-o-jã.
cook-SS.PST child-CLS:COL-OBJ feed-2/3S.F.PST.N.ASS-REP
'After he had killed (the game), he brought (it) and his wife plucked (it), cut (it) to pieces and cooked (it) and gave (it) to the children to eat.'

Five out of the six verbs in Example (21) are dependent verbs. Dependent verbs are marked for switch reference. This marking is determined by the subject of the following verb and not the main verb. For instance, the verb *waani* '(he) killed' is marked for 'same subject.' This means that its subject is the same as that of the following verb *daaɨna* 'he brought.' The subject of *waani* is, however, different to that of the independent verb *ãõjã* 'she fed.' The subject of *daaɨna* 'he brought' is different from that of *soeni* '(she) plucked,' and hence a different-subject marker *-na* is used.

Dependent verbs are not semantically subordinate to the main verb. Clause-chains describe a series of actions that occur simultaneously or sequentially. There is generally no indication that the action expressed by the independent verb is more salient than the other action. Generally the independent verb is the chronologically final action that occurs in the sequence and therefore, it fills the final slot in the sentence. Dependent verbs are, however, syntactically dependent. These verbs require an independent verb for their anchoring in real time and for assigning a clause type to a sentence. Explicitly, we know that all actions expressed by the verbs in (20) occurred in the past and that the speaker found out through hearsay, since the independent verb *ãõjã* is marked for past tense and as a report.

Similarly to the assertive and non-assertive verbs, dependent clauses have their own subject agreement paradigm. Same-subject and different-subject verb forms are the same, except for the different subject suffix *-na*, which is not present in same subject clauses. The suffixes are provided in the example below. One should note that only the different subject combinations are possible in the translations.

(22) a. *Kaa-ko-na sa-i-jɨ.*
say-S.F.PRS.DEP-DS go-IMPF-N3S.PRS.ASS
'While I (F)/ you (F)/she spoke, I/ you (S)/we/you (PL)/they left.'

b. *Kaa-kɨ-na sa-i-jɨ.*
say-S.M.PRS.DEP-DS go-IMPF-N3S.PRS.ASS
'While I (M)/you (M)/he spoke, I/ you (S)/we/you (PL)/they left.'

c. *Kaa-hɨ-na* *sa-i-jɨ.*
say-PL.PRS,DEP-DS go-IMPF-N3S.PRS.ASS
'While we/ you (PL)/they spoke, I/ you (S)/we/you (PL)/they left.'

The dependent verb paradigm consists of three suffixes just as in the cases of the assertive and the non-assertive subject agreement morphology. The suffixes *-ko* and *-kɨ* are identical in form to the non-assertive suffixes, yet they mark a different subject agreement category. The dependent suffixes do not mark person, only number and gender: *-ko* marks singular feminine subjects, *-kɨ* marks singular masculine subjects, and *-hɨ* marks plural subjects.

3.2 Tense and subject agreement morphology

Ecuadorian Siona marks tense on the verb by changing the subject agreement paradigms. The paradigms that were shown in the previous subsections are the present tense paradigm. Past tense verbs show different, yet related subject agreement suffixes for the three clause type categories presented in the previous subsection. The agreement categories do not change within a clause type, the only change is in the form of the suffixes. This is illustrated for the non-assertive verbs in Table 1 below:

Table 1. Non-assertive subject agreement in present & past tense

Tense	Person / number / gender	Subject agreement
PRESENT	2/3S.F	-ko
	2/3S.M	-kɨ
	N2/3S	-je
PAST	2/3S.F	-o
	2/3S.M	-ɨ
	N2/3S	-de

It can be observed that the past tense suffixes are different from the present tense. However, there are striking similarities as well. The second or third person singular feminine suffixes both contain the vowel *-o*, the second or third person singular masculine suffixes the vowel *-ɨ*, and the non-second or third person singular the vowel *-e*. One can conclude from this that the vowel remains stable, while the differences apply to the onset of the suffix. In the second or third person singular suffixes, there is a difference with regard to the presence of an onset, and with the non-second or third person singular, the consonants are distinct.

Tense is also indicated in dependent verbs, albeit relative tense. Tense marking in dependent clauses does not anchor the described event to speech time, but

it anchors it to the time line of the episode described in the main clause. This is illustrated in the example below:

(23) a. *Jɨ̃'ɨ̃ jãã-ko-na sa-ha'i.*[2]
1s see-S.F.PRS.DEP-DS see-3S.M.PST.ASS
'While I saw (him), he left.'

b. *Jɨ̃'ɨ̃ jãã-o-na sa-ha'i.*
1s see-S.F.PST.DEP-DS see-3S.M.PST.ASS
'After I saw (him), he left.'

The action of the first person seeing, as expressed by *jããkona* in (23a), is present tense with respect to the moment in time of the third person going, as expressed by *saha'i*. This relative present tense semantics yields an interpretation of simultaneity of both actions. The action of the first person seeing in (23b), as expressed by *jããona*, occurred in the past from the perspective of the third person going. This relative tense semantics generates a sequential interpretation.

There are more distinctions in the expression of switch reference in the past tense than in the present tense. In present tense, both same-subject and different-subject verb forms carry a subject agreement morpheme, as discussed in the previous subsection. In past tense, only the different-subject forms contain a subject agreement suffix. The verb in dependent clauses with the same-subject in the past tense is marked by the suffix *-ni*, as illustrated in the example below:

(24) *Hãɨ̃ bõ'sɨ̃ jõhkʷa nee-ni daa-ni sede-ni kʷa'ko-ni*
that young.man chambira make-SS.PST bring-SS.PST strip-SS.PST cook-SS.PST
kʷẽna-ni ĩ [jõhkʷa oja-kɨ̃]$_{NP}$ *ba-ã'-kɨ̃-jã.*
dry-SS.PST he chambira roll-S.M.PRS.DEP be-REM.PST-2/3S.M.PST-REP
'The young man made chambira (palm fiber (twine)), he brought (it), stripped (it), cooked (it) and twisted the chambira.'

The episode in (24) describes a sequence of actions all carried out by the same subject, namely by *hãĩ̃ bõ'sɨ̃* 'that young man.' Therefore, almost all the verbs are marked with the past tense same subject suffix *-ni*.

Table 2 presents the suffixes that are used to mark past and present dependent verbs both in same-subject and different-subject contexts:

2. The verb *saha'i* belongs to a different verb class than the ones that are presented in this paper. This verb class shows different but related subject agreement suffixes than the verb class that is used here to illustrate the historical development of clause types. The verb class of *saha'i* does, however, follow the same principal of development. For lack of space the discussion on verb classes has been left out of this paper.

Table 2. Dependent verb subject agreement in present & past tense

Tense	Gender & number	Same Subject	Different Subject
PRESENT	S.F	-ko	-ko-na
	S.M	-kɨ	-kɨ-na
	PL	-hɨ	-hɨ-na
PAST	S.F	-ni	-o-na
	S.M		-ɨ-na
	PL		-de-na

The subject agreement suffixes that are used in dependent clauses show roughly the same pattern with respect to present and past tense. The singular feminine and singular masculine forms have the same vowel and there is only a difference in the onset. For the plural, there is both a mismatch in the consonant and in the vowel; the suffixes *-hɨ* and *-de* do not demonstrate similarities in either consonant or vowel. This mismatch can be explained historically; the suffixes have developed form distinct nominalizing suffixes.

In summary, we have observed that the subject agreement suffixes are actually portmanteau suffixes that also mark clause-type and tense. The subject agreement is marked by different suffixes. Distinct clause types are marked by distinct sets of suffixes and the distinct distribution of subject agreement categories. Tense is, broadly speaking, marked by a change in the onset of the suffixes. This change can be reconstructed as a morphophonological effect on the consonant for Proto-Western Tukanoan: some paradigms had fortis and others lenis consonants. This reconstructed morphophonological effect is more apparent in other Western Tukanoan languages, such as Colombian Siona (Wheeler 1987: 155–156) and Máíhɨ̃kì (Velie & Velie 1981: 123–125; Michael 2012a).

4. The source of the subject agreement suffixes

It was already alluded to in the introduction that several subject agreement forms in the non-assertive and dependent verb paradigms show a striking similarity in form to the nominal classifiers that are used to nominalize verbs. Forms identical to the feminine classifier *-ko* and the masculine classifier *-kɨ* are found throughout all the non-assertive and dependent verb paradigms with the same gender distribution. The suffix that is used when the subject is not second or third person singular in the non-assertive paradigm, *-je*, which is used in the present tense, is identical to the general nominal classifier that is used to create event nominaliza-

tions. Additionally, the plural subject suffix for present tense dependent verbs, *-hɨ,* is found as a classifier as well, namely, *-hɨ* is a collective classifier.

Two suffixes from the non-assertive and dependent verb paradigms do not have a classifier counterpart in Ecuadorian Siona. These suffixes are *-de[-re]*, which is used for non-second or third person singular past in non-assertive clauses and for plural past in dependent clauses, and *-ni*, which is used for dependent verbs with the same subject as the adjacent next verb. Nevertheless, there are related suffixes within the language family with nominalizing functions. Ecuadorian Siona belongs to the western branch of the Tukanoan family. The Tukanoan family consists of two branches: the western and the eastern branch. Languages in the eastern branch possess nominalizing suffixes with identical or similar forms to *-re* and *-ni*.

The suffix *-re* is found throughout the Eastern Tukanoan languages as some type of nominalizing suffix or nominal classifier. There is a suffix *-re* that is labelled as an infinitive marker in Barasana (Gomez-Imbert 1997: 235). Both Tuyuka (Malone & Barnes 2000: 445) and Yurutí (Kinch & Kinch 2000: 476) are described as having a nominalizer *-re*. Kotiria has a generic nominal classifier *-re* that is also found in specific nominalizations (Stenzel 2013: 335).

The suffix *-ni* does not exist in this exact form in the sets of nominal classifiers or nominalizers in Eastern Tukanoan languages. There is, however, an oral counterpart *-ri* that is found throughout the eastern branch. One indication that *-ri* is the oral counterpart of *-ni* is found in the Western Tukanoan language Sekoya. This language is closely related to Ecuadorian Siona and it has a very similar system of subject agreement forms. One of the few differences is that the suffix *-ni* is only used in nasal contexts. When an oral verb is marked for same subject in past tense, the form *-ri* is used (Vallejos 2015). Therefore, it is likely that the Ecuadorian Siona suffix *-ni* is a cognate of the nominalizing suffix *-ri* that is found in many Eastern Tukanoan languages such as Barasana (Gomez-Imbert 2004: 62; Jones & Jones 1991: 29–30), Desano (Miller 1999; Silva 2012), Kotiria (Stenzel 2013), Kubeo (Chacón 2012), and Tatuyo (Gomez-Imbert 2011).

Since all non-assertive and dependent subject agreement suffixes have a formally identical or similar counterpart that functions as a nominal classifier or a nominalizer, it is conceivable that the subject agreement function of these suffixes derived from a nominalizing function. These suffixes used to be nominalizers that ended up developing into subject agreement suffixes in reports, questions and dependent clauses. I am not the first to observe that the subject agreement suffixes in non-assertive and dependent verbs are related to the nominalizing suffixes in Western Tukanoan languages (Idiatov & van der Auwera 2004, 2008; Michael 2012b; Schwarz 2012). Idiatov and van der Auwera (2004, 2008) propose that nominalizations in Tukanoan languages were first introduced in the evidential paradigms, which then obtained a mirative reading. The erstwhile nominalizers

then developed interrogative semantics in this proposal. I propose a different explanation for the emergence of nominalizers in the subject agreement paradigms in Ecuadorian Siona in Section 5.

5. Reconstructing the non-assertive and dependent constructions

Consequently, the next question that arises is: how did these nominalizers develop into subject agreement suffixes? In order to provide a feasible account of how this happened, one needs to provide a source construction and a grammaticalization pathway. In this section, I will provide separate but related accounts for the reportative in Subsection 5.1, the interrogative in Subsection 5.2, and the dependent verbs in Subsection 5.3.

5.1 The origin of the reportative

Reported speech constructions are a cross-linguistically common origin for reportative evidentials (Aikhenvald 2004: 271–273). The Ecuadorian Siona reportative seems to have originated as a reported speech construction as well. Because of the presence of the non-assertive subject agreement suffixes that are likely to cognates of nominalizing suffixes, the reportative is reminiscent of an indirect speech construction. In this analysis, this construction comprised a verb that introduced an indirect speech complement. I will refer to this verb as the SPEECH VERB. The reportative suffix *-jã* may be a trace of this verb. The nominalizations, in this reconstruction, functioned as the complement clauses that expressed the indirect report. The reconstruction of this indirect speech construction is provided in (24):

(25) *
[[VERB ROOT-NOMINALIZER] SPEECH-AGREEMENT]
[[je'je-kɨ] jã-jɨ.]
[[study-NMZR.M] SPEECH-N3S.PRS]
'They say that he studies.'

The reconstructed template, as illustrated in (25), consists of two clauses: the main clause and the complement clause. It can also be observed that the reconstructed verb **jã* is an independent verb that is fully inflected.

The next phase in the proposed development of the reportative is clause union. Clause union can be described as a process in which two clauses become one complex verb phrase, following Givón (2009: 61-63). When the Ecuadorian Siona reported speech construction underwent clause union, the speech verb **jã* lost its independent properties and became part of a complex predicate. This

process probably consisted of at least three stages. First the verb must have lost its original subject agreement morphology. The speech verb *jã* may have been used as a particle. Then *jã* lost its independence and was reanalyzed as a reportative suffix. Finally, the nominalizer was reanalyzed as subject agreement morphology. The process of clause union is illustrated in (26):

(26) a. *
[[VERB ROOT-NOMINALIZER] SPEECH-AGREEMENT]
[[je'je-kɨ] jã-jɨ.]
[[study-NMZR.M] SPEECH-N3S.PRS]
'They say that he studies.'
↓

b. *
[[VERB ROOT-NOMINALIZER] SPEECH]
[[je'je-kɨ] jã.]
[[study-NMZR.M] SPEECH]
'He studies, they say.'
↓

c. *
[[VERB ROOT-NOMINALIZER]-REP]
[[je'je-kɨ]-jã.]
[[study-NMZR.M]-REP]
'He studies.' (They say.)
↓

d. [[VERB ROOT-AGREEMENT]-REP]
[[je'je-kɨ]-jã.]
[[study-M]-REP]
'He studies.' (They say.)

Example (26a) illustrates the earliest stage during which there was still an indirect speech construction. This is a reconstructed construction that is not available in the language anymore. Example (26b) illustrates the form of the construction at the stage when the original subject agreement morphology of the speech verb had reduced and disappeared. The construction in (26c) illustrates the stage when the speech verb became a reportative suffix. The final stage of the process of clause union, during which the original nominalizer starts functioning as subject gender agreement morphology, is illustrated in (26d).

The final result of the process of clause union, as presented in (26d), resembles the current reportative construction greatly. However, an additional change needed to take place for the synchronic construction to come into being. The nominalizing suffixes did not express person, only gender and number to some extent. The

suffix *-ko* marks feminine and the suffix *-kɨ* masculine. The suffixes *-je* and *-re* were general nominalizers that did not mark gender or number. The suffixes *-ko* and *-kɨ* were reinterpreted as second or third person singular markers for their respective gender. The general suffixes *-je* and *-re* came to express all other categories. This final change is illustrated in (27):

(27) *
[[VERB ROOT-(GENDER) AGREEMENT]-REP]
[[je'je-kɨ]-jã.]
[[study-M]-REP]
'He studies.' (They say.)
↓
[[VERB ROOT-(GENDER, PERSON, NUMBER) AGREEMENT]-REP]
[[je'je-kɨ]-jã.]
[[study-2/3S.M]-REP]
'He studies.' (They say.)

In summary, the reportative developed from an indirect speech construction consisting of a speech verb that formed the verb phrase in the main clause and a complement clause that was marked with a nominalization. The two separate clauses evolved into one complex clause as a result of the loss of the original subject agreement morphology on the speech verb, the reanalysis of the speech verb as a suffix and the reanalysis of the nominalizers as subject agreement morphemes. Finally, reanalysis of the subject agreement categories took place. These changes did not necessarily take place in this chronological order; it is possible that there was considerable overlap between the changes.

There are various facts that corroborate this scenario of emergence of the reportative. First of all, the subject agreement suffixes can all be traced back to nominalizing suffixes, as discussed in Section 4. Second, reported speech constructions are a widespread source for reportative morphology constructions, as mentioned at the beginning of this subsection. A final possible line of evidence consists of the possible cognates of the speech verb found in some Eastern Tukanoan languages. Kubeo has a very similar reportative clitic *=ja* (Chacón 2012), so it is possible that there may have been a verb with the shape of *ja* or *jã* in Proto-Tukanoan. There are some speech verbs in the eastern branch that may be cognate with the Ecuadorian Siona reportative suffix *-jã*. For instance, Barasana has a speech verb *jãgo* 'to speak' (Jones & Jones 1991: 28). A possible speech verb cognate in Kubeo is *jáwa* 'to speak.' Another possibility is that it was a copula that introduced the speech complement. The copula *jã* in Makuna (Smothermon et al. 1995: 43) may be cognate to the Ecuadorian Siona reportative. Another possible candidate is the

copula *ja* in Barasana that can be used to introduce speech and thought reports (Jones & Jones 1991: 28).

5.2 The origin of the interrogative

The interrogative verb forms are very similar to the reportative verb forms, as shown in Section 3. Here I propose that the reportative and the interrogative clause types resemble each other, not only in form, but also in their behavior. Just as the reportative, the interrogative subject agreement suffixes developed from nominalizers that were used to mark complement clauses. These complement clauses were probably not introduced by a speech verb. The nominalizations in the case of the interrogative probably functioned as complement clauses in (pseudo-) cleft constructions. This type of construction was probably used in order to put the requested new information in focus at an early stage. It is cross-linguistically not uncommon to find (historical) (pseudo-)cleft constructions in questions (Givón 2001: 307–308). For instance, the question particle *est-ce que* in French developed from a cleft construction (Harris 1978). The use of cleft constructions in questions is also reported for the Bantu language Kihung'an (Givón 2001: 308–309), for Dravidian and Tibeto-Burman languages (Bhattacharya & Devi 2004), and for others (Shibatani this volume).

Since (pseudo-)cleft constructions are commonly found in interrogative contexts, it would not be implausible for the Ecuadorian Siona to have developed from a similar construction. A reconstruction of this pseudo-cleft construction in questions is presented in (28):

(28) *

[INT PRONOUN	[VERB ROOT-NOMINALIZER]	COPULA-AGREEMENT]
[ke-de	[je'je-kɨ]	a-bi.]
[what-OBJ	[study-NMZR.M]	COP-3S.M.PRS]

'What is it that he studies?'

As illustrated in (28), it was probably a copula, such as the verb *a-*, which currently functions as a copula, that introduced a complement clause in questions. A (pseudo-)object cleft is used here to illustrate this construction; the object of his studying is here in focus.

The use of (pseudo-) cleft construction probably became conventionalized at some stage and the construction lost its additional focus interpretation on a particular constituent in the clause. This is illustrated in (29):

(29) a. *
[INT PRONOUN [VERB ROOT-NOMINALIZER] COP-AGREEMENT]
[ke-de [je'je-kɨ] a-bi.]
[what-OBJ [study-NMZR.M] COP-3S.M.PRS]
'What is it that he studies?'
↓

b. *
[INT PRONOUN [VERB ROOT-NOMINALIZER] COP AGREEMENT]
[ke-de [je'je-kɨ] a-bi.]
[what-OBJ [study-NMZR.M] COP-3S.M.PRS]
'What does he study?'

It is shown in (29b) that the pseudo-cleft construction at this stage does not add any additional meaning to the sentence; it has become the conventionalized form to ask a question.

A difference between the reportative and the interrogative is that in the case of the latter there is no trace of any verb in the reconstructed main clause. There is no suffix that can be reconstructed as the copula. As a consequence, it is not possible to construe an identical path of development for the interrogative. The process that probably resulted in the emergence of the interrogative is insubordination. The term insubordination was defined by Evans (2007: 367) as "the conventionalized main-clause use of what, on prima facie grounds, appear to be formally subordinate clauses." Evans establishes that many languages use constructions that are subordinate in form, in main clause contexts. Within the process of insubordination, first, the main clause verb is deleted. Then the deletion is conventionalized. During the final stage the formerly subordinate construction is reanalyzed as a main clause. The new main clause construction typically obtains a specific use.'

The reconstructed Ecuadorian Siona interrogative that contained a conventionalized cleft construction underwent insubordination. During this process the Ecuadorian Siona interrogative constructions underwent copula deletion and the nominalization was reanalyzed as a main clause verb form. This is illustrated in (30):

(30) a. *
[INT PRONOUN [VERB ROOT-NOMINALIZER] COP-AGREEMENT]
[ke-de [je'je-kɨ] a-bi.]
[what-OBJ [study-NMZR.M] COP-3S.M.PRS]
'What does he study?'
↓

b. *
[[INT PRONOUN VERB ROOT-NOMINALIZER] Ø]
[[ke-de je'je-kɨ] Ø]
[[what-OBJ study-NMZR.M] Ø]
'What does he study?'
↓

c. *
[[INT PRONOUN VERB ROOT-AGREEMENT] Ø]
[[ke-de je'je-kɨ] Ø]
[[what-OBJ study-M] Ø]
'What does he study?'
↓

d. [[INT PRONOUN VERB ROOT-AGREEMENT] Ø]
[[ke-de je'je-kɨ] Ø]
[[what-OBJ study-2/3M] Ø]
'What does he study?'

In (30b), it is shown that the copula was deleted. The next stage would be the reanalysis of the nominalizer as subject agreement morphology, as shown in (30c). Finally, as in the case of the reportative the subject agreement morphology is reanalyzed and starts to express person and number. For instance, the erstwhile masculine nominalizer *-kɨ* comes to express second or third person singular masculine, as shown in (30d).

5.3 The origin of the dependent verb marking

The subject agreement suffixes in dependent clauses can be reconstructed as nominalizing suffixes as well. From a Tukanoan perspective, this is not a surprising change. Many Eastern Tukanoan languages, such as Barasana (Jones & Jones 1991: 113) Kotiria (Stenzel 2013), and Makuna ((Smothermon et al. 1995: 68–70), use nominalizations in clause chaining-contexts (See Shibatani this volume). Additionally, there is cross-linguistic support for this reconstruction. It is observed for some Asian languages that some subordinators used in clause-chaining show the same form as nominalizers, such as in Classical Chinese (Yap & Wang 2011) and in the Formosan language Budai Rukai (Sung 2011). It is, therefore, not unfeasible that this use already existed in Proto-Tukanoan, and it is conceivable that nominalizations grammaticalized as clause chain-markers.

Furthermore, it is only a small shift from nominalizing suffix to dependent subject agreement suffix. These functions are similar and most of the suffixes maintained their reference to a closely related category. Specifically, the feminine marker *-ko* evolved into a singular feminine subject marker, the masculine marker

-kɨ into a singular masculine subject marker, and the animate collective marker *-hɨ* into a plural subject marker. These meanings are not conceptually far apart. It was a larger leap for the generic classifier *-re* to turn into a plural subject marker. The suffix *-re* was probably introduced as a default category marker in the dependent paradigm as a result of analogy with respect to the non-assertive paradigms. The suffix *-re* marks a very similar default category in these paradigms as well, namely non-second or third person singular. The only difference between the subject agreement categories is that the non-assertive suffix marks first person singular as well, whereas the dependent form marks only plural subjects.

The introduction of switch reference requires some further explanation. Ecuadorian Siona developed a same-subject marker *-ni*. This suffix is probably cognate with the general nominalizer *-ri* that is found in Eastern Tukanoan languages, as mentioned in Section 4. Since *-ri* does not mark any specific subject, its subject probably depended on the subject of the following verb for its interpretation. So it seems that the leap from general classifier that was used in clause-chaining to same subject suffix is not that far. The reanalysis of the nominalizer as a same-subject marker probably went hand-in-hand with the emergence of the different-subject constructions.

The different-subject construction consists of a former nominalizer that now functions as a subject agreement marker and the different-subject suffix *-na*. This construction probably derived from a nominalized verb that was marked for oblique case. The different subject suffix *-na* can be reconstructed as an oblique case marker; there is an oblique goal case marker *-na* in Ecuadorian Siona that is identical in form. The template of this reconstructed form (VERB-NMZR-CASE) is still a very productive template in the Ecuadorian Siona, although this template is mostly used to form (headless) relative clauses and it is mostly found in combination with the subject case marker *-bi* or the object case markers *-de [-re]* and *-ni*. One can imagine that the nominalized verb marked for oblique case was used adverbially. A change that must have taken place is that the nominalized verb lost its role as oblique argument and became a dependent verb with a clause-chaining function. Then the construction became a conventionalized different-subject construction and developed together with the same-subject marker as a grammaticalized switch reference system.

Ecuadorian Siona is not the only language in which a nominalization in combination with a case marker came to mark a different subject. The Imbabura Quichua different-subject marker *-kpi,* as described by Cole (1982), is most likely a combination of the agentive nominalizer *-k* and the locative case suffix *-pi*. A similar development was described by Overall (2011) for the Jivaroan language Aguaruna in which the locative case in combination with a subordinator developed into a different subject marker. This author found similar developments

in Panoan languages as well (Overall 2011). Aymaran languages show a similar development of a case marker into a switch reference marker (Cerrón-Palomino 2000: 244–245). In all these languages, it has been shown that an oblique case marker has developed into a switch reference marker. The Ecuadorian Siona development does not seem to be exceptional.

6. Summary

The historical reconstruction of the reportative and interrogative constructions and the dependent verbs as presented above helps to understand why there is considerable overlap in the subject agreement morphology in these three clause types. The existence of similar paradigms is due to the rise of the nominalizers that obtained the function of subject agreement suffixes. The reportative developed from an indirect speech construction in which a copula or a speech verb introduced a complement clause that was formed by a nominalized clause. This construction underwent clause union. The suffix *-jã* is analyzed here as the remains of the verb that used to introduce the complement clause. The interrogative developed from (pseudo-)cleft construction that became conventionalized. It probably consisted of a copula and a complement clause that was again formed by a nominalization. As a consequence of insubordination the interrogative construction developed. The switch reference system in dependent clauses developed as a reinterpretation of nominalizations that were used in clause-chaining environments. In the case of all three clause types, the nominalizing suffixes were reinterpreted as subject agreement suffixes. The reportative and interrogative suffixes gained the features of person and number agreement and the dependent suffixes that of number agreement.

These developments explain the high degree of homophony in the subject agreement paradigms. The former nominalizations may even have found their way into the assertive paradigm. This paradigm contains a suffix that marks third person singular feminine *-ko* that is identical in form to a nominalizing suffix and subject agreement suffixes in the non-assertive and dependent verb paradigms that also mark feminine gender in some way. The suffix *-ko* probably entered the assertive marking in analogy to the cognate forms in other clause types. The other assertive subject agreement suffixes can be reconstructed as the original main clause subject agreement morphemes. For instance, the third person singular masculine suffix *-bi* can be reconstructed to Proto-Tukanoan, according to Chacón (2014). The non-third person singular suffix *-wɨ* may also have existed in Proto-Tukanoan, since it is also found in Eastern Tukanoan languages, such as Tuyuka (Barnes 1984; Malone & Barnes 2000; Malone 1988).

The developments described in this paper have produced a complex verbal morphology and they led to grammaticalization of clause types in Ecuadorian Siona. The language now has a typologically rare system in which the reportative and the interrogative can be classified as one non-assertive category.

Acknowledgements

First of all, I would like to thank Ligia Criollo, Rogelio Criollo, Wilian Criollo, Eva Merino, Alicia Suale, Herberina Piaguaje, and Miguel Angel Piaguaje for teaching me about their language. Secondly, I would like to thank the Foundation for Endangered Languages, the Hans Rausing Endangered Languages Project (SM0067; IPF0211) and the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) (sub-project of "The Nature of Evidentiality") for funding my research. Finally, I would like to thank Willem Adelaar, Johan Rooryck, Elsa Gomez-Imbert and the participants of the Workshop on the Nominalization in the Languages of the Americas for their useful comments on previous versions of this paper.

Abbreviations

1	first person	IMPF	imperfective
2	second person	M	masculine
3	third person	MED	medial
ANIM	animate	N	non
ASS	assertive	NEG	negation
AUX	auxiliary	NMZR	nominalizer
BEN	benefactive	OBJ	object
CLS	classifier	PL	plural
CMPL	completive	PRP	purpose
CNTEXP	counterexpectational	PRS	present
COL	collective	PST	past
COP	copula	REM	remote
DEM	demonstrative	REP	reportative
DEP	dependent	S	singular
DS	different subject	SBJ	subject
F	feminine	SS	same subject
GEN	general	TOP	topic
IMP	imperative		

References

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: OUP.

Barnes, Janet. 1984. The evidentials in the Tuyuca verb. *International Journal of American Linguistics* 50(3): 255–271. https://doi.org/10.1086/465835

Bhattacharya, Tanmoy & Thangjam Hindustani Devi. 2004. Why cleft? In *SALA XXIII Procedings*, Doug Bigham, Mark Brown, Q Wan Kim & Sadia Rahman (eds), 1–10. Stanford CA: CSLI. <http://people.du.ac.in/~tanmoy/papers/Cleft_Meiteilon.pdf>

Bruil, Martine. 2014. Clause-typing and Evidentiality in Ecuadorian Siona. PhD dissertation, Leiden University.

Bruil, Martine. 2015. When evidentials are not evidentials: The case of the Ecuadorian Siona reportative. *Linguistic Typology* 19(3): 385–423. https://doi.org/10.1515/lingty-2015-0012

Cerrón-Palomino, Rodolfo. 2000. *Lingüística aimara*. Cuzco: Centro "Bartolomé de Las Casas".

Chacón, Thiago Costa. 2012. The phonology and morphology of Kubeo: The documentation, theory, and description of an Amazonian language. PhD dissertation, University of Hawai'i.

Chacón, Thiago Costa. 2014. A revised proposal of Proto-Tukanoan and Tukanoan family classification. *International Journal of American Linguistics* 80(3): 275–322. https://doi.org/10.1086/676393

Cole, Peter. 1982. *Imbabura Quechua* [Lingua Descriptive Series 5]. Amsterdam: North-Holland.

Evans, Nicholas. 2007. Insubordition and its uses. In *Finiteness: Theoretical and Empirical Foundations*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.

Faller, Martina T. 2002. The semantics and pragmatics of evidentials in Cuzco Quechua. PhD dissertation, Stanford University.

Givón, T. 2001. *Syntax. An Introduction*, Vol. II. Amsterdam: John Benjamins.

Givón, T. 2009. *The Genesis of Syntactic Complexity: Diachrony, Ontogeny, Neuro-cognition, Evolution*. Amsterdam: John Benjamins.

Gomez-Imbert, Elsa. 1997. Morphologie et phonologie Barasana: Approche non-linéaire. PhD dissertation, Unversité de Paris 8.

Gomez-Imbert, Elsa. 2004. Fonología de dos idiomas tukano del Piraparaná: Barasana y Tatuyo. *Amerindia* 29: 43–80.

Gomez-Imbert, Elsa. 2011. Le Tatuyo. In *Dictionnaire de langues*, Emilio Bonvini, Joëlle Busuttil & Alain Peyraube (eds), 1548–1555. Paris: Presses Universitaires de France.

Harris, Martin. 1978. *The Evolution of French Syntax: A Comparative Approach*. London: Longman.

Idiatov, Dmitry & van der Auwera, Johan. 2004. Nominalization as a question formation strategy in Tucanoan. *Santa Barbara Papers in Linguistics* 15: 9–24.

Idiatov, Dmitry & van der Auwera, Johan. 2008. From nominalizations to questions – Evidence from Tucanoan. *Jezikoslovlje* 1(1–2): 35–51.

Jones, Wendell H. & Jones, Pamela Simmons. 1991. *Barasano Syntax* [Studies in the Languages of Colombia 2]. Dallas TX: The Summer Institute of Linguistics and The University of Texas at Arlington.

Kinch, Rodney A. & Kinch, Pamela G. 2000. El yurutí. In *Lenguas indígenas de Colombia: Una visíon descriptiva*, María Stella González de Pérez & María Luisa Rodríguez de Montes (eds), 469–488. Santafé de Bogotá: Caro y Cuervo.

König, Ekkehard & Siemund, Peter. 2007. Speech act distinctions in grammar. In *Language Typology and Syntactic Description*, Vol. 1: *Clause Structure*, 2nd edn, Timothy Shopen (ed.), 276–324. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619427.005

Malone, Terrell. 1988. The origin and development of Tuyuca evidentials. *International Journal of American Linguistics* 54(2): 119–140. https://doi.org/10.1086/466079

Malone, Terrell & Barnes, Janet. 2000. El tuyuca. In *Lenguas indígenas de Colombia: una visíon descriptiva*, María Stella González de Pérez & María Luisa Rodríguez de Montes (eds), 437–452. Santafé de Bogotá: Instituto Caro y Cuervo.

Matthewson, Lisa, Davis, Henry & Rullmann, Hotze. 2007. Evidentials as epistemic modals: Evidence from St'át'imcets. *Linguistic Variation Yearbook* 7(1): 201–254.

Michael, Lev D. 2012a. *Máíjìkì finite verbal inflectional paradigms*. Ms, University of California, Berkeley.

Michael, Lev D. 2012b. *Tone in Máíhĩ̃kì*. Ms, University of California, Berkeley.

Miller, Marion. 1999. *Desano Grammar* [Studies in the Languages of Colombia 6]. Dallas TX: Summer Institute of Linguistics.

Overall, Simon E. 2011. Clause-chaining and nominalisations in Aguaruna (Jivaroan). Paper presented at the SLE, Universidad de la Rioja, Logroño.

Peterson, Tyler Roy Gösta. 2010. *Epistemic modality and evidentiality in Gitksan at the semantics-pragmatics interface*. PhD dissertation, University of British Columbia.

Portner, Paul. 2009. *Modality*. Oxford: OUP.

Sadock, Jerrold M. & Zwicky, Arnold M. 1985. Speech acts distinctions in syntax. In *Language Typology and Syntactic Description*, Vol. 1: *Clause Structure, Timothy Shopen* (ed.), 155–196. Cambridge: CUP.

Schwarz, Anne. 2012. Speaker confidence and speculation. The epistemic dichotomy in the verb inflection in Secoya. Paper presented at the Friday Afternoon Lectures, Leiden University.

da Silva, Wilson. 2012. A descriptive grammar of Desano. PhD dissertation, University of Utah.

Smothermon, Jeffrey R., Smothermon, Josephine H. & Frank, Paul Stephen. 1995. *Bosquejo del Macuna*. Santafé de Bogotá: Editorial Alberto Lleras Camargo.

Stenzel, Kristine. 2013. *A Reference Grammar of Kotiria (Wanano)*. Lincoln NE: University of Nebraska Press. https://doi.org/10.2307/j.ctt1ddr99n

Sung, Li-May. 2011. Clausal nominalization in Budai Rukai. In *Nominalization in Asian Languages: Diachronic and Typological Perpsectives* [Typological Studies in Language 96], Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds), 523–559. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.96.19sun

Vallejos, Rosa. 2015. On the fuzzy bounderies between complex predicates and multiclause constructions: The case of Secoya. Paper presented at the Western Tukanoan Workshop, University of California, Berkeley.

Velie, Daniel & Velie, Virginia. 1981. *Vocabulario Orejón*. Lima: Ministerio de Educación, Instituto Lingüístico de Verano.

Waldie, Ryan J., Peterson, Tyler Roy Gösta, Rullmann, Hotze & Mackie, Sott. 2009. Evidentials as epistemic modals or speech act operators: Testing the tests. Paper presented at the Workshop on Structure and Constituency in the Languages of the Americas 13, Purdue University. <http://evidentialchicago.pbworks.com/w/file/fetch/46624592/Waldie_Peterson_Rullmann_Mackie_2009_Evidentials.pdf>

Wheeler, Alva. 1987. *Gantëya Bain, el pueble siona del río Putumayo, Colombia*, Tomo I. Lomalinda: Editorial Townsend.

Yap, Foong Ha & Wang, Jiao. 2011. From light noun to nominalizer and more: The grammaticalization of *zhe* and *suo* in Old and Middle Chinese. In *Nominalization in Asian Languages: Diachronic and Typological Perpsectives* [Typological Studies in Language 96], Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds), 61–108. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.96.02yap

CHAPTER 11

Form and functions of nominalization in Wampis

Jaime Peña
Pontificia Universidad Católica del Perú

This paper examines the main morphosyntatic and functional properties of nominalization in Wampis, a Jivaroan (or Chicham) language spoken in Peru. In addition to having derivational properties, nominalized structures are used extensively in Wampis for a variety of functions, including relativization and complementation. This paper also shows that nominalized stems are expanding their functions to those equivalent to predicating structures, such that they are able to predicate on their own and receive finite verbal morphology.

1. Introduction

This paper discusses key aspects of the morphosyntactic and functional dimensions of nominalization in Wampis, a Jivaroan (or Chicham)[1] language spoken in Peru. One of the salient morphosyntactic features of Wampis grammar is the use of nominalizations in a wide range of constructions. These include structures in relativization, complementation and adverbial functions, as well an in copy-verb constructions, chained constructions and other complex constructions which function at the level of the TAM system. Within the frame of this paper, nominalization is understood as a metonymic process that derives a noun-like element which then functions as nominals do in the language (see Shibatani this volume). Following a preliminary discussion of relevant typological and grammatical aspects of Wampis in § 2, § 3 introduces the Wampis nominalizers to be analyzed in this work. Next, § 4 describes the lexical derivational properties of Wampis nominalizations. The main section of this paper is § 5, which contains a description and analysis of several Wampis constructions where nominalizations are used, as well as their functions. In § 6, the development of nominalizations into finite structures that can predicate on their own is described. Finally, § 7 offers brief conclusions

1. The term *Chicham* 'speech, word, language' has been proposed by Katan Jua (2011).

https://doi.org/10.1075/tsl.124.11pen

and comments on possible implications for the study of nominalization arising from the present analysis.

2. Background and grammatical sketch

Wampis is an endangered indigenous language spoken between the eastern foothills of the Andes and the lowland Amazon forest in northeastern Peru. There are approximately 10,000 speakers of Wampis, most of whom live along the Santiago and Morona Rivers, in the Peruvian departments of Amazonas and Loreto. Wampis has remained largely undescribed and underdocumented until very recently (Peña 2015).

Wampis is nominative-accusative, both in head-marking and in dependent-marking. The language is polysynthetic and agglutinating, with some degree of fusion. A preference for SV/APV order[2] is observed, although some specific constructions favor post-verbal subjects and objects. Adjectives and relative clauses generally follow the head noun. Wampis distinguishes between subject and non-subject; that is, all notional objects (direct, indirect, object of applicative) are marked with *=na* or its high tone (ˊ) allophone. Wampis nominal and verbal morphology is very complex, particularly the verbal morphology, which has a large inventory of suffixes and clitics that mark valence-change, aspect, tense and modality, among other categories. There are pervasive processes of vowel elision and suprasegmental features (nasality and tone) that have grammatical values. Like other languages of the family, Wampis has a sophisticated switch-reference system that distinguishes same from different subjects and indexes grammatical person.

2.1 Parts of speech

There are two major open classes in Wampis: nouns and verbs. The distinction between simple nouns and verbs is nearly categorial. Except for a few examples of zero derivation, in general noun roots cannot be used as predicates without a verbalizer, and verb roots cannot function as nominals without being nominalized first. In addition to nouns and verbs, there are around 30 underived roots that constitute the class of adjectives, a few of them being trans-categorial words. Adjectives in Wampis constitute a semi-closed class, as creation of new members is limited; instead, modifying expressions are achieved via nominalizations wherein

2. I use Comrie's (1978) notation for core arguments: single argument of an intransitive construction (S), most agent-like argument of a transitive construction (A), and most patient-like argument (P) of a transitive construction.

the derived noun assumes an attributive function. In fact, adjectival modification of nouns is infrequent (adjectives are mostly used in predicative constructions with a copula); instead, speakers typically make use of nominalizations to modify nouns, often in very elaborate ways.

There are a number of other semi-closed and closed word classes in Wampis: pronouns, quantifiers, demonstratives, interrogative words, adverbs, interjections, particles, and ideophones. Peña (2015: 329 ff.) provides a thorough analysis of closed and semi-closed word classes in Wampis, including classes composed of unique items. I briefly discuss the most important features of adjectives, nouns and verbs in Wampis, as they are relevant for this paper.

2.1.1 *Adjectives*

Morphosyntactically, adjectives in Wampis act as heads of adjective phrases, are gradable, and are the only word class that can take the derivational deintensifier suffix *-taku* 'Partly X', where X is an adjective root. Adjectives are readily distinguished from verbs in that they do not take any of the morphological markers for categories associated with verbs (see § 2.1.3). Unlike nouns, adjectives do not head NPs and cannot take case marking as independent constituents, although they can host the object marker and adpositional clitics when the adjective is the last element of a NP.

2.1.2 *Nouns and noun phrases*

Morphosyntactically, nouns in Wampis are heads of NPs, can be possessed, and can be possessors. Grammatical categories uniquely associated with nouns are: possession (i.e., marking of a possessed noun), attribution with the suffix *-tinu* 'Attributive'[3] and the nominative, genitive and vocative cases. In addition, a number of information structure clitics often occur with nouns: *=ka* 'Focus', *=kɨ* 'Restrictive', *=ʃa* 'Additive'. The object marker *=na* frequently occurs attached to the last element of the NP, and oblique relations are marked with postpositional clitics, including typical categories such as the locatives *=nVma* and *=(n)ĩ*,[4] ablative *=ia*, allative *=ni*, and comitative *=haĩ*. There is no morphological number-marking for nouns; rather a copula marked for plural number is used when speakers feel the need to be explicit about plurality. Demonstratives trigger a type of case-marking agreement among certain elements of the noun phrase, including the demonstrative itself, the head noun and its attributive modifiers. The morphemes that follow

3. The attributive is not to be confused with the future nominalizer *-tinu*, which is described later in this paper.

4. There is a third locative that is marked with a high pitch on the last surface vowel of a subset of nouns.

this pattern of agreement are the object marker *=na*, the locatives *=numa* and *=(n) ĩ*, the focus *=ka* and the additive *=ʃa.* Example (1) illustrates this pattern of agreement with *=na* – if no demonstrative is present in the NP, no agreement between head and modifier is established, as in (2). The preferred order for attributive modification of the noun is post-nominal, as can be observed in (1)–(2).[5]

(1) *nu=na kutʃi=na tsakari=na waina-ka-ma-ha-i*
NON.VIS=OBJ knife=OBJ sharp=OBJ see-INTENS-REC.PST-1SG-DECL
'I found that sharp knife.'

(2) *kutʃi tsakari=na waina-ka-ma-ha-i*
knife sharp=OBJ see-INTENS-REC.PST-1SG-DECL
'I found the sharp knife.'

Though not very common in natural texts, modification of a noun by another noun is possible, as shown by (3).

(3) *auhumatu-sa-tata-ha-i ʃuara hɨmpɨ=na=ka*
tell-ATT-DEF.FUT-1SG-DECL person hummingbird=OBJ=FOC
'I'm gonna tell (the story of) the hummingbird person.'

Relativization follows the typical Head-Modifier structure previously observed in (1)–(2). There are two relativization strategies in Wampis: one makes use of nominalizers and is described in § 5.2, the other makes use of a demonstrative or the intensifier *ima* encliticized to a copula, as in (4).

(4) *nu ʃuara* [*irinku a=nu*] *haka-ma-ji*
NON.VIS person gringo COP=NON.VIS die-REC.PT-3.PT + DECL
'That person that was a *gringo* died.'

In this paper, I will analyze relativization accomplished with nominalizers; for more details about relativization with a demonstrative or intensifier plus copula, see Peña (2015: 851–853).

5. The transcription used in examples represents underlying phonemic forms using IPA symbols. In addition, some particular conventions are used. A plus "+" symbol is used for morphemes whose first or last vowel, depending on the specific context, is lost to the vowel of a contiguous morpheme due to a morphophonological process where moras are reduced. For instance, a stem like *puha* (from *puhu* 'live', *-a* 'Imperfective') is glossed as **live + ipfv**. A backslash symbol "\" is used to introduce a morpheme that is instantiated via nasalization, high tone (i.e. a high pitch) or apophony. For instance, *utʃîrî* (from *utʃi* 'child', *rĩ* '3.poss' and ´ 'Genitive') is glossed as **child-3.poss\gen**. Note that in some instances the phonetic realization of words may be very different from the underlying forms due to extensive processes of vowel elision.

2.1.3 *Verbs*

Morphosyntactically, a finite verb in Wampis is defined by the marking of the following categories: aspect, tense, person and mood/modality. Structurally, the verb is the most complex element in a Wampis clause, in that it encodes numerous grammatical categories. There are several morphological positions in the verb piece; a simplified version for the purposes of this paper is shown in Figure 1. The underlined numbers represent positions whose categories are obligatory for a verb to be finite.

Figure 1. Morphological positions in the Wampis verb

−1	Causative *V-*
0	Root
1	a. Valence-decreasing suffixes b. Valence-increasing suffixes c. Applicatives *-tu*, *-ru**
2	Object suffixes
3	a. Aktionsart suffixes b. Imperfective *-a*, plural imperfective *-ina* c. Durative *-ma* d. Present habitual *-na* e. Potential *-mai*
4	Negative *-tʃa*, *-tsu*
5	Non-imperfective plural *-ara*
6	a. Tense b. Desiderative *-tah*, Imperative *-tá*, Jussive *-ti*, Apprehensive *-ai*, hortative *-mi*
7	Person (mostly Subject)
8	Declarative *-i*, Exclamative ∅
+	Narrative *timaji* and mood/modality clitics Mirative = *hama*, Interrogative = *ka*, Sudden realization/Tag question = *api***

* In Wampis, the applicative forms *-ru* and *-tu* are not only valence-increasing devices, but they can also rearrange the argument structure of the clause.

** The narrative marker *timaji* occurs as its own phonological word, it is not a suffix. The clitics = *hama* 'Mirative', =*ka* 'Interrogative' and = *api* 'Sudden realization/Tag question' may occur with other word classes. However, when these morphemes mark the verb, no other mood/modality marker occurs.

The morphological position 3 shown in Figure 1 is shared by a group of suffixes that enter in a system of oppositions: aktionsart, imperfectives, durative, present habitual and potential. The occurrence of one or another suffix corresponding to any of these categories is conditioned morphosyntactically. Table 1 lists all of these suffixes, including aktionsart.

Table 1. Wampis suffixes that occupy verbal morphological position 3 (expanded)

Suffix	Gloss
1. *-á(w)*	'High affectedness (of Patient or Location)'
2. *-i*	'Low affectedness (of Patient or Location)'
3. *-ka*	'Intensive (action of Agent)'
4. *-ki*	'Do while moving away'
5. *-ra*	'Distributed action'
6. *-sa*	'Attenuative'
7. *-ri*	'Do in proximity/while coming'
8. *-u*	'Do at a distance'
9. *-a*	'Imperfective'/*-ina* 'Plural Imperfective'
10. *-ma*	'Durative'
11. *-na*	'Present habitual'
12. *-mai*	'Potential'

For suffixes 1–8 in Table 1, Peña (2015) adopts the name *aktionsart* as a term of convenience following Overall's (2007) analysis of a similar set that exists in the related language Awajun. Most verbs select a preferred aktionsart suffix with which they occur in specific morphosyntactic contexts; for the purpose of this paper it is important to mention that this aktionsart stem is used, among other things, in perfective contexts. However, aktionsart suffixes do not occur in imperfective, durative, present habitual or potential contexts. This information is important because it will be shown that Wampis nominalizers differ in whether they can attach to a verbal root or to specific types of verb stems. There are five types of verb stems in Wampis that are formed with suffixes 1–12 in Table 1. Table 2 lists these stems.

Table 2. Wampis verb stems

Type of stem	Formed with	Morphosyntactic context
Aktionsart	Aktionsart suffixes	Perfective and most past tenses, future tenses, imperative, jussive, hortative, apprehensive, prohibitive
Imperfective	*-a* (SG), *-ina* (PL)	Imperfective, present tense
Durative	*-ma*	Durative, imperative
Present Habitual	*-na*	Present habitual
Potential	*-mai*	Potential

Examples (5)–(6) with the root 'eat' illustrate different morphosyntactic environments with aktionsart, aspectual and potential verb stems. The verb 'eat' selects the high affectedness aktionsart suffix *-á* as its preferred aktionsart suffix and it appears with it in perfective contexts, such as in (5). In imperfective (6) and potential (7) environments the aktionsart suffix does not occur.[6]

(5) *ju-á-ha-i*
eat-HIAF-1SG-DECL
'I just finished eating.'

(6) *ju-a-ha-i*
eat-IPFV-1SG-DECL
'I am eating.'

(7) *ju-mai-tsu-ha-i*
eat-POT-NEG-1SG-DECL
'I cannot eat it.'

2.2 Finiteness and structure of the clause

Although there is a long tradition of studies on finiteness,[7] there is not one definition of this category agreed upon in the broader literature (Cristofaro 2007). Thus, I define finiteness for Wampis using the language's own particular terms (i.e., using relevant categories found in Wampis). I follow Givon's (2002: 25) view of finiteness as a scalar clausal phenomenon, having a prototypical transitive clause as a reference point. A prototypical transitive clause in Wampis is defined as a clause with a main declarative verb and two arguments instantiated respectively as the grammatical subject and object of the clause. I briefly discuss the following categories traditionally associated with finiteness and their application for the Wampis prototypical transitive clause:

- Tense, Aspect, Mood/Modality
- Argument indexation
- Case assignment and word order
- Clause structure

6. In Examples (5) and (6), the high affectedness aktionsart suffix *-á* and the imperfective *-a* are distinguished by the ability of the aktionsart suffix to attract a high tone.

7. Nikolaeva (2007) constitutes a good introduction to the various ways finiteness has been analyzed within different linguistic traditions.

With regard to TAM categories, all simple verb forms in Wampis that occur in an independent clause obligatorily receive marking for aspect, tense, person and mood/modality. Dependent verb forms are never marked for tense[8] or mood/modality, but they can be marked for aspect and person (only in switch-reference constructions). In addition, dependent verbs are marked with a nominalizing suffix or a switch-reference suffix. For instance, in (8) the main verb 'cook' is marked for the categories of aspect (with the intensive action aktionsart *-ka*), tense, person (which is also obligatory, see discussion below) and declarative mood. The subordinate verb 'arrive' is not marked for tense or mood. In addition, notice the suffix *-nu* in the subordinate verb, which belongs to a paradigm of switch-reference markers and which marks the verb as dependent. In (9), the main verb 'stop doing' is marked for the same categories as the main verb in (8) (notice that the aktionsart selected by the verb in this case is *-sa* 'attenuative'); the subordinate verb 'work' is nominalized with the nominalizer *-ta*, which creates an action nominalization.

(8) Dependent Main
hĩa=nama ta-á-nu *inara-ka-ma-ha-i*
house=LOC arrive-HIAF-1SG.SS cook-INTENS-REC.PST-1SG-DECL
'Having arrived at the house, I cooked.'

(9) Dependent Main
hĩa=nama taka-ta=na *inai-sa-ma-ha-i*
house=LOC work-NMLR=OBJ stop.doing-ATT-REC.PST-1SG-DECL
'I stopped working in the house.'

Person marking on the verb follows a rather complicated pattern of hierarchical argument indexation in Wampis. For the sake of brevity, it can be stated that there are two positions in the verb that are reserved for the marking of person: with regard to Figure 1, position 2 is reserved for object marking (third-person objects are unmarked) and position 7 is reserved mostly for subject marking.[9] The sentence in (10) illustrates the marking of object and subject in these two different verbal slots. However, when a third person acts upon a 1PL or 2PL person, the verb agrees with the object (the 1PL or 2PL person), and not with the subject, in position 7, as shown in (11). In the examples below, the relevant morphemes and their glosses appeared in boldface.

8. With the exception of the future nominalizer *-tinu*, which, as the name indicates, provides a future temporal grounding to the nominalized form. See discussion in § 3.

9. This is a simplified description of argument indexation in Wampis that is valid for the goals of the present work. See Peña (2015: 656 ff.) for a more detailed analysis of argument indexation in Wampis.

(10) *amu-**tu**-ka-mia-**ji***
finish-**1SG.OBJ**-INTENS-DIST.PST-**3.PST + DECL**
'He killed me.'

(11) *amu-**tama**-ka-mia-**hi***
finish-**1PL.OBJ**-INTENS-DIST.PST-**1PL + DECL**
'He/They killed us.'

With regard to case and word order, Wampis exhibits a nominative-accusative profile and the preferred order in the clause is SV/APV. Overt subjects are always possible; however, subordinate clauses forming chains frequently omit their subjects – co-referential arguments are generally omitted. The nominative case is zero, and objects are marked with = *na,* as shown in (12). However, if the object is a third person and the subject is first person plural or second person, the object is not marked with = *na*, as in (13).

(12) *wii jawaã=na mã-á-ma-ha-i*
1SG jaguar=OBJ kill-HIAF-REC.PST-1SG-DECL
'I killed the jaguar.'

(13) *ii jawaã mã-á-ma-hi*
1PL jaguar kill-HIAF-REC.PST-1PL + DECL
'We killed the jaguar.'

With regard to clause structure, clauses can be verbal and verbless. Verbless clauses are possible but not frequent, and are limited to specific juxtaposition constructions conveying the functions of equation, proper inclusion and attribution. The order of juxtaposition constructions is strictly [NP NP/Adj], where the first element is the subject and the second element is the predicate. Importantly, these verbless constructions can always be turned into a verbal clause by adding a copula. Examples of semantically equivalent verbless and verbal constructions (with a copula) are given in (14)–(15).

(14) *ʃuara pɨnkɨra*
person good
'He/She is a good person.'

(15) *ʃuara pɨnkɨr=aiti*
person good=COP.3 + DECL
'He/She is a good person.'

3. Overview of Wampis nominalizations

The Wampis nominalizers to be analyzed in this paper belong to two sets. Both Set I and Set II nominalizers are used for lexical and grammatical nominalizations. Lexical nominalization (§ 4) "creates new lexical items belonging to the noun class of the language", whereas grammatical nominalization (§ 5) "creates new referring expressions that have no lexical status" (Shibatani 2009: 187). Wampis nominalizers are presented in Table 3.

Table 3. Nominalizing suffixes in Wampis

Set I	Set II
-inu 'Agentive S/A nominalizer'	*-u* 'Subject nominalizer'
-tinu 'Future nominalizer'	*-mau* 'Non-subject nominalizer'
-taĩ 'Non-Agentive nominalizer'	
-ta 'Action nominalizer'	

Wampis nominalizations may retain much of the verbal structure, being able to attach to a verbal stem marked for any grammatical category (or categories) within positions −1 to 5 in Figure 1. The main distinction between Set I and Set II nominalizers is that Set I nominalizers (with the exception of *-tinu*) do not attach to verbal stems with position 3 filled, while Set II nominalizers can attach to verbal stems with position 3 filled (cf. Table 2).

The main morphosyntactic properties of nominalizations derived with Set I and Set II nominalizers are presented in Table 4. Table 4 also includes properties of nouns and adjectives so that nominalization properties can be compared within the language's nominal-adjectival continuum.

Table 4. Morphosyntactic properties of nouns, nominalizations and adjectives

Properties	Noun	Set I NMLZ	SetII NMLZ	Adjective
Receives case and oblique clitics	Yes	Yes	Yes	not on its own**
Can be possessor	Yes	Yes	Yes	No
Can be possessed	Yes	Yes*	Yes*	sometimes†
Can head an NP	Yes	Yes	Yes	No
Can be pluralized adding *a-ina* (COP-PL.IPFV)	Yes	Yes	Yes	No

* Nominalizations with *-inu* and *-u* cannot be possessed.

** Demonstratives trigger case agreement in certain elements of a noun phrase, including adjectives.

† Adjectives that are possessed function as nominals, e.g. *ʃiirama* 'beautiful', when possessed with *-rĩ* '3.poss', means 'her/his beauty'.

At first glance, from the semantics and properties seen above, there seems to be some redundancy in the nature of these nominalizing morphemes. However, they present particular morphosyntactic differences. Morphologically, as pointed out above, Set I and Set II differ in one important property: Set I – with the exception of the future nominalizer *-tinu* – can attach to roots, but do not attach to aktionsart or imperfective stems, whereas set II morphemes attach to aktionsart and imperfective stems. In addition, the agentive nominalizer combines with the potential stem marked by *-mai*; whereas the Set II subject nominalizer *-u* and Set I future nominalizer *-tinu* and non-agentive nominalizer *-taĩ* can combine with the verbal negative suffix *-tʃa*. Table 5 lists the combining properties of the individual nominalizers of Wampis.

Table 5. Combining features of Wampis nominalizers

Properties/ Suffixes	Can attach to bare root	Can attach to a verbal stem (after verbal slot 3 in Figure 1)	Can attach to potential stem *V-mai*	Can attach to negative stem *V-tʃa*
-inu	Yes	No	Yes	No
-tinu	Yes	Yes	No	Yes
-taĩ	Yes	No	No	Yes
-ta	Yes	No	No	No
-u	No	Yes	No	Yes
-mau	No	Yes	No	Yes

4. Lexical nominalization

This section presents the Wampis nominalizers in their lexical derivation function.

4.1 Set I Agentive S/A nominalizer *-inu*

The agentive S/A nominalizer *-inu* creates a noun that refers to the subject of a verb. Semantically, *-inu* denotes an animate referent who is prototypically a human "doer". The examples in (16) show instances of nouns derived with *-inu*.

(16) *iwiʃi* 'heal' → *iwiʃi-inu* 'shaman'
kutama 'spin' → *kutama-inu* 'spinner'
iwara 'make joke' → *iwara-inu* 'joker'
hintina 'teach' → *hintina-karata-inu* 'teacher' (teach-1PL.OBJ-NMLR)
aʃama 'be afraid' → *aʃama-inu* 'coward'

4.2 Set I Future nominalizer *-tinu*

As the name indicates, the future nominalizer creates a noun that refers to participants of events that have not occurred at the time of the speech act. Historically, *-tinu* is a conflation of the suffix *-ta* 'Immediate future' plus the Agentive S/A nominalizer *-inu*. However, *-tinu* is not exclusively an S/A nominalizer, as it creates nouns referring to the subject or object of the verb:

(17) *umu* 'drink' → *uma-ra-tinu*
drink.PFV-DISTR-FUT.NMLR
'one who will drink' or 'what will be drunk (i.e. a drink)'
ju 'eat' → *ju-á-tinu*
eat-HIAF-FUT.NMLR
'one who will eat' or 'what will be eaten'

The future nominalizer *-tinu* can also derive an abstract action/stative noun that sometimes serves as a citation form of the verb. However, *-tinu* is not as commonly employed for this use as the action nominalizer *-ta* (§ 4.3). The use of *-tinu* for citation forms probably derives from an eventive reading of object nominalizations achieved with this morpheme, as explained above. Since subjecthood is no longer part of the meaning of *-tinu* (unlike *-inu*), it is possible that *-tinu* has extended its uses to events or states.

(18) *iya* 'fall' → *iya-tinu* 'to fall'
wɨ 'go' → *wɨ-tinu* 'to go'

4.3 Set I Action nominalizer *-ta*

The action nominalizer *-ta* derives an abstract noun that denotes an event or state expressed by the verb. A *-ta* nominalization is often translated as an infinitive, thus many speakers and previous works, e.g. Jakway et al. (1987), use *-ta* nominalizations as citation forms for Wampis verbs. The following examples illustrate the derivational use of *-ta*.

(19) *puhu* 'live' → *puhu-ta* 'life'
mɨsɨ 'ruin, die' → *mɨsɨ-ta* 'death, war'
akiina 'be born' → *akiina-ta* 'birth'
kahɨra 'hate' → *kahɨra-nai-ta* 'to become enemies'
hate-RECP-NMLR

4.4 Set I Non-agentive nominalizer *-taĩ*

In contrast with the agentive nominalizer *-inu*, the nominalizer *-taĩ* creates a noun that denotes a patient of a transitive verb, a location or an instrument. The following examples illustrate derivations that denote patients of transitive verbs ('drink', 'food'), a location ('resting place') and an instrument ('oar').

(20) *umu* 'drink' (V) → *umu-taĩ* 'drink' (N)
ju 'eat' → *ju-taĩ* 'food'
ajama 'rest' → *ajama-taĩ* 'resting place'
anunta 'row' → *anun-taĩ*[10] 'oar'

4.5 Set II Subject nominalizer *-u*

The subject nominalizer *-u* derives nouns that encode a variety of semantic roles: agent, experiencer, undergoer, force – all of which map onto S/A arguments of the verb. Unlike nominalizations with *-inu*, which prototypically denote a human participant, nominalizations with *-u* can be animate or inanimate. As mentioned previously, Set II nominalizations differ from Set I nominalization in that Set II nominalizations are formed from verbal stems that possess aspectual information. The following examples illustrate instances of the use of *-u* with imperfective (21) and aktionsart stems (22).

(21) *ɨsa* 'burn' → *ɨsa-u*
burn+IPFV-NMLR
'one/thing that burns'
kana 'sleep' → *kana-u*
sleep+IPFV-NMLR
'sleeper; i.e. one who is sleeping'
tɨpɨ 'lie down' → *tɨpa-u*
lie.down+IPFV-NMLR
'one who lies down'
mina 'melt' → *mina-u*
(a candle) melt+IPFV-NMLR
'thing that melts down'

10. In this word, the entire last syllable of the verb *anunta* is deleted because of a word-internal elision process.

(22) *ɨsa* 'burn' → *ɨsa-ka-u*
burn-INTENS-NMLR
'one/thing that burned'
kana 'sleep' → *kana-ka-u*
sleep-INTENS-NMLR
'one who slept'
tɨpɨ 'lie down' → *tɨpɨ-sa-u*
lie.down-ATT-NMLR
'one who lied down'
mina 'melt' → *mina-ra-u*
melt-DISTR-NMLR (e.g. wax from melted candle)
'melted thing'

4.6 Set II Non-subject nominalizer *-mau*

The suffix *-mau* creates a noun that denotes the patient of a transitive verb, a location of an intransitive verb, or an action/event. The most common derivations achieved with *-mau* are nouns referring to the patient of a transitive verb; but nouns referring to a location associated with the verb is also quite possible, e.g. *akiina-mau* (be.born+IPFV-NMLR) means 'birth' or 'place where one is born'; *arakama-mau* (sow. + IPFV-NMLR) means 'what is planted (i.e. seeds)' or 'place where one sows seeds (i.e. sown field)'.[11] The nominalizer *-mau* can also create an action nominalization, cf. *akiina-mau* (be.born+IPFV-NMLR) 'birth' above. Like Set II subject nominalizer *-u*, *-mau* attaches to a verbal stem that includes aspectual (imperfective/perfective) information. The following examples illustrate instances of the use of *-mau* with imperfective and aktionsart stems.

(23) *ihi* 'wrap' → *ihi-a-mau* 'package'
wrap-IPFV-NMLR
ɨsa 'burn (V)' → *ɨsa-mau* 'burn (N)'
burn-IPFV-NMLR

(24) *atʃi* 'capture' → *atʃi-ka-mau* 'prisoner'
grab-INTENS-NMLR
ɨa 'swell' → *ɨa-ra-mau* 'swelling (of injury)'
swell-DISTR-NMLR

11. However, a *-mau* nominalization occurs with a locative postposition = *nVma* when used for adverbial subordination; i.e. for adverbial subordination *-mau* creates event nominalizations (cf. § 5.4.1).

Historically, *-mau* is composed of an old nominalizer *-ma* and Set II nominalizer *-u* (Peña 2015: 838).

5. Grammatical nominalization: Structure and functions

In what follows, I discuss the structure of grammatical nominalizations (§ 5.1), and their varied functions in Wampis: relativization (§ 5.2), argument realization (§ 5.3), adverbial function (§ 5.4), as well as their uses in copy-verb constructions (§ 5.5) and in constructions that function at the level of the TAM system (§ 5.6).

5.1 Internal and external structure

Structurally, grammatical nominalizations share a number of characteristics with NPs while maintaining some clausal properties. In terms of their external syntax, grammatical nominalizations are treated as nominals. The most clear evidence of this is the fact that grammatical nominalizations, in addition to being able to head NPs, receive case marking when they function as an argument of a main verb, as in (25).

(25) [*mina-u*]$_{\text{NMLZ}}$ *=na=ʃa* *neka-hak-u=iti*
come+IPFV-NMLR = OBJ = ADD know-HAB.PST-NMLR = COP.3 + DECL
'He used to know the one who was coming too.'

In terms of their internal structure, a nominalized verb may retain several morphological categories with clausal information, depending on whether the nominalizer belongs to Set I or Set II (see § 3). These verbal categories may include valence, argument indexation and aspectual information. In addition, constituent order in main clauses is typically verb final: grammatical nominalizations show the same pattern, with the nominalized verb in final position.

On the other hand, there are some important differences between grammatical nominalizations and main clauses. Verbal morphology such as mood/modality, tense (with the exception of nominalizations with *-tinu* 'Future nominalizer') and subject markers, for instance, are restricted to main clauses. Another important difference between grammatical nominalizations and main finite clauses is the marking of NP arguments. Case-marking of an argument of a nominalized verb is optional – in fact, in the majority of examples from a textual database (calculated on a subset of about 1,200 clauses), the object argument of a nominalized verb is not marked for case. In contrast, the NP argument of a finite verb generally receives object marking, except for those cases of hierarchical argument marking

described in § 2.2. An example of the non-marking of the object of a nominalized verb can be seen in (26), where 'fish' does not receive the object marker = *na*.

(26) *Nayapí puhutīn=ka* [*namak ihu-ra-mau*]$_{\text{NMLZ}}$=*a*
Nayap.GEN life.3.POSS=FOC fish stab-DISTR-NMLR=COP
'As for Nayap's life, he fished.' (Literally: 'As for Nayap's life, what (he) stabbed (was) fish.'

One important point of analysis in Wampis grammatical nominalizations has to do with the instantiation of the targeted argument of the nominalization. There are different ways in which an argument is expressed in Wampis grammatical nominalizations. The argument may be manifested through a gap with the overt argument itself being external to the nominalization (traditionally considered to be "externally-headed"), the argument may not be overtly expressed at all ("headless"), or the argument may be overtly expressed within the nominalization ("internally-headed"). For instance, a nominalization with Set II *-u* denoting an A/S argument uses a gapping strategy, with some exceptions discussed below. That the coreferent argument is external or internal in this type of nominalization may not be apparent upon initial inspection, because the subject is zero-marked in Wampis, and object NPs are typically marked on the last element of the NP with =*na*. Therefore, given the order [N Modifier], it is not always clear whether the common argument is internal or not to the nominalization. The best syntactic test to determine whether the coreferent argument is external or internal to the nominalization is via the addition of a demonstrative in the NP. As seen in § 2.1.2, demonstratives trigger agreement-marking among demonstratives, head noun and attributive modifiers. Thus, with an object NP (marked with =*na*), the following structure is predicted to occur when a demonstrative and a nominalization are present in the NP, and an example is given in (27).

[Dem=*na* N_{head} =*na* [V-*NMLR*]=*na*]

(27) [*nu=na* *apuupu=na*$_j$ [$\varnothing_j$ *jukuma-u*]$_{\text{NMLZ}}$ =*na*]
NON.VIS=OBJ dolphin=OBJ swim+IPFV-NMLR=OBJ
waina-ka-mia-ha-i
see-INTENS-DIST.PST-1SG-DECL
'I saw that dolphin that was swimming.'

In (27), the referent of the grammatical nominalization is external, otherwise the head noun *apuupu* 'dolphin' would not be marked with = *na*; rather, it would be unmarked (recall that the nominative is zero in Wampis) as it is the coreferent subject of the nominalized verb.

In other instances, the targeted argument does not occur, or it occurs internal to the nominalizaton structure. That is, we have grammatical nominalizations

equivalent to what traditionally are called "headless" and "internally-headed" relative clauses, such as in (28) and (29), respectively.

(28) [*kisara=numa kaʃi eakma-u*]$_{\text{NMLZ}}$ *kuntina=na waina-wa-i*
cliff=LOC night look.for+IPFV-NMLR animal=OBJ see+IPFV-3-DECL
'(The one) who searches (for game) at the cliff at night finds animals.'

(29) [*wii arutama aniaku-a-mau*]$_{\text{NMLZ}}$ *= ka kamɨ a-wa-i*
1SG spirit.power remember-IPFV-NMLR = FOC INTERJ exist-3-DECL
muuka = na
head = OBJ
'*Arutam* that I remember, well, there is Head-*Arutam*.'[12]

Within the frame of this paper, grammatical nominalizations equivalent to headless and internally-headed relative clauses are considered to be NP uses of nominalizations (Shibatani 2009; this volume). In other words, these types of nominalizations fulfill the function of a referring expression and are typically instantiated as arguments (§ 5.3).

5.2 Relativization

Traditionally, relativization is seen as involving a subordinate clause (a relative clause) that functions as a modifier of a head noun (Keenan 1985). In Wampis, this modifying function is accomplished via nominalization – the close relationship between relativization and nominalization is cross-linguistically well-attested (DeLancey 1986; DeLancey 1999; Genetti et al. 2008; Shibatani 2009; Shibatani this volume; Givón 2012). There is non-future and future relativization in Wampis.

5.2.1 *Non-future relativization*

All nominalizers can be used in Wampis non-future relativization, but the most commonly used ones are set II nominalizers *-u* and *-mau*. Either A/S (30) or non A/S (31)–(32) arguments can be relativized.

(30) [*ʃuara* [*mina-u*]$_{\text{NMLZ}}$] *ami=na neka-wa-i*
person come+IPFV-NMLR 2SG=OBJ know+IPFV-3SG + DECL
'The person that is coming knows you.'

(31) [*jurumaka* [shiir *inara-ka-mau*]$_{\text{NMLZ}}$] *waina-ka-hi*
cooked.manioc well cook-INTENS-NMLR see-INTENS-1PL + DECL
'We have found the manioc that has been well cooked.'

12. *Muuk Arutam* 'Head-Arutam' in the Wampis worldview is a monstrous head with human and feline features.

(32) [*najapí nuku-rĩ puha-mau*]$_{\text{NMLZ}}$ = *nama hĩã̄*
Nayap\GEN mother-3.POSS live+IPFV-NMLR=LOC house\LOC
puhu-sa-ara
live/be-ATT-3PL.SS
'[They] Having been in the house where Nayap's mother lived…' (Literally: 'where Nayap's mother lived, in the house, they having been…')

Locative relativizations employ Set II non-subject nominalizer *-mau* plus the locative *=nVma* ([V-*mau* =*nVma*]), as in (32) above.

5.2.2 *Future relativization*

Relativization with a future reference in Wampis is done with Set I future nominalizer *-tinu*. As we saw previously, *-tinu* may refer to an A/S (an example with A in (33)) or P argument (34) of the nominalized verb.

(33) [*nu ʃuara* [*ju-á-tinu*]$_{\text{NMLZ}}$] *nunká*
NON.VIS person eat-HIAF-FUT.NMLR ground\LOC
ɨkɨma-sa-ara-ma-ji
sit-ATT-PL-REC.PST-3.PST + DECL
'Those people who were going to eat sat on the ground.'

(34) *nĩn=ka arakama-mia-ji* [*tikitʃi numi* [*hatɨ-á-tinu*]$_{\text{NMLZ}}$] = *na*
3SG=FOC plant-DIST.PST-3.PST + DECL other tree cut-HIAF-FUT.NMLR=OBJ
'She planted other trees that are going to be cut.'

Interestingly, all examples of oblique relativization with *-tinu* occur in a "double" relativization structure: the *-tinu* nominalization is postposed by a copula, which is in turn relativized with an encliticized demonstrative, as shown in (35).

(35) [*utʃi* [*ta-á-tinu*]$_{\text{NMLZ}}$ *a=nu=haĩ*] *tʃĩtʃa-sa-tinu*
child arrive-HIAF-FUT.NMLR COP=NON.VIS=COM speak-ATT-FUT.NMLR
a-ji
COP-3.PST + DECL
'She/he was going to speak with the child that was going to come.'

5.3 Nominalizations as arguments

Cross-linguistically, grammatical nominalizations can function as arguments of verbs and nominal predicates. Traditionally, grammatical nominalizations in the role of an argument in the clause have been analyzed as headless relative clauses, but as Shibatani points out: "Grammatically nominalized forms are referring expressions by themselves without any head nominal or pronoun, either abstract or concrete, and they fill the syntactic noun function as arguments or predicate

nominals" (2009: 192). The following examples illustrate grammatical nominalizations of the type otherwise known as headless relatives in different argument positions as well as in nominal predicates.

(36) *As subject*
[nĩi uha-tu-ka-mau]$_{\text{NMLZ}}$ *nuĩ nankani-a-wa-i*
3SG inform-1SG.OBJ-INTENS-NMLR there finish-IPFV-3SG-DECL
'What she informed me finishes there.'

(37) *As object*
[intaʃi-rĩ tsupi-ka-u]$_{\text{NMLZ}}$ *= na awatu-i-ma-ji*
hair-3.POSS cut-INTENS-NMLR = OBJ hit-LOAF-REC.PST-3.PST + DECL
'She hit the one who cut her hair.'

(38) *As complement of copula (notice the switch-reference within the nominalization)*
nu = aiti *[auhumatu = hak-mau tsamarain = na*
NON.VIS = COP.3 + DECL tell-HAB.PST-NMLR Tsamarain = OBJ
patʃi-sã]$_{\text{NMLZ}}$
mention-SUB\3.SS
'That is what used to be told about Tsamarain.' (Literally: 'This is what used to be told mentioning Tsamarain).

(39) *As an oblique*
[puerto galilea = ʃa a-tinu]$_{\text{NMLZ}}$ *= numa*
Puerto Galilea = ADD COP-FUT.NMLR = LOC
ta-á-mia-ji
arrive-HIAF-DIST.PST-3.PST + DECL
'He arrived at the place that would later become Puerto Galilea.'

5.3.1 *Complementation*

While there is a long tradition that views a complement clause as one that functions as an argument of another clause (Noonan 1985), in Wampis, there is no unique construction or set of related constructions with the internal syntax of a clause used only for complementation. Following Dixon (2006), it is better to analyze Wampis as a language that employs various strategies to achieve the function of complementation. The main strategies are:

- Nominalization
- Switch-reference
- Quotative construction
- Switch-reference + Quotative construction
- Desiderative construction

Which strategy is chosen seems to be fairly lexical, though the semantics of the matrix verb may also play a role. More than one strategy may be available for the same verb, but usually the number of available strategies is limited. For Wampis, it is also useful to distinguish whether the subject of the matrix verb and the complementation structure are co-referential or not. Table 6 provides a sample of common complement-taking verbs with their respective attested strategies.

Table 6. Sample of complement-taking verbs and their complementation strategies*

Verb	Gloss	Strategy used	
		Co-referent subjects	**Non-coreferent subjects**
nɨka	know	Set I *-ta*	Set II *-u, -mau*
ii	think	Q + SR	Set II *-u*
ii	see	–	Set II *-u*, *-mau*
antu	hear	–	Set II *-mau*; SR
waina	see	–	Set II *-u*, *-mau*
wakɨru	want	DES	SR
nɨkapɨ	feel	Set II *-u*	Set II *-u*, *-mau*
nankama	begin	Set I *-ta*	–
amu	finish	Set I *-ta*	–
umi	complete	Set I *-ta*	–
tu	say	Q; SR + Q	Q; SR + Q
inii	ask	–	Q

* Q = Quotative construction, SR = Switch-reference constructions, DES = Desiderative constructions

The remainder of this section is dedicated to exemplifying different complementation strategies with selected verbs from Table 6. I will focus here on strategies involving nominalizations, and discuss switch-reference and quotative construction only tangentially, illustrating them with few examples. For more details about switch-reference, quotative constructions, and desiderative constructions, consult Peña (2015: especially 861–866).

A verb of cognition such as *nɨka* 'know' takes a complement with the Set I action nominalizer *-ta* (an event nominalization) if the subject of the main verb and the nominalization functioning as a complement are co-referent. As with any regular NP, the complement takes object marking, as in (40). However, note that the complement clause does not take object marking in (41) because of the hierarchy (1PL/2 → 3) explained in § 2.2. Thus, in terms of its external sysntax, a grammatical nominalization is treated in exactly the same manner as underived nominals with respect to case marking.

(40) Object (P) Matrix Verb
[*kuntina mã-ta*]$_{\text{NMLZ}}$ = *na nɨka-ha-i*
animal kill-NMLR=OBJ know+IPFV-1SG-DECL
'I know how to hunt.'

(41) Subject (A) Object (P) Matrix Verb
amɨ [*jukuma-ta*]$_{\text{NMLZ}}$ *nɨka-mɨ*
2SG swim-NMLR know+IPFV-2SG + DECL
'You know how to swim.'

In contrast, when the complement of *nɨka* 'know' involves different subjects, Set II nominalizers are used (an event nominalization).

(42) Object (P) Matrix Verb
[*amɨ mina-u*]$_{\text{NMLZ}}$ *=na=ʃa* *nɨka-wa-i*
2SG come+IPFV-NMLR=OBJ=ADD know+IPFV-3-DECL
'He knows that you are coming too.'

Other verbs, like *antu* 'hear', take complements in which the subject of the matrix and the subject of the complement are different. When the complement refers to an event or to the object of the dependent verb, *antu* 'hear' takes a complement with Set II nominalizer *-mau*. This is illustrated in (43) and (44), respectively. As in the previous examples, the nominalization acting as the complement of the verb receives the object marker *=na*.

(43) Object (P) Matrix Verb
[*utʃi-rĩ uuta-mau*]$_{\text{NMLZ}}$ = *na* *antu-ka-mia-ji*
child-3.POSS cry+IPFV-NMLR=OBJ hear-INTENS-DIST.PST-1.PST + DECL
'[The bear's mother] heard her cub crying.'

(44) Object (P) Matrix Verb
[*afi ta-mau*]$_{\text{NMLZ}}$ *=na* *antu-ka-mia-ji*
all say+IPFV-NMLR=OBJ hear-INTENS-DIST.PST-3.PST + DECL
'She/he heard all that was being said.'

On the other hand, when the nominalization functioning as the complement is oriented toward the subject of the dependent verb, a completely different strategy is used. In this case, the structure in complement function is marked with a subordinator and a switch-reference marker. In (45), the subordinate verb *ta* 'arrive' receives the suffix *-ku* 'simultaneous action', plus a person marker *-mi*,[13] and the switch-reference marker *-nĩ* 'different subject'. Thus, rather than a nominalization,

13. The person markers used in switch-reference constructions are different from the ones used in finite verbs.

a typical Wampis switch-reference construction is used. Notice that the switch-reference construction that marks the dependent verb is not treated as a nominal: it does not receive object marking, unlike the nominalizations in the previous examples. Also compare (45) with (46), which illustrates that when a simple noun is the object of this same verb, it does take the object marker *=na*.

(45) Dependent
kaʃi-tin ikama=numa=ia ta-a-ku-mi-nĩ
night-TIME forest=LOC=ABL arrive-IPFV-SIM-2SG-DS
Main
antu-ka-ma-ha-i
hear-INTENS-REC.PST-1SG-DECL
'Last night, I heard you arriving from the forest.'

(46) *pampaina=na antu-ka-ma-ha-i*
noise=OBJ hear-REC.PST-1SG-DECL
'I heard the noise.'

Complement-taking verbs such as *nankama* 'begin', *amu* 'finish', *umi* 'complete', *inai* 'stop doing' all take complement clauses involving co-referent subjects formed with the Set I action nominalizer *-ta*. Examples (47) and (48) illustrate this strategy with *nankama* 'begin' and *umi* 'complete'.

(47) [*ikám hu-ta*]$_{\text{NMLZ}}$ *=na nankama-a timaji*
forest\LOC take-NMLR=OBJ begin-IPFV NARR
'[He] started to take him to the forest.' (Lit. 'he would start the taking to the forest.')

(48) [*aʃi taka-ta*]$_{\text{NMLZ}}$ *=na umi-ka-ma-ĩ*
all work-NMLR=OBJ complete-INTENS-REC.PST-3.PFV + DECL
'He just completed all the work.'

The verb *ii* 'see' is interesting because it can convey the meaning of 'see' or it can convey the idea of a thought, and it follows partially different strategies for these two meanings. For the meaning of 'see', the verb *ii* takes a complement formed with either of the Set II nominalizers *-u* or *-mau*, depending on whether the nominalized argument is A (subject of transitive) or P (object), as in (49) and (50), respectively.

(49) wii [*mina ʃiampi-ru mãn-tu-áw-ar-u*]$_{\text{NMLZ}}$ *=na ii-sa-ha-i*
1SG 1SG.GEN hen-1SG kill-APPL-HIAF-PL-NMLR=OBJ see-ATT-1SG-DECL
'I saw the ones who killed my hen.'

(50) *wii* [*mina jatsu-ru mãn-tu-á-mau*]$_{\text{NMLZ}}$ *=na ii-sa-ha-i*
1SG 1SG.GEN brother-1SG kill-APPL-HIAF-NMLR=OBJ see-ATT-1SG-DECL
'I saw my dead (killed to my detriment) brother.'

When the same verb *ii* is used to express a thought, it follows two strategies depending on whether the subject is coreferential or not. With non-coreferential subjects, nominalizations formed with Set II nominalizers *-u* or *-mau* are used, similarly to the previous examples. An example with *-u* is provided in (51).

(51) *nĩn=ka ii-a-ha-i* [*kanu-haĩ mina-u*]$_{\text{NMLZ}}$ *=na=ʃa*
3SG=FOC see-IPFV-1SG-DECL canoe-COM come+IPFV-NMLR=OBJ=ADD
'As for him, I think he's coming by canoe too.'

With coreferential subjects, the strategy changes completely. In this case, a mix of a quotative construction (the direct speech report) and switch-reference is used. In (52), there is a direct speech report (Lit. 'I filled it!') introduced by the verb *tu* 'say'. At the same time, *tu* 'say' is subordinated by the non-temporal manner subordinator *-sa* and takes the switch-reference marker *-nu*.

(52) *aima-ka-mia-ha tu-sa-nu ii-sa-ma-ha-i*
fill-INTENS-DIST.PST-1SG + EXCL say-SUB-1SG.SS see-ATT-REC.PST-1SG-DECL
'I thought that I had filled it (i.e. the pot with water)' (Lit. 'Saying "I filled it!", I saw.')

5.3.2 *Complements of verbs of movement*

There are two constructions involving nominalizations that are used for oblique complements of verbs of movement. The most frequent one uses the Set II nominalizer *-mau* or the Set I nominalizer *-tinu,* marked with a locative postposition (cf. (32) and (39), respectively).

The other, less frequent construction, is formed with the Set I non-agentive nominalizer *-taĩ*. As seen in § 4.4, the nominalizer *-taĩ* can derive a noun that denotes, among other things, the location of a verb. The use of the nominalizer *-taĩ* in such a capacity can be seen in (53). In this case, a locative postposition is not required.[14]

(53) [*mitaju wɨka-taĩ-rĩ*]$_{\text{NMLZ}}$ *hĩã̂*
hunting walk-NMLR-3.POSS arrive+HIAF.3SG.SS
'Having arrived where she walked to hunt…' (Literally: 'Having arrived to her hunting walking-place…')

14. Note that *-taĩ* is likely historically composed of the Set I action nominalizer *-ta* plus a the locative morpheme *-ĩ*.

5.4 Adverbial modification

Nominalized constructions with *-mau* plus a locative marker are also used for adverbial functions, typically temporal, although there are also examples of nominalizations conveying a reason sense. In addition, nominalizations with *-tinu* are used to convey purpose.

5.4.1 *Temporal and reason modification with* -mau =nVma

A nominalized adverbial temporal construction is illustrated in (54). It involves an event nominalization plus a locative marker.

(54) [*kaʃi has-mau*]$_{\text{NMLZ}}$ = *nama hĩá-u timaji*
night become-NMLR=LOC arrive+HIAF-NMLR NARR
'When it became night, he arrived.'

The next example was translated with an interpretation of reason rather than with a temporal interpretation. The construction is basically the same as in (54).

(55) [*tʃitʃama iwara-mau*]$_{\text{NMLZ}}$ *=nama awaruna=ka kamɨ huĩn=ka*
problem fix.PFV-NMLR=LOC Awajun=FOC INTERJ here=FOC
utsaaná-u=aiti
enter+HIAF-NMLR=COP.3 + DECL
'Because the problem (i.e. the war) was fixed, the Awajun have entered here (i.e. in this region).'

In discourse, adverbial temporal constructions are especially frequent with resumptive verbs. Resumptive verbs in Wampis are used in a tail-head-like strategy in which the resumptive verb connects a previous clause with the next one by making reference to the last action of the preceding clause, as can be seen in (56b):

(56) a. *ɨhama=kɨ=ʃa hĩã́*
midnight=RESTR=ADD arrive+HIAF.3.SS
'Having arrived at midnight,
b. *nuĩ* [*nuni-á-mau*]$_{\text{NMLZ}}$ *=nama*
there do.that-HIAF-NMLR=LOC
'there, when they did that,'
c. *ʃuara a-ina aa irun-u awaruni a-ina*
enemy COP-PL.IPFV outside get.together-NMLR Awajun COP-PL.IPFV
'the enemies, the ones that were outside, the Awajun (spoke…)'

5.4.2 *Purpose modification with Set I future nominalizer* -tinu

The future nominalizer *-tinu* is also used in purpose constructions. The connection between the categories of future and purpose is well established in the typological literature (Bybee, Perkins & Pagliuca 1994: 274; Schmidtke-Bode 2009). In

(57), a Wampis elder speaks to a group of people to encourage them to work in order to create a community and educate their children, conveying purpose senses with a *-tinu* nominalization in lines (b) and (c).

(57) a. *atumi=ʃa umintsa-tá-rumɨ jaakata nahana puhu-sa-mi*
2PL=ADD prepare-IMP-2PL town make+IPFV live/be-ATT-HORT
'You too, prepare! Let us create a town (Lit. 'Let us live creating a town').

b. [*iina utʃi-rĩ nunka umi-ru-ka-tinu*]$_{NMLZ}$
1PL.GEN child-3.POSS.GEN land prepare-APPL-INTENS-FUT.NMLR
'to prepare the land of our children'

c. [*iina utʃi-rĩ papi auhuma-mitika-sa-tinu*]$_{NMLZ}$
1PL.GEN child-3.POSS\GEN book read-CAUS-ATT-FUT.NMLR
'and to get our children educated.' (Literally: and to make read the books of our children').

5.5 Copy-verb construction

In Wampis, a particular construction involving Set I non-agentive *-taĩ* nominalizer is comparable to what in other languages is analyzed as 'copy-verb' (Good 2003), 'tautological infinitive' (Goldenberg 1998) or 'fronted infinitive' (Güldemann 2003) constructions. In this Wampis copy-verb construction, a *-taĩ* nominalization is fronted and followed by a semantically main verb. The semantically main verb may occur in its bare form or as a nominalized form (with Set II *-u*); thus the event nominalization with *-taĩ* is followed by an argument nominalization with *-u* (see (58) below). The fronted element (the nominalized structure with *-taĩ*) frequently occurs focused with *=kɨ* 'Restrictive' plus *=ʃa* 'Additive'.[15] The copy-verb construction typically occurs in key moments of Wampis narratives, and serves to intensify an action that has been recently introduced in the discourse. In (58), an enemy party is searching for the protagonist of the story. A key point in the story is when they find him with his family and lay siege to his house. This event is emphasized with the copy-verb construction.

(58) *ta-ru-i-ara* [*ɨpɨna-**taĩ**=kɨ=ʃa*]
arrive-APPL-LOAF-3PL.SS fence.in-**NMLR** =RESTR=ADD
[*ɨpɨna-ara-**u*** *hɨĩ=nĩ*]
fence.in-PL-**NMLR** house\3.POSS=LOC
'Having arrived (i.e. where the man lived with his family), as shutting him in, they shut him in his house.'

15. The combination [Noun=*kɨ=ʃa*] is used to indicate temporal or locational continuity in Wampis (Peña 2015: 773). This meaning is extended in the copy-verb construction to express continuity or imminency of events.

5.6 Main clause constructions involving nominalizations

In cross-linguistic perspective, it is not unheard of for nominalizations to be used in predicative functions in several types of main clauses (DeLancey 2011). This is a fascinating phenomenon that involves the use of referential expressions as a predicate or main element in a predicate. Wampis exhibits a very rich TAM system that distinguishes various categories. Apart from the categories conveyed via the suffixes that can occupy position 3 of the verbal piece (cf. Table 1), there are also numerous morphemes used for tense and mood/modality. Table 7 lists the morphemes used for indicating tense and mood/modality in Wampis.

Table 7. Morphemes used in the TAM system of Wampis

Tense	**Mood/Modality**
-tata 'Definite future'	*-i* 'Declarative'
-ta 'Immediate future'	∅ 'Exclamative'
∅ 'Present'	*-tá* 'Imprerative'
Aktionsart stem + ∅ 'Just done action'	*-pa* 'Prohibitive'
-ma 'Present past'	*-ti* 'Jussive'
-ɨmia 'Intermediate past'	*-ai* 'Apprehensive'
-mia 'Distant past'	*-mi* 'Hortative'
-ia 'Remote past'	*-tai* 'Inferential'
=hak 'Habitual past'	*timaji* 'Narrative'
	=ʃa 'Speculative'
	=api 'Sudden realization' and 'Tag Question'
	=ka 'Interrogative'
	=hama 'Mirative'
	=ʃa 'Speculative'

The Wampis TAM system is expanded with the use of nominalizations in predicative function. It is widely attested that nominalization is one of the major ways via which languages obtain new finite structures (Noonan 1997; Himmelmann 2005; Gildea 2008; DeLancey 2011; Ahland 2015). In this regard, in order for different nominalizations to achieve a predicative function, the most common strategy is to add a copula to the nominalization. In this way, the complex construction functions as the main predication as it becomes finite with the use of the copula. However, it will be seen that in some cases the copula is not necessary and that the nominalization can stand alone in predicative function. Table 8 introduces the predicative structures involving nominalizations, which are described in the next subsections.

Table 8. Predicative structures involving nominalizations

Semantics	Structure
Indefinite/Distant future	V-***tinu***=cop
Habitualness (with Remote Past)	V-***inu*** cop
Habitual past	V=*hak*-***u***/-***mau*** (=cop)
Narrative	V=*hak*-***u***/-***mau*** *timaji*
Normative	V-***taĩ*** (=cop)
Potential and deontic	V-*mai*-***inu*** (=cop)

5.6.1 *Indefinite/Distant future with Set I future nominalizer* -tinu

The future nominalizer *-tinu*, when used with a copula, refers to a distant, indefinite future. Compare the sentence with an embedded nominalization in (61) with the simple sentences in (59) and (60).

(59) *paki=na* *mã-á-ta-wa-i*
peccary=OBJ kill-HIAF-IMM.FUT-3-DECL
'He's gonna hunt peccaries.'

(60) *tuminku paki=na* *mã-á-tata-wa-i*
Sunday peccary=OBJ kill-HIAF-DEF.FUT-3-DECL
'He's gonna hunt peccaries on Sunday.'

(61) *mina* *jatsu-tʃi-ru=ka* *paki=na*
2SG.GEN brother-DIM-1SG=FOC peccary=OBJ
*mã-á-**tinu** = **iti*** *tsaka-rã*
kill-HIAF-**FUT.NMLR=COP.3+DECL** grow-DISTR/3.SS
'My brother will hunt peccaries when he grows up.'

In (59), the verb is marked with the immediate future suffix *-ta*. As the name indicates, the immediate future refers to actions that are going to happen relatively soon. In (60), the verb carries the definite future *-tata*. The definite future refers a time frame which is when the speaker understands the predicated action is going to be carried out. Usually there is an overt indicator of the time the action will happen (e.g., 'Sunday'). On the other hand, a [V-*tinu*=cop] construction as in (61) has a distant, typically indefinite, future reference.

5.6.2 *Habitualness with Set I S/A nominalizer* -inu

An *-inu* nominalization plus a copula with remote past tense marking assumes a habitual sense. These constructions are used very frequently for providing background information, usually from a second-hand source. The most typical example of this use of *-inu* in texts usually occurs in the introductory lines of a narrative:

(62) *mina papa-ru uha-tu-**inu** **a-ia-ji** utʃi*
1SG.GEN father-1SG tell-1SG.OBJ-**NMLR COP-REM.PT-3.PT + DECL** child
a-sa-mataĩ
COP-SUB-1SG/3.DS
'My father used to tell me (this story) when I was a child.'

(63) *hu=na=ka mina apatʃi-ru-ka auhumatu-**inu***
PROX=OBJ=FOC 1SG.GEN grandfather-1SG-FOC tell-**NMLR**
a-ia-ji
COP-REM.PST-3.PST+DECL
'As for this (story), my grandfather used to tell it.'

Contrast the above examples with the following simple, non-nominalized structure with the remote past tense in (64), where no habitual sense is obtained. This time, a speaker tells of his one-time adventure with his family in a beach town in the coast of Peru.

(64) *tura najantsa=nama jukuma-ia-hi*
and sea=LOC swim-REM.PST-1PL+DECL
'And we swam in the sea.'

5.6.3 *Habitual Past*=hak *with Set II* -u *and* -mau

Another way to mark habitualness in the past is via the habitual past. The habitual past = *hak* always carries a Set II nominalization (*-u* or *-mau*, depending on the nature of the argument of the nominalized verb) and occurs with either a copula clitic ((65)–(66)) or with the narrative marker *timaji* (67).[16]

(65) *iʃitʃi=ki umu= **hak-ara-u=iti***
little.bit=RESTR drink=**HAB.PST**-PL-**NMLR=COP.3+DECL**
'They used to drink (a special banana drink) only a little bit.'

(66) *au ʃuara a-ina-wa-i amaini=ia ʃuara*
DIST person COP-PL.IPFV-3-DECL opposite.bank.of.river=ABL person
*jauntʃuki=ka tu=**hak-mau=waiti***
long.ago=FOC say=**HAB.PST-NMLR=COP.3+DECL**
'Those people, long ago, used to be called "the people of the opposite bank of the river"'.

(67) *jauntʃuki paki hĩã kaunu **hak-u** **timaji***
long.ago peccary house\LOC come.many **HAB.PST-NMLR NARR**
'Long ago, it is said that the peccaries used to come to the house.'

16. The habitual past occurs as a separate phonological word in (67).

Interestingly, a habitual past nominalization can predicate on its own; that is, there is no need for the copula to be present for the nominalization to function as a finite verb:

(68) *nu=ka uun=ka mãa-nai-kã=ʃa ʃuara=haĩ* [*tukɨ́*
NON.VIS=FOC elder=FOC kill-RECP-INTENS\3.SS=ADD person=COM always
nɨpɨtma=hak-u]$_{NMLZ}$
win=HAB.PST-NMLR
'That elder, when he fought another person, he always defeated him.'

5.6.4 *Potential and deontic sense V-mai + Set I* -inu

A potential stem is frequently formed with the Set I Agentive S/A nominalizer *-inu*, and as with the habitual past constructions described in the preceding section, the potential+ – *inu* can occur with a copula (69) or stand alone (70). Similarly to Awajun (cf. Overall 2007: 299), a potential stem in Wampis makes a transitive verb ambitransitive S=P.

(69) *urutma=ka hu-**mai-inu=aita***
how.many=Q take-**POT-NMLR=COP.Q**
'How many could be taken?' or 'How many could he take?'

(70) *tsarur hi-á-mau kaa-sa-u-tʃi=na*
young.fish roast-HIAF-NMLR become.crispy-ATT-NMLR-DIM=OBJ
[*nu=na=ka yu-**mai-inu***]$_{NMLZ}$
NON.VIS=OBJ=FOC eat-**POT-NMLR**
'Young fish roasted and crispy, that can be eaten.'

A potential form nominalized with *-inu* can also assume a deontic sense, as in (71).

(71) *apatʃi puhuta nu=ʃa mai mɨtɨka unuima-mai-inu=aita-hi*
mestizo life NON.VIS=ADD both equal learn-POT-NMLR=COP-1PL+DECL
'The mestizo way of living, that too, both (the mestizo culture and the Wampis culture) equally we must learn.'

This deontic sense of *V-mai-inu* is not reported in other Jivaroan varieties like Shuar (Pellizaro & Náwech 2005: 28), and Achuar and Shiwiar (Fast et al. 1996), where the deontic construction is based on the cognate of the Wampis future nominalizer *-tinu*.

5.6.5 *Normative with Set I non-agentive* -taĩ

As do other languages of the family (Fast et al. 1996: 47; Pellizaro & Náwech 2005: 29; Overall 2007: 357), Wampis forms a normative construction with the

Set I nominalizer *-taĩ* plus a copula. Following Overall (2007: 357), a normative construction is understood as a construction used to describe "how we do things" normally. Interestingly, even though the copula clitic is always a third person copula, the semantic interpretation of the normative construction is in the first person plural. This can be made explicit, as in (72), where the first plural, same subject suffix *-ri* is attached to the subordinated verbs meaning 'slice' and 'scrape'.

(72) a. *pita=ka kaapi=haĩ taka-**taĩ=aiti***
traveling.basket=FOC tamshi.vine=COM work-**NMLR=COP.3+DECL**
'The traveling basket, we make it with tamshi vine.'

b. *kaapí ɨtsɨntsɨ-ra-ri*
tamshi.vine\OBJ slice-DISTR-1PL.SS
'Having sliced the tamshi vine,'

c. *ʃiira-tʃi masɨ-ka-ri pɨnkɨra hawi_hawintu*
well-DIM scrape-INTENS-1PL.SS good flexible
*nahana-**taĩ=aiti***
make-**NMLR=COP.3+DECL**
'having scrapped it well, we make it very flexible.'

As in previous constructions with the potential+ – *inu* or the habitual past, *-taĩ* nominalizations with a normative sense can also predicate without the need of a copula. In the next example, the speaker explains that in order to have a vision of *Arutam* 'spirit of power', one ought not to be afraid. A *-taĩ* nominalization with no copula acts as the main predicate and subordinates concessive and manner clauses.

(73) [*iʃama-cha-ku-ri-ʃa kamɨ arantu-sã ii-taĩ*]$_{\text{NMLZ}}$
be.afraid+IPFV-NEG-SIM-1PL.SS-CONCESS INTERJ respect-SUB\3.SS see-NMLR
'but without (we) being afraid, respecting it, we see it (i.e. *Arutam*).'

6. From nominalizations to finite structures

In the previous sections I have shown that some nominalizations can function as main predicates alone. These constructions are equivalent to what are described as 'stand-alone' nominalizations (Matissof 1972; Noonan 1997; see also Evans 2007); that is, nominalized constructions that are not embedded and constitute a sentence on their own. Stand-alone nominalizations are frequently used for specific discourse functions, such as speaker mood or stance in Asian languages (Yap et al. 2011: 7). For instance, in Wampis, stand-alone nominalizations can occur in evaluative comments, among other uses. In (74), there is a parenthetical comment ('it can be said (he was) respected'): the speaker is describing that a person who

drank ayahuasca would obtain a certain power and become strong, and adds '(he was) respected, it can be said'.

(74) *nu ʃuara ʃiira sɨntʃi a-hak-u=iti naa*
NON.VIS person very strong COP-HAB.PST-NMLR=COP3.PST+DECL HESIT
[*arantu-taĩ*]$_{NMLZ}$ [*tu-mai-inu*]$_{NMLZ}$
respect-NMLR say-POT-NMLR
'That person used to be very strong, mmm, (he was) respected, it can be said.'

Perhaps the most striking feature of Wampis is the fact that often a nominalization may function as the main verb in a sentence (75) or even as main verb in chains of switch-reference clauses (76). This is primarily done with set II nominalizers, especially with *-u* in narratives.

(75) *ʃuara=na hatɨ-a-kũ puha-u=na tsɨrɨ*
person=OBJ cut-IPFV-SIM\3.SS live/be+IPFV-NMLR=OBJ Tsere
waina-ka-u
see-INTENS-NMLR
'Tsere saw him cutting people.'

(76) *kintamɨ-a-ka-ĩ waina-ka-tahkamã tsɨrɨ=na tapit*
become.late-IPFV-SIM-DS see-INTENS-FRUST\3.SS Tsere = OBJ IDEO
atʃi-ka-u *iwa*
grab-INTENS-NMLR Iwa
'When it was becoming late, having been unable to find it (i.e. his axe), Iwa grabbed Tsere.'

In fact, it seems that in Wampis these nominalizers are being used as final-verb markers with past tense reference. This fact brings up the issue of re-analysis. Previous studies of other Jivaroan languages have considered *-u* as a reportative marker. For instance, Fast et al. (1996: 45) indicates that in Achuar-Shiwiar *-u* is a reportative past (i.e. non-first hand) when attached to the verb *tu* 'say'. Likewise, for Awajun, Overall (2014) offers a convincing analysis of *-u* as re-analyzed for marking non-first hand information source. Data from Wampis shows, however, that in this language *-u* has extended its functions to mark past tense, but not exclusively for reportative or non-first hand information. The following is an excerpt from a conversation recorded at the dinner table. The speaker has arrived from hunting and is bringing some food to my host's house. He is telling that he saw a deer when he was in the forest, but the event took him by surprise and he was not quick enough to kill it.

(77) *ɨntsa-ki* *mina-u-nu*
carry.on.back-WHILE.MOVING come+IPFV-NMLR-1SG.SS
panka-i-nu *hapa uaha-ma=hama* ***uaina-ka-u***
raise.head-LOAF-1SG.SS deer stand-DUR=MIR **see-INTENS-NMLR**
'[…] I was coming carrying (the game), having raised my head, I saw a deer that was standing! […]'

In the above example, the verb form that stands in the function of a finite main verb is the nominalization *waina-ka-u*, which is translated as 'I saw'. The nominalized verb has past temporal grounding, but it is definitely not remote past (moreover, it cannot be legendary or mythical past) as the events are recent (the speaker had been hunting the previous night). Because this is a first person passage that tells about the speaker's own experience, it cannot be claimed that the nominalization is having a reportative or non-first hand evidential sense either.

To finish this section, I briefly report on a more infrequent phenomenon: the use of nominalized verbs as verb stems in Wampis. Some examples in the data show that (erstwhile?) nominalizations are sometimes treated as verb stems; that is, nominalized stems can receive morphology associated with verbal finite categories (cf. § 2.2) without the presence of a copula.

(78) *mɨtɨka mãa-nai-**inu**-tʃa-ara-mia-ji*
equal kill-RECP-**NMLR**-NEG-PL-DIST.PST-3.PST+DECL
'They were not fighting equally.'

(79) *amaini* *katirpisa=numa katin-ka-**u**-mia-ji*
other.side.of.river Katirpis=LOC cross-INTENS-**NMLR**-DIST.PST-3.PST+DECL
'He crossed to the other side of the Katerpisa river.'

(80) *nuku-ru=na* *taka-ta=na* *jai-nai-a*
mother-1SG=OBJ work-NMLR=OBJ help-RECP-IPFV
*wɨ-**u**-mia-ha-i*
go-**NMLR**-DIST.PST-1SG-DECL
'I used to go help my mother at work.'

In (78)–(80), the nominalizers occupy a position before tense markers, person (subject) and mood/modality – that is, positions 6, 7 and 8 in Figure 1. Although more data is needed to examine what purpose the nominalizers in these structures serve, a current hypothesis is that the nominalizers are being reanalyzed as aspectual markers. For instance, when the nominalizer *-u* combines with an aspectual stem (cf. (79) with the aktionsart *-ka* 'intensive') the verb clearly has a perfective sense (in (79), a bounded, one-time event). On the other hand, in (78) and (80) the action exhibits an imperfective sense (plus habitual in (80)): in both of these cases the nominalizer occupies a position where any of the suffixes in morphological

position 3 of the verb (cf. Figure 1) is expected to occur. This process of reanalysis is apparently ongoing, as (79) shows that the nominalizers can still occur with other suffixes of such position.[17] In fact, (79) looks like a 'normal' nominalized stem which has nevertheless received finite morphology; (80), on the other hand, looks more like a 'normal' verb stem in the sense that *-u* seems to be filling position 3 (there is neither aktionsart or imperfective suffixes present, unlike in (79)) and providing aspectual information to the clause. Thus, it seems that new morphological pieces are being "recruited" into the verbal structure, but this process is not yet completely grammaticalized.

7. Conclusion

The two sets of Wampis nominalizers studied in this paper are derivational in nature; i.e. they constitute class-changing devices. In this way, Wampis speakers create referring expressions that are noun-like in their morphosyntactic properties, being able to head NPs, receive case marking, as well as oblique and discourse-related clitics that are normally carried by nominals. At the same time, Wampis nominalizations may occur retaining a good portion of verbal categories.

There are a fair number of constructions in Wampis that employ grammatical nominalizations. Grammatical nominalizations can be used for different functions, among them relativization, complementation and adverbial subordination. While nominalizations generally do not carry tense or mood markers, it has been shown that they also occur in constructions that function at the level of the Wampis TAM system, conveying tense, modality and other grammatical and discourse distinctions. Even more interestingly, stand-alone nominalizations in Wampis function as verbal predicates, heading independent clauses. There are at least two strategies for the reanalysis of nominalizations into predicative constructions. The first is to add a copula verb that receives finite morphology. This strategy is found in several languages around the world and seems to be one that commonly motivates the emergence of new inflectional morphology in a language. However, in Wampis copulas are not necessary for nominalizations to function as predicates; that is, nominalizations can stand alone in predicative function. The second pattern of reanalysis of nominalizations in Wampis is more intriguing. It involves the use of a nominalization as a verb stem, which thus receives inflection proper of a typical Wampis finite verb. While this phenomenon needs to be more fully investigated, the fact that a nominalized structure can so readily be used as

17. Recall from § 2.1.3 that aktionsart, imperfective, durative, present habitual and potential suffixes do not co-occur with each other.

a "normal" verb stem poses further questions about the nature of nominalization. In that regard, this paper has also contributed to a growing body of literature that shows that constructions involving a nominalization may, and often do, develop functions that go beyond that of simple referring expressions.

Abbreviations

1	First person	HESIT	Hesitation
2	Second person	HIAF	High affectedness of P or change of location of S
3	Third Person	HORT	Hortative
ABL	Ablative	IDEO	Ideophone
ADD	Additive	IMM.FUT	Immediate future
APPL	Applicative	INTENS	Intensive action of A/S
ATT	Attenuative	INTERJ	Interjection
ATTRIB	Attributive	IPFV	Imperfective
CAUS	Causative	LOAF	Low affectedness of P or S
COM	Comitative	LOC	Locative
COP	Copula	NMLR	Nominalizer
DECL	Declarative	NON.VIS	Non-visible demonstrative
DEF.FUT	Definite future	OBJ	Object
DIST	Distal demonstrative	PL	Plural
DIST.PST	Distant past	POSS	Possessed
DISTR	Distributed action	POT	Potential
DS	Different Subject	Q	Question marker
DUR	Durative	REC.PST	Recent past
GEN	Genitive	RECP	Reciprocal
FOC	Focus	REM.PST	Remote past
FRUST	Frustrative	SG	Singular
FUT.NMLR	Future nominalizer	SIM	Simultaneous action
HAB.PST	Habitual past	SS	Same subject

References

Ahland, Michael. 2015. The development of finite verbs from nominalized structures in Northern Mao. Paper presented at the 46th Annual Conference on African Linguistics, University of Oregon.

Bybee, Joan, Perkins, Revere & Pagliuca, William. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago IL: University of Chicago Press.

Comrie, Bernard. 1978. Ergativity. In *Syntactic Typology: Studies in the Phenomenology of Language*, Winfred P. Lehmann (ed.), 329–334. Austin TX: University of Texas Press.

Cristofaro, Sonia. 2007. Deconstructing categories: Finiteness in a functional-typological perspective. In *Finiteness: Theoretical and Empirical Foundations*, Irina Nikolaeva (ed.), 91–114. Oxford: OUP.

DeLancey, Scott. 1986. Relativization as nominalization in Tibetan and Newari. Paper presented at the 19th International Conference of Sino-Tibetan Languages and Linguistics.

DeLancey, Scott. 1999. Relativization in Tibetan. In *Studies in Nepalese Linguistics*, Yogendra Yadava & Warren Glover (eds), 231–249. Kathmandu: Royal Nepal Academy.

DeLancey, Scott. 2011. Finite structures from clausal nominalization in Tibeto-Burman. In *Nominalization in Asian Languages: Diachronic and Typological Perspectives* [Typological Studies in Language 96], Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds), 343–360. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.96.12del

Dixon, Robert M.W. 2006. Complement clauses and complementation strategies in typological perspective. In *Complementation: A Cross-linguistic Perspective*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 1–48. Oxford: OUP.

Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and Empirical Foundations*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.

Fast, Gerhard, Warkentin de Fast, Ruby & Fast Warkentin, Daniel. 1996. *Diccionario Achuar-Shiwiar~Castellano*, 2 edn. Yarinacocha: Ministerio de Educación/Instituto Lingüístico de Verano.

Genetti, Carol, Coupe, Alexander R., Bartee, Ellen, Hildebrandt, Kristine & You-Jing, Lin. 2008. Syntactic aspects of nominalization in five Tibeto-Burman languages of the Himalayan area. *Linguistics of the Tibeto-Burman Area* 31(2): 97–143.

Gildea, Spike. 2008. Explaining similarities between main clauses and nominalized phrases. In *La structure des langues amazoniennes*, Ana Carla Bruno, Frantomé Pacheco, Francesc Queixalós & Leo Wetzels (eds). *Amerindia* 32: 57–75.

Givón, T. 2002. *Syntax: An Introduction*, Vol. II. Amsterdam: John Benjamins.

Givón, T. 2012. Toward a diachronic typology of relative clause. In *Relative Clauses in Languages of the Americas: A Typological Overview* [Typological Studies in Language 102], Bernard Comrie & Zarina Estrada-Fernández (eds), 3–26. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.102.01giv

Goldenberg, Gideon. 1998. Tautological infinitive. In *Studies in Semitic Linguistics: Selected Writings*, 36–85. Jerusalem: Magnes Press, Hebrew University.

Good, Jeff. 2003. Clause combining in Chechen. *Studies in Language* 27(1): 113–170. https://doi.org/10.1075/sl.27.1.05goo

Güldemann, Tom. 2003. Present progressive vis-à-vis predication focus in Bantu: A verbal category between semantics and pragmatics. *Studies in Language* 27(2): 323–360. https://doi.org/10.1075/sl.27.2.05gul

Himmelmann, Nikolaus P. 2005. The Austronesian Languages of Asia and Madagascar: Typological characteristics. In *The Austronesian languages of Asia and Madagascar*, Alexander Adelaar & Nikolaus P. Himmelmann (eds), 110–181. London: Routledge.

Jakway, Martha, Horacio, Lorenzo C. & Arturo, Antonio A. 1987. *Vocabulario huambisa*, 2 edn. Yarinacocha, Pucallpa: Ministerio de educación/Instituto Lingüístico de Verano.

Katan Jua, Tuntiak. 2011. Ii chichame unuimiamu / Investigando nuestra lengua / Investigating our language: Shuar Chicham. In *Endangered languages: Voices and images. Proceedings of the FEL XV Annual Conference*, Marleen Haboud & Nicholas Ostler (eds), 103–105. Bath: Foundation for Endangered Languages.

Keenan, Edward L. 1985. Relative clauses. In *Language Typology and Syntactic Description, Vol. II: Complex Constructions*, Timothy Shopen (ed.), 141–170. Cambridge: CUP.

Matissof, James A. 1972. *Lahu nominalization, relativization, and genitivization*. In *Syntax and Semantics*, John P. Kimball (ed.), 237–357. New York NY: Seminar Press.

Nikolaeva, Irina (ed.). 2007. *Finiteness: Theoretical and Empirical Foundations*. Oxford: OUP.

Noonan, Michael. 1985. Complementation. In *Language Typology and Syntactic Description, Vol. II: Complex Constructions*, Timothy Shopen (ed.), 42–140. Cambridge: CUP.

Noonan, Michael. 1997. Versatile nominalizations. In *Essays on Language Function and Language Type: Dedicated to T. Givón*, Joan Bybee, John Haiman & Sandra A. Thompson (eds), 373–394. Amsterdam: John Benjamins. https://doi.org/10.1075/z.82.21n00

Overall, Simon. 2007. A grammar of Aguaruna. PhD dissertation, La Trobe University.

Overall, Simon. 2014. Nominalization, knowledge, and information source in Aguaruna (Jivaroan). In *The Grammar of Knowledge: A Cross-linguistic Typology*, Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds), 227–244. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198701316.003.0011

Pellizaro, Siro M. & Fausto Oswaldo Náwech. 2005. *Chicham: Diccionario shuar-castellano*. Quito: Abya-Yala.

Peña, Jaime. 2015. A grammar of Wampis. PhD dissertation, University of Oregon.

Schmidtke-Bode, Karsten. 2009. *A Typology of Purpose Clauses* [Typological Studies in Language 88]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.88

Shibatani, Masayoshi. 2009. Elements of complex structures – where recursion isn't: The case of relativization. In *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution*, [Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 163–198. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.07ele

Yap, Foon Ha, Karen Grunow-Hårsta & Janick Wrona. 2011. Introduction: Nominalization strategies in Asian languages. In *Nominalization in Asian Languages: Diachronic and Typological Perspectives* [Typological Studies in Language 96], 1–57. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.96.01yap

CHAPTER 12

Nominalization in Harakmbut

An Van linden
University of Liège & University of Leuven

This paper focusses on verb-based nominalization in Harakmbut (isolate, Peru), which falls into two formal types on the basis of the prefix used. The first type, using the nominalizing prefix *wa(ʔ)-*, is restricted to participant nominalization and is predominantly used to produce nouns for NP-use. The second type, using the nominalizing prefix *e(ʔ)-*, is mainly used for event nominalization and typically produces multi-word nominalizations. Depending on the constructions they occur in and additional suffixation they take, nominalizations with *e(ʔ)-* can serve complementation as well as adverbial functions. Across the two formal types, multi-word nominalizations combine NP-like external syntax with verb-like internal syntax. The two nominalizing prefixes also serve a basic function in noun-based nominalization, lending independent status to obligatorily bound nouns.

1. The Harakmbut language and collection of data

Harakmbut is an underdescribed language from the Peruvian Amazon, spoken in a number of 'native communities' in the departamentos of Madre de Dios and Cusco. The communities are located on the Madre de Dios River and its upper tributaries, such as the Colorado River. Before I go into more detail about my own fieldwork, I will first summarize what has already been written about the language and its speakers.

The genetic affiliation of Harakmbut has been a topic of debate. The language has formerly been classified as an Arawak or Maipuran language by McQuown (1955) (see Hart 1963: 6) and Matteson (1972), but this has found little acceptance (Adelaar 2007: 39). Wise (1999: 307) states that Harakmbut is commonly accepted to be an isolate (cf. Dryer & Haspelmath 2013 in WALS). More recently, drawing on mainly lexical evidence, Adelaar (2000, 2007) has proposed that it is genetically related to the Brazilian Katukina family, which may be further linked to Macro-Jê. In addition, language contact should also be reckoned with, as Harakmbut

https://doi.org/10.1075/tsl.124.12lin

exhibits a number of Western Amazonian grammatical features, as well as features characteristic of the Guaporé-Mamoré linguistic area in southwest Brazil and eastern Bolivia (Crevels & van der Voort 2008), close to the border with the Peruvian departamento of Madre de Dios, with one member language, Ese Ejja (Tacanan), also being spoken in Madre de Dios. Harakmbut has already been noted to share some grammatical features with Ese Ejja (Pozzi-Escot 1998: 93), and I also believe it shares features with other languages in the area like Cavineña and Kwaza, e.g. in the domains of aspect, associated motion, and, more pertinently to this paper, noun-based nominalization (see Section 5.3).

Previous discussions have also focussed on relations within the Harakmbut group. The main question is whether Harakmbut should be regarded as a single language with a number of dialectal variants or rather as a small language group or family consisting of distinct, related languages. The distinction of seven ethno-linguistic groups (Amarakaeri, Watipaeri, Arasaeri, Sapiteri, Kisambaeri, Pukirieri and Toyoeri) by the anthropologist Andrew Gray (1996: 7–9) might hint at the latter option, while linguists such as Helberg (1984, 1990), Wise (1999) and Adelaar (2007) agree on the former. The proposal that Harakmbut is in fact a single language is consistent with my language consultants' assessments. By now, the most vital varieties are the first two listed above. Of many of the other dialects only a handful of speakers – if any – are left, and very little information is available.

Earlier linguistic work on Harakmbut has mainly concentrated on the most vital dialect, i.e. Amarakaeri (Hart 1963; Helberg 1984, 1990; Tripp 1976, 1995). It should be noted that 'vital' is a relative term, as the number of speakers has been estimated at 1,000 by Moore (2007: 46), and I noted during my field stays that young parents are reluctant to pass on the language to their children, as it is felt to socially stigmatize them. Children are thus mainly brought up in Spanish, and acquire only a passive competence in Harakmbut. Young adults and speakers up to the age of fifty generally are bilingual in Harakmbut and Spanish. Speakers older than fifty are mainly monolingual in Harakmbut. My own fieldwork also focusses on the Amarakaeri dialect.[1] The data presented in this paper are drawn from audio recordings made in the native communities of Puerto Luz, San José del Karene and Shintuya, all with Amarakaeri informants, in the summers of 2010, 2011 and 2016. So far, I have mainly transcribed and analysed elicitation sessions

1. I would like to point out that the speakers of this variety regard the label *Amarakaeri* as a depreciating term; it is adapted from *wa-mba-arak-a-eri* (NMZR-VPL-kill-TRNS-AN), a verb-based nominalization meaning '(fierce) killer/murderer', which goes back to an ancient story about the origin of the different ethno-linguistic groups of the Harakmbut people. They prefer to call their variety 'Arak(m)but', as distinct from the Watipaeri variety, towards whose speakers they generally entertain feelings of enmity rather than brotherhood.

with bilingual speakers, which implies that the bulk of the data used in Sections 3 to 5 does not represent spontaneous speech. If it does, this has been indicated in the example. The practical orthography used is IPA-based, and different from the community spelling.

2. Nominalization in Harakmbut

South American languages generally show a rich diversity of nominalization structures and functions. This also holds for Harakmbut, as this paper aims to show. Its main focus will be on verb-based nominalizations that do not feature person or mood marking. It is beyond the scope of this paper to discuss the nominalization of finite verb forms by the relativizing suffix *-niŋ* (or 'nominalizing' suffix, cf. Shibatani this volume), which is detailed to some extent in Van linden (Forthcoming).

The data available in the literature (Tripp 1976, 1995; Helberg 1984) and my own fieldnotes indicate that (non-finite) verb-based nominalization in Harakmbut falls into two formal types, which can be distinguished on the basis of the prefix used. A first type uses the nominalizing prefix *wa(ʔ)-*, illustrated in (1), while the second type features the nominalizing prefix *e(ʔ)-*, exemplified in (2), which is also used in the citation form of verbs and other non-finite verb forms.[2] I will show that these prefixes serve a basic function in noun-based nominalization as well. In the examples given, grammatical nominalizations (cf. Shibatani this volume) are rendered between square brackets.

(1) *Jonas-tewapa o-niŋ-ka wa-wedn griŋgo-a*
Jonas-BEN 3SG.IND-BEN-make NMZR-lie foreigner-NOM
'The foreigner makes a bed for Jonas.'

(2) *ndak õʔ-ẽ [e-mbaʔ-tiak, ãnĩ, keme]$_{NMLZ}$*
good 3SG.IND-be NMZR-VPL-come FILLER tapir
'It is good that (, eh,) the tapirs have come.'

In (1), the prefix *wa-* attaches to the verb root *-wedn* 'lie' to form the noun 'bed'. This resultant nominalized form functions as an argument participant, viz. it constitutes the direct object of the finite verb form *oniŋka*; it is left unmarked as is generally the case for inanimate O-participants (see Section 3). In (2), the nominalized form *embaʔtiak* consists of the nominalizing prefix *e-* and the verb base *-mbaʔtiak*. It serves as the verb phrase of the complement clause functioning as

2. The glottal stop has no phonemic value in Harakmbut (pace Helberg 1984: 22), but rather a suprasegmental one: it is optionally used to demarcate syllable boundaries when these lack consonantal onsets or codas (see Van linden Forthcoming).

the subject of the commentative predicate *ndak õʔẽ* 'is good'. Thus, while nominalization with *wa(ʔ)-* derives a noun from a lexical verb and realizes participant nominalization in (1), nominalization with *e(ʔ)-* yields an 'action nominal' (cf. Comrie & Thompson 2007: 343) from a predicate, containing also a noun phrase that corresponds to the subject of the verb stem (i.e. *keme*), and realizes event nominalization in (2). Both participant and event nominalization are common in South-American languages (van Gijn et al. 2011: 10–13).

While (1) and (2) illustrate the predominant functions and uses of the two formal types of verb-based nominalization available in Harakmbut, they do not exhaust them. Specifically, nominalizations with *wa(ʔ)-* are found to sometimes modify other nouns, in which function they are equivalent to relative clauses, having their own notional argument participants. Similarly, nominalization with *e(ʔ)-* is not limited to deriving action nominals from predicates, since it is sometimes also used to derive participant nominalizations that function in the same manner as underived nouns (see (53) in Section 5.3 for an example). Together the two formal types of verb-based nominalization realize all three subtypes of subordinate clauses traditionally distinguished, with nominalizations with *wa(ʔ)-* coding relative relations, while nominalizations with *e(ʔ)-* are used to code both complement and adverbial relations.

The discussion is organized as follows. Section 3 discusses the basic features of Harakmbut grammar that are needed to analyse the nominalization data. Section 4 focusses on nominalization with *wa(ʔ)-*, while Section 5 homes in on nominalization with *e(ʔ)-*. Each section will discuss further subtypes of these formal types, with a focus on the internal and external syntax of the nominalized forms; it will be investigated to what extent these are verb-like or NP-like. It will also be examined to what degree the nominalized forms retain verbal categories and adopt nominal ones. With regard to the latter, it will be shown that both formal types of nominalized forms use suffixes that are also used on underived nouns. Section 6, finally, recapitulates the major findings, and proposes some questions for further research.

3. Basic features of Harakmbut grammar

Before we delve into the analysis of nominalized forms, this section discusses some basic features of Harakmbut grammar that are crucial to determining their internal and external syntax. Specifically, it will concentrate on nominal categories, verbal categories and the coding of grammatical relations (based on Van linden Forthcoming).

Harakmbut nouns can be marked for a number of categories; Table 1 visualizes the morphological template of the head of a noun phrase (NP). Table 2 presents

the (rather extensive) set of case markers, two of which are polysemous with the instrumental function. Many of (the functions of) these markers have already been noted before by Helberg (1984: 436–444) and/or Tripp (1995: 194–200); this is also indicated in Table 2.

Table 1. Morphological template of the nominal head

(complex) (pro)noun stem	Collective	Case	$Focus_1$	$Focus_2$
	-(o)mey COLL	(see Table 2)	*-nãỹõ* COND	*-nda* FOC
			-yo REST	

Table 2. Inventory of case suffixes ([H]: Helberg (1984: 436–444); [T]: Tripp (1995: 194–200))

Suffix	**Case**
-ʔa~-a	nominative[H, T]
	instrumental[H, T]
-ere	comitative[H, T]
	instrumental[T]
-ta(h)	accusative[H, T]
-en~-edn~-wedn~-ʔedn	genitive[H, T]
-tewapa	benificiary[T]
-(o)niŋ	similative
-apo	reason
-mbayo	privative
-yo, -ya, -taʔ, -te, -yon, -pen	locative[H, T]

Furthermore, nouns lack the category of number and Harakmbut lacks articles, which would express definiteness or specificity. Instead, nouns pattern with a number of adnominal modifiers, like indefinite and demonstrative modifiers (e.g. *ken* in (4b)), as well as indefinite and cardinal quantifiers. The syntactic relation of adnominal possession is expressed by genitive case marking on the possessor (pro)noun, cf. (3).

(3) *ndoʔ-edn siro*
1SG-GEN machete
'my machete'

Another type of adnominal construction that is pertinent to this paper (see Section 5.2.1) is modification by adjectives. In my data, adjectives appear in both

continuous and discontinuous NPs.[3] In the first subtype – the only relevant one here – they occur in prenominal (4a) as well as postnominal position (4b).

(4) a. *aʔ-yok-i* *sal* *uru-wettone-ta-nda*
1SG.IMP-give-1.IMP salt beautiful-woman-ACC-NDA
'I (should) give salt to the beautiful woman.'
b. *ih-yok-i* *sal* *ken* *wettone-tewapa* *uru-nda*
1SG-give-1.IND salt DIST woman-BEN beautiful-NDA
'I give salt to that beautiful woman.'

(4a) and (4b) are translations of the same stimulus, but they show some interesting differences. For one, they show that R-participants of ditransitive events can receive either accusative marking (4a) or beneficiary marking (4b) (see below). A more important difference lies in the adjectival construction type. While in (4b) the adjective follows the head noun and the adjectival root is suffixed by *-nda*, in (4a) the adjectival root precedes the noun, and the *-nda* suffix is appended only after the case-marked noun. The NP in (4a) also shows phonological fusion; the stressed syllable nucleus is underlined. Both constructions feature the suffix *-nda*, whose function in (4) I am uncertain about, but it seems to be different from that in *Lupeanda* in (5) below, in which it is used as a focus marker appended to a nominative-marked noun (not modified by any adjective) (see Table 1). In (4), *-nda* does not mark focus; it seems to be required by the modification construction. In *mbiʔigŋtonenda* in (5), it does not mark focus either; nouns suffixed by derivational affix *-tone* 'adult, old, big' are also found without suffix *-nda* in contexts similar to (5). I hypothesize that *-nda* basically serves to produce (independent) modifiers or to signal a modification relation (this hypothesis especially targets examples like (4) and infinitival nominalizations, see Sections 5.2.1 and 5.2.2).

(5) *Lupe-a-nda* *oʔ-teŋ-me* *mbiʔigŋ-tone-nda*
Lupe-NOM-FOC 3SG.IND-cut-REC.PST fish-big-NDA
'Lupe herself cut the big fish.'

Harakmbut verbal morphology comprises inflectional as well as derivational categories. The former involve tense, (types of) aspect, mood, modality, evidentiality, and verbal argument marking. The latter include valency-changing categories like transitivizers and applicatives, as well as spatial elements and (types of) aspect. A number of these are illustrated in (6).

3. I am not sure whether NPs whose elements are not adjacent are 'merely' discontinuous or rather appositional.

(6) *o-ma-niŋ-toʔ-tiak-me-te* *aypo wa-mbet-ta*
3SG.IND-VPL-BEN-SOC-come-REC.PST-NVIS food NMZR-family-ACC
Puerto-lus-yo
Puerto-Luz-LOC
'She took food to Puerto Luz for her family (members).'

The verb stem in (6) is intransitive (*-tiak* 'come'), but its valency is increased by the sociative causative prefix *toʔ-*, which promotes *aypo* 'food' to direct object status, and by the benefactive applicative *niŋ-*, which promotes *wambet* 'family' to beneficiary status, receiving accusative marking. These derivational prefixes also appear on nominalized verb forms (e.g. (26), (35)). In (47), even the (inflectional) tense marker *-me* coding recent past, like in (5)–(6), is retained in the nominalized form.

Verbal plural marking by *mba-~ma-~mã-* (phonologically conditioned allomorphs) is also retained in nominalized forms. This category serves to signal plurality of the action denoted by the verb or plurality of participants engaged in the event. In the latter function it works ergatively, indicating plurality of the S-participant in intransitive events (cf. nominalized form in (40)) and of the (applied) O-participant in (extended) transitive events, like in (6), which need not be expressed by external NPs. In (6), the A-participant brought food for more than one family member. In (12), the verbal plural marker indicates that the action of the nominalized verb is performed several times.

Furthermore, Harakmbut verbs show all four types of noun incorporation identified in Mithun (1984). It will become clear that incorporated nouns (types I to III) and verbal classifiers (type IV) are retained in nominalized forms, e.g. (19), (31), and (51).

Finally, grammatical relations are reflected by both head and dependent marking. The head marking system involves hierarchical indexation resulting in a configuration-based split (without direction marking), based on the position of the patient participant on the person hierarchy 1/2 > 3: while third-person patients are never indexed, speech act participant patients trigger distinct relational prefixes, viz. portmanteau prefixes indexing both agent and patient. This split amounts to accusative alignment in non-local configurations (involving a third person acting on another third person) and direct configurations (involving a speech act participant acting on a third person), as agents acting on third person patients on the one hand and the sole participants of intransitive clauses on the other are cross-referenced on the verb by the same set of prefixes (A > 3-markers = S-markers).

The dependent marking system is different, but no less complex, as the three argument roles (S, A and O) show differential or optional marking in independent clauses (case vs. zero exponence). The marking of O-participants is animacy-based. Human and higher order animate Patient-like arguments carry accusative

case marking (e.g. *Lisbet-ta* in (8)), while inanimate and lower order animate Os go unmarked (e.g. *mbiʔigŋtonenda* in (5) and *aypo* in (6)). As illustrated in (4a) and (6), accusative case is also marked on human Recipient-like arguments in (applied) ditransitive clauses. However, (applied) R-participants are also found to carry beneficiary case marking (see (4b)).

The marking of A-participants is governed by both animacy and focus. Non-focal animate As are typically left unmarked, e.g. *mboerek* in (7), while inanimate As are marked, specifically by a case suffix analysed as nominative by Helberg (1984) and Tripp (1995), e.g. *kurudn-a* in (8). This type of differential A-marking is cross-linguistically recurrent (cf. Fauconnier 2011).

(7) *sik-yo-edn-nda ānī, mboerek o-n-ka, ānī, […] pera*
black-LOC-?-NDA FILLER man 3SG.IND-SPAT-do FILLER pear
'Early in the morning, eh, a man is picking, eh, pears.' (spontaneous speech)

(8) *kurudn-a o-seŋ-pak-a Lisbet-ta*
thunder-NOM 3SG.IND-crazy-VBZ-TRNS Lisbeth-ACC
'The thunder drives Lisbeth crazy.'

Animate A-participants that are in argument focus also tend to go marked, e.g. *Lupe-a-nda* in (5), just like As that are in focus within the broader discourse context, cf. (9). Like (7) and (10)–(11) below, (9) is taken from my Pear story data, and is assumed to represent spontaneous speech. In the first (dependent) clause, the A-argument is the boy who stole the pears (see (10)), while in the next one, there is a switch in A-participant to *muneyosiʔpo-a*.[4]

(9) *o-k-to-wa-po bisikleta-te; ken ānī o-ndeh*
3SG.IND-SEPARATION-SOC-go-DEP bicycle-LOC then FILLER 3SG.IND-meet
ānī muneyo-siʔpo-a, ānī, ndaŋ-no-po-te
FILLER girl-DIM-NOM FILLER path-(vital.centre-CLF:round)$_{middle}$-LOC
'He goes away with them [i.e. pears] on his bike; then, eh, a little girl crosses him, eh, in the middle of the road.' (spontaneous speech)

S-participants are typically left unmarked, whether they have human referents, e.g. *wasiʔpo* in (10), or inanimate ones, e.g. *widn* in (11). Only very rarely are S-participants marked by 'nominative' case.

4. In addition, the girl is newly introduced in (9). However, first-mention use does not suffice to trigger case marking on animate As; (7) instantiates the first mention of the pear-picker participant in the story, but here A is left unmarked.

(10) *ken wa-siʔ-po o-tiak-po ānī bisikleta-te*
then NMZR-(peel-CLF:round)$_{child}$ 3SG.IND-come-DEP FILLER bicycle-LOC
ānī o-ta-mbere ānī kanasta
FILLER 3SG.IND-APPL-steal FILLER basket
'Then a child comes, eh, on his bike, eh, and he steals his [i.e. the pear picker's], eh, basket.' (spontaneous speech)

(11) *ken adnte ānī oʔ-wedn ānī widn, widn-tone-nda*
then far.away FILLER 3SG.IND-lie FILLER stone stone-big-NDA
'Then, further down, eh, there lies, eh, a stone, a big stone.'
(spontaneous speech)

While the Harakmbut case marking system has been analysed as showing nominative-accusative alignment in earlier work (Helberg 1984; Tripp 1995), the observed patterns of optional A- and S-marking point to an optional ergative-accusative system of alignment, in which formal marking of S is highly constrained (cf. McGregor 2007, 2010) (but in this paper I have not yet adapted glossing of *-a* accordingly).

4. Nominalization with prefix *wa(ʔ)-*

The first formal type of nominalization in Harakmbut discussed here is characterized by affixation of the nominalizing prefix *wa(ʔ)-* to the verb stem. It is functionally restricted to participant nominalization, and it mainly produces heads of NPs that can occur in any participant slot in the higher clause. In addition, it also yields forms that show modification uses. Within this first formal type, a further distinction can be made on the basis of the presence of an additional suffix, viz. *-eri* 'animate'. While *wa(ʔ)*-nominalizations suffixed by *-eri* refer to animate entities, *wa(ʔ)*-nominalizations without *-eri* have inanimate referents. The first subtype invariably involves agentive nominalization, while the second one realizes instrumental or objective nominalization (see Comrie & Thompson 2007).

4.1 Animate referents

Harakmbut morphology caters for a derivational process whereby verbs can be made into nouns denoting an animate entity that can be described as 'one who "verbs"'. This type of participant nominalization is traditionally termed 'agentive nominalization' (cf. Comrie & Thompson 2007: 336). In Harakmbut, it involves affixation of the nominalizing prefix *wa(ʔ)-* in conjunction with the suffix *-eri*

'animate' to the verb stem (see also Tripp (1976: 1), who labels *wa-* (<hua->) as a classifier). Examples are in (12) and (13).

(12) *wa-mba-yorok-eri*
NMZR-VPL-dream-AN
'dreamer'

(13) *wa-mationka-eri*
NMZR-hunt-AN
'hunter/one who hunted/one who hunts'

As can be gathered from these examples, the resultant noun need not be in an 'Agent' relationship with the verb it is derived from (cf. Comrie & Thompson 2007: 336). In (12), the noun is in an Experiencer relationship with the verb 'dream'. Also, (12) retains verbal plural marking, which indicates here plurality of action; in Harakmbut culture a dreamer is a shaman-like figure who regularly receives dreams from the spirits.

Agentive nominalizations are found to serve two functions. Firstly, they can serve as head of an NP realizing any type of argument function (cf. Tripp 1976: 2; 1995: 194). In (14), for example, the form *wamationkaeri-ta* functions as direct object of the verb *oketea*, as signalled by the accusative case marker *-ta*.

(14) *apetpet-a o-ket-e-a wa-mationka-eri-ta*
jaguar-NOM 3SG.IND-run-ITER-TRNS NMZR-hunt-AN-ACC
'The jaguar makes the hunter run.'

Secondly, agentive nominalizations can also be used to modify other nouns, and are thus functionally equivalent to relative clauses (this function is not described by Tripp 1976, nor Helberg 1984).[5] In (15) the phrase *henpu wambakaerita* is functionally equivalent to a right-adjoined relative clause, restricting the reference of the head noun it modifies, viz. *arakmbutta* (cf. Andrews 2007: 214–217); note that basic word order in Harakmbut is (not strictly) OVS (Van linden Forthcoming). In (16), the form *wamanokoteri* functions as a headless relative clause (lacking a nominal head like *arakmbutta* in (15)). It thus shows NP-use of a grammatical nominalization; it realizes a non-referential NP functioning as complement to the subject in a predicational copular clause.

(15) *arakmbut-ta iʔ-uk-i [henpu wa-mba-ka-eri-ta]$_{NMLZ}$*
person-ACC 1SG-search-1.IND string.bag NMZR-VPL-make-AN-ACC
'I am looking for the person who makes string bags.'

5. The main relativization strategy in Harakmbut involves suffixation of the finite verb form of the relative clause by *-niŋ* (Van linden Forthcoming).

(16) *wa-mba-yorok-eri* *õʔ-ẽ* *[wa-ma-no-kot-eri]*$_{NMLZ}$
NMZR-VPL-dream-AN 3SG.IND-be NMZR-VPL-(vital.centre-fall)$_{realize}$-AN
'The dreamer is one who knows many things.'

Example (15) is the most interesting one with regard to the external and internal syntax of agentive nominalization. In (15), the nominalized form is marked for accusative case. This use of the nominal category of case suggests that the action nominal has an NP-like external syntax. The notional direct object of the nominalized form (*henpu*) goes unmarked, just like inanimate direct objects in independent clauses (see Section 3). In addition, the nominalizations in (15) and (16) both retain verbal plural marking. This suggests that the internal syntax of agentive nominalizations is verb-like rather than NP-like; however, I have no examples with animate notional direct objects to bolster the argument.

Finally, it should be noted that the suffix *-eri* is also used in a productive noun-to-noun derivational process:[6] nominal bases suffixed by *-eri* come to refer to animate entities living in or coming from the place denoted by the nominal base, which can be a common noun (17a) or a proper noun (17b). The derived nouns often serve as demonyms or gentilics (see also Tripp 1995: 193). In (17b), for example, the suffix is added to the name of the native community of Puerto Luz to denote its members. Note that the names of the Harakmbut ethno-linguistic groups mentioned in Section 1 also end in *-eri*.

(17) a. *pato numba-eri*
duck forest-AN
'duck, one that is in the forest/bush'
[Sp. *pato de monte*; Lat. *Sarkidiornis melonotos*]

b. *Porto-lus-eri*
Puerto-Luz-AN
'the people from Puerto Luz'

4.2 Inanimate referents

The second type of verb-based nominalization with *wa(ʔ)-* uses no further marking, and produces nouns that refer to inanimate entities. These nouns can bear two different relationships to the verb they are derived from. In one type, the resultant nouns denote the instrument for "verbing", and in the other, they denote the result or the typical or 'cognate' object of the action designated by the verb. The

6. Helberg (1984: 445) discusses the *-eri* suffix in a section on noun-based nominalization only, describing its meaning as 'group of persons'; yet his examples include both verb-based and noun-based nominalizations. Example (17a) shows that *-eri* is not restricted to humans, but applies more generally to animate entities.

associated processes are traditionally called 'instrumental' and 'objective' nominalization respectively (cf. Comrie & Thompson 2007: 338–342). This semantic distinction does not correlate with a formal one.

The derivational process of forming "non-personal" nouns by adding the nominalizing prefix *wa(ʔ)-* to a verb stem has also been noted by Tripp (1976: 1). His examples all constitute instrumental nominalizations, although his paraphrase for (18) reads "thing that discharges, i.e. a rifle" (1976: 2). Another example is (19). Note that (19) shows noun incorporation of type I; 'head-dressing' qualifies as a "name-worthy" activity (Mithun 1984: 849). However, this type of lexical compounding does not make the nominalized verb more 'finite'.

(18) *wa-potoŋ*
NMZR-discharge
'a rifle/something for the purpose of discharging or firing' (cf. Tripp 1976: 2)

(19) *wa-ku-ot*
NMZR-head-get.dressed
'a hat/something to dress your head with'

Like agentive nominalizations, instrumental nominalizations can serve two different functions. In (20), for example, *wawedn* functions as the head of an NP realizing the direct object of the verb form *oniŋka* (see (1) in Section 2). In (21), by contrast, the nominalization serves to modify the indefinite pronoun *kate(pi)* 'something', and thus is functionally equivalent to a relative clause. This modification use is also described by Helberg (1984: 455) (unlike the NP-use). He analyses examples like (21) as purposive clauses, although the nominalized forms clearly modify nouns and thus function at NP-level. In my data, purposive subclauses operating at clause level invariably use finite verb forms marked for imperative mood and suffixed by *-po*, which signals the dependent status of the clause (see Van linden Forthcoming).

(20) *Jonas-tewapa o-niŋ-ka wa-wedn griŋgo-a*
Jonas-BEN 3SG.IND-BEN-make NMZR-lie foreigner-NOM
'The foreigner makes a bed for Jonas.'

(21) *kate=pi [wa-ka hak]*$_{NMLZ}$
something=INDET NMZR-make house
'something to make a house with, like a beam'

In (21), the notional subject of the nominalized form is left implicit (generic reference), but the notional direct object is expressed (*hak*). Like in (15) above, its lack of extra markers points to the verb-like nature of the internal syntax of this nominalization.

In addition to instrumental nominalizations, prefixation of *wa(ʔ)-* to a verb stem also produces objective nominalizations, i.e. nouns denoting the result or the typical or 'cognate' object of an action (cf. Comrie & Thompson 2007: 340–341). In (22a), the word for 'language' is construed as the 'result of saying' (there is no other lexical item available), and in (22b), the term 'alphabet' is construed as the 'result of writing' or the 'cognate object of writing'. Arguably, these examples could also be analysed as instrumental nominalizations.

(22) a. *arakmbut-en waʔ-aʔ*
person-GEN NMZR-say
'the language of the people', 'the Harakmbut language'
b. *or(oʔ)-edn wa-ma-ndoya*
1PL-GEN NMZR-VPL-write
'our letters', 'our alphabet'

Examples (22a) and (22b) form true noun phrases with genitive-marked (pro) nouns functioning as possessive determiners of lexicalized nominalizations. Although semantically these genitive forms correspond to the subjects of the verbs from which the head nouns are derived, they arguably do not form part of the nominalization. This can even be upheld for examples like (23), which features a productive indirect causation construction in which the causing event is coded as the fixed phrase "Causer-GEN *waʔ-aʔ-te*" and the caused event is coded as a main clause.

(23) *Luis-en waʔ-aʔ-te Yesika o-mba-wedn-a mbiʔigŋ mbarak-te*
Luis-GEN NMZR-say-LOC Yesica 3SG.IND-VPL-lie-TRNS fish ground-LOC
'Luis makes Yesica put the fish on the ground.' (Literally: 'At Luis's words/speech, Yesica puts the fish on the ground.')

While in cases like (23) the nominalization does refer to what the referent of the genitive-marked noun said, the fixedness of the expression and its high productivity support an analysis in terms of a one-word nominalization, as suggested by the literal translation. Note that the locative case suffix *-te* is also found on (multi-word) nominalizations with *e(ʔ)-*, with such forms functioning as an adverbial time clause (see Section 5.2.1).

The affix used to derive nouns referring to inanimate entities from verbs also serves a function in noun-based nominalization, in fact a very basic one. Harakmbut common nouns divide into two groups on the basis of their morphological status, viz. potentially free nouns and obligatorily bound nouns (Van linden Forthcoming). Whereas potentially free nouns can stand on their own as a word form, obligatorily bound ones require a nominalizing prefix to obtain independent nominal status, e.g. *wa(ʔ)-* in (24) (but also *e(ʔ)-*, see Section 5.3).

(24) *wa-mbaʔ*
NMZR-hand
'hand' (Helberg 1984: 437)

The set of bound nouns is semantically restricted to nouns that refer to parts of entities, such as (human or animal) body parts, plant parts, and landscape parts (cf. the class of *e*-nouns in Cavineña as described by Guillaume (2008: 409–416)), as well as basic shapes or qualities of entities. This set has been identified as "shape morphemes" by Hart (1963) (and adopted as such by Helberg (1984: 243)), and analysed as classifiers by Payne (1987: 35–37). I will come back to nominalization of bound nouns in Section 5.3, where it will be placed in an areal perspective.

5. Nominalization with prefix *e(ʔ)-*

The second formal type of nominalization in Harakmbut features the prefix *e(ʔ)-* appended to the verb stem. This type is predominantly used for event nominalization and only marginally for participant nominalization. In event nominalization, *e(ʔ)-*nominalizations are found to code complement as well as adverbial relations, which will be discussed in Sections 5.1 and 5.2 respectively. It will become clear that this semantic distinction has a formal correlate, in that – unlike in complement relations – the forms coding adverbial relations all feature an extra marker signalling the type of adverbial relation. However, what is shared in both types of subordination relations is that the nominalized forms retain very few – if any – inflectional verbal categories, and that they combine NP-like external syntax with verb-like internal syntax. In the latter respect, they pattern like nominalizations in Kakataibo (Valle & Zariquiey this volume) and Cahita (Álvarez this volume). As the prefix *e(ʔ)-* is also used in the citation form of verbs, nominalization with *e(ʔ)-* used in event nominalization functions will be termed 'infinitival' nominalization. Section 5.3, in turn, will concentrate on participant nominalization functions of *e(ʔ)-*nominalizations.

5.1 Complement relations

This section takes a closer look at infinitival nominalization used to code complement relations in Harakmbut. This type of infinitival nominalization is found in syntactic environments in which a core argument noun phrase is called for. Not unexpectedly, therefore, these forms either show no further morphology, or they are marked for case, specifically accusative case, flagging the direct object of the main clause verb phrase. The discussion below is organized according to the

semantic types of the complement-taking predicates that pattern with infinitival nominalization (based on Noonan 2007: 120–145).

5.1.1 *Commentative predicates*

Commentative predicates provide a comment on their complement proposition in that they express the speaker's attitudinal evaluation of the propositional content coded in the complement (Noonan 2007: 127–128). Examples include predicates expressing judgements (*be odd*, *be significant*, *be important*) or emotional reactions (*regret*, *be sorry*, *be sad*) (Noonan ibid.). These two types of commentative predicates are also found in Harakmbut, as illustrated in (25), which repeats (2), and (26) respectively.

(25) *ndak õʔ-ẽ [e-mbaʔ-tiak, ãnĩ, keme]*$_{NMLZ}$
good 3SG.IND-be NMZR-VPL-come FILLER tapir
'It is good that (, eh,) the tapirs have come.'

(26) *nduruŋ-nda ĩʔ-ẽ-ỹ, [tareʔ e-niŋ-to-tiak opudn-a]*$_{NMLZ}$
happy-NDA 1SG-be-1.IND manioc NMZR-BEN-SOC-come 2PL-NOM
'I am very happy that you (pl) have brought manioc (for us).'

In both examples, the nominalized forms constitute the head of an action nominal, containing also a noun phrase that corresponds to the subject of the verb stem. In (25) and (26), these notional subjects are not coreferential with the matrix subjects; I have no examples in which coreference does obtain. In (25), the notional subject of the nominalization, *keme*, is left unmarked, just like S-participants of independent clauses; verbal plural marking is used here to indicate plurality of the S-participant. In (26), the notional A-participant *opudna* (which itself is optional in this case according to my informant) is marked for nominative case, which contrasts with the genitive case markers found on the notional subjects of the one-word participant nominalizations in (22)–(23). We can conclude that as far as argument marking is concerned, Harakmbut infinitival action nominals do not accommodate themselves to noun phrase syntax, but have a verb-like internal syntax instead.

If we focus on the retention of inflectional verbal categories in the infinitival action nominals in (25)–(26), we are led to assume that neither tense nor aspect is retained. Even though the propositional contents denoted by the nominalizations refer to events that reached completion before the moment of the attitudinal assessment, and thus are presupposed true (Van linden 2012: 54–62, cf. Noonan 2007: 128), neither infinitival form is marked for past tense or any aspectual category. Nevertheless, (25)–(26) illustrate that infinitival nominalizations do allow for (derivational) valency-increasing morphology, as (26) has two more arguments

than (25), viz. an applied direct object *tare?* through sociative causative prefix *to-*, and an applied indirect object (implied, not overtly expressed) through the benefactive applicative prefix *niŋ-*, cp. (6) for an independent clause (Section 3).

5.1.2 *Ability predicates*

Ability predicates indicate the ability of the subject participant to carry out a particular State of Affairs. They are traditionally included in the class of modal predicates, and in many languages they take part in clause or lexical union, functioning as auxiliary verbs rather than complement-taking predicates (cf. Noonan 2007: 138–139). The type of ability that is expressed by means of infinitival nominalization in Harakmbut is that of 'learned' or 'acquired' ability; the expression of 'intrinsic' ability does not use nominalization (cf. distinction between constructions with *saber* versus *poder* in Spanish). Examples are given in (27)–(28).

(27) *ndo ĩh-nõ-põ-ẽ-ỹ* *[e-ndopih]*$_{NMLZ}$
1SG 1SG-(vital.centre-CLF:round-be)$_{know}$-1.IND NMZR-swim
'I am able to swim.'

(28) *ĩh-nõ-põ-ẽ-nde-y* *[e-mba-tuk-e?*
1SG-(vital.centre-CLF:round-be)$_{know}$-ALREADY-1.IND NMZR-VPL-plant-ITER
tare?]$_{NMLZ}$
manioc
'I already learned how to sow (a field of) manioc.'

As can be expected from the semantics of the complement relation, the notional subject of the action nominal is coreferential with the syntactic subject of the matrix, and has been equi-deleted (cf. Noonan 2007: 75–76). It can even be stated that in constructions like (27)–(28) finite forms of *ẽnõpõẽ* can transfer their subject selection to the infinitival nominalization. This possibility signals that *ẽnõpõẽ* in its acquisition of ability sense has moved already some way on the auxiliation path proposed by Heine (1993: 58–66). A more detailed discussion of whether *ẽnõpõẽ* should be analysed as a complement-taking predicate or auxiliary here is beyond the scope of this paper.

The observation that *ẽnõpõẽ* constructions like (27)–(28) are semantically restricted to acquired ability contexts can be explained by the verb's polysemy, which is illustrated in (29). The first finite form of *ẽnõpõẽ* in (29) functions as a knowledge or acquisition of knowledge predicate that patterns with a sentence-like complement featuring a different subject and finite verb form (see Van linden Forthcoming). This second finite form, by contrast, functions as an ability predicate and patterns with a nominalized form whose notional subject has been equi-deleted. It is probable that the (acquisition of) knowledge sense of *ẽnõpõẽ* blocks

further semantic extension to the sense of intrinsic ability. In addition to '(get to) know', *ẽnõpõẽ* is also observed to express meanings like 'think' (propositional attitude predicate), 'hope' (desiderative predicate) and 'feel' (immediate (sensory) perception predicate).

(29) *wa-siʔ-po* *õ-nõ-põ-ẽ-mẽ*
NMZR-(peel-CLF:round)$_{child}$ 3SG.IND-(vital.centre-CLF:round-be)$_{know}$-REC.PST
kuwa [e-ndopih]$_{NMLZ}$ *õ-nõ-põ-ẽ*
dog NMZR-swim 3SG.IND-(vital.centre-CLF:round-be)$_{know}$
'The child learned that dogs are able to swim.'

Taking into account the differences in semantic properties of modal versus commentative complement relations, I assume that the nominalized forms patterning with *ẽnõpõẽ* will tend to show retention of fewer verbal categories than those patterning with commentative predicates. As described in more detail in Van linden (2012: 203–207),[7] drawing on Noonan (2007) and Cristofaro (2003), modal relations involve a higher degree of semantic integration and semantic dependency than commentative relations, which explains why they are frequently observed to combine with reduced complement types across languages. In Harakmbut, the main formal distinction lies in the treatment of the notional subject of the infinitival nominalization (equi-deleted vs. overtly expressed). The retention of tense is not expected in the case of modal relations, and this is borne out by the Harakmbut data. The same is expected for inflectional aspectual categories. Note that iterative aspect, coded by suffix *-e*, is a derivational type of aspect in Harakmbut, as evidenced by its influence on word stress (see Van linden Forthcoming). In (28), the presence of the iterative marker in *embatukeʔ* can be explained by the nature of the activity learned, which was not the planting of just one manioc seed, but the sowing of a whole field of manioc (i.e. planting repeatedly), which is even enhanced by the presence of verbal plural marking.

5.1.3 *Immediate perception predicates*

Immediate perception predicates name the sensory mode by which the main clause subject participant directly perceives the event referred to in the complement (cf. Noonan 2007: 142). Harakmbut has three different formal strategies to code complements of this type of predicate, one of which involves infinitival nominalization. This strategy is illustrated in (30) and (31).

7. See also Van linden & Davidse (2009: 178–180).

(30) *mboerek o-tiaway-tuy apetpet-ta [e-arak ken-en*
man 3SG.IND-see-DIST.PST.NVIS jaguar-ACC NMZR-kill;hit 3-GEN
wã-tõ-ẽ-ta]$_{NMLZ}$
NMZR-SOC-be-ACC
'The man saw the jaguar attacking his wife.'

(31) *ken-en nãŋ-a o-pẽ-ẽ-tuy apetpet-ta*
3-GEN mother-NOM 3SG.IND-ear.canal-be-DIST.PST.NVIS jaguar-ACC
[e-kwiriʔ-min ken-en wa-yombu-ta]$_{NMLZ}$
NMZR-brain-suck 3-GEN NMZR-daughter-ACC
'The mother heard the jaguar sucking the brains of her daughter.'

While in (30), the subject participant becomes aware of the (horrible) event coded in the complement clause by seeing it happen, in (31) the subject participant relies on auditory perception. Nevertheless, the examples share the same syntactic structure; in both instances the subject of the complement proposition is marked for accusative case (*apetpet-ta*), which suggests that it is syntactically treated as the (animate) direct object of the complement-taking predicate, although semantically it is the entire event that is perceived (cf. Kirsner & Thompson 1976). While the notional subjects of the nominalized forms receive their marking from the matrix clause, the direct objects of the dependent clauses seem to receive it from the nominalized forms; in both examples they have animate referents and are marked for accusative case as well (*kenen wãtõẽ-ta* in (30); *kenen wayombu-ta* in (31)). It can be noted that the latter participant is treated as direct object by virtue of noun incorporation of type II in Mithun's (1984) typology. Incorporation of the bound noun *wa-kwiriʔ* 'brain(s)' vacates the position of object of *e-min* 'suck', to which the possessor of the incorporated noun is advanced, i.e. *kenen wayombu-ta* (cf. Mithun 1984: 857–858). This type of NI is comparable to the applicative marking observed in (26), as it basically serves as a valency-increasing mechanism. Finally, as can be expected from a complement relation that shows a high degree of semantic integration and dependency (cf. Noonan 2007: 142–144), the nominalized forms show no retention of inflectional verbal categories.

Comparing the infinitival nominalization strategy with the other two strategies available in Harakmbut, I hypothesize that the former is dedicated to contexts of non-deliberate perception that do not involve first person singular matrix subjects that are coreferential with the direct object of the complement proposition. For these special first person contexts, Harakmbut speakers produced constructions with sentence-like complements. The third strategy, in which the complement takes the form of a finite relative clause with the direct object of the matrix as its antecedent, seems to be restricted to contexts of deliberate perception, focus-

sing on how exactly the perceived action proceeds. As these alternative strategies do not feature infinitival nominalization, no further examples are provided.

5.1.4 *Desiderative predicates*

Desiderative predicates express a desire that the State of Affairs coded in the complement be actualized (Noonan 2007: 132). Of the three subtypes distinguished by Noonan (2007: 132–135), it is only the *want*-class that patterns with infinitival nominalization in Harakmbut. What may strike the reader immediately when considering examples (32)–(34) is that the nominalized form is marked for accusative case (by *-ta*, cf. Table 2). Among the complement relations coded by infinitival nominalization, the desiderative relation is the only one in which the nominalized complement occurs with the nominal category of case, and thus most clearly features NP-like external syntax. However, this coding pattern is unexpected in view of the animacy constraint on O-marking, as events are inanimate entities. I have no explanation for this (but see Section 5.2.1). Incidentally, desiderative relations form the only complement relation described by Tripp (1976: 3; 1995: 216) and Helberg (1984: 360, 451–452). Both provide examples with *eʔpak* 'want, love' as complement-taking predicate, cf. (32)–(34), and note that in such cases the infinitive functions as object.[8]

(32) *[e-kore-ta]*$_{NMLZ}$ *on-pak-me* *o-wi-nok*
NMZR-turn.back-ACC 3PL.IND-want-REC.PST 3SG.IND-rain-BECAUSE
'They wanted to go back because it was raining.'

(33) *pa [eʔ-wa-ta]*$_{NMLZ\_1}$ *iʔ-pak-Ø* *[e-mationka-ta* *ndoʔ-ere]*$_{NMLZ\_2}$*?*
Q NMZR-go-ACC 2SG-want-DUB NMZR-hunt-ACC 1SG-COM
'Do you (sg) want to go and hunt with me?'

(34) *[mbuttinda e-ma-n-a-ta]*$_{NMLZ\_part1}$ *ih-pak-i* *[opudn-ta]*$_{NMLZ\_part2}$
truth NMZR-VPL-SPAT-say-ACC 1SG-want-1.IND 2.PL-ACC
'I want to tell you (pl) the truth.'

In examples (32)–(34), the notional subject of the nominalization is coreferential with the matrix subject, and has been equi-deleted. Equi-deletion also obtains in my other examples with identical subjects in main and complement clause. In addition to these contexts, my data include two examples in which the matrix and complement proposition have different subjects, one of which shows infinitival nominalization, cf. (35).

8. Like in the case of complements of perception predicates, however, it should be noted that semantically it is the whole event coded by the complement clause that functions as direct object of the desiderative predicate.

(35) *mboerek oʔ-pak-me* *[e-niŋ-to-tiak-ta* *keme]*$_{NMLZ}$
man 3SG.IND-want-REC.PST NMZR-BEN-SOC-come-ACC tapir
'The man wanted them to bring along tapir for him.'

In (35), the matrix subject is coreferential with the (applied) indirect object of the complement proposition. The same situation holds in the other example, which uses a sentence-like complement instead, whose verb is marked for imperative mood. However, similarly to the distribution of complementation strategies used to code immediate perception relations, this case involves a first person singular matrix subject that is coreferential with a non-subject argument in the complement proposition. More systematic research is needed to check whether this first person singular context is the odd one out in other areas of grammar as well.

While the external syntax of infinitival nominalizations functioning as desiderative complements is NP-like, their internal syntax looks verb-like. Example (34) is most informative in this respect: the notional primary object of the nominalized verb is accusative-marked (*opudn-ta*), like human primary objects are treated in independent clauses. Note that this constituent of the multi-word nominalization is separated from the others by the main verb (hence *NMLZ_part1* and *NMLZ_part2* in subscript). The notional object of the nominalized verb in (35) is left unmarked, just like objects referring to dead (and lower order) animals are in independent clauses.

Finally, with desiderative relations showing a high degree of semantic integration and dependency, much like the modal and immediate perception relations (cf. Noonan 2007: 142–144), no retention of inflectional verbal categories is expected. This expectation is borne out in my data. Again, derivational categories prove different than inflectional ones, with valency-increasing morphology appearing in example (35) (cp. (26)).

5.2 Adverbial relations

While the previous section focussed on infinitival nominalization serving to code complement relations (NP-use of event nominalizations), this section concentrates on infinitival nominalization used to signal adverbial relations (modification use of event nominalizations), thus carrying information on the circumstances of the main event in which the core argument participants engage. This function is formally reflected in the presence of extra markers on the nominalized form. Importantly, all of the markers observed are also found to occur on underived nouns, though they are not all case markers. This feature points to the external syntax of the nominalized forms being NP-like. In what follows, I will discuss

the use of infinitival nominalization to code temporal, concessive, conditional and locative relations respectively.

5.2.1 *Temporal relations*

In Harakmbut all types of temporal relations are expressed through non-finite clause types, and a number of them use infinitival nominalization followed by the suffixes *-te* or *-anda*. I hypothesize that these suffixes are in complementary distribution, with *-te* used in contexts with different subjects in matrix and subordinate clause, cf. (36)–(37), and *-anda* being restricted to same-subject contexts, cf. (38)–(39). As the function of *-te* in the nominal domain, i.e. that of a locative case marker (see *bisikleta-te* in (9), Section 3), is semantically related to that in infinitival nominalization, both being dedicated to spatio-temporal location, *-te* has kept its locative case gloss in the examples below. Since the function of *-anda* in the nominal domain (see Section 3) seems at first sight to be unrelated to the function it has in (38)–(39), I first tentatively provide a separate gloss (but see below).

(36) *siʔnoŋ oʔ-tay-on-me [apane*
baby 3SG.IND-sleep-PFV.NVOL-REC.PST grandfather
e-n-mba-tiapak[9]*-te]*$_{NMLZ}$
NMZR-SPAT-VPL-narrate-LOC
'The baby fell asleep while the grandfather told him a story.'

(37) *noŋ-ok-a 2010 [e-tiak-te An]*$_{NMLZ}$
other-period-ADV 2010 NMZR-come-LOC An
o-to-k-ka-me-y mbiʔigŋ
1PL.EXCL-SOC-SEPARATION-do-REC.PST-1.IND fish
'Last year, in 2010, when An came, we fished with her.' (spontaneous speech)

(38) *[on-a oroʔ-ta eʔ-uk-anda]*$_{NMLZ}$ *tiaway-we õʔ-ẽ-nẽ*
2SG-NOM 1PL-ACC NMZR-search-SIMUL.SS see-NEG 1<>2SG-be-IND
'While you (sg) are looking for us, you (sg) don't find us.'

(39) *[pomelo e-tipit-anda]*$_{NMLZ}$ *ih-mbaʔ-teŋ-me-y*
grapefruit NMZR-peel-SIMUL.SS 1SG-hand-cut-REC.PST-1.IND
'I cut my hand while I was peeling a grapefruit.'

Both types have been noted by Tripp (1976), who does not define the distribution of the suffixes as explicitly, but does hint at it. Specifically, Tripp's (1976: 9–10) description of the *-anda* type mentions its relation to the subject of the main

9. Helberg (1984: 461) proposes a further morphological analysis of *-tiapak*.

clause; the verbal noun[10] in *-ada* "gives the circumstance or condition related to the subject". Tripp (1976: 8–9) treats infinitival nominalization with *-da* separately from that with *-ada*,[11] and is less precise about its function, saying it indicates "the circumstance of the lexical predicate". However, what is striking is that all the examples he provides feature verbs whose root ends in *a* (e.g. *-ka* 'do', *-wa* 'go'). This suggests that suffixation of *-anda* to stems ending in *a* involves loss of one vowel segment. On infinitival nominalization with *-te*, Tripp (1976: 4) merely states that it yields a temporal noun denoting a simultaneous event. In the grammatical notes accompanying his dictionary, Tripp (1995: 216) does state explicitly that *-te* is used in different-subject contexts. As illustrated in (40) below, we will see that the *-te* type is not restricted to simultaneous events, but can also be used to code anterior events. Incidentally, Helberg (1984: 451) considers this to be the main function of infinitival nominalization with *-te* (he does not discuss forms in *-anda*).

While all my examples with *-anda* and the majority of those with *-te* denote simultaneous events, some of those with *-te* refer to events that are anterior to the main clause event. As exemplified in (40), the nominalized form carries an extra marker, viz. *-nde*, which has the specific temporal adverbial meaning of 'already'.

(40) *arakmbut on-mba-uk-me wandey-ta, heridos,*
person 3PL.IND-VPL-search-REC.PST wounded-ACC wounded
wandey-ta [taka e-mba-wa-nde-te]$_{NMLZ}$
wounded-ACC Taca NMZR-VPL-go-ALREADY-LOC
'The Harakmbut looked for the wounded after the Taca (people) had gone.'

Example (40) thus indicates that infinitival nominalization with *-te* retains markers that express temporal adverbial meanings. On forms with *-anda*, by contrast, no such markers are attested in my data. The verbal plural category is also retained in (40), just like in other types of verb-based nominalization.

If we take a closer look at how the notional arguments of the nominalized forms are marked, we can observe that these use the same markers as canonical main clause arguments. An interesting example in this respect is (38), in which the matrix verb form is highly ambiguous (on referential obscurity in the verbal

10. While Tripp (1976) uses the Spanish term *sustantivo verbal* ('verbal noun'), it would be more accurate to describe the cases in Sections 5.1 and 5.2 as *action nominals* (cf. Comrie & Thompson 2007: 343).

11. Tripp (1976) writes <(a)da>, while I represent the suffix as <(a)nda>. Tripp (1995: 12) does recognize the presence of a nasal sound in the suffix, but prefers not to represent this in writing. He states that <d> is pronounced as [^{n}d], with the plosive being prenasalized, in syllable-initial position. The same holds for <tada> in Section 5.2.2 below. In my analysis, [nd] is a post-stopped allophone of /n/ preceding oral vowels (see Van linden Forthcoming).

argument marking system in Harakmbut, see Van linden 2014). It appears that this referential ambiguity is resolved in the nominalized clause, with two case-marked personal pronouns preceding the nominalized form, which is in turn followed by the main clause verb phrase. The notional subject of the nominalized form is marked for nominative case (*on-a*), while the direct object is marked for accusative case (*oroʔ-ta*). In all other examples, no case marking is found on the notional arguments of the nominalized forms, but this absence of marking is no different from what would be the case in independent clauses: S-participants go unmarked (cf. (37), (40)), just like non-focal animate A-participants (cf. (36)) and inanimate O-participants (cf. (39)) (see Section 3). All of this leads to the conclusion that the internal syntax of infinitival nominalization with *-te* and *-anda* is more verb-like than NP-like.

Yet, the distribution of the 'temporal' suffixes in other domains of Harakmbut grammar points to the nominal character of the external syntax of the nominalized types looked at here. Both suffixes are also found to occur on underived nouns. Within the nominal domain the suffix *-te* functions as a locative case marker, cf. (9) above (see also Helberg 1984: 439; Tripp 1995; 196); in addition to spatial location, it is also used to express temporal location, e.g. *agosto-te* 'in August'. The suffix *-anda* is also found on underived nouns, specifically as a morphologically complex suffix combining nominative case suffix *-a* with the focus marker *-nda*, cf. *Lupe-a-nda* in (5) above. This morphological complexity may be key to an alternative analysis of *-anda* suffixed to infinitival nominalizations. Taking into account the same-subject restriction of infinitival nominalizations with *-anda* on the one hand, and the function of *-nda* in adnominal modification contexts (see Section 3) on the other, the suffix might as well be analysed as a complex suffix comprising the nominative suffix *-a* and the general modifier suffix *-nda*, as presented in (41).

(41) *[e-waʔ-e-a-nda]*$_{NMLZ}$ *ih-kot-me-y*
NMZR-go-ITER-NOM-NDA/MOD 1SG-fall-REC.PST-1.IND
'I fell while I was walking.' (Literally: 'I, who was walking, fell', or 'I fell walking.')

In this alternative analysis, the infinitival nominalization is nominative-marked so as to indicate that the event denoted by the nominalized form specifies a circumstance related to the subject of the main event. In Shibatani's (this volume) terms, the parsing of *-anda* in (41) points to modification use of a participant nominalization, while the one in (38)–(39) points to modification use of an event nominalization. The alternative analysis is corroborated by what is observed for the suffix *-tanda* in (44), Section 5.2.2. More generally, this proposal assumes that the external syntax of nominalizations differs from that of underived nouns in terms of the differential/optional nature of case marking. While S-participants in independent

clauses tend to go unmarked, infinitival nominalizations with *-anda* are used to modify any main clause subject, whether it be an S-participant (cf. (41), or a (focal or non-focal) A-participant (cf. (38)). Similarly, nominalizations functioning as O-participants of desiderative predicates also flout the animacy constraint on case-marking; as discussed in Section 5.1.4, these do carry accusative marking, in spite of their inanimate nature.

5.2.2 *Concessive relations*

A second type of adverbial relation that is coded by infinitival nominalization in Harakmbut is that of concession. In this type, the nominalized forms are suffixed by *-tanda*, as is illustrated in (42) and (43). Example (43) shows that this suffix can be used in different-subject contexts (*ẽʔẽtanda*) as well as same-subject contexts (*embaukpaktanda*).

(42) *[sik-yo ẽʔ-ẽ-tanda*[12]*]*$_{\text{NMLZ}}$ *o-mbewik-me-ne Porto-lus-yo*
black-LOC NMZR-be-CONC 1PL.INCL-go.up-REC.PST-IND Puerto-Luz-LOC
'Although it was already dark/night, we went up(river) to Puerto Luz (by canoe).'

(43) *[lus ẽʔ-ẽ-tanda i e-mba-uk-pak-tanda]*$_{NMLZ}$
light NMZR-be-CONC and NMZR-VPL-hot-VBZ-CONC
wa-siʔ-po on-mba-tay-mbedn ndak-a
NMZR-(peel-CLF:round)$_{\text{child}}$ 3PL.IND-VPL-sleep-ALL.NIGHT good-ADV
'In spite of the light and the heat (although they felt hot), the children slept well all night.'

The examples above do not show retention of any inflectional verbal category, nor does any other example in my data. With regard to argument marking, example (43) suggests that the internal syntax of infinitival nominalization with *-tanda* is verb-like, as the notional subject of the first form (*lus*) is unmarked, like the subject of independent existential constructions (S-participants, see Section 3).

The availability of infinitival nominalization with *-tanda* in both different-subject and same-subject contexts is also observed by Tripp (1976: 10–12), but he does not attribute concessive semantics to it. Instead, he argues that the construction indicates "the circumstance or condition of the related predication"; his examples do not lend themselves well to a concessive interpretation either. Similarly, Helberg (1984: 471–472) analyses infinitival nominalization with *-tanda* as expressing the temporal relation of simultaneity. His examples include both different-subject and same-subject contexts, but do not seem to involve concessive linking. However, in later work, Tripp (1995: 216) notes that *-tada* (and *-ada*, cf.

12. The *eʔ-* prefix has nasal quality here through nasal spreading from the verb root *-ẽ* 'be'.

Section 5.2.1) frequently has an "adversative" meaning. Some of the examples adduced allow for a concessive interpretation as well.

In addition, Tripp (1976: 10–11) notes that "verbal nouns" in *-tanda* can also "describe the circumstance of the object of the related predication." In his example (44), I believe that *-tanda* should be analysed further into *-ta-nda* (-ACC-NDA/MOD) along the same lines as my alternative analysis for *-a-nda* in (41), Section 5.2.1. This analysis is not proposed by Tripp, although he calls *-tada* a "complex suffix" further below (1976: 12).

(44) *ken y-ok-wek-po ndo*
then 1SG-SEPARATION-wound.with.arrow-DEP 1SG
[eʔ-ti-mon-an-ta-nda]$_{NMLZ}$
NMZR-UP-flee-PFV.VOL-ACC-NDA/MOD
'Thus I pierced the animal that he had lost.' ['Then I pierced the escapee, viz. an animal/person that fled (from him) from high up.' AVL]
(Tripp 1976: 11, Example 57; adapted spelling; my morpheme breaks and glosses)

Tripp (1976: 10–11) gives two more examples that could be analysed similarly to (44); I have not encountered any comparable example in my own data so far. Semantically, in cases like (44) – like in the other examples given by Tripp – no concessive relation holds between the nominalized event and the main clause event, but rather a general modification relation targeting a main clause participant. The nominalized form in (44) thus differs from those in (42)–(43) in involving participant nominalization rather than event nominalization. Finally, it should be noted that the *-tanda* suffix is only to be found on underived nouns when it is indeed further analysable into *-ta-nda* (-ACC-NDA), but not in its monomorphemic form (see (4a) in Section 3). This implies that for concessive infinitival nominalization with *-tanda* we cannot be as confident about its external syntax being NP-like as for, e.g., infinitival nominalization with *-te*.

5.2.3 *Conditional relations*

Infinitival nominalization is also used in Harakmbut to signal conditional relations between events. In these cases the nominalized forms are suffixed by *-nãỹõ*, irrespective of their semantic subtype. The construction in (45) exemplifies a reality condition, while that in (46) instantiates an unreality condition of the predictive subtype (the semantic classification adopted here is the one proposed by Thompson et al. (2007: 254–262)).

(45) *[eʔ-wi-nãỹõ]*$_{NMLZ}$ *mbaʔa-we ĩh-ẽ-ãpo-y*
NMZR-rain-COND work-NEG 1SG-be-FUT-1.IND
'If it rains, I won't work.'

(46) *[aya-nda, aya-nda e-mba-pe-nãỹõ]*$_{NMLZ}$ *o-yok-i* *gayeta*
all-NDA all-NDA NMZR-VPL-eat-COND 1>2SG.IMP-give-1.IMP biscuit
'If you (sg) eat (up) everything, I'll give you (sg) a biscuit.'

The nominalized forms in (45) and (46) share the same formal make-up featuring the nominalizing prefix *e(ʔ)-* and the conditional suffix *-nãỹõ*, but the verb forms in the main clauses are different. The reality condition construction in (45), which refers to a habitual situation, has a main clause verb form marked for indicative mood and future tense (*ĩhẽãpoy*), while the predictive (unreality condition) construction in (46) contains an imperative verb form (*oyoki*, literally 'I should give you'). However, other predictive examples in my dataset also show indicative future forms. In addition to reality conditions and predictive ones, my data include one counterfactual situation, which is given in (47).

(47) *[eʔ-wi-me-nãỹõ]*$_{NMLZ}$ *mbaʔa-we ĩh-ẽ-ỹ* *taʔmba*
NMZR-rain-REC.PST-COND work-NEG 1SG-be-1.IND swidden
'If it had rained (yesterday), I would not have worked (on) the swidden.'

In (47), the nominalized form has an extra marker compared to the forms in (45) and (46), viz. the recent past tense marker *-me* (cf. (5)–(6) in Section 3). Of all the nominalized forms available in Harakmbut, formed with *wa(ʔ)-* (Section 4) or *e(ʔ)-* (Section 5), this subtype is the only one that shows retention of the verbal category of tense, or at least the value of recent past. The use of a past tense marker in a counterfactual construction can readily be explained in terms of the origins of counterfactuality (cf. Van linden & Verstraete 2008), and it also forms a cross-linguistically recurrent formal feature of this type of conditional construction (cf. Van linden 2004).

A third type of unreality condition, viz. hypothetical conditions, is illustrated in Helberg (1984: 464). As can be seen in (48), the nominalized form has the same formal make-up as the forms used in reality and predictive conditions, but the main clause verb form contains the modal suffix *-ipot*, which denotes future-oriented possibility and invariably combines with verbal argument markers of the dubitative mood paradigm.[13]

(48) *[eʔ-ti-kot-nãỹõ]*$_{NMLZ}$ *oʔ-mon-ipot*
NMZR-UP-fall-COND 1PL-flee-FUT.POSS
'If they discovered us, we would flee.'
(Helberg 1984: 464; adapted spelling; my glosses)

13. Tripp (1995: 222) characterizes *-iput* as a future subjunctive form restricted to 1SG subjects. However, my data include examples with all person and number combinations. Helberg (1990: 239) attributes desiderative meaning to *-ipot* in addition to the meaning of possibility.

Conditional infinitival nominalization has also been described by Tripp (1976: 6) as taking the form of *e(ʔ)-* + verb stem + *-nãỹõ*. One of his examples shows retention of the temporal adverbial marker *-nde*, viz. (49) below, which we have also observed for infinitival nominalization with *-te* (cf. (40) above).

(49) *[on e-k-waʔ-nde-nãỹõ]*$_{NMLZ}$ *mo-mba-arak-aʔ-∅*
2SG NMZR-separation-go-already-COND 3>1/2PL-VPL-kill-TRNS-DUB
'If you (sg) go (and leave us) now, they will kill us.'
(Tripp 1976: 6, Example (31); adapted spelling; my morpheme breaks and glosses)

Example (49) is not only informative with regard to the formal properties of the nominalized form, it also shows what form the notional subject of the nominalized form takes. Like in the case of other types of infinitival nominalization, the notional subject is unmarked here. As S-arguments are typically left unmarked in independent clauses as well (see Section 3), we can conclude that conditional infinitival nominalization has verb-like internal syntax.

In addition, example (49) is also interesting as it challenges my earlier generalizations on the main clause verb phrase in predictive conditional constructions. Specifically, its main clause verb form is not marked for imperative or indicative mood, but for dubitative mood (see also Tripp 1976: 7, ex. (33)); an imperative form in (49) would end in *-e*, while an indicative one would end in *-ne*; cf. Van linden 2014). On the basis of Tripp's and my own examples, I hypothesize that mood marking in the main clause verb phrase of predictive conditionals is determined by the person category of the grammatical subject, with third person subjects being restricted to dubitative mood forms, while first person subjects take indicative or imperative verb forms. This hypothesis ties in with the circumstance that we are intrinsically unable to predict how others will react if a certain condition obtains, while we can be rather confident of what we would or should do. Of course, more data are needed to verify this hypothesis.

Finally, like the suffixes used in temporal infinitival nominalization, the suffix *-nãỹõ* is also found on underived nouns (and pronouns). An example is given in (50) below, in which the conditional suffix attaches to the noun *wambo* 'youngster'. The elicitation stimulus was intended to yield a modal expression (of permission), which it did, but my language consultant additionally used a conditional construction to spell out the modal agent of the (permitted) action.

(50) *wambo-nãỹõ mã-õʔõ-ndik õnʔ-ẽ noŋ-ti-a wẽʔẽỹ-ỹõ*
youngster-COND VPL-bathe-POT 3PL.IND-be (other-on)$_{one}$-ADV water-LOC
'Youngsters can bathe in the river on their own.' (Literally: 'If they are youngsters, they can bathe in the river on their own.')

In conclusion, the data analysed so far suggest that conditional infinitival nominalization fits the generalization of combining verb-like internal syntax with NP-like external syntax.

5.2.4 *Locative relations*

The last type of infinitival nominalization to be discussed here codes the adverbial relation of location. For this type of adverbial relation, nominalization constitutes only one of the three strategies available in Harakmbut (in addition to relative and main clause constructions). Unfortunately, I can only reproduce a case in point from Tripp (1976), as my own data do not include any instances. In (51) below, the nominalized form is marked for locative case by the suffix *-yo*.

(51) *[Kereto e-n-pa-wedn-yo]*$_{NMLZ}$ *oʔ-ey-wa-po*
Kereto NMZR-SPAT-CLF:stick-lie-LOC 1PL-ʔ-GO-DEP
'We went to the place where Kereto was lying ill.'
(Tripp 1976: 6, Example (27); adapted spelling; my morpheme breaks and glosses)

The locative marker found on the nominalized form in (51) is also commonly appended to underived nouns, as illustrated in example (50) above (*wẽʔẽỹ-ỹõ*) (see also Table 2). This observation testifies to the external syntax of this locative infinitival nominalization being NP-like. The nominalized form shows no retention of inflectional verbal categories, but it does feature noun incorporation of type IV (cf. Mithun 1984), with verbal classifier *-pa* '(shape of a) stick' categorizing the S-participant of the nominalized form; the sick person is stick-like in that he cannot move anymore. Its notional subject goes unmarked (*Ketero*) (like S-participants in independent clauses, see Section 3), which, like in the other types of infinitival nominalization, points to verb-like internal syntax.

5.3 Participant nominalization

While the majority of cases of nominalization with *e(ʔ)-* form action nominals from predicates or propositions, which in turn function as complements or circumstantial adjuncts to the main clause event, some cases just form nouns from lexical verbs. The most straightforward case is the formation of the citation form of verbs. In eliciting verbs using Spanish infinitives like *correr* 'run' to gather vocabulary items, cf. (52), I invariably obtained forms in *e(ʔ)-* from all my consultants.[14] Another example is in (53), which is similar to Tripp (1976: 3, Example (11)).

14. In view of this observation and further analysis of the verbal paradigms, I believe that all verb roots are obligatorily bound morphemes.

(52) *eʔ-ket*
NMZR-run
'(to) run'

(53) *oʔ-sot-me* *eʔ-wi-a*
3SG.IND-make.wet-REC.PST NMZR-rain-NOM
'The rain made him wet.'

In (53), the nominalized form *eʔwi-a* functions as the (head of the) subject NP of the clause; it is even marked for nominative case, which is to be expected on the basis of its semantic properties and grammatical role, i.e. it is an inanimate A-participant (see Section 3). My data include a similar example with *e-digŋpak* '(have) fever'. In any case, nominalized forms like *eʔwi-a* in (53) refer to inanimate entities, which can be conceived of as the result of the action denoted by the verb stem. We can therefore conclude that nominalization with *e(ʔ)-* can also result in objective nominalization, just like nominalization with *wa(ʔ)-* (see Section 4.2 above).

Also, like in the case of nominalization with *wa(ʔ)-*, the affix used to derive nouns referring to inanimate entities from lexical verbs serves a basic function in noun-based nominalization. In the same way as *wa(ʔ)-*, *e(ʔ)-* also attaches to bound nouns to produce independent nouns. Example (54a) forms a minimal pair with (24) above. While (24) yields the independent noun 'hand', the nominalization in (54a) yields the independent noun 'leaf of a plant or tree' (see also Helberg 1984: 254, 437), which has of course a shape very similar to that of a hand and also forms an upper extremity of a living body. In its noun-based nominalization function, then, *e(ʔ)-* serves to produce the "citation form" of (a small set of) bound nouns; another example is *eʔ-pu* 'bamboo'. Note that in specific construction types, i.e. those featuring adnominal modifiers that obligatorily precede the nominal head in continuous noun phrases, bound nouns can phonologically fuse with their modifier – with the nominalizing prefix being dropped (see Van linden Forthcoming). Absence of the nominalizing prefix is also observed in word formation processes like compounding; in (54b) the bound noun root *-mbaʔ* attaches to the adjectival root *pay* 'bitter' to form a(n) (independent) compound noun denoting 'tobacco'.

(54) a. *e-mbaʔ*
NMZR-hand;leaf
'leaf of a plant or tree' (Helberg 1984: 437)

b. *pay-mbaʔ*
bitter-hand;leaf
'tobacco'

The two nominalizing prefixes studied here thus share the function of lending independent status to a set of inalienably possessed nouns (however, note that the set of morphologically bound nouns does not exhaust the set of inalienably possessed nouns, as for instance a number of kinship terms do not constitute bound nouns, e.g. *nãŋ* 'mother', *pagŋ* 'father'). The very basic nature of this function possibly suggests that it may have formed the diachronic source for its function in verb-based nominalization.[15] The morphological boundedness of verb roots and bases may have facilitated this development.

The prefixes *wa(ʔ)-* and *e(ʔ)-* thus are competitors in noun-based nominalization, but at this stage I cannot say what factors exactly determine their distribution within the set of bound nouns. Cases like *-mbaʔ* which combine with the two prefixes seem to be very infrequent. In general, the prefix *wa(ʔ)-* is used for many more bound nouns than *e(ʔ)-*, but their respective host classes do not seem to differ in terms of semantic properties like having animate versus inanimate possessors. Rather than *wa(ʔ)-*, it is the less frequent prefix *e(ʔ)-* that is more interesting from an areal perspective. Specifically, it is formally and functionally similar to the dummy noun prefix *e-* in Cavineña and other Tacanan languages (Guillaume 2008: 409–416). In addition, it is comparable to the semantically empty noun formative *e-* in Kwaza, which serves as "a noun formative to lend independent status to classifiers" (van der Voort 2005: 397). In fact, Crevels & van der Voort (2008) identify the availability of a semantically empty noun formative root taking the form of *e-/i-* as an areal feature characteristic of the Guaporé-Mamoré region.

6. Conclusion

This paper has investigated (non-finite) verb-based nominalization in the Amarakaeri/Arakmbut variety of Harakmbut, with some excursions to noun-based nominalization. Within these two types, two formal subtypes have been distinguished. While in noun-based nominalization the two formal types share the same function but show a skewed distribution of frequency (or size of host class) (Section 5.3), in verb-based nominalization they show skewed distributions of

15. The functions of the nominalizing prefixes can be compared to two functions exhibited by classifiers in multiple classifier systems in North West Amazonian languages. In Bora-Miraña, for instance, classifiers can transform mass nouns into countable nouns as well as derive nouns from verbs (Seifart 2007). Aikhenvald (2000: 220–221) further lists Guahibo, Tucano and Tariana as examples in point. It is important to note, however, that the classifiers in these languages carry a specific semantic load (and are used in up to five classifier environments) whereas the Harakmbut prefixes are semantically empty (they only have the functional value of turning a bound noun into an independent nominal, or a verb base into a nominalization).

functions. That is, whereas nominalization with *wa(ʔ)-* is restricted to participant nominalization and is predominantly used to produce nouns for NP-use (typically one-word nominalizations), nominalization with *e(ʔ)-* is mainly used for event nominalization (typically multi-word nominalizations). Table 3 summarizes the main findings on verb-based nominalization; the numbers refer to the examples given in Sections 4 and 5 above.

Table 3. Forms and functions of verb-based nominalization in Harakmbut (-X refers to any ending that an underived noun can take)

Formal type (prefix)	**Extra suffix**	**One-word nominalization [type of participant NMLZ]**	**Multi-word nominalization [subordinate type]**	**Use**	**Event/ Part. NMLZ**
wa(ʔ)-	*-eri-X*	Noun with animate referent: (12)–(13)–(14) [agentive]	(15) [relative]	Mod	Part.
			(16) [relative]	NP	Part.
	-X	Noun with inanimate referent: (18)–(19)–(20) [instrumental] (22)–(23) [objective]	(21) [relative]	Mod	Part.
e(ʔ)-	–	(52) [citation form of verb]	(25)–(26) [commentative complement]	NP	Event
			(27)–(28)–(29) [ability complement]	NP	Event
			(30)–(31) [perception complement]	NP	Event
	-X	Noun with inanimate referent: (53) [objective]	–		
	-ta	–	(32)–(33)–(34)–(35) [desiderative complement]	NP	Event
	-te/-yo (LOC)	–	(36)–(37)–(40) [temporal adverbial]	Mod	Event
			(51) [locative adverbial]	Mod	Event
	-a-nda	–	(38)–(39)–(41) [temporal adverbial]/[relative]	Mod	Event/ Part.
	-tanda	–	(42)–(43) [concessive adverbial]	Mod	Event
	-nãỹõ	–	(45)–(46)–(47)–(48)–(49) [conditional adverbial]	Mod	Event

Table 3 bears out that the formal and semantic categories distinguished cross-cut each other. However, there is no 'complete' mismatch between formal and semantic categories, as the formal categories nicely carve up the domain of subordination, with *wa(ʔ)*-nominalizations coding relative relations while *e(ʔ)*-nominalizations are used to code complement and adverbial relations (but see the discussion of *-a-nda* in Section 5.2.1).

Further generalizations that could be made pertain to the internal and external syntax of verb-based nominalization. All types of one-word participant nominalization showed NP-like external syntax (perhaps apart from the citation form); the use of genitive-marked nouns to signal the notional subjects of the nominalized forms in (22) and (23) in fact testified to these forms constituting one-word nominalizations, and hence lacking any internal syntax. Multi-word nominalizations, in turn, were all found to combine NP-like external syntax with verb-like internal syntax, just like, for example, nominalizations in Kakataibo (Valle & Zariquiey this volume) and Cahita (Álvarez this volume). Specifically, if notional subjects were expressed, they appeared either unmarked or marked for nominative case, just like S- or A-participants of independent clauses do. Similarly, notional direct objects appeared unmarked when referring to inanimate entities, and marked for accusative case when referring to animate entities, thus adhering to the principles governing differential O-marking in independent clauses (see Section 3). With respect to the external syntax of event nominalizations, some reservation was already made for the concessive subtype (see Section 5.2.2), and it was hypothesized more generally that infinitival nominalizations – unlike underived nouns – disrespect the differential/optional character of case marking (see Section 5.2.1). Here, I would also like to question the status of the nominalized forms that take no extra suffix and are used to code the complement clauses of commentative, ability and immediate perception predicates. While commentative clauses still feature notional subjects taking the same form as S- or A-participants in independent clauses, the forms used in modal and immediate perception complements might be better analysed as infinitives rather than nominalizations. In the case of ability predicates, the subject of the complement proposition proved equi-deleted (see Section 5.1.2), which is typical of infinitival complements (see Noonan 2007: 67), and in the case of immediate perception predicates, the subject of the complement proposition has its case assigned by the main clause verb phrase (see Section 5.1.3). In fact, this construction comes close to an *Accusativus-Cum-Infinitivo* construction found in the complementation system of, for example, a fair number of Indo-European languages. Further investigation is needed here, also with regard to the status of the predicate *ẽnõpõẽ*, which – in its acquired ability sense – seems to be moving along the auxiliation pathway proposed by Heine (1993).

A final topic that this paper touched upon only briefly is comparison with Harakmbut's neighbouring languages. Areality was brought into the discussion in Section 5.3 on noun-based nominalization, but the processes of verb-based nominalization were not placed in any comparative perspective. Comparison with other Peruvian Amazonian languages and with the languages included in the Guaporé-Mamoré region (Crevels & van der Voort 2008) will reveal to what extent the Harakmbut system stands out. Comparison with the Katukina-Kanamari system (Anjos Gonçalves da Silva 2011) will contribute to the debate on the hypothesized genetic link between the Katukina family and Harakmbut. Needless to say, these form interesting avenues for further research.

Acknowledgements

The research reported on in this paper has been made possible by mobility grants and a postdoctoral grant from the Research Foundation – Flanders (FWO), as well as a postdoctoral grant from the Fund for Scientific Research – FNRS and a research project grant from the research council of the University of Leuven (KU Leuven) (GOA/12/007). I thank an anonymous reviewer as well as the editors for their insightful comments on an earlier version of this paper. Of course, any errors of fact or interpretation remain my own responsibility. Finally, my sincere thanks go to the Harakmbut people, who warmly welcomed me in their communities and patiently taught me their beautiful language.

Abbreviations

1	1st person	INS	instrumental
2	2nd person	ITER	iterative
3	3rd person	LOC	locative
>	'acts on'	MOD	modifier
ACC	accusative	NEG	negation
ADV	adverb(ializer)	NMZR	nominalizer
AN	animate	NOM	nominative
APPL	applicative	NVIS	non-visual evidential
BEN	beneficiary/benefactive	NVOL	non-volitional
CLF	classifier	PFV	perfective
COLL	collective	PL	plural
COM	comitative	POT	potential
CONC	concessive	PRIV	privative
COND	conditional	Q	question particle
DEP	dependent verb form	REAS	reason

DIM	diminutive
DIST	distal
DIST.PST	distant past
DUB	dubitative
EXCL	exclusive
FOC	focus
FUT	future
FUT.POSS	future-oriented possibility
GEN	genitive
IMP	imperative
INCL	inclusive
IND	indicative
INDET	indeterminate
REC.PST	recent past
REST	restrictive
SG	singular
SIM	similative
SIMUL	simultaneity
SOC	sociative causative
SPAT	spatial prefix
SS	same subject
TRNS	transitivizer
VBZ	verbalizer
VOL	volitional
VPL	verbal plural

References

Adelaar, Willem. 2000. Propuesta de un nuevo vínculo genético entre dos grupos lingüísticos indígenas de la Amazonía occidental: Harakmbut y Katukina. In *Actas del I Congreso de Lenguas Indígenas de Sudamérica*, Vol. 2, Luis Miranda Esquerre (ed.), 219–236. Lima: Universidad Ricardo Palma.

Adelaar, Willem. 2007. Ensayo de clasificación del katawixí dentro del conjunto harakmbut-katukina. In *Lenguas indígenas de América del Sur: Estudios descriptivo-tipológicos y sus contribuciones para la lingüística teórica*, Andrés Romero-Figueroa, Ana Fernández Garay & Angel Corbera Mori (eds), 159–169. Caracas: Universidad Católica Andrés Bello.

Aikhenvald, Alexandra Y. 2000. *Classifiers. A Typology of Noun Categorization Devices*. Oxford: OUP.

Andrews, Avery. 2007. Relative clauses. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, 2nd edn, Timothy Shopen (ed.), 206–236. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.004

Anjos Gonçalves da Silva, Zoraide dos. 2011. Fonologia e gramática Katukina-Kanamari [LOT Series 281]. PhD dissertation, Vrije Universiteit Amsterdam. Utrecht: LOT.

Comrie, Bernard & Thompson, Sandra A. 2007. Lexical nominalization. In *Language Typology and Syntactic Description*, Vol. 3: *Grammatical Categories and the Lexicon*, 2nd edn, Timothy Shopen (ed.), 334–381. Cambridge: CUP. https://doi.org/10.1017/CBO9780511618437.006

Crevels, Mily & van der Voort, Hein. 2008. The Guaporé-Mamoré region as a linguistic area. In *From Linguistic Areas to Areal Linguistics* [Studies in Language Companion Series 90] Pieter Muysken (ed.), 151–179. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.90.04cre

Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.

Dryer, Matthew S. & Haspelmath, Martin (eds). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info>

Fauconnier, Stefanie. 2011. Differential agent marking and animacy. *Lingua* 121: 533–547. https://doi.org/10.1016/j.lingua.2010.10.014

Guillaume, Antoine. 2008. *A Grammar of Cavineña* [Mouton Grammar Library 44]. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110211771

Gray, Andrew. 1996. *The Arakmbut: Mythology, Spirituality and History*. Oxford: Berghahn.

Hart, Raymond. 1963. Semantic components of shape in Amarakaeri grammar. *Anthropological Linguistics* 5(9): 1–7.

Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. Oxford: OUP.

Helberg Chávez, Heinrich A. 1984. *Skizze einer Grammatik des Amarakaeri*. PhD dissertation, Tübingen University.

Helberg Chávez, Heinrich A. 1990. Análisis functional del verbo amarakaeri. In *Temas de lingüística amerindia*, Rodolfo Cerrón Palomino & Gustavo Solís Fonseca (eds), 227–249. Lima: Concytec.

Kirsner, Robert & Thompson, Sandra A. 1976. The role of pragmatic inference in semantics: A study of sensory verb complements in English. *Glossa* 10: 200–240.

Matteson, Esther. 1972. Proto Arawakan. In *Comparative Studies in Amerindian Languages*, Esther Matteson, Alva Wheeler, Frances L. Jackson, Nathan E. Waltz & Diana R. Christian (eds), 160–242. The Hague: Mouton. https://doi.org/10.1515/9783110815009.160

McGregor, William B. 2007. Ergative marking of intransitive subjects in Warrwa. *Australian Journal of Linguistics* 27(2): 201–229. https://doi.org/10.1080/07268600701531351

McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120: 1610–1636. https://doi.org/10.1016/j.lingua.2009.05.010

McQuown, Norman A. 1955. Indigenous Languages of Latin America. *American Anthropologist* 57: 501–570.

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60: 847–879. https://doi.org/10.1353/lan.1984.0038

Moore, Denny. 2007. Endangered languages of lowland tropical South America. In *Language Diversity Endangered*, Matthias Brenzinger (ed.), 29–58. The Hague: Mouton.

Noonan, Michael. 2007. Complementation. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, 2nd edn, Timothy Shopen (ed.), 52–150. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.002

Payne, Doris L. 1987. Noun classification in the Western Amazon. *Language Sciences* 9(1): 21–44. https://doi.org/10.1016/S0388-0001(87)80007-4

Pozzi-Escot, Inés (ed.). 1998. *El Multilingüismo en el Perú*. Cuzco: CBC-SID.

Seifart, Frank. 2007. The prehistory of nominal classification in Witotan languages. *International Journal of American Linguistics* 73(4): 411–445. https://doi.org/10.1086/523772

Thompson, Sandra A., Longacre, Robert E. & Hwang, Shin Ja J. 2007. Adverbial clauses. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, 2nd edn, Timothy Shopen (ed.), 237–300. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.005

Tripp, Robert. 1976. Sustantivos verbales y frases de sustantivos verbales en Amarakaeri. In 14. Datos Etno-Lingüísticos: Colección de los archivos del ILV 50. Lima: Instituto Lingüístico de Verano.

Tripp, Robert. 1995. *Diccionario amarakaeri-castellano*. Yarinacocha: Min. de Educación & Instituto Lingüístico de Verano.

van der Voort, Hein. 2005. Kwaza in a comparative perspective. *International Journal of American Linguistics* 71 (4): 365–412. https://doi.org/10.1086/501245

van Gijn, Rik, Haude, Katharina & Muysken, Pieter. 2011. Subordination in South America: an overview. In *Subordination in native South American languages* [Typological Studies in Language 97], Rik van Gijn, Katharina Haude & Pieter Muysken (eds), 1–24. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.97.01van

Van linden, An. 2004. A typological study of counterfactuality. MA dissertation, University of Leuven. (KU Leuven). https://doi.org/10.13140/RG.2.2.22541.59369

Van linden, An. 2012. *Modal Adjectives: English Deontic and Evaluative Constructions in Diachrony and Synchrony* [Topics in English Linguistics 75]. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110252941

Van linden, An. 2014. Argument marking in Harakmbut: Looking for referential transparency. Paper presented at Syntax of the World's Languages VI, 8–10 September 2014, University of Pavia. https://doi.org/10.13140/2.1.5077.4400

Van linden, An. Forthcoming. Harakmbut. In *Amazonian Languages. An International Handbook*, Patience Epps & Lev Michael (eds). Berlin: de Gruyter Mouton.

Van linden, An & Verstraete, Jean-Christophe. 2008. The nature and origins of counterfactuality in simple clauses: Cross-linguistic evidence. *Journal of Pragmatics* 40: 1865–1895. https://doi.org/10.1016/j.pragma.2008.03.008

Van linden, An & Davidse, Kristin. 2009. The clausal complementation of deontic-evaluative adjectives in extraposition constructions: A synchronic-diachronic approach. *Folia Linguistica* 43(1): 175–217. https://doi.org/10.1515/FLIN.2009.005

Wise, Mary R. 1999. Small language families and isolates in Peru. In *The Amazonian languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 307–340. Cambridge: CUP.

CHAPTER 13

Nominalization in Shawi/Chayahuita

Luis Miguel Rojas-Berscia
Radboud Universiteit Nijmegen/Max Planck Institute for Psycholinguistics/
University of Queensland

This paper deals with the Shawi nominalizing suffixes *-su*'~*-ru*'~*-nu*' 'general nominalizer', *-napi/-te*'/*-tun*'performer/agent nominalizer', *-pi*'patient nominalizer', and *-nan* 'instrument nominalizer'. The goal of this article is to provide a description of nominalization in Shawi. Throughout this paper I apply the Generalized Scale Model (GSM) (Malchukov, 2006) to Shawi verbal nominalizations, with the intention of presenting a formal representation that will provide a basis for future areal and typological studies of nominalization. In addition, I dialogue with Shibatani's model to see how the loss or gain of categories correlates with the lexical or grammatical nature of nominalizations. STRONG NOMINALIZATION in Shawi correlates with LEXICAL NOMINALIZATION, whereas WEAK NOMINALIZATIONS correlate with GRAMMATICAL NOMINALIZATION. A typology which takes into account the productivity of the nominalizers is also discussed.

1. Introduction

The Shawi language ministers to nominalization for the creation of new nouns and the production of dependent clauses. The latter can be used to create relative clauses, to stand for arguments in the form of headless relative clauses, or to create purposive constructions. Here I deal with the Shawi nominalizing suffixes *-su*'~*-ru*'~*-nu*' 'general nominalizer', *-napi/-te/-tun* 'performer/agent nominalizer', *-pi* 'patient nominalizer', and *-nan* 'instrument nominalizer'. The goal of this contribution is to provide a description of nominalization in Shawi. I apply the Generalized Scale Model (GSM) (Malchukov 2006) to Shawi verbal nominalizations, with the intention of presenting a formal representation that will provide a basis for future areal and typological studies of nominalization.

Shawi distinguishes between two types of nominalization: strong and weak. In the case of STRONG NOMINALIZATION, most verbal features are lost. In WEAK NOMINALIZATION, on the contrary, several verbal features are retained. The discussion is closed with a dialogue with Shibatani's account on nominalization.

https://doi.org/10.1075/tsl.124.13roj

The loss or gain of categories correlates with the lexical or grammatical nature of nominalizations. In Shawi, STRONG NOMINALIZATION correlates with LEXICAL NOMINALIZATION, whereas WEAK NOMINALIZATION correlates with GRAMMATICAL NOMINALIZATION.

2. The Shawi language

The Shawi or Chayahuita language, a descendant of Southern Mayna or *lengua de Cerros de Maynas* (Rojas-Berscia 2015), is a member of the Kawapanan language family, together with Shiwilu, with which it shares 60% of its lexicon (Valenzuela 2012: 1). According to a national census, there were approximately 21,000 speakers of Shawi in 2007; however, this number amounts to people who identify themselves as Shawi, but do not necessarily speak the language. The actual number of speakers is closer to 15,000 (Barraza 2005a). As is evident from the regular use of the language, monolingualism in women and children, and language attitude, Shawi is still a very vital language compared to other adjacent Andean or Amazonian languages in Peru.

In typological terms, Shawi is a head-marking language with a strong tendency to agglutination and synthesis, meaning that morphemes are easily distinguishable from each other and that there are a large number of morphemes per word. There are some instances, however, in which we can find isolated monosyllabic morphemes, such as the negative marker *ku* or the first person pronoun *ka*, as in (1) . Suffixation as well as prefixation can be found in the language, although prefixation is not so frequent. Below I provide an example of agglutination and synthesis, where multiple suffixes are added to the main verb *tepa* 'to die', as well as a prefix, the indirect causative *a-* (Rojas-Berscia 2013):

(1) *Nunu ka a-tepa-ra-we-su' ku ka'-n-awe.*
monkey 1 CAUS-kill-NON.FUT-1-NMZR NEG eat-NON.FUT-1
'I do not eat the monkey which I made someone kill'.

Shawi is a Predicate-Final language (Seuren 2018). This is reflected in the most frequent constituent order in the language, which is AOV. However, there are some instances in which AVO order seems to be preferred, perhaps due to intense contact with Amazonian Spanish. OAV order is also sometimes preferred, mostly when a third person NP subject acts upon a third person NP object (3 > 3), in transitive clauses in which the ergatively-marked argument comes after the object (Rojas-Berscia & Bourdeau 2018). For intransitive clauses, the preferred word order is SV, which mirrors the predicative clause word order, S(Pred).

Person suffixes obligatorily occupy the final position in the verb. Both subject and object are marked on the verb, as in (2) :

(2) *Ka-ri nanian-te-r-**au-nke**.*
1-ERG forget-VAL-NON.FUT-1A-2O
'I forgot you.'

As can be deduced from the previous example, subject markers always precede object markers, i.e. in transitive verbs the subject suffix will precede that of the patient/object, see (2), while in ditransitivized verbs – only existent after a valency changing operation takes place – the subject suffix will precede that of the recipient (equally marked as the object), see (3) .

(3) *Ina ke-te-r-**in-ku** a'nara*[1] *kayu'.*
3-ERG give-VAL-3A-1R one egg
'He gives me one egg'.

Nouns and verbs are open classes in the language, both occurring with rich morphological derivation and inflection paradigms. There is a closed class of adjectives with very few members. Other property concepts are coded by verbs. The modifier precedes the modified noun and there cannot be more than one modifier in a single noun phrase. For instance:

(4) **[panka$_{ADJ}$ [atari$_{N}$ kayu'$_{N}$]] → [[panka$_{ADJ}$ kayu'$_{N}$] [atari$_{N}$ kayu'$_{N}$]]- ø*
big chicken egg big egg chicken egg-3.COP
'the big chicken egg/The big egg is a chicken egg'.

In (4) , the occurrence of two modifiers before the noun *kayu*' 'egg' is infelicitous. An alternative construction is presented following the arrow, in which two noun phrases, both headed by 'egg' are conjoined by means of a copula construction.[2]

In addition, the language has pronouns, deictics, adverbs, classifiers, conjunctions and interjections as closed word classes (*cf.* Schachter & Shopen 2007: 3). There is a first person singular exclusive *ka* and a first person singular inclusive *kanpu'*. Their pluralized forms are *kiya* and *kanpua*' respectively. The second person is *kema*, and the third, *ina.* The third person in the language is also a third person demonstrative. Both the second and third person pronouns can be pluralized by means of the pluralizer *-pita.*

1. The numeral 'one' *a'na* in Balsapuerto Shawi is commonly lexicalized with the 'stones; grains' classifier *-ra*.

2. It is noteworthy that the third person copular marker in Shawi is *-ø*. This becomes of great importance in Section 5.1, regarding the 'relativization of pronouns and deictics'.

The use of causative constructions is frequent in the corpus. These are commonly marked by means of the prefix *a-*, for indirect causation, and *ichi-*, for sociative causation. The reciprocal *ni-* is likewise a prefix and occupies the same slot as the causative markers. There is a valency-changing affix *te-* which can increase or diminish valency, as well as verbalize non-verbal roots. There are no passive constructions in Modern Shawi, although the modern ergative marker *-ri* might be an innovation based on a historical passive construction (Rojas-Berscia 2015).

In terms of grammatical relations, Shawi exhibits different alignment types. For case marking, the language shows an ergative-absolutive alignment with a sometimes pragmatically driven use (*cf.* Bourdeau 2015; Rojas-Berscia & Bourdeau 2018). As for verbal person-number agreement, Shawi shows a nominative-accusative alignment in transitive clauses, while a quite obscure active-stative system (cf. Barraza 2005b)[3] when indexed verbal objects, as in (6) , and predicative clauses, as in (5) , are compared.

(5) *Ka shawi-**ku.***
1 Shawi-1O
'I am a Shawi.'

(6) *Kema tepa-r-an-**ku.***
2 kill-NON.FUT-2A-1O
'You are killing me.'

As noted from the previous examples, nominal predication suffixes as well as object markers are identical.

Table 1. Shawi person agreement suffixes

	Nominal predication	**Intransitive subject**	**Transitive subject**	**Transitive object**
1	***-ku***	*-aw(e)*	*-aw(e)*	***-ku***
1 + 2	***-npu***	*-e'*	*-e'*	***-npu***
2.sg	***-ken/nke***	*-an*	*-an*	***-ken/-nke***
3.sg	***-ø***	*-in*	*-in*	***-ø***
1 + 3	***-kui***	*-ai*	*-ai*	***-kui***
1 + 2.pl	***-npuwa'***	*-ewa'*	*-ewa'*	***-npuwa'***
2.pl	***-kenma'/nkema'***	*-ama'*	*-ama'*	***-kenma'/nkema'***
3.pl	***-ø***	*-pi/-na*	*-pi/-na*	***-ø***

The language has a large inventory of case markers, including genitive, ablative, locative, perlative, comitative, comparative, benefactive, and limitative.

3. This claim awaits further investigation.

As in many Amerindian languages, nominalizations are common in Shawi. They are mostly marked by means of:

Table 2. List of nominalizers in Shawi

Nominalizer	Function
-su'	General nominalizer (typically found in relative clauses and purposive clauses)
-te'	Performer agent 1
-napi	Performer agent 2
-tun	Performer agent 3
-pi	Patient
-nan	Instrument

All these suffixes and their meanings in the predicative domain are analyzed in the subsequent sections. In the current state of documentation of the language, these suffixes exhaust all nominalizing morphemes. No de-adjectival or de-ideophonic nominalizations have been found.

In § 3, I present the Generalized Scale Model, henceforth GSM, as a formal baseline for the analysis of nominalization in the language. § 4 is devoted to the analysis of strong nominalization markers. § 5 deals with weak nominalization markers, namely the general nominalizer *-su'*, which also occurs with pronouns and demonstratives. § 6 presents an overview of the findings, sketching a dialogue with Shibatani's (this volume) own account of nominalization.

3. The Generalized Scale Model (GSM)

In typology, nominalization is often seen as both the loss of verbal properties and the acquisition of nominal properties (Comrie & Thompson 1985; Lehmann 1988; Croft 1991; Koptjevskaja-Tamm 1993 et al.). In Malchukov's Generalized Scale Model (2004, revised 2006), deverbal nominalization is conceived as a twofold transcategorial process involving decategorization – loss of verbal properties – and recategorization – acquisition of nominal properties.[4] Although the two definitions do not seem to differ substantially, the latter encompasses a formalization of well-known ideas in typology, as well as a constraint-based enriched model, inspired by Prince and Smolensky (2004)'s Optimality Theory, in which functionally based hierarchies of nominal and verbal categories have a major role: it is

4. The same would apply to a de-adjectival nominalization: the loss of adjectival properties and the acquisition of verbal properties.

the interplay of constraints in this transcategorial process which triggers a formal output in a given language.[5,6]

According to this model there are two main constraints which contribute to lexical categorization:

> *FuncFaith:* Assign (morphological) categories to a lexical item in accordance with its discourse function;
> *LexFaith:* Assign (morphological) categories to a lexical item in accordance with the semantic class of a lexical root (Malchukov 2006: 975)

The ranking of these constraints is language-dependent. If FuncFaith is ranked first, a strong nominalization will be produced, while a LexFaith prominence will produce a weak nominalization,[7] since more lexical properties will be maintained rather than lost. In a nutshell, it is FuncFaith that is in charge of assigning features related to the new expected category, i.e. a noun, and removing features related to the original category; that is, a verb. LexFaith, by contrast, shields the semantic class of the root; that is, the verbal root, and avoids full recategorization impeding the acquirement of some noun features (Diagram 1). It is remarkable, following this reasoning, that, across many languages, outermost features – features which are not at the core of the category, like illocutionary force for verbs – tend to be affected/lost first during nominalization processes, while innermost ones – core features related to a determined category, like voice or valency in verbs – tend to be kept.

Thus, "the set of verbal categories retained depends on the point at which LexFaith constraints are interpolated in the FuncFaith hierarchy" (ibid 2006: 982). In the same fashion, the set of nominal categories acquired depends on where LexFaith blocks the acquiring of categories by the demotion of FuncFaith. In this way, LexFaith prevents the process from leading to full recategorization. In Diagram 1, we see the FuncFaith hierarchy for deverbalization (a) and substantivization (b) respectively:

---

5. The general insight also applies to verbalization (Malchukov 2004).

6. This is model intends to explain the internal-to-external syntax of nominalizations from a formal perspective. From a functional perspective, agreeing with Shibatani (this volume), a nominalization is "a metonymic process yielding constructions associated with a denotation comprised of entity (thing-like) concepts that are metonymically evoked by the nominalization structures such as events, facts, propositions, and resultant products ("event nominalization") and event participants ("argument nominalization")".

7. FuncFaith >> LexFaith LexFaith >> FuncFaith, respectively.

Diagram 1. A hypothetical deverbalization and substantivization cline

Deverbalization[8] (a) *IF >> *AGR_S >>**FuncFaith**>> **LexFaith** >> Mood >> Tense >> Aspect >>
Voice >> Valency
Substantivization[9] (b) -Case >> -Det >>>**FuncFaith**>>**LexFaith** >> *-Pos >> *-Nb >> *-CL

In the hypothetical case above, FuncFaith leads to the loss of illocutionary force and subject agreement, but the ranking of LexFaith protects mood, tense, aspect, voice and valency from being lost. Similarly, during substantivization: FuncFaith triggers the acquisition of case and determiners, while the ranking of LexFaith prevents the acquisition of possessive markers, number and noun classifiers. The insertion of these constraints involves a certain amount of abstraction from the surface data, but Malchukov's model helps us shed light on the syntactic processes behind the formation of nominalizations in the languages of the world.[10]

In the following sections, I will present a typology of weak and strong nominalizations in Shawi. There are some bases which are very much deverbalized, leading to the formation of STRONG NOMINALIZATIONS. Meanwhile, other bases acquire many nominal features without losing most verbal features, which leads to the formation of WEAK NOMINALIZATIONS.

4. Strong nominalization

In this class of nominalizations in Shawi all or most verbal features are lost. The class strong nominalization includes performer/agent nominalization, patient nominalization, and instrument nominalization.

8. The degree of deverbalization is measured by the loss of verbal properties depending on the demotion of FuncFaith and LexFaith's shielding of other verbal properties. The asterisk in this formalization indicates either the loss or the prevention of acquisition of certain categorial properties.

9. The degree of substantivization is measured by the degree to which something becomes more "nouny" by the process of acquiring nominal properties depending on the scope of FuncFaith and its blocking by LexFaith.

10. Whether this model sheds light on anything related to deep structure/universal grammar related topics was not my main question (cf. Chomsky 1970 on how the study of nominalizations can shed light on UG et alia.). I chose this model due to its powerful explanatory power and its usefulness as a descriptive tool for many languages in the world (cf. Malchukov 2004), and due to the typological implications this can have for future studies.

4.1 Performer/Agent nominalizer *-napi*[11] (very frequent)/ *-te'* (less frequent)/ *-tun* (rare)

The three derivative nominalizing suffixes that convey the meaning of agent or performer are in competition with each other. In contemporary Shawi, *-napi* is used much more frequently than *-te* or *-tun,* even in cases where they produce synonymous nominalizations. All my consultants agreed that these three suffixes can be used for the same purposes. However, *-napi* can be used with almost any base, while *-te'* is restricted to a limited number of bases, and *-tun* to even fewer. Apparently, *-napi* is more grammatical and productive, while *-te'* and *-tun* are more lexicalized. Table 3 presents some examples in which the suffixes convey similar meanings.

Table 3. Performer/agent nominalizations

Nominalization	Meaning
sha'wi-**te'**/sha'wi-**napi**	gossiper
chimin-**napi**	dead one
nitu-**napi**/nitu-**te**	watcher
nitu-**tun**-**napi**/nitu-**te'**	wise one
ni-**napi**/ni-**te'**	maker
u'u-**napi**/u'u-**te'**	drinker
yanuwi-**napi**/yanuwi-**te'**	player
saka-**tun**-**napi**/saka-**te'**/saka-**tun**	worker
ma'pa-**tun**-**napi**/ma'pa-**tun**	receiver
mapi-**tun**-**napi**/mapi-**tun**	the hunter

The reader may have noted that 'wise one', 'worker, 'receiver', and 'hunter' contain two nominalizing suffixes, *-tun* and *-napi.* One possible interpretation of this is that *-tun* nominalizations are lexicalized. Nominalizations with *-napi* would just be the productive counterpart. What they have in common is that they form strong nominalizations: they do not allow verbal morphology to precede them. We can capture this in terms of the Generalized Scale Model by the following ranking:

Substantivization: -Case >> -DET >> -POS >> -NB >> -CL >> **FuncFaith** >> **LexFaith**
Deverbalization: *IF >> *(AGRS/AGRo)AGR >> *Mood >> *Tense >> *Aspect >> *Voice >> *Valency >> **FuncFaith** >> **LexFaith**

11. The element *-pi* in this suffix is possibly related to the human classifier *-pi.*

These suffixes produce nominalizations that can behave morphosyntactically like a noun, allowing case-marking, definiteness markers, possessive markers, number markers and classifiers, since FUNCFAITH is demoted nonstop without the block of LEXFAITH, as shown in (7) , in which the nominalized verbal root *sha'wi-*, carries a pluralizer and an ergative suffix. However, all verbal features are lost, as captured by the demotion of the FUNCFAITH constraint in the deverbalization cline which LEXFAITH does not block, thus resulting in a strong nominalization, as illustrated in (7) :

(7) *Pawara-rawa sha'wi-te'-ru'sa-ri tepa-pi tanan-ke.*
tapir-DIM tell-AG.NMZR-PL-ERG kill-3.PL woods-LOC
'The gossipers hunted a tapir in the woods.'

Nominalizations formed with *-napi* are less strong than those formed with *-te'* or *-tun*. There are cases in which *-napi* allows some verbal morphology. For instance, the nominalized base may still accept a reflexive/reciprocal marker (8) , a volitive prefix (9) and/or a valency-changing operator (10) .

(8) ***ni**-tepa-**napi***
REFL-kill-AG.AG.NMZR
'the one that commits suicide'

(9) ***ya-ni**-tepa-**napi***
VOL-REFL-kill-AG.NMZR
'the one that wants to commit suicide'

(10) *atari nusha **a**-ka'-**napi***
chicken meat CAUS-eat-AG.NMZR
'The one that makes someone eat chicken meat.'

The preceding examples illustrate well the nature of the grammatical nominalizer *-napi*. In Examples (8)–(10) , the slots for valency changing operations are not blocked by LEXFAITH, thus acquired in the demotion of FUNCFAITH. Nominalizations with *-te'* do not allow this.[12] In terms of deverbalization, *-napi* still maintains some verbal properties, unlike *-te'* or *-tun*, which can be captured with the following ranking:

*IF >> *(AGR$_S$/AGRo)AGR >> *Mood >> *Tense>> *Aspect >> **FUNCFAITH** >> **LEXFAITH** >>>> Voice >> Valency

12. I still tried to get examples of nominalizations with *-te'* holding different kinds of valency changing operators, such as **a-tepa-te'* 'the one that causes someone to kill'. In every case, it was infelicitous.

-Napi allows for the use of a causative, hence maintaining the valency feature. It also allows for the use of a reflexive/reciprocal, hence maintaining the voice feature.

4.2 Patient nominalizer *-pi*

The patient nominalizer *-pi* conveys the patient of an event; that is, the nominalized construction refers to the patient of a given event. It creates a strong nominalization, comparable to nominalizations formed with-*te*'. Table 4 presents some examples.

Table 4. Patient nominalizations

Nominalization	Meaning
pasa-**pi**	minced one
ayu-**pi**	cooked one
wiki-**pi**	burnt one
pumu-**pi**	one put inside

The instrumental *-nan* can be suffixed to all these nominalized constructions and the meaning remains stable. Examples (11) and (12) show the nominalizing suffix *-pi* in a sentence.

(11) *Kema kara ayu-**pi** ka'-n-an.*
2 three cook-PAT.NMZR eat-NON.FUT-2
'You ate three cooked things.'

(12) *Pumu-**pi** chimin-pi.*
put-PAT.NMZR die-3.PL
'The (living) things that were put inside died.'

Nominalization involving *-pi* can be formalized as follows in line with the GSM:

Substantivization: -Case >> -DET >> -POS >> -NB >> -CL >> **FuncFaith** >> **LexFaith**
Deverbalization: *IF >> *(AGRS/AGRo)AGR >> *Mood >> *Tense >> *Aspect >> *Voice >> *Valency >> **FuncFaith** >> **LexFaith**

In this case, FUNCFAITH is demoted nonstop, allowing for a complete loss of verbal properties. Thus, LEXFAITH did not block any of the noun properties, allowing for a full recategorization. In terms of productivity, *-pi* is very productive. Constructions with *-pi* are not lexicalized but they still form new nouns.

4.3 Instrument nominalizer *-nan*

The instrumental nominalizer *-nan* conveys what an element is used for; that is, the nominalized construction refers to an object that has a particular use. It creates a very strong nominalization, like *-te*' and *-pi.* Table 5 presents some examples.

Table 5. Instrumental nominalizations

Verbal Root	Nominalization	Meaning
weyute[a] 'to sweep'	weyu-**nan**	broom
perate'- 'to call'	perate-**nan**	telephone
nanpe- 'to climb'	nanpe-**nan**	stairs
we'e- 'to sleep'	we'e-**nan**-te'[b]	bed
wense- 'to sit'	wense-**nan**	chair
pe'tu- 'to perforate'	pe'tu-**nan**	blowgun

a In this case, there is no synchronic evidence for the existence of a verb **weyu.* In modern Shawi, it needs the valenciator *-te* in order to indicate that it is a verb. *Weyu* is apparently a noun which means 'waste'. The evidence for this is the fossilized use of the classifier *-ru*' with the noun, meaning 'garbage'.
b Hart (1988: 81) documents this suffix *-te*', as "instrument", which seems to be another but very infrequent instrument nominalizer in the language.

As in the case of *-te*' and *-pi,* nominalized constructions with *-nan* are very strong, as can be observed in the GSM representation:

Substantivization: *-Case >> *-DET >> *-POS >> *-NB >> *-CL >> **FuncFaith** >> **LexFaith**
Deverbalization: ~~*IF~~ >> ~~*(AGRS/AGRo)AGR~~ >> ~~*Mood~~ >> ~~*Tense~~ >> ~~*Aspect~~ >> ~~*Voice~~ >> ~~*Valency~~ >> **FuncFaith** >> **LexFaith**

We see that, as for the substantivization cline, FuncFaith is demoted nonstop. In the case of the deverbalization cline, the same happened, allowing for a full recategorization.

Historically, this may be another use of the classifier *-nan* in Shawi, meaning 'long sharp thing'. However, its use for 'stairs' or 'chair' would be problematic for this meaning. It is probable that, like many other constructions in Kawapanan languages (Valenzuela 2015), an Andean construction was copied in Shawi. In Quechuan languages, there is a productive instrumentalizer *-na,* which is used for the same purposes as *-nan* in Shawi. See for example Quechua: *picha-* 'to sweep'→ *picha-na* 'broom', *puñu-* 'to sleep'→ *puñuna* 'bed'. Today, although most Shawi *-nan* nominalizations are lexicalized, the suffix is still productive. The word *perate-nan* (call-NMZR) 'telephone' was recently created.

In a nutshell, one consequence of the low ranking of LexFaith for these nominalizations is that deverbalization is complete: these strong nominalizations are formally so different from verbs that they recategorized entirely.

5. Weak nominalization

Weak nominalizations retain several verbal features. In terms of the GSM, the loss of verbal features is blocked by the ranking of the LexFaith constraint, and thus FuncFaith cannot be demoted completely. The general nominalizer *-su'* discussed in § 4.1 provides a good example of weak nominalization in Shawi.

5.1 General nominalizer *-su'*

The general nominalizer *-su'* is widespread in the corpus.[13] Unlike most of the previously examined nominalizers, it can be used with all types of verbs. *-Su'* is mainly used in relativization functions, as in (13) .

(13) *[Nara [wania-r-in-**su'**]* $_{NMLZ}$*]* *ira-ke-ø.*
tree stand-NON.FUT-3-NMZR path-LOC-3
'The standing tree is on the path.'

-Su' is also very frequent in so-called headless relative clauses, as expansions of participant slots in matrix clauses. In (14), I present a simple sentence with the verb *chimin-* 'to die'. In (15) , the same verb is nominalized with the general nominalizer *-su'*. The resulting clause refers to the one that died:

(14) *Tata i'-ke* ***ya**-chimin-n-in.*
dad water-LOC want-die-NON.FUT-3
'My father drowns'.

(15) *Pei-ne-we-ke* *[chimin-n-in-**su'**]* $_{NMLZ}$ *na'ku-r-in.*
house-ALIEN-1-LOC die-NON.FUT-3-NMZR pass.near-NON.FUT-3
'The one that died passed near my house.'

The same happens with other intransitive verbs:

(16) *A'nara ni'ni nara nii-te-r-in.*
one dog tree jump-VAL-NON.FUT-3
'The dog jumps over he tree.'

13. My corpus consists of data presented in Hart (1988) and examples extracted by elicitation as well as text collection sessions with consultants from Balsapuerto, Sillay, Cahuapanas and Pueblo Chayahuita.

(17) *[Nii-ra-r-in-**su'**]* $_{NMLZ}$ *imin-ke-ø.*
jump-PROG-NON.FUT-3-NMZR field-LOC-COP
'The one who jumps is in the field.'

In (17) , the intransitive verb *nii-* 'jump' is nominalized. It is clear that *nii-* is an intransitive verb in the language since no other argument than the agent is allowed. To add the goal of the jump as an argument, the valency changing suffix *-te* with applicative functions needs to be used, as in (16) .

Both nominalizations of the intransitive verbs in (15) and (17) refer to the subject of the clause. Subject nominalizations, in this case, compete with *-napi* nominalizations. The latter, however, seems to be more lexicalized than those formed with *-su'*.

Below I present some examples with transitive verbs, such as (18) , where the transitive verb *pu'mu-* 'put inside' is nominalized, once more in the form of a headless relative:

(18) *Wa'washa-rawa [pu'mu-r-in-su']* $_{NMLZ}$ *inape wa-te-a-r-in.*
child-DIM put.inside-NON.FUT-3-NMZR up stand-VAL-NON.FUT-3
'The child was lifting what he (the dog) had put inside.'

The nominalized clause in this case, unlike (14) and (17) refers to the object of the verb *pu'mu-* 'put', i.e. 'what is put'. In addition, the nominalized clause was used as object of the verb *wa-te-* 'stand-VAL'. However, it can also be used as a subject.

In this case, the transitive verb *ipu-* 'hug' is nominalized:

(19) *[Ina ipu-r-in-**su'**]* $_{NMLZ}$ *nansa-pu-n.*
3 hug-NON.FUT-3-NMZR dance-FUT-3
'The one he hugged is going to dance.'

The nominalized clause in (19), as in (18) , refers to the object of the verb *ipu-*, not to its subject. From the point of view of grammatical relations, the alignment of the formation of nominalization with relativizing functions in Shawi is ergative (see Chapter 10 for a similar phenomenon in Kakataibo): the nominalized intransitive clause refers to the subject of the verb, while the relativized transitive clause refers to the object. It is, however, a very weak nominalization, since it allows for the appearance in the verb of person agreement markers, tense/mood markers,[14] aspect markers and valency markers.

In different parts of the Shawi speaking area, *-su'* can be lenited to *-ru'*, as in (20) [15]. Below I present an example:

14. Non-future and future markers are the tense markers in the language.

15. The same phenomenon occurs with the progressive aspect marker *-sa→-ra* in the language.

(20) *[Ni'nira ka-ri ni'-na-w-**ru'**]*$_{NMLZ}$ *chimin-n-in.*
dog 1-ERG see-NON.FUT-1-NMZR die-NON.FUT-3
'The dog I saw died.'

It can also be nasalized to *-nu'*[16], if the modified theme ends with a nasal. Below I present an example:

(21) *[Pewara ta'a-ra-r-in-**nu'**]*$_{NMLZ}$ *ka-ken-ø.*
lizard run-PROG-NON.FUT-3-NMZR 1-GEN-1
'The lizard that is running is mine.'

In terms of the GSM, nominalizations with *-su'* are weak, since several verbal features are still maintained. In (22) and (23) , a valency changing operator – the indirect causative and the sociative causative respectively – is allowed, as well as the non-future marker, and the subject/object agreement markers.

(22) *[Nunu **a**-tepa-**r**-**a**-**nke**-su']*$_{NMLZ}$ *ka'-n-awe.*
monkey CAUS-kill-NON.FUT-1S-2O-NMZR eat-NON.FUT-1
'I eat the monkey I made you hunt.'

(23) *[Nunu ka **ichi**-tepa-**r**-**aw**-su']*$_{NMLZ}$ *ka'-n-awe.*
monkey 1 SOC.CAUS-kill-NON.FUT-1-NMZR eat-NON.FUT-1
'I eat the monkey I killed with you.'

This can be captured by the following ranking:

Substantivization: -Case >> **FuncFaith** >> **LexFaith** >> *-DET >> *-POS >> *-NB >> *-CL
Deverbalization: **FuncFaith** >> **LexFaith** >> IF[17] >> (AGRS/AGRo)AGR >> Mood >> Tense >> Aspect >> Voice >> Valency

As for substantivization, the only nominal characteristic acquired is case-marking. This characteristic is important, because only nouns and nominalized constructions can take an ergative marker in the language. In Example (24) , the nominalization is marked in the ergative case:

(24) *[Kema ka'-n-an-su'-**ri]*$_{NMLZ}$ *tepa-r-in wa'an.*
2 eat-NON.FUT-2-NMZR-ERG kill-NON.FUT-3 chief
'What you ate killed the chief.'

The fact that *-su'* nominalizations are very verbal is not surprising. Shibatani (this volume) already noted that grammatical nominalizations, like those with *-su'*,

16. Similarly, this also occurs with the ergative marker *-ri→-ni* in the language.

17. Illocutionary force is also protected by LexFaith in *-su'* nominalizations. This is shown in § 5.1.2.

have mostly verbal internal syntax, hence showing structural similarities to a full-clause. The extent to which they are verbal, however, depends on the type of nominalizations. *-Napi* nominalizations in this case are stronger, therefore more nouny, than *-su'* nominalizations.

Moreover, *-su'* is used for **argument nominalizations**, typically showing a gap in one of the argument positions. It is the entire nominalization that denotes the non-overtly expressed argument. (25) and (26) illustrate this phenomenon:

(25) *[ka-ri kusina-r-awe-**su'**]*$_{NMLZ}$
-ERG cook-NON.FUT-1-NMZR
'what I cooked'

(26) *[i'-ke pa-i'-te-r-in-**su'**]* $_{NMLZ}$
water-LOC go-water-VAL-NON.FUT-3-NMZR
'what fell into the water'

In (25), for example, the missing argument is the object, represented by the nominalization. In (26), likewise, the missing argument is the intransitive subject, represented by the whole structure, as was argued above for (15) and (17) .

An interesting detail about the behavior of this suffix is that it can be suffixed to pronouns or nouns. Superficially, this appears to be the nominalization of a pronoun or a noun. Functionally, however, it seems to be used to emphasize what is being said among other topics, as in (27), or the speaker's illocutive force, as in (28).

(27) *Ina atate-r-in kankan pe'pete-**su'**.*
3 push-NON.FUT-3 wasp hive-NMZR
'It is the wasp nest that he knocks down.'

(28) *Ka-**su'***[18] *tupi-ra-w ku ama a-nanian-t-awa-ru' wa'wa-ru'sa.*
1-NMZR say-NO.FUT-1 NEG PROH CAUS-forget-IMP-2PL-NMZR child-PL
'It is me who says that we do not have to make our children forget that' [On Shawi: 19].[19]

The Shawi copula marker for the third person singular is *-ø,* see (17). The cases presented in (27) and (28) are simply nominalizations of copula constructions which function as cleft constructions. Since copula constructions in the language cannot carry TAM markers, the perception of these constructions as nominalizations of NPs is just superficial. Below I present a couple more examples:

18. Pronouns in Shawi are commonly suffixed with a *-su'* nominalizer in discourse.

19. "On Shawi" is a natural conversation from our corpus, the source of many of our examples.

(29) *Ina-ke yu-sha-ri atate-r-in. [Wa'washa-**ø-su**]$_{NMLZ}$ ni'ni-re'*
3-LOC deer-DIM-ERG push-NON.FUT-3 child-3.COP-NMZR dog-COM
ahpuiachin tawan-ke pa'-sa-r-in-we.
together cliff-LOC go-PROG-NON.FUT-3-NEG
'The deer pushed [them]. It was the child, together with the dog , who falls from the cliff.'

(30) *Na'pite=aipi ma'su wa'washa wania-t-a-r-in, a'atu'pa*
stone=on.top HESIT child stand-VAL-PROG-NON.FUT-3 there
*[pu'u-**ø-su**]$_{NMLZ}$ wense-a-r-in.*
owl-3.COP-NMZR sit-PROG-NON.FUT-3
'The child was standing on top of the stone. There, it was the owl who was sitting.'

Moreover, demonstratives in the language also include the contemporary nominalizer *-su*' as part of their root. Below I present the contemporary determiners in Shawi:

Table 6. Demonstrative pronouns in Shawi

Deictic	Meaning
*i-**su***'	this
*pa-**su***'	that
*a-**su***'	that one over there

Although it would be farfetched to say that this *-su*' is still a nominalizer in the language, this seems to have been the case historically. The roots *i-* 'here', *pa-* 'there next to you', and *a-* 'over there' were more independent – possibly demonstrative adjectives – in previous stages of the language. This can be inferred from the fact that, to date, the third person singular pronoun in shawi is *ina,* which could be segmented into *i-na* 'this-PRON'.[20] Moreover, *asu*' 'that one over there', pronounced [ah.su], can still be found without *-su*' in the Paranapura varieties as *ha.* Their original meanings as cleft constructions could tentatively be inferred as **i-ø-su*' 'It is the one here', *pa-ø-su*' 'it is the one there', *a-ø-su*' 'it is that one over there', always including the third person singular copula empty morpheme *-ø.*

-Su' can also be found attached to question words like *in-* 'who', *ma-* 'what', and *unpu-* 'when'. These seem to have originated in cleft constructions as well.[21]

20. *-Na* survived in modern Shiwilu in the third person singular pronoun *nana.* In Shawi it can be also found as the third person plural subject suffix in verbal morphology, § 2, Table 1.

21. It is very possible that *-su*' is lexicalized with these questions words, thus forming indefinite pronouns and losing the cleft construction meaning. This question awaits further investigation.

In the case *-su'* is suffixed to these question words, they acquire an indefinite pronominal meaning as in the following examples:

(31) *Kema in-**su'*** *nuwan-te-r-an?*
2 which-NMZR want-VAL-NON.FUT-2?
'Do you like something?'

(32) *Kema ma-**su'*** *nuwan-te-ra-n?*
2 what-NMZR want-VAL-NON.FUT-2
'Do you want something?'

(33) *Ka unpu-**su'*** *saka-te-r-aw.*
1 how-NMZR work-VAL-NON.FUT-1
'I work somehow.'

These questions words, however, are not just nouns. They can be classified within different categories. *Ma-* 'what' is very nouny. It allows for case markers to be added to it without a nominalization, e.g. *ma-ke-ta',* 'on what'. On *in-* 'who', however, this is not possible. *In-* can only take a case marker if it was previously nominalized with *-su'*, as in, *in-**su'**-ke,* 'where/in something which is', or via other means as in *in-**tu***[22]*-pa,* 'towards where'. These case-marked nominalized bases can be nominalized once again, as in (34)–(35) :

(34) *Ina [in-**tu**-pa-ø-**su'**]* $_{NMLZ}$ *pa'-sa-mara-ø.*
3 which-land-PERL-3COP-NMZR go-PROG-DUB-3
'It is towards somewhere that he might go.'

(35) *Ina ya'we-r-in [in-**se***[23]*-ke-ø-**su'**]*$_{NMLZ}$ *nimara, nipirinwe ku*
3 live-NON.FUT-3 which-NMZR-LOC-3.COP-NMZR maybe but NEG
[ka-ø-su']$_{NMLZ}$ *nuwi-te-r-awe.*
1-3.COP-NMZR know-VAL-NON.FUT-1
'It is somewhere that he lives, but it is me who doesn't know where.'

(36) *Ina [in-**tu**-pa-ke-ran-**ø**-**su'**]* $_{NMLZ}$ *nimara we'-n-in.*
3 which-NMZR-PERL-LOC-ABL-3COP-NMZR maybe come-NON.FUT-3
'It is from somewherethat he might come.'

Once nominalized, these constructions can become *ever*-free-relatives by means of the suffixation of *-na* (*cf.* Ulloa ms.). Below I present some examples:

22. *-tu* is originally the Shawi *-nu',* meaning 'land', which loses its nasality due to an apparent phonological restriction: two contiguous nasals are not allowed.

23. This is an allomorph of *-su'*. Before the central vowel <e>, *-su'* becomes *-se'*.

(37) *[In-**su**'-pei-ke-**ø**-**su**']* $_{NMLZ}$ *-**na*** *ma'sha ya'we-r-in.*
which-NMZR=house-LOC-3COP-NMZR-FR animal exist-NON.FUT-3
'No matter in which house it is, there are animals.'

(38) *In-**su'-na*** *nanan [ke-te-r-in-ku-su']*$_{NMLZ}$ *ku*
which-NMZR-FR word take-VAL-NON.FUT-3A-1R NEG
sha'wi-te-r-a-nkema-we, *ite-r-in* *Kisusu-ri.*
tell-VAL-1.SBJ-2.PL.OBJ-CBJ-NEG say-NON.FUT-3 Jesus-ERG
'Whoever (lit. no matter who that is) gave me the authority I won't tell you, said Jesus.' (Matthew 21:27, adapted from Ulloa ms.:11)

Moreover, one sees that that the nominalized version of *ma'*- 'what' can be theticalized (*cf.* Kaltenböck, Heine & Kuteva, 2011, for a survey of thetical grammar) into which shows a certain similarity to the English hesitation particle 'ehm' or the Peruvian Spanish 'este', as in (39) :

(39) *I-su'-ita* ***ma'su'*** *i-su'* *wa'washa-ru'sa ventana-ke-ran*
this-NMZR-PL THET this-NMZR boy-PL window-LOC-ABL
newe-ta-pi.
look.for-VAL;PROG-3.PL
'They, ehm, (and) these children were looking for it from the window.'

5.1.1 *Purposive* -ka-*1/2/3*-su'*(= mare')*

The purposive construction is commonly used as a complement in a predicate headed by a verb. The Shawi purposive marker *-ka* behaves formally in the same fashion as other modal particles (occupying the same slot in a verb), even though the semantics of modality are taken from the main clauses that precede it.

In Shawi, the purposive is always followed by a person suffix and the nominalizer *-su'*. The nominalized forms in (40)–(42) also display the third person singular copula suffix *-ø*. In (43)–(44) , common person-marking suffixes for the purposive are used. Table 7 presents the verbal endings in the purposive construction followed by examples of each ending:

Table 7. Verbal endings in the purposive

	Singular	**Plural**
1	*-wa*[a]*-su'/-ka-su'*[b]	**Incl.:** *-wa-i-su'*
		Excl.: *-wa-su'*
		Dual: *¿?*
2	*-ka-ma-su'*	*-ka-ma-su'*
3	*-ka-ø-su'*	*-ka-i-su'*

a This seems to be a historically fused form of the purposive marker *-ka* and the first person singular suffix *-we*.

b Something interesting in the purposive verbal themes is that they always carry the canonical NMZR marker *-su.'*

(40) *Unpu pi'i-ta' '[ya'we-te-r-in-su']* $_{NMLZ}$ *[sha'wi-ta-**ka-ø-su**']* $_{NMLZ}$*?*
how.long-INT exist-VAL-NON.FUT-3-NMZR say-VAL-PURP-3.SG-NMZR
'How old are they to talk?' [On Shawi: 5].

(41) *[Ka'tu' nitu-ta-**ka-ø-su**']* $_{NMLZ}$ *wa'wa-ru'sa nun-pi-su'.*
two know-VAL-PURP-3-NMZR child-PL talk-3.PL-NMZR
'That the children learn to speak two (languages)' [On Shawi: 9]

(42) *Kunpanama' iraka ya'we-r-in, a'chi-n-in nani [ma'sha*
Cumpanama long.ago live-NON.FUT-3 teach-NON.FUT-3 already thing
*[nika-**ka-ø-su**']* $_{NMLZ}$*].*
do-PURP-3-NMZR
'Long time ago, Cumpanama lived and taught (people) to do many things.'

(43) *Ina-pita pa'an-pi ka'tu' awarinchi [kanpuwa'*
3-PL buy-3.PL 2 moonshine 1.INCL.PL
*[weka-**wa-i-(su)**']* $_{NMLZ]}$ *=mare'.*
come-PURP;1.INCL-PL-NMZR=BEN
'They buy two (bottles) of moonshine so that we come.'

(44) *Kanpuwa pei pa'a-n-ewa' [a'na taweri ina-ke*
1.INCL.PL house buy-NON.FUT-1.INCL-PL one day 3-LOC
*[ya'we-**ka-wa-su**']* $_{NMLZ]}$ *=mare'.*
live-PURP-1.INCL-NMZR=BEN
'We buy a house so that, one day, we live there.'

As can be inferred from these examples, internally, the nature of these nominalizations in the GSM is the same as that of *-su*', that is, most verbal properties are maintained. This construction is not used in the same fashion as the normal weak nominalization constructions. It is commonly used to convey **event nominalizations** or, more precisely, scope nominalizations.

Finally, the nouniness of these constructions when the benefactive case enclitic is added, as in (43) or (44) , must be noted. This is expected, since we are dealing with a nominalization.

5.1.2 *Prohibitive ama V-IMP-(su')-we*

Prohibitives in Shawi are formed by means of the prohibitive particle *ama* plus an optional nominalization and a negative suffix indexed in the verb. Nominalization is common among negative constructions in the world (*cf.* Singerman, 2016 for Tupari). Below I present prohibitive constructions both with a nominalization, as in (45) , and without, as in (46) . Both were unreservedly accepted by consultants.

(45) ***Ama*** *inute-ran* *[ka'-sa-**ke-su'-we**[24]]*$_{NMLZ}$*
PROH vagina-ABL eat-PROG-2.IMP-NMZR-NEG
'Do not think of having sex.'

(46) ***Ama*** *wenu* *u'u-**ke-we**.*
PROH manioc.beer drink-2.IMP-NEG
'Do not drink manioc beer.'

It should be noted as well that, just like any other *-su*' nominalization, this construction retains most verbal features. The prohibitive type of nominalization in Shawi falls under the same GSM typology as common *-su*' nominalizations, § 5.1.

6. Discussion

As was shown in the previous sections, Shawi has very strong nominalizations such as the agentive nominalizer *-te'/-tun*[25], the instrument nominalizer *-nan,* and the patient nominalizer *-pi,* which lose most verbal features in the deverbalization cline and gain most features in the substantivization cline. There are, however, other less strong nominalizations which can still carry some verbal features, like the agentive nominalizer *-napi*. As it was presented in § 3, some of these nominalizations are synchronically lexicalized. Other nominalizing suffixes, like *-pi* and *-napi,* are not restricted to lexicalized stems and are highly productive, combinable with almost any verb. These nominalization constructions can be classified as LEXICAL NOMINALIZATIONS, in line with Shibatani (this volume). However, nominalizations formed with *-napi* are not as strong as with the other nominalizers, and thus it serves to create both lexical and grammatical nominalizations.[26]

24. From a historical point of view, these constructions might have been introduced into Shawi through contact with Quechua. This language shows a parallel construction for prohibitives, see (a), as also noted in Rojas-Berscia (2013), Valenzuela (2015), and Ulloa (ms.).

(a) *Ama upya-y-chu* *yaku-ta!*
PROH drink-2.IMP-NEG water-ACC
'Do not drink water!' (Chanka Quechua)

It is true, however, that no nominalizer can be found in Quechua for these constructions. That might have been an innovation in Shawi.

25. They might be historically related. A comparative study involving the sister language Shiwilu, remains in process.

26. The latter only when voice or valency markers are kept.

There are also weak nominalizations in the language, as shown with the GSM. Nominalizer *-su',* which functions as a relativizer, a purposive marker after suffix *-ka,* or a prohibitive, in *ama* constructions, lacks most noun properties and retains many of its original verbal properties. *-Su'* nominalizations can be found as ARGUMENT NOMINALIZATIONS, as expansions of argument slots in the form of headless relative clauses, and EVENT NOMINALIZATIONS, as purposive constructions. The case of the prohibitive is special, but not typologically uncommon. In line with Shibatani's proposal, these three types of nominalization can be considered GRAMMATICAL NOMINALIZATIONS. A typology of Shawi nominalization is given below.

As can be seen in Table 8, STRONG NOMINALIZATION tends to correlate with Shibatani's LEXICAL NOMINALIZATION. Nominalizers like *-te*, *-tun*, and *-nan* were possibly also used for grammatical nominalizations at some point in the history of the language. Apparently, the stronger a nominalization, the greater the likelihood of it becoming a lexicalized nominalization. Lexicalization correlates well with Shibatani's proposal in the sense that very lexicalized nominalizations are always lexical nominalizations.

Table 8. Typology of nominalization in Shawi

Nominalizer	Formal type	Function	Lexicalized
-tun	STRONG	LEXICAL	+
-te	STRONG	LEXICAL	+
-nan	STRONG	LEXICAL	+/−
-pi	STRONG	LEXICAL	−
-napi	STRONG (less strong)	LEXICAL	−
-su	WEAK	GRAMMATICAL	−*

* It could be still argued that *-su* is lexicalized in deictics.

Malchukov's (2004, 2006) Generalized Scale Model is well-suited to describe and explain what occurs when nominalizations are formed in the language. In previous literature, there was a lack of specification in the dichotomy of strong vs. weak nominalizations. Malchukov's model helps to capture differences in the behavior of nominalizers by locating the FUNCFAITH and LEXFAITH constraints on a hierarchy of features associated with verbhood and nounhood. Shibatani´s proposal captures well the typology of the constructions found in the language in terms of lexicalization and productivity.

Typologically, it would be interesting to know if other languages of northwestern Amazonia or regions close to the Shawi speaking area also show a similar or identical interplay of the LEXFAITH and FUNCFAITH constraints, as well as a similar correlation between levels of lexicalization and lexicality/grammaticality

of the nominalizations concerned. If more areal commonalities were found, we would be able to deepen our understanding of this large area which involves not only Kawapanan, but also Chicham, Candoan, Arawakan, Tupian and Quechuan languages. This is still, however, an open question.

Last but not least, from a theoretical point of view, it would be useful to seek more in-depth interactions of more formal and ecological models of language description – e.g. the GSM and Shibatani's respectively for this case[27] – in order to generate richer descriptions of identified categories of a language. This could also help to enrich future descriptions of any given language in other aspects besides nominalization and therefore provide better accounts of certain phenomena for future typologically and theoretically-oriented studies.

Acknowledgments

This article would not have been possible without the help of my once assistant and eternal best friend Corentin Bourdeau. His help and accurate linguistic eye in the field helped me a great deal. Additionally, I wish to thank my supervisors, Pieter Muysken and Stephen Levinson, as well as my former advisors Gunter Senft and Simeon Floyd, for checking previous drafts of this article and for their constant patience and support. My special thanks go to my field advisor Roberto Zariquiey for his support after the first final draft of this article was lost in the waters of the Paranapura River. I am also thankful to Mark Dingemanse, my current daily supervisor, for providing me with many insightful comments which led to many improvements in this article. Pieter Seuren also played a significant role, providing me with constant theoretical support. Last but not least, I must thank my Shawi friends and consultants. Their patience and smiles were my best tools when trying to provide a proper account of their language.

List of conventions used

1	first person	IMP	imperative
2	second person	INCL	inclusive
3	third person	INT	interrogative
A	subject of transitive clause	LOC	locative
		NEG	negative
ACC	accusative	NMLZ	nominalization
ABL	ablative	NMZR	nominalizer
ADJTZ	adjetivization	NON.FUT	non-future
AG	agent	OBJ	object
AGR	agreement	PAT	patient

27. See Seuren (2009) for a historical overview and definition of formalism vs. ecologism.

AGR$_S$	subject-agreement
AGR$_O$	object-agreement
ALIEN	alienable
BEN	benefactive
CAUS	causative
COMP	comparative
COP	copula
DIM	diminutive
DUB	dubitative
ERG	ergative
EXCL	exclusive
FR	free relative
FRUST	frustrative
HIP	hipothetical
PERL	perlative
PL	plural
PROG	progressive
PROH	prohibitive
PURP	puposive
R	recipient
REFL	reflexive
S	subject of intransitive clause
SOC.CAUS	sociative-causative
SBJ	subject
THET	thetical
VAL	valenciator

References

Barraza de García, Yris. 2005a. El sistema verbal en la lengua shawi. PhD dissertation, Federal University of Pernambuco.

Barraza de García, Yris. 2005b. ¿Es la lengua Shawi una Lengua Activa? In *Memorias del Congreso de Idiomas Indígenas de Latinoamérica-II*. Austin TX: University of Texas at Austin.

Bourdeau, Corentin. 2015. Ergativity in Balsapuerto Shawi. MA thesis, Radboud Universiteit Nijmegen.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English Transformational Grammar*, Roderick Jakobs & Peter S. Rosenbaum (eds), 184–221. Boston MA: Ginn and Company.

Comrie, Bernard & Thompson, Sandra A. 1985. Lexical nominalization. In *Language Typology and Syntactic Description*, Vol. 3: *Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 349–398. Cambridge: CUP.

Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Chicago IL: University of Chicago Press.

Hart, Helen. 1988. Diccionario Chayahuita-Castellano Castellano Chayahuita: Canponanquë nisha nisha nonacaso'. Lima: SIL International.

Kaltenböck, Gunther, Heine, Bernd & Kuteva, Tania. 2011. On thetical grammar. *Studies in Language* 35(4): 848–893. https://doi.org/10.1075/sl.35.4.03kal

Koptjevskaja-Tamm, Maria.1993. *Nominalizations*. London: Routledge.

Lehmann, Christian. 1988.Towards a typology of clause linkage. In *Clause Combining in Grammar and Discourse* [Typological Studies in Language 18], John Haiman & Sandra A. Thompson (eds), 181–226. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.18.09leh

Malchukov, Andrei L. 2004. *Nominalization/verbalization, Constraining a Typology of Transcategorial Operations*. Munich: Lincom.

Malchukov, Andrei L. 2006. Constraining nominalization: Function/form competition. *Linguistics* 44: 973–1009. https://doi.org/10.1515/LING.2006.032

Payne, Thomas E. 1997. *Describing Morphosyntax*. Cambridge: CUP.

Prince, Alan & Smolensky, Paul. 2004. *Optimality Theory, Constraint Interaction in Generative Grammar*. Oxford: Blackwell. https://doi.org/10.1002/9780470756171.ch1

Rojas-Berscia, Luis Miguel. 2013. La sintaxis y semántica de las construcciones causativas en el chayahuita de Balsapuerto. Licentiate thesis, PUCP, Lima.

Rojas-Berscia, Luis Miguel. 2015. Mayna, the lost Kawapanan language. *LIAMES*. Campinas: UNICAMP. https://doi.org/10.20396/liames.v15i2.8642308

Rojas-Berscia, Luis Miguel & Bourdeau, Corentin. 2018. Optional or syntactic ergativity in Shawi? Distribution and possible origins. *Linguistic Discovery* 15(1): 50–65. https://doi.org/10.1349/PS1.1537-0852.A.481

Schachter, Paul, and Timothy Shopen. 2007. Parts-of-speech systems. In *Language Typology and Syntactic Description*. Timothy Shopen, ed. CUP: CUP.

Seuren, Pieter A. M. 2009. *Language in Cognition: Language from Within*, Vol. 1. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199559473.001.0001

Seuren, Pieter A. M. 2018. *Semantic Syntax*, 2nd rev. edn. Leiden: Brill.

Shibatani, Masayoshi. 2009. Elements of complex structures – Where recursion isn't: The case of relativization. In *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution* [Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 163–198. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.07ele

Singerman, Adam. 2016. *Negation as nominal privation in Tupari*. Talk for AMAZONICAS VI. Leticia: Colombia.

Ulloa, Luis. n.d. Interrogative pronouns and interrogative-like constructions in Shawi. Ms, University at Buffalo.

Valenzuela, Pilar. 2012. La lengua shawi. Ms. <http://www.chapman.edu/wilkinson/_files/pdf/Shawintro.pdf>

Valenzuela, Pilar. 2015. ¿Qué tan amazónicas son las lenguas cahuapana? Contacto con las lenguas centro-andinas y elementos para un área lingüística intermedia. *Lexis* 39(1): 5–56. Lima: Pontificia Universidad Católica del Perú.

CHAPTER 14

Clausal nominalization in Kakataibo (Panoan)

Daniel Valle and Roberto Zariquiey
University of Mississippi / Pontificia Universidad Católica del Perú

This paper describes the meaning, functions and structure of grammatical nominalizations in two dialects of Kakataibo, a Panoan language spoken in the Peruvian Amazon basin. Grammatical nominalization is a pervasive feature of Kakataibo connected speech, which is explained by the wide array of functions they are used for, which include relativization and complementation. In terms of their morphosyntax, grammatical nominalizations show a mixed behavior in that they combine internal clausal properties with external nominal features. These synchronic properties shed some light into the historical development of some of the switch-reference suffixes in the language.

1. Introduction

This paper describes the meaning, functions and structure of grammatical nominalizations in two dialects of Kakataibo, the Lower Aguaytía dialect (LA) and the San Alejandro dialect (SA). Following Shibatani (this volume), we understand grammatical nominalization as a metonymic process by means of which a clause-like constituent is derived into a construction capable of denoting an entity-like concept. Grammatical nominalizations, which are the result of the process mentioned above, are one of the most pervasive constructions in spontaneous Kakataibo speech. Given the wide array of functions that these constructions are used for (see § 3.4), they are highly frequent in connected speech.

The extensive use of grammatical nominalizations seems to be a shared feature among Panoan languages. However, this construction has only been studied thoroughly for a few of them, such as Shipibo-Konibo (Valenzuela 2003: Ch. 10) and Matses (Fleck 2003: 1011–1048, also this volume). For other languages of the family, short descriptions of nominalization are available (e.g. Shanenawa, Cândido 2005). As for the Kakataibo language, Winstrand (1968) studied grammatical nominalizations focusing on their relativizing function whereas Zariquiey (2011a: Ch. 20) and Valle (2017) provides a more general picture of the properties

https://doi.org/10.1075/tsl.124.14val

of both lexical and grammatical nominalizations in the LA and SA dialects, respectively. This paper provides the first comparative analysis of grammatical nominalization in the LA and the SA dialects of Kakataibo.

As will be demonstrated, grammatical nominalizations in both dialects show a mixed behavior in that they combine internal clausal properties with external nominal features (§ 3.3). These synchronic properties shed some light into the historical development of some of the switch-reference suffixes in the language (§ 4).

The data for this paper mainly come from firsthand fieldwork on naturalistic data that included different kinds of texts (e.g. traditional stories, procedures, etc.) and conversations. Each author has worked on a different dialect (Valle on SA and Zariquiey on LA) and they together have produced a database of approximately 42 hours of fully annotated speech. In addition, hundreds of sentences targeting grammatical nominalizations have been elicited in both dialects.

This paper is organized as follows. In the next section, we introduce the Kakataibo language within its social context. In Section 3, we discuss the main topic of this paper: we begin by presenting the nominalizing suffixes in § 3.1. Then, we distinguish the different types of nominalizations attested in Kakataibo (§ 3.2), their internal syntax (§ 3.3) and their external syntax (§ 3.4). Section 4 briefly explores some of the historical links between nominalizing and switch-reference suffixes in Kakataibo. Our final remarks are presented in § 5.

2. The Kakataibo language and its dialects

The Kakataibo language (also known as Kashibo) belongs to the Panoan language family, within which it represents the westernmost branch and, therefore, the closest one to the Andes mountains. Shell (1965, 1975) states that Kakataibo is the phonologically most divergent Panoan language in her database (see Zariquiey 2011a: 9–10 for a brief discussion). In accordance with Shell (1965, 1975), D'Ans (1973) includes Kashibo and Kakataibo as the only two "languages" of his Pre-Andean subgroup, and Loos (1999) claims that Kakataibo is an ungrouped Panoan language. Fleck (2013) recognizes a "Kashibo" subgroup in his proposed mainline branch (which in his Panoan classification is opposed to the Mayoruna branch, that groups the Northern Panoan languages). Thus, as briefly summarized here, regardless of its different proposed positions within the internal structure of the Panoan family, there is general agreement that Kakataibo represents an independent subgroup/branch within it.

According to the most recent Census of Indigenous Communities of the Peruvian Amazon (INEI: 2007), currently the Kakataibo number about 1879. However, the Kakataibo's political organization (FENACOCA) considered that

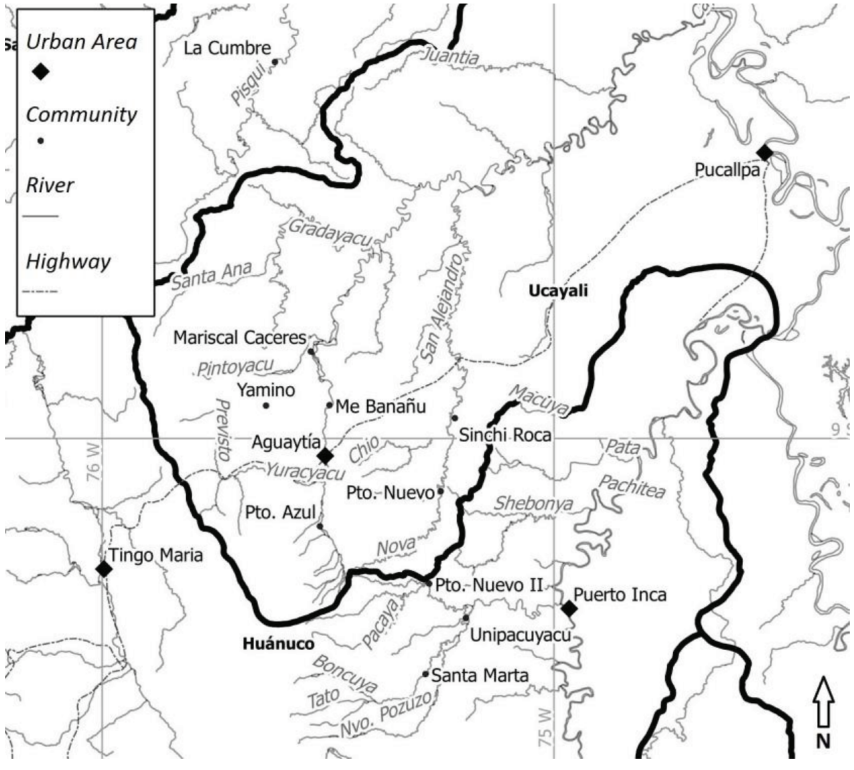


Figure 1. Kakataibo native communities (Zariquiey 2013)

their population was around 3,000 or 3,500 in 2007 (Fernando Estrella, p.c.). This figure is the one that we assume here to be more accurate.

One salient aspect of the Kakataibo language is its dialectology, which includes (at least) four distinct extant dialects. Such dialectal complexity is even more interesting if we take into consideration the relatively small number of Kakataibo speakers and their geographical proximity (see Figure 1). The dialect from the San Alejandro River is clearly the most divergent within the entire language, and it is possible to argue that there are two main Kakataibo subgroups (as was also proposed in Cortez-Mondragón 1998). All this information is summarized in the following dialectal tree (adapted from Zariquiey 2011b).

In terms of its morphosyntactic properties, Kakataibo is a (mainly) post-positional language, with a clear tendency towards agglutinating and synthetic structures, found mostly in verbs. However, Kakataibo also exhibits a closed set of prefixes associated with body-parts (see Zariquey & Fleck 2012) and there are cases of fusion and *portmanteau* morphemes in the language. Kakataibo has a split alignment system that varies in its details in the two dialects studied here. Nouns in both dialects follow an ergative alignment where the A participant is marked

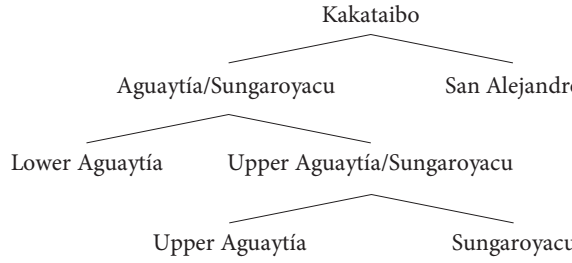


Figure 2. Extant dialects of Kakataibo

by *=n*[1] while the S and O participants are unmarked. The Lower Aguaytía dialect shows a tripartite alignment in pronouns where A is marked by *=n*, S is marked by *=x* and O is left unmarked. The San Alejandro dialect has an accusative alignment in pronouns, in which the nominative (A and S) is marked by *=n* whereas the O argument is left unmarked. Notice that this basic alignment system may be overruled in connected speech by information structure requirements related to topicality and focus, which lead to other types of alignments (e.g. neutral, Zariquiey 2011a: 724–728; Valle 2014).

The clause in Kakataibo usually contains two main elements, the main inflected verb and a second-position enclitic complex, although these two elements may be omitted in connected speech. This second-position enclitic complex contains a series of obligatory markers that encode notions such as register, evidentiality and subject agreement with slight differences among the dialects considered here. The first position of the clause or pre-field, right before the second-position enclitic complex, may be occupied by any constituent except the main verb, or it can also be left empty. The medial position, between the clitic complex and the main verb, can be occupied by any number of phrases. The presence and position in which arguments occur in the clause is strongly influenced by information structural factors such as focus and topic (Valle 2014), AOV and SV being the preferred constituent orders in out-of-the-blue sentences. Note that the verb is obligatorily clause-final in the SA dialect, whereas in the LA dialect the verb may be followed

1. Although the phonological systems of the two dialects examined here share various features, some differences are found in the fricative and approximant series as well as in the number of vowels. Both dialects have the same series of stops and nasals and one flap. The graphemic form, in square brackets, is followed by the corresponding phonological form, in between slashes, as customary: <p> /p/, <t> /t/, <k> /k/, /kw/ [kʷ], <ʻ> /ʔ/, <m> /m/, <n> /n/, <ñ> /ɲ/, <r> /ɾ/. The LA dialect has, in addition, two affricates <ts> /ʦ/, <ch> /ʧ/, three fricatives <s> /s/, <sh> /ʃ/, <x> /ʂ/, <b > and one approximant /β̞/. In turn, the SA dialect has one fricative <s> /s/, two affricates <x> /dʐ/ and < ch> /ʧ/ and three approximants <b> /w/, y /j/ and < d> /ɻ/. Both dialects show the vowels <i> /i/, <a> /a/, <o> /o/, <u> /u/ and <ë> /ɨ/. The LA dialect has an additional vowel <e> /e/.

by a focused constituent (Zariquiey 2011: 715–717). Verbs are lexically transitive or intransitive, with only a few instances of ambitransitive verbs in the language, and exhibit interesting processes of transitivity harmony and transitivity agreement, which constitute one of the core properties of the syntax of Kakataibo and the Panoan family in general.

3. Grammatical nominalizations

3.1 Markers

Grammatical nominalizers in Kakataibo are verbal suffixes that attach after derivational morphology and are usually mutually exclusive with most verbal inflectional suffixes (see Zariquiey 2011: 620–623 and Valle 2017: 389–394 for a list of verbal inflectional categories in LA and SA, respectively) and switch-reference markers. The table below shows a simplified verbal template that indicates which inflectional suffixes can be combined with nominalizations (see also example (2) for a case of a verb carrying a tense marker and a nominalizer). Notice that nominalizing suffixes occur instead of verbal suffixes of slots 4, 5, and 6 (and this is also true for switch-reference suffixes).

Table 1. Kakataibo verbal template

−1	0	1	2	3	4	5	6
Body-part	Root	Valency adjusting	Associated-motion	Tense	Aspect	Third person	Proximity/ Affectiveness
					Nominalizer		
					Switch reference		

Kakataibo shows five nominalizing morphemes *-kë* 'non-future NMLZ', *-ti* 'future NMLZ', *-a* 'remote past NMLZ', *-ai* 'present non-habitual NMLZ' and *-tibu* 'present habitual NMLZ'. This last nominalizer has not been attested in the San Alejandro dialect. The example sentences in (1) vary only in the nominalizer used, which in turn changes the temporal perspective of the nominalized clause. Grammatical nominalizations are enclosed in between square brackets in the morpheme breakdown line and are boldfaced in the free translation in order to facilitate their identification.

(1) a. *uni* *[ʻinu* *rëtë-kë]=ka=a* *kwan-i*
man jaguar kill-N.FUT.NMLZ=VAL=3A/S go-IPFV
'**The man who killed/kills the jaguar** goes away.' (LA)

b. *uni [ʻinu rëtë-a]=ka=a kwan-i*
ʻthe man who killed the jaguar years ago goes away.ʼ
c. *uni [ʻinu rëtë-ti]=ka=a kwan-i*
ʻthe man who is going to kill jaguar goes away.ʼ
d. *uni [ʻinu rëtë-tibu]=ka=a kwan-i*
ʻthe man who always kills jaguars goes away.ʼ
e. *uni [ʻinu rëtë-ai]=ka=a kwan-i*
ʻthe man who occasionally kills jaguars goes away.ʼ

3.1.1 *Nominalizer* -kë

Grammatical nominalizations with *-kë* indicate that the event has a non-future interpretation, which covers both past and present events. Notice that, as is the case with the remaining Kakataibo nominalizers, the grammatical function that the grammatical nominalization codes in relation to the main predicate is not indicated by the nominalizing suffix itself, but by the internal syntax of the nominalized clause (see § 3.2). The verb of the grammatical nominalization in (2) receives a recent past interpretation that is triggered by the presence of the suffix *-ëxan* ʻrecent pastʼ in the verb of the nominalized clause. In contrast, the nominalized clause in (3) could be interpreted as referring to a past or current event since no tense-aspect marker restricts its interpretation.

(2) *[a=n ëd-bait-ëxan-kë]* *a=n=ka=a* *ki-i-a*
3=A/S see-DUR-REC.PST-N.FUT.NMLZ 3=A/S=VAL=3A/S say-IPFV-N.PROX
ʻ**Those who saw (it) the other day**, they are telling (it).ʼ (SA)

(3) *[a=n ñui-kë]* *a=x=ka=a* *ë=n* *xukë* *ʻikën*
3=A tell-N.FUT.NMLZ 3=S=VAL=3A/S 1=POSS brother be.IPFV.3
ʻ**The one who is telling (something)/told (something)**, he is my brother.ʼ (LA)

The suffix *-kë* ʻnon-future nominalizerʼ can also indicate that the nominalized event occurs habitually in the present and that it is not possible to assign a temporal frame to it. Example (4) presents a sentence in which the marker triggers a habitual interpretation. Notice that *-kë* has this habitual interpretation by default when used in the present tense in the San Alejandro dialect since this nominalizer does not contrast with anything else for this function e.g. *-tibu* in the LA dialect).

(4) *pia=nun* *a=nun* *[ñuina* *ʻa-kë]*
arrow=INST 3=INST animal do-N.FUT.NMLZ
ʻWith an arrow, **with (what usually we) kill animals**ʼ (LA)

3.1.2 *Future nominalizer* -ti

The suffix *-ti* 'future nominalizer' covers the full spectrum of future events, as shown in (5). No further distinctions in the future are morphologically encoded in the verb. Given the temporal interpretation of the nominalizer *-ti*, grammatical nominalizations having it are usually found functioning as adverbial purposive clauses (see § 3.4.3), as in (6).

(5) *bëri=ka=na [ë=n bëchikë 'i-ti] ka-isa-tan-i-n*
today=VAL=1A/S 1=POSS son be-FUT.NMLZ say-IRR-go.to-IPFV-1/2
'Today I want to tell **(what) our sons will be**.' (LA)

(6) *a=nun [ñu=ira=bi maru-ti] nu=n ñu 'arakat-i*
3=INST thing=DIM=EMPH trade-FUT.NMLZ 1PL=A/S thing
'We raise animals **in order to buy things with (the earnings from it)**.' (LA)

3.1.3 *Remote past nominalizer* -a

The suffix *-a* 'remote past nominalizer' covers the temporal spectrum from approximately a year before the time of the utterance to the remote past including the historical and mythical past. Within this timeframe, the nominalizer *-a* may also indicate that the event occurred habitually or not. In example (7), the event of the nominalization occurs a year prior to the time of speaking (this is indicated by the noun phrase *bësi bari* 'other year'). The nominalization in (8) expresses that the being into existence of the speaker's ancestors took place many years or even centuries ago.

(7) *uni a=n [bësi bari Iquitos=nu kwan-a]=ka=a a=nu 'ikë*
man 3=A/S other year Iquitos=LOC go-PFV=VAL=3A/S 3=LOC be.IPFV.3
'The man **who went to Iquitos last year** is there.' (SA)

(8) *bëri=ka=na [nukën rara 'i-a]*
today=NAR=1A/S 1PL.O/POSS ancestors be-REM.PST.NMLZ
ka-isa-tan-i-n
say-IRR-go.to-IPFV-1/2
'Today I want to tell **(what) our ancestors were**.' (LA)

3.1.4 -tibu *'present habitual nominalizer'*

The nominalizer *-tibu* seems to be the result of combining the nominalizer *-ti* plus the adverbial enclitic *=bu* 'imprecise reference' in the LA dialect. The marker *-tibu* is used for present tense nominalizations with habitual aspect, which were always translated by Kakataibo speakers by means of the Spanish adverb *siempre* 'always'. One example follows.

(9) *[a=n ënu-xun pi-**tibu]* *a=x=ka=a* *Lima=nu* *kwan-i-a*
3=A here-PA:A eat-PRS.HAB.NMLZ 3=S=NAR=A/S Lima=LOC go-IPFV-N.PROX
'**The one who always eats here** is going to Lima.' (LA)

3.1.5 -ai *'present non-habitual nominalizer'*

The marker *-ai* 'present non-habitual nominalizer' requires further study. Semantically, it expresses present nominalizations with a non-habitual aspectual meaning, which was systematically translated by Kakataibo speakers as 'occasionally'. One example of this form in a grammatical nominalization is shown in (10):

(10) *[a=n ënu-xun pi-ai]* *a=x=ka=a* *Lima=nu* *kwan-i-a*
3=A here-PA:A eat-N.HAB.NMLZ 3=S=NAR=3A/S Lima=LOC go-IPFV-N.PROX
'**The one who eats here ocasionally** is going to Lima.' (LA)

3.2 Argument and event-nominalizations

Grammatical nominalizations in Kakataibo can be of two semantic types: event-nominalizations and argument-nominalizations. The former denotes "a state of affairs characterized by an event denoted by the clause" (Shibatani 2009: 191), as in *my buying of that book*. The latter denotes "an entity characterized in terms of the denoted event in which it has crucial relevance" (Shibatani 2009: 191), as in *the book which I bought* (which in Kakataibo would be expressed by a nominalization).

Argument nominalizations in Kakataibo are mainly produced by positing a gap corresponding to the targeted argument within the internal structure of the nominalization. That is, in many cases, the argument denoted by the nominalization will not be overtly expressed within its internal structure. This is illustrated in (11) and (12), where the targeted arguments denoting the subject and object, respectively, are absent in the construction. Notice that the gapped argument is indicated by Ø in the examples below.

(11) *[Ø Lima=nu kwan-ti]=ka=a* *in* *papi-i-a*
Lima=LOC go-FUT.NMLZ=VAL=3A/S wood carry-IPFV-N.PROX
'**The one who is going to go to Lima** is carrying wood.' (SA)

(12) *mi=n cada uno [mi=n Ø kwën-kë]* *a ñu* *'a-i*
2=A/S each one 2=A/S want-N.FUT.NMLZ 3 thing do-IPFV
'You, everyone, do **what you like**, that thing.' (SA)

However, it is also possible for an argument of the grammatical nominalization to be overtly expressed within it and still be denoted by the construction, as shown in (13). In this example, the subject NP *uni* 'people' forms part of the nominalization even though that is the argument being denoted. A reading targeting a different

argument, i.e. the object, is not licensed for this construction even though the subject is overt.

(13) *[uni-n ñu ʻunan-kë] ë=n ainsi bana-i*
man=A/S thing know-N.FUT.NMLZ 1SG=POSS relative speak-A/S > S:SE
kua-akë
hear-REM.PST
'I have listened years ago to **the people who knew things**, to what my ancestors used to say.'
*'I have listened years ago to **the things that people knew**, to what my ancestors used to say.' (SA)

Another possibility is to have all the arguments occurring non-overtly. This construction is usually interpreted as targeting the object of the clause, but it could also be read as targeting the subject. The nominalizations in (14) are interpreted as referring to the O argument, the things that are being cooked and cut, not the entity cooking or cutting those things. In turn, the two arguments of the grammatical nominalization in (15) are non-explicit, but, in this case, the nominalization corresponds to the A argument. Notice that the following third person pronoun makes the A-oriented interpretation clearer by having the *=n* case marker. The reading by which the nominalization stands for the O argument is blocked by the presence of such pronoun.

(14) *ʻasa [Ø Ø ʻaru-kë] nodi [Ø Ø ʻarukë] ñu nami*
manioc cook-N.FUT.NMLZ banana cook-N.FUT.NMLZ thing meat
[Ø Ø tëa-pat-kë] a nukën ainsi nukën chaiti
cut-down-N.FUT.NMLZ 3 1PL.POSS relatives 1PL.POSS ancestors
piakëxa
eat-REM.PST-3.N.PROX
'Our ancestors, our relatives ate **cooked manioc, cook plantains, cut meat**.' (SA)

(15) *[Ø Ø uan-kë] a=n chuna=n tita akiribi*
take-N.FUT.NMLZ 3=A/S spider.monkey=POSS mother again
mëra-a-x-a
find-PFV-3-N.PROX
'**(The one) who took (it)**, that (one) found again King Kong.' (SA)
*'**(What was) taken**, that (one) found again King Kong.'

Instances of nominalizations in which only one lexical item occurs are problematic in that they could be analyzed as a case of lexical nominalization or an example of a grammatical nominalization where the arguments are not overtly expressed. Given their isomorphic syntactic structure; rather, we rely on their semantics to

distinguish them. That is, lexical nominalizations behave as nouns in that they have a more 'stable' denotation comprising a fairly uniform range of concepts. In contrast, grammatical nominalizations have a more open denotation and may evoke a variety of entity concepts. Thus, we will treat nominalizations such as those in (15) and (14) involving only one lexical item as grammatical nominalizations since they may have a more open-ended denotation, not just being able to denote a unique entity concept.

Grammatical nominalizations may also target non-core arguments. In this case, a third person pronoun marked by the specific non-core case obligatorily occurs at the beginning of the nominalization for it to target that argument. For instance, consider the sentence below (16) in which the nominalization stands for the commitative argument of the transitive verb *bë-* 'bring'. A third person pronoun *a* receives the case marker = *bëtan* 'commitative.A' for the grammatical nominalization to refer to that non-core argument. Notice that the third person pronoun in that nominalization takes the transitive commitative case marker and not the intransitive version of it suggesting that the case-marked pronoun is part of the nominalization and not an argument of the main (intransitive) clause, which would require the intransitive form of the case marker.

(16) *[a=bëtan 'asa bë-kë] kwan-a-x-a*
3=COM.A manioc bring-N.FUT.NMLZ. go-PFV-3-N.PROX
'**(The one) I brought manioc with** left.' (LA)

The use of a third person pronoun within the nominalization itself is optional for S, A and O arguments. In the example in (17), as an illustration, we find an A-nominalization expressed by an internal pronoun.

(17) *[a-n-is=a_j ain xanu 'a-kë]$_{NMLZj}$ uni$_j$*
3=A=REP=3A/S 3.POSS woman do-N.FUT.MNLZ person
'The man **who it is said that he used to have sex with (other man's) wife**' (LA)

Event nominalizations, in turn, do not require **internal gaps** and they usually occur with all their arguments overtly express. It is not uncommon to find some words of general reference within event nominalizations, such as *ñu* 'thing' or *uni* 'man'. Thus, both participant and event nominalizations may show the same structure when having all their arguments overt, although this is infrequent in spontaneous discourse. The examples in (18) and (19) contrast an argument nominalization with an event nominalization, respectively.

(18) *[nukën rara=n ø$_j$ 'a-a]$_{NMLZj}$ a=ka=na is-akë-n*
1PL.POSS ancestor=A/S do-REM.PST.NMLZ 3=NAR=1A/S see-REM.PST-1/2
'I saw **(the things) that our ancestors did a long time ago.**' (LA)

(19) *[nukën rara=n ñu 'a-a]$_{NMLZ}$ =ka=na is-akë-n*
1PL.POSS ancestor=A/S thing do-REM.PST.NMLZ =NAR=1A/S see-REM.PST-1/2
'I saw **that/how our ancestors did the things a long time ago**.' (LA)

3.3 Structure of grammatical nominalizations

In terms of their morpho-syntactic characteristics, grammatical nominalizations share a number of properties with NPs while at the same time they have clausal properties. This makes them 'look' clausal internally and nominal externally. As argued by Shibatani (this volume), in the study of grammatical nominalizations, it is necessary to make a distinction between the internal and the external syntax of the construction. In Shibatani's terms, "grammatical nominalizations have mostly verbal internal syntax, i.e. show structural resemblances to a full clause, but the extent to which they are verbal depends on the types of NMLZs; e.g. infinitival and *-ing* participial NMLZs differ from *that* NMLZs in typically lacking subject nominals". In turn, regarding their external syntax, grammatical nominalizations "head an NP like any other nouns and can function as a verb complement (Subject/Object)." In this section, we explore the internal and external syntax of grammatical nominalizations in the two dialects of Kakataibo studied in this paper.

3.3.1 *Internal syntax*

In terms of its internal syntax, grammatical nominalizations share a number of properties with main clauses. Constituent order in main clauses is pervasively verb final with the caveat that the verb may be followed by a focus constituent in the Lower Aguaytía dialect. In grammatical nominalizations the verb is always final, following a SOV pattern.

In addition, verbs in both grammatical nominalizations and main clauses may take verbal morphology that includes associated movement, valency-changing and TAM suffixes as well as body-part prefixes; but person and addressee's perspective markers are restricted to main clauses. An instance of a paradigm showing the different combinatorial possibilities of tense markers and the 'non-past nominalizer' *-kë* was given in (1) above.

Verbs in grammatical nominalizations are able to assign case to its arguments as main clause verbs do. Verbal arguments in grammatical nominalizations follow the same alignments as in main clauses retaining the dialectal differences pointed out in § 2. However, notice that the assignment of case markers is optional in grammatical nominalizations in the San Alejandro dialect.

Although grammatical nominalizations exhibit similar properties to main clauses, some differences remain. For instance, grammatical nominalizations optionally have a reduced subset of the second-position clitic complex that is

obligatory in main clauses, as can be seen in (17), in which we find the evidential marker = *is* within the nominalization. A semantic motivation may be behind this characteristic since the meanings that second-positon clitics convey include register and mood, which have sentential scope. Their absence in nominalizations is expected given their non-finite status.

3.3.2 *External syntax*

Nouns function as the nuclei of noun phrases, and as such they may be marked for case and be pluralized. In contrast to main clauses, grammatical nominalizations may function as the head of noun phrases and, therefore, host nominal morphology. Example (20) shows an instance of a nominalization marked by = *n* 'A/S' functioning as the A argument of the main clause. In (21), we see a grammatical nominalization being pluralized by = *kama* 'PL'.

(20) *a=n [uni 'asa bë-kë]=**n**=ka=a kwan-a-x-a*
3=A/S man manioc bring-N.FUT.NMLZ=**A/S**=VAL=3A/S go-PFV-3-N.PROX
'(A/the) man who brought manioc left'. (SA)

(21) *(…) no a [ri-kwasin-kë=**kama]* ki-a*
foreigner 3 go.PL-come.INTR-N.FUT.NMLZ=**PL** say-A/S/O > O:SE
kua-ti ri-kwasin-ti 'a-i
listen-FUT.NMLZ go.PL-come.INTR-FUT.NMLZ do-IPFV
'(You all) have to come to listen to what the mestizo people who came are saying.' (SA)

3.4 Functions of grammatical nominalizations

In what follows, we describe the different functions that grammatical nominalizations may code when appearing inside larger constituents (relativization § 3.4.1, complementation § 3.4.2 and verb modification § 3.4.3) or by themselves (standalone § 3.4.4). A dedicated brief section on the diachronic relation between nominalization and switch-reference can be found in § 4).

3.4.1 *Relativization*

Participant nominalizations can appear either after or before an NP in a construction which is functionally equivalent to relative clauses in other languages. Nominalizations can be heads of NPs by themselves and do not need an external nominal head to appear in discourse. This fact strongly suggests that they are not equivalent to relative clauses, which at least in their prototypical cases are dependent elements that require such a nominal head.

In Kakataibo, grammatical nominalizations may appear with (but do not grammatically depend on) an NP in appositional constructions, which are to be analyzed as $[\text{NMLZ}]_{\text{NP}}$ $[\text{N}]_{\text{NP}}$.[2] The appositional analysis is supported by the fact that modifying nouns in $[\text{N N}]_{\text{NP}}$ constructions are always pre-head. If grammatical nominalizations were NP-internal modifiers, they would be expected to appear exclusively before the head – but they can be either pre- or post-nominal. This appositional analysis also finds support in prosodic facts, particularly in the prosodic independence between the grammatical nominalization and the NP, which can be separated by a pause and even by other elements, such as a highlighting pronoun. This is shown by the following example:

(22) *[xanu]$_{jNP}$ a=x [ain$_j$ bënë is-kë=ma]$_{j\ NMLZ}$*
woman 3=S 3.POSS husband see-N.FUT.NMLZ=NEG
'The woman, she, whose husband did not see (was blind).'(LA)

In the appositional construction proposed here, an NP happens to denote the same entity as a participant nominalization (which by definition denotes one of its participants). The result is a construction with a relativizing function according to which the grammatical nominalization constrains the interpretation of the NP it co-occurs with. For instance, in the example above, the woman denoted by the NP is meant to be the same woman denoted by the nominalization: 'the one whose husband did not see'. This is also what we find in the following example (note that these constructions have been analyzed as relative clauses in the literature about Kakataibo; see, particularly, Winstrand 1968):

(23) *a=x=kais=a [uni]$_{iNP}$ [a-n=is=a$_j$ uni=n xanu*
3=S=NAR.REP=3A/S person 3=A=REP=3A/S person=POSS woman
'a-kë]$_{iNMLZ}$ […] 'i-akë-x-ín
do-N.FUT.NMLZ be-REM.PST-3-PROX
'It is said that he was a man, someone who was said to have habitually had sex with the wife of (another) man.' (LA)

3.4.2 *Complementation.* It is cross-linguistically common that not all the verbs of a given language can take (all types of) complement clauses. Usually,

2. According to Shibatani (this volume), grammatical nominalizations differ from appositives in their referential status. Grammatical nominalizations in this function restrict the denotation of the nominal element they are combined with; but are not referential by themselves. This is different from prototypical appositional constructions, where both members are referential NPs. In this context, it may be important to mention that preliminary research indicate that it is not possible to use two nominalizations in these appositional constructions and this fact seems to give support to Shibatani's analysis.

only a subset of verbs triggers complementation, and complement clauses are thus prototypically associated with some types of verbs (see Dixon and Aikhenvald 2006 for a summary of these verb classes). In Kakataibo, given the right pragmatic context, any verb can potentially occur with a grammatical nominalization as one of its core arguments or adjuncts, producing a type of construction which is reminiscent of complementation constructions in other languages. This distribution shows that the process discussed here goes beyond what is usually defined as prototypical complementation. We exemplify this fact by presenting one verb in (24) which is not included among the types of complement-taking verbs listed by Dixon, but which can take nominalizations as one of its arguments in Kakataibo.

(24) *[[María-nën ʻaru-kë]* $_{NMLZ}$*]*$_{NP}$ *a=ka=na pi-a-n*
María=A cook-N.FUT.NMLZ 3=NAR=1A/S eat-PFV-1/2
'I ate what María cooked.' (LA)

Similar grammatical nominalizations are used with verbs that are considered prototypical complement-taking verbs, such as verbs of perception (25), desire (26) or speaking (27):

(25) *[[María=n nami ʻaru-kë]*$_{NMLZ}$*]*$_{NP}$ *a=ka=na is-a-n*
María=ERG meat cook-N.FUT.NMLZ 3=NAR=1A/S see-PFV-1/2
'I saw **that María cooked meat**.' (LA)

(26) *[[María=n ʻaru-ti]*$_{NMLZ}$*]*$_{NP}$ *=kana kwëën-i-n*
María=ERG cook-N.FUT.NMLZ =NAR=1A/S want-IPFV-1/2
'I want **María to cook**.' (LA)

(27) *[[María=n nu=n pi-ti ʻaru-kë]*$_{NMLZ}$*]*$_{NP}$ *a=ka=na*
María=ERG 1PL=POSS food-FUT.NMLZ cook-N.FUT.NMLZ 3=NAR=1A/S
ñui-i-n
tell-IPFV-1/2
'I (will) tell **that María cooked our food**.' (LA)

Note that equivalent constructions with the verb *sinan-* 'think', another prototypical complement-taking verb according to Dixon and Aikhenvald (2006), require the grammatical nominalization to carry the oblique 'comparative' marker *=sa* in the LA dialect, as shown in (28), but not in the SA one.

(28) *[[María=n ʻaru-kë]*$_{NMLZ}$*]*$_{NP}$*=sa=kana sinan-i-n*
María=ERG cook-N.FUT.NMLZ=COMP=NAR=1A/S think-IPFV-1/2
'I think **that María cooked**.' (LA)

3.4.3 *Verb modification (adverbial function).* Grammatical nominalizations in Kakataibo are also used to modify verbs in an adverbial clause-like function.

Grammatical nominalizations with *-ti* 'future NMLZ' are frequently found in connected speech as purposive clauses modifying verbs, as in (29). Grammatical nominalizations marked by *-kë* may also function as adverbial modifiers being broadly translated as 'when…', as in (30).

(29) *[bësi 'a-ti] a-ribi nisi-n-tëkën-nun pëi*
other do-FUT.NMLZ 3=also stand.up-TRZR-again-A/S > A:FE leaf
tada-kin nisi-n-i
nail-A/S > A:SE stand.up-TRVZ-IPFV
'**In order to make another one**, going to (make) it also to stand up, nailing the leaf, I make (it) to stand up.' (SA)

(30) *a=n […] kwan-ti=dapi=ka=a medio día [JP u-kë]*
3=A/S go-FUT.NMLZ=DUBT=VAL=3A/S half day JP come-N.FUT.NMLZ
kwan-i-a
go-IPFV-N.PROX
'Those who will go are going at midday, **when JP comes**.' (SA)

3.4.4 *Stand-alone nominalizations.* In addition to functioning as the nucleus of an NP and being able to modify both nouns and verbs, grammatical nominalizations may also stand in for a whole sentence without an external main verb, a phenomenon that has been attested for a wide variety of languages (Evans 2007). When this occurs, the nominalized verb carries the predicative semantic load of the matrix clause. Crucially, the verb of the nominalized clause has to occur at the end of the sentence, as main clause verbs do. This is shown in example (31) where the verb *ka-* 'to say' is part of the nominalized clause and occurs sentence-finally.

(31) *[…] aunque sea algo=bi mi 'inan-mainun=ka=a*
at least something=EMPH 2 give-A/S≠A/S:SE=VAL=3A/S
pi-ti ki-xun [ë hasta ahora ka-ñá-kë=ma]
eat-FUT.NMLZ say-A/S > A:PE 1SG until now say-keep-N.FUT.NMLZ=NEG
'[…] **(he) keeps not telling me even now**, "I am giving at least something to you for you to eat."' (SA)

4. Some notes on the historical relationship between nominalizers and switch-reference suffixes

One of the most striking features of Panoan languages is their large and complex set of switch-reference suffixes. With some language-specific differences, switch-reference suffixes in Panoan languages encode at least three dimensions: (i) co-referentiality or not between one argument of the main clause and one of

the subordinate structure, (ii) temporal relation between the events (e.g. previous, simultaneous) and (iii) the grammatical function of the targeted argument in the main clause. The inventory of switch-reference suffixes varies from one language to the other. There are 21 switch-reference markers in the LA dialect of Kakataibo and 19 in the SA dialect, as shown in the table below.[3]

Table 2. Kakataibo switch-reference suffixes*

Subject to subject	X argument to object	Object to subject	No co-referentiality
-i 'S/A > S:SE'	-këtia(n)* 'S/A/O > O:PE'	-kë(x) 'O > S:PE'	-nu(n) 'DS/A:FE'
-kin 'S/A > A:SE'	-ia 'S/A/O > O:SE'	-kë(x)=bi 'O > S:SE'	-a(n) 'DS/A/O:PE'
-a(x) 'S/A > S:PE, SE'		-këxun 'O > A:PE'	-këbë 'DS/A/O:SE.INTR'
-xun 'S/A > A:PE, SE'		-këxun=bi 'O > A:SE'	-këbëtan 'DS/A/O:SE.TR'
-tankë(x) 'S/A > S:PE'			-mainun 'DS/A/O:SE'
-tankëxun 'S/A > A:PE'			
-nu(x) 'A/S > S:FE'			
-nuxun 'S/A > A:FE'			
-tanan* 'S/A > S/A:SE'			
-ana(n) 'S/A > S/A:SE.DO'			

* Segments in parenthesis are not attested in the SA dialect switch-reference suffixes.

A historical pathway for the development of a set of switch reference markers including *-xun*, *-ax*, *-kin*, *-nux* and *-nun*, which are common in Panoan languages, has been proposed by Valenzuela (2003: Ch. 20, 2013). Her proposal suggests that those switch-reference suffixes derive from instances of case stacking in which the first layer of case indicated the temporal basis while the second layer of case provided the grammatical function orientation. For instance, *-nux* and *-nun* might have derived from the fusion of **-nu* 'locative' plus *-x* 'S argument' and *-n* 'A argument', respectively.

In this section we explore the historical process by which a different set of switch-reference suffixes might have arisen. This set of switch-reference suffixes includes those that contain the form **kë* in it, such as *-tankëx*, *-tankëxun*.

3. In the San Alejandro dialect, the suffix *-tanan* 'S/A > S/A:SE' and *-nun* 'DS/A:FE' have not been attested. In this dialect *-anan* 'S/A > S/A:SE.DO' also covers the function that *-tanan* has in the LA dialect. In addition, the suffixes in the object to subject column in the chart above may target any participant S/A/O instead of only O in the San Alejandro dialect. In this way, these suffixes behave as those in the X argument to object column in this dialect. The function that the *-nun* suffix covers in the LA dialect is mainly covered by *-mainun* and *-këbëtan*.

-këtian, *-këx*, *-këxun*, *-këxbi* and *-këxunbi*, *-këbë* and *-këbëtan*. Our proposal is that these suffixes emerged by the fusion of the nominalizer **-kë* plus a case marker. Additionally, the first two suffixes might have had a formative **tan,* as Valenzuela (2013) suggests, whose function remains unknown. In a first stage of the process, a clause might have been nominalized with the suffix *-kë*. In a second step, this *-kë*-marked grammatical nominalization received a case marker in order to indicate the grammatical function that that nominalization was playing in the main clause. In other words, this case-marked grammatical nominalization was functioning as an argument of the matrix clause verb. Notice that these properties, namely nominalization by *-kë* and the ability of nominalization to be marked for case, do occur in the synchronic language, as has been described in § 3.3.2. Our proposal is schematically represented in (32) below, showing that the nominalizer and case suffixes were reanalyzed as a single switch-reference suffix with the corresponding change of meaning. Notice that core case markers *-x* 'S' and *-n* 'A' are used when the switch-reference suffix indicates same subjects while oblique cases are used when the suffix indicates different subjects. This situation is expected if we assume that the nominalization was once an argument of the main clause, as we do here.

(32) [_____ -kë]$_{\text{NMLZ}}$-CASE _____ V > [_____ -kë-CASE]$_{\text{SR}}$ _____V.

Unfortunately, we do not have historical records that enable us to confirm this hypothesis. However, Kakataibo synchronic data provide a bridging context that gives support to what has been argued here. As we have seen above, nominalizations may be marked for case and, thus, when the nominalizer *-kë* receives a case marker, that sequence may look exactly like some of the switch-reference suffixes that contain the form *-kë*, the only difference being their diverging readings. For instance, the sentences below are ambiguous between a reading in which the constituent in brackets is treated as an argument of the main verb or as a switch-reference construction. This may have led to a reanalysis of the combination of the nominalizers and the case markers as single switch-reference suffixes. The table below summarizes this process of syntactic change by showing the current switch-reference suffixes and their diachronic source components.

(33) *[mi Ø uan-kë]=ka=a abat-ti nuan-i*
2 take-**kë**=VAL=3A/S run-N.FUT.NMLZ fly-IPFV
'When you bring (it), (it) flies away.' or
'What you brought flies away.' (SA)

(34) *['asa pi-kë=bë]=ka=na rëú=mi kwan-i*
manioc eat-**kë**=**bë**=VAL=1A/S tip=DAT go-IPFV
'I go upwards with the one who eats manioc.' or
'While (he) eats manioc, I go upwards.' (SA)

Table 3. Nominalizer *-kë* and switch-reference suffixes in Kakataibo

Proposed segmentation of switch-reference suffix	Nominalizer	Case marker	Meaning
-tan-kë-x	-kë	-x 'S'	S/A > S:PE
-tan-kë-xun		-xun 'A'	S/A > A:PE
-kë-tian		-tian 'TEMP'	S/A/O > O:PE
-kë-x		-x 'S'	O > S:PE
-kë-xun		-xun 'A'	S/A > A:PE, SE
-kë-bë		-bë 'COM.S'	DS/A/O:SE.INTR
-kë-bëtan		-bëtan 'COM.A'	DS/A/O:SE.TR

While this section focuses only on the development of switch-reference suffixes in Kakataibo, the analysis proposed here seems to account for similar patterns in other Panoan languages. In other words, it seems that some Panoan languages had a similar process of syntactic change where nominalizers and case markers derived into switch-reference suffixes. Such cases are found in the Shipibo-Konibo and Yaminawa languages. In both languages, the nominalizers used are different from the ones found in Kakataibo, *-ai* 'incompletive' and *-a* 'completive' in Shipibo-Konibo and *-ai* 'progressive' and *-a* 'completive' in Yaminawa. Notice that these nominalizers seem to be cognate between these two languages. The other component of synchronic switch-reference suffixes in these languages were apparently case markers cognate with those found in Kakataibo: *=n* 'A argument', *=x* 'S argument',[4] *-tian* 'temporal' and *= no* 'locative'. Table 4 summarizes the source and derived morphemes of this process.

Table 4. Nominalizers and switch-reference in Shipibo-Konibo and Yaminawa

	Nominalizer suffix	Case marker	Switch reference suffix
Shipibo-Konibo (Valenzuela 2013)	{*-ai*} 'incompletive NMLZ'	{*=n*} 'ERG'	{*-ain*} 'simultaneous event, DS/A'
		{*=tian*} 'temporal'	{*-aitian*} 'simultaneous inmediate event, DS/A'
	{-a} 'completive NMLZ'	*{-x} 'S'	-ax 'A/S > S:PE'
Yaminahua (Faust and Loos 2002)	{*-ai*} 'PROG NMLZ'	{*=no*} 'LOC'	{*-aino*} 'simultaneous event, DS'
	{*-a*} 'completive NMLZ'		{*-ano*} 'previous event, DS'

4. *-*x* is not a synchronic case marker in Shipibo, but it might had been present in the language at the time when this process of grammaticalization took place (see Valenzuela 2003:Ch. 20).

The historical development of the switch-reference suffixes in Panoan languages is a complex topic and a full coverage of it goes beyond the goals of this paper. In this section, we aimed to offer a historical path for some of these suffixes based on their synchronic similarities that clauses having them have with grammatical nominalizations. We hope that this pattern will help to understand how switch-reference and nominalizing suffixes relate diachronically as more data becomes available in other Panoan languages.

Conclusions

This paper described the synchronic properties of grammatical nominalizations. In line with Shibatani (this volume), we argued that grammatical nominalizations denote entity-like concepts by which they resemble nouns. By having this property, grammatical nominalizations may head NPs, which in turn, allows them to occur in the same environments as NPs headed by lexical nouns do, e.g. verbal argument. By occurring in apposition with another NP, a grammatical nominalization may restrict the reference of an NP, which is functionally equivalent to the prototypical function of a relative clause. Other functions of grammatical nominalizations, such as verb modification and standing-alone, push the boundaries of what a prototypical 'noun' is able to do since these functions do not seem to have a core denotational meaning. While Kakataibo nominalizing suffixes also encode tense and aspect, they are neutral with regard to the participant they target. In order to signal the argument the nominalizations stands for, a gap occurs in its place instead. By combining nominal and clausal properties, grammatical nominalizations in Kakataibo suggests that a clear-cut distinction between lexical categories does not account for the phenomena described here. Finally, the understanding of the functions and structure of nominalizations might help to elucidate the diachronic development of a subset of the switch-reference suffixes in this language.

Acknowledgments

Thanks to the audience of the workshop 'Nominalization in indigenous languages of the Americas' and to the Kakataibo speakers who taught us the use of nominalizations in their language.

Abbreviations

1	first person
2	second person
3	third person
A	subject of transitive verb
COM	commitative
DIM	diminutive
DS	different subjects
DUBT	dubitative
DUR	durative
EMPH	emphatic
ERG	ergative
FUT	nominalizer
INTR	intransitive
INST	instrumental
IPFV	imperfective
IRR	irrealis
LOC	locative
NAR	narrative
NEG	negation
N.FUT	non-future nominalizer
NMLZ	nominalizer
N.PROX	non-proximate
O	object of transitive verb
PA	participant agreement
PE	previous event
PFV	perfective
PL	plural
POSS	possessive
PROX	proximate
PRS.HAB	present habitual
REC.PST	recent past
REM.PST	remote past
REP	reportative
S	subject of intransitive verb
SE	simultaneous event
TEMP	temporal
TR	transitive
TRZR	transitivizer
VAL	validational

References

Cândido, Gláucia. 2005. Construções relativas na língua shanenawa (pano). *Estudos Lingüísticos* 34: 575–579.

Cortez-Mondragón, María. 1998. Cashibo-Cacataibo. In *Multilinguismo en el Perú*, Inés Pozzi-Escott (ed.). Cuzco: Centro de Estudios Regionales y Andinos Bartolomé de Las Casas.

d'Ans, Andre-Marcel. 1973. *Estudio glotocronológico sobre nueve hablas pano*. Lima: UNMSM.

Dixon, Robert M.W & Aikhenvald, Alexandra Y. (eds). 2006. *Complementation: A Cross-linguistic Typology*. Oxford: OUP.

Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and Empirical Foundations*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.

Faust, Norma, and Eugene Loos. 2002. Gramática del idioma yaminahua. Lima: Instituto Lingüístico de Verano.

Fleck, David. 2003. A grammar of Matses. PhD dissertation, Rice University.

Fleck, David. 2013. Panoan languages and linguistics. *Anthropological Papers of the American Museum of Natural History* 99.

INEI 2007: http://censos.inei.gob.pe/cpv2007/tabulados/

Loos, Eugene. 1999. Pano. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald, 227–250.

Shell, Olive. 1965. Pano reconstruction. PhD dissertation, University of Pennsylvania.
Shell, Olive. 1975. *Estudios Pano, III: Las lenguas Pano y su reconstrucción* [Serie Lingüística Peruana, No. 12]. Yarinacocha: ILV.
Shibatani, Masayoshi. What is nominalization? Towards the theoretical foundations of nominalization. In *Nominalization in Languages of the Americas* [Typological Studies in Language 124], Roberto Zariquiey, Masayoshi Shibatani & David W. Fleck (eds). Amsterdam: John Benjamins. (this volume)
Shibatani, Masayoshi. 2009. Elements of complex structures, where recursion isn't. In *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution* [Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 163–198. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.07ele
Valenzuela, Pilar. 2013. El sistema de concordancia del participante en las lenguas Pano y sus implicancias para el conocimiento de la Proto-Lengua. *Revista Brasileria de Lingüística Antropológica* 5(1): 117–157.
Valenzuela, Pilar. 2003. Transitivity in Shipibo-Konibo grammar. PhD dissertation, University of Oregon.
Valle, Daniel. 2014. Focus marking in Kakataibo. *Oklahoma Working Papers in Indigenous Languages* 1: 55–78.
Valle, Daniel. 2015. Modality in Kakataibo. *Interdisciplinary Studies on Information Structure* 19:111–137.
Valle, Daniel. 2017. Grammar and information structure of San Alejando Kakataibo. PhD dissertation, University of Texas at Austin.
Winstrand, Lila. 1968. Kashibo relative clause construction. MA thesis, University of Texas at Austin.
Zariquiey, Roberto. 2011a. A grammar of Kashibo-Kakataibo. PhD dissertation, La Trobe University.
Zariquiey, Roberto. 2011b. Aproximación dialectológica a la lengua cashibo-cacataibo. *Lexis* XXXV(1): 5–46.
Zariquiey, Roberto & Fleck, David. 2012. Prefixation in Kashibo-Kakataibo: Synchronic or diachronic derivation? *International Journal of American Linguistics* 78: 385–409.
Zariquiey Biondi, Roberto. 2013. Tessmann's <Nokamán>: a linguistic investigation of a mysterious Panoan group. *Cadernos de Etnolingüística* 5. 1-48.

CHAPTER 15

Nominalization and switch-reference in Iskonawa (Panoan, Peru)

Roberto Zariquiey
Pontificia Universidad Católica del Perú

The present chapter presents a characterization of grammatical nominalizations in Iskonawa, an obsolescing language in the Panoan family. It is shown here that grammatical nominalizations are highly isomorphic with independent (non-nominalized) clauses in terms of their internal syntax, but that the external syntax of nominalizations is highly nominal. A characterization of their relativization function is also offered, showing that the type of relativizing construction in which a grammatical nominalization may appear is highly dependent upon the participant with which it is coreferential. Finally, a discussion of switch-reference and its interactions with nominalization is presented, arguing that the situation found in Iskonawa, which has a relatively small inventory of switch-reference makers, is crucial for understanding the development of the category in the family.

1. Introduction

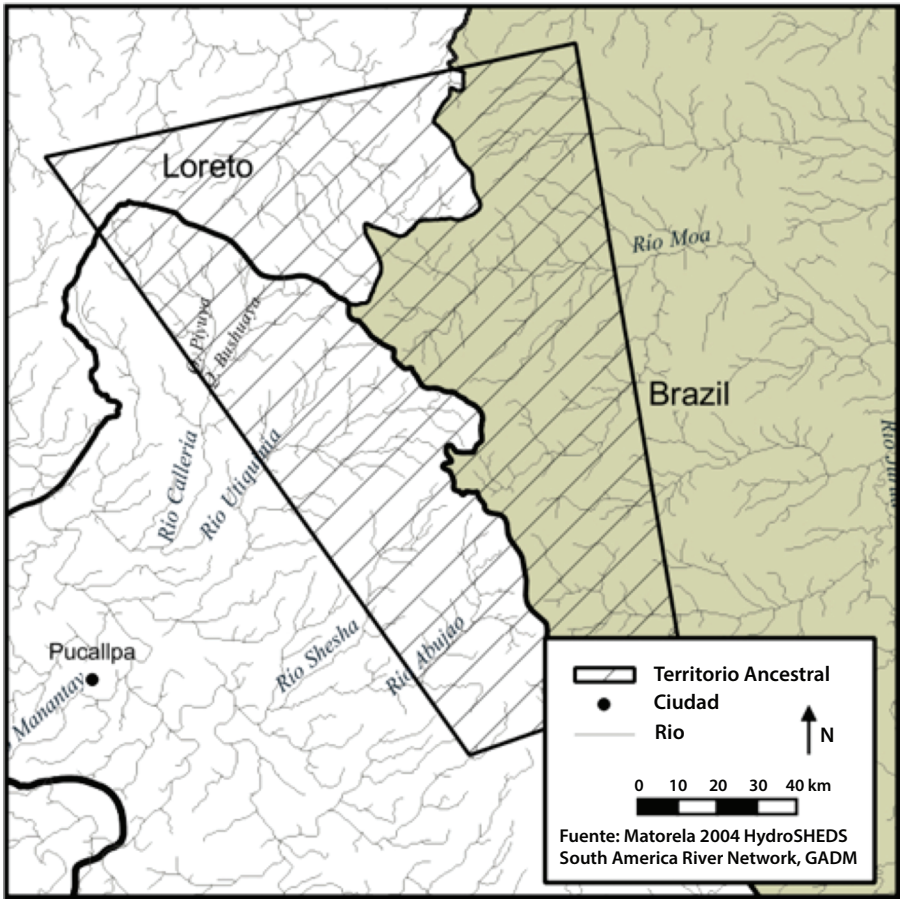
The present chapter offers a characterization of grammatical nominalizations in Iskonawa, an obsolete language in the Panoan language, elaborating on what has been previously presented in Zariquiey (2015) and paying particular attention to some of their most salient properties and functions. A detailed discussion of switch-reference and its interactions with nominalization is also presented in this chapter. This discussion will show that the situation found in Iskonawa, a language in which switch-reference and grammatical nominalization largely overlap, may be crucial for understanding the historical development of the complex switch-reference systems usually found in Panoan languages.

This chapter has been organized as follows: in § 2, I present some basic information about the Iskonawa language and its speakers; in § 3, I introduce grammatical nominalizations in Iskonawa. In § 4, I discuss in detail the interaction between switch-reference and nominalization, offering synchronic (§ 4.1) and diachronic (§ 4.2) considerations. Some conclusions and topics for further research are listed in § 5.

https://doi.org/10.1075/tsl.124.15zar

2. The Iskonawa language and its speakers

It is believed that the Iskonawa people traditionally lived as a small band of nomads in a large geographic area, which covers the headings of the Callería, Utiquinia, Shesha and Abujao rivers), as indicated in Map 1.



Map 1. Traditional geographic area of the Iskonawa people

The iskonawa people were contacted in 1959 by American missionaries, and since then they have been living in intense cultural and linguistic contact with speakers of Shipibo-Konibo (a fairly closely-related Panoan language). Nowadays, there are only five people who still speak the Iskonawa language fluently, all of them also speak Shipibo-Konibo (three know some Spanish as well). In fact, the last speakers of Iskonawa had not spoken in their language on a regular basis for a long time before they were contacted by the author of this chapter in the context of a

research project entitled *Documentation and Revitalization of Iskonawa: an interdisciplinary project.*[1] This project has produced a partially available text database with approximately 30 hours of recordings (7 of which have been transcribed and translated using the ELAN software and parsed using the Toolbox software), a vocabulary and grammatical sketch of the language (Zariquiey 2015 and 2017). There was almost no descriptive work on Iskonawa, prior to the outcomes of the "Iskonawa" project.

The classification of Iskonawa within the Panoan family is open to debate: D'Ans (1973) places his "Isconahuano" grouping within his "Pano de las cabeceras" subgroup, which also includes Panoan languages from the Yuruá and Purús rivers. For Loos (1999), Iskonawa belongs to the "Capanahua" subgroup, which also includes Shipibo-Konibo and Huariapano, among other languages. Finally, Fleck (2013) considers that Iskonawa belongs to the "Poyanawa" subgroup (and therefore is highly related to the Poyanawa language, spoken in Brazil), although he considers it as an intermediate between the members of this subgroup and languages from Purús, Yuruá and Ucayali Rivers.

Tables 1 and 2 illustrate the Iskonawa phoneme inventory (see Zariquiey 2015 for a detailed description). The language exhibits some grammatical characteristics that may be considered unusual from a Panoan perspective. These include (i) large inventories of ambitransitive verbs and polyfunctional words (i.e. words that follow the distributional patterns that correspond to nouns and verbs at the same time); (ii) an unstable and highly optional case system, and (iii) a morphologically simple switch-reference paradigm, mainly based on nominalizations (see Section 4).

Table 1. Iskonawa consonants

	Bilabial	**Alveolar**	**Palatal**	**Velar**	**Glottal**
Stops	p	t		k	
Nasals	m	n			
Flaps		ɾ <r>			
Affricates		t͡s	t͡ʃ <ch>		
Fricatives		s	ʃ <sh>		h
Approximants	β <b > w		j < y>		

1. This project was funded by the National Science Fundation (NSF) and allocated to two universities: Pontificia Universidad Católica del Perú and Tufts University.

Table 2. Iskonawa vowels

	Front	Central	Back
High	i	ɨ <e>	ụ <o>
Low		a	

3. Nominalization in Iskonawa

Definitions of nominalization vary in the literature, but generally the different definitions available point toward the idea of deriving expressions that can yield nominal uses from verbs, adjectives or larger constituents. In that line, Comrie and Thompson (1985: 334) state that "[t]he term nominalization means in essence turning something into a noun", but Payne (1997: 223) uses the label nominalization for "operations that allow a verb to function as a noun." Malchukov (2004) postulates that, in fact, the term "nominalization" conflates two different (and to some extent independent) processes: deverbalization and substantivization. These two processes may operate to different degrees in specific cases of nominalization. Shibatani (this volume) offers a summary and highlights some problems in association with the different approaches to nominalization found in the literature. Basically, these different definitions of nominalization exhibit one or more of the following characteristics: (i) they pay exclusive attention to the output or the input of the process (but not to both); (ii) they are only built on morphosyntactic and not semantic properties (although the former may vary radically among languages); and (iii) sometimes they do not make a clear distinction between the structure and its uses. In this paper, I follow Shibatani's definition of nominalization, understood as a twofold phenomenon that can be seen both as a process or as a product. As a process, nominalization relates to structures "denoting substantive or entity concepts that are metonymically evoked by the nominalization structures themselves" (Shibatani this volume). These entity-concepts may include events, facts, and propositions as well as event participants and entities conventionally associated with specific events (instruments, locations, etc.). In turn, as products, "nominalizations are like nouns […] by virtue of their association with an entity-concept denotation, a property that provides a basis for the referential function of a noun phrase headed by such nominalizations" (Shibatani this volume). Thus, nominalization is a metonymic process that produces constituents which exhibit entity-concept denotations and, due to this, can usually accomplish NP-like functions. The semantics of nominalizations (i.e. their entity-concept denotations) allow to classify them as event nominalizations (when denoting the event as a whole) and participant nominalizations (when denoting more specifically one of their participants).

The process of nominalization can apply to single lexemes or phrases, or to whole clauses. Following Shibatani (2009; this volume), I use the terms **lexical nominalization** and **grammatical nominalization** to refer, respectively, to these two situations. Lexical and grammatical nominalization do not only differ regarding their formal properties, but also in relation to their semantics. According to Shibatani (this volume), many lexical nominalizations tend to have more uniform denotations, whereas grammatical nominalizations do not exhibit fixed denotations and are deeply dependent upon speech context and pragmatics. Although there may be problematic cases (i.e., grammatical nominalizations composed of a single verb without any overt participants, or lexical/lexicalized nominalizations that involve two or more words, as the ones discussed for Matses by Fleck, this volume), the distinction between lexical and grammatical nominalizations is often straightforward and it is useful to describe nominalized expressions in Iskonawa.

This chapter focuses on grammatical nominalizations, but some mentions to lexical nominalization will also be offered in the following discussion. As we will see, grammatical nominalizations in Iskonawa are radically similar to non-nominalized clauses in terms of their internal structure and morphology (what Shibatani, this volume, calls **internal syntax**). Therefore, in most cases there is no overt or obvious morphological derivation of any sort and grammatical nominalizations are often isomorphic with finite clauses. What distinguish these two clause types is, primarily, their semantics, in the sense that grammatical nominalizations exhibit the metonymic entity-concept denotations previously mentioned and exhibit nominal properties, which make them non-clausal in terms of some crucial behavioural properties. These properties correspond to what is usually found in other languages (see Malchukov 2004): grammatical nominalizations in Iskonawa can be used as **heads of NPs** in different syntactic positions and can be marked for **case** and **number** (properties that correspond to the **external syntax** of the nominalized construction in Shibatani's terms). Additionally, grammatical nominalizations in the Iskonawa language can accomplish relativizing (Comrie & Thompson 1985) and complementation (Noonan 1985) functions and the data show that they also participate in switch-reference. The interaction between nominalization and switch-reference is crucial for understanding the development of the complex switch-reference systems typically found in Panoan languages (see Section 3, and Valle and Zariquiey, this volume).

3.1 Internal syntax and extreme isomorphism with independent clauses

Grammatical nominalizations in Iskonawa (which can be both participant and event nominalizations, see below) exhibit an extreme isomorphism in terms of their internal morphosyntax with (non-nominalized) independent clauses.

Non-nominalized clauses and grammatical nominalizations exhibit basically the same possibilities regarding the overt expression of arguments and obliques and show identical case-marking frames (although only grammatical nominalizations are obligatory verb-final). This isomorphism is also appreciated in the morphology of the verb: predicates in both nominalized and non-nominalized constructions carry basically the same morphological endings. There are no overt indicators of nominalization and, as summarized below in Table 3, basically the same list of bound morphemes is found at the end of verbs in independent clauses and grammatical nominalizations. Note that, as far I as know, all the markers in Table 3 may function as both participant or event nominalizers.

Table 3. Some verbal final markers in Iskonawa

Form	Gloss	Independent clause	Grammatical nominalization
-a	'perfective'	YES	YES
-i	'imperfective'	YES	YES
-shina	'past, a few days'	YES	YES
-ni, -ti	'remote past'	YES	YES

Although this paper focuses on grammatical nominalization, it may be relevant to mention that in our corpus, there are some dedicated lexical nominalizers, which never produce grammatical nominalizations. These include *-rasi* and *-(k) ewan*, which are nominalizers equally used to express irrepressible tendencies, like in *wini-rasi* 'irrepressible crier (*wini* 'cry')' and *pi-ewan* 'irrepressible eater (*pi* 'eat')'. Note that the markers *-a* and *-i* from Table 3 are also recruited for lexical nominalizations.

In what follows, I illustrate the distribution of the marker *-a* in Table 3 in independent clauses and grammatical nominalizations. All the examples in (1)–(4) come from the same narrative and feature the same verb *ewe* 'life'. The marker *-a* is glossed as 'PERF' in (1) and (4), and as 'PERF.NOMLZ' in (2) and (3). The example in (1) is an independent clause whose only verb is *ewea*, which is therefore interpreted as a finite predicate. This becomes clear from the free translation, worked in collaboration with speakers and semi-speakers of Iskonawa. The examples in (2) and (3) are different from the example in (1) since in (2) and (3) we find a copula verb, which is the finite predicate of the clause. The Iskonawa form *eah ewea* in (2) and (3) is a grammatical nominalization: it has an entity-concept denotation, based on a metonymic principle: it refers to a particular place, which is defined by the event 'I live'. This semantic configuration supports its use in a nominal function: in both cases, we find that *eah ewea* 'where I live' is an argument of the copula

riki.[2] Note that the example in (4) is structurally ambiguous since *ani ewea* may also be analized as a grammatical nominalization. The interpretation given here corresponds to the one preferred by the speakers. If this interpretation is accurate the example in (4) shows the property that more clearly distinguishes between grammatical nominalizations and independent clauses: while the former are always verb-final (see (2) and (3)), independent clauses are freer in terms of their constituent order and may exhibit post-verbal constituents (4).[3]

(1) Eah ani ewea.
eah ani ewe-a
1SG.S there live-PERF
'I lived there.' (NC-my.life.2013.008)

(2) [Eah ewea] Calleria riki.
eah ewe-a Callería riki
1SG.S live-PERF.NMLZ Calleria COP
'Where I lived is Calleria.' (NC-my.life.2013.012)

(3) Awen kahen riki [eah ewea].
awen kahe-n riki eah ewe-a
3.GEN house-LOC COP 1SG.S live-PERF.NMLZ
'In her house is where I lived.' (NC-my.life.2013.017)

(4) Ani ewea eah.
ani ewe-a eah
there live-PERF 1SG.S
'I lived there'. (but also 'I (am) the one who lives there')
(NC-my.life.2013.008)

3.2 External syntax of grammatical nominalizations

It has been argued here that in (2) and (3) the nominalized form *eah ewea* functions as one of the constituents of the copula *riki*. However, it is possible to present examples in which grammatical nominalizations like the ones in (2) and (3)

2. Note that, as highlighted by one external reviewer, the copula functions as an auxiliary combined with non-finite verbal forms in periphrastic constructions. The examples in (2) and (3), however, are clearly not instances of these periphrastic constructions. Verbal periphrasis in Iskonawa always exhibit the auxiliary immediately following the non-finite predicate. Furthermore, the form of the auxiliary is always *iki* and not *riki.* Thus, the examples in (2) and (3) are instances of grammatical nominalizations functioning as copula arguments. This interpretation is also supported by the translations given by speakers in fieldwork sessions.

3. Nominalizations appear in brackets in all the examples presented in this paper.

appear in other nominal functions. This evidence has to do with the fact that, as other nominal constituents, grammatical nominalizations can carry case and number markers and can be modified, for instance, by numerals or adjectives. All this is illustrated in the following examples.

In the example in (5), the grammatical nominalization *Germanan hawe bea* 'the tortoise that German brought' carries the ergative marker *-nan* and is the A argument of the transitive predicate *kiw* 'bite'. Notice that the whole grammatical nominalization *Germanan hawe bea* is coreferential with the P argument of the construction, in this case, the tortoise.

(5) Germanan hawe beaton Jeberson kiwa.
[German-nan hawe be-a]-ton Jeberson kiw-a
Germán-ERG tortoise.ABS bring-PERF.NMLZ-ERG Jeberson.ABS bite-PERF
'The tortoise that Germán brought bit Jeberson.' (NR-EE-2013)

In (6), the same grammatical nominalization featured in (5), *Germanan hawe bea* 'the tortoise that German brought', carries the comitative marker *-betan*.

(6) Germanan hawe beabetan ewen piro nirui.
[German-nan hawe be-a]-betan ewen piro niru-i
Germán-ERG tortoise.ABS bring-PERF.NOMLZ-COM 1SG.GEN dog walk-IMPF
'My dog walks with the tortoise that Germán brought.'

The examples in (7) and (8) illustrate how the plural marker *-bo* and numeral *wistewan* 'one' can be combined with grammatical nominalizations. The grammatical nominalization is again *Germanan hawe bea* 'the tortoise that German brought'.

(7) Germanan hawe beabo.
[German-nan hawe be-a]-bo
German-ERG tortoise.ABS bring-PERF.NMLZ-PLUR
'the tortoises that German brought'

(8) Germanan hawe bea wistewan.
[German-nan hawe be-a] wistewan
German-ERG tortoise.ABS bring-PERF.NMLZ one
'one tortoise that German brought'

The examples in (5)–(8) show that, regardless of the extreme similarity between independent clauses and the internal structure of grammatical nominalizations, nominalizations are highly nominal in terms of their external syntax. Among the external nominal properties of grammatical nominalizations in Iskonawa, we find the possibility of appearing in appositional constructions in which they accomplish a relativizing function. This is illustrated in Section 3.3.

3.3 Adpositions and the relativization function of grammatical nominalizations

Grammatical nominalizations may appear in combination with an external nominal constituent that constrains their interpretation. This type of construction, appositional in nature since we find two nominal constituents (a grammatical nominalization and a noun phrase) in a chain, reminds one of cases of prototypical ('externally-headed') relativization, as shown in the following examples. Note that the variable position of the two nominal constituents in relation to each other supports the appositional analysis (modifying nouns are always prehead in Iskonawa, see Valle and Zariquiey, this volume, for a similar situation in Kakataibo). The appositional analysis is in line with the metonymic approach to nominalization presented in Shibatani (this volume) and followed here. The grammatical nominalization *tapas maroa* 'who bought a house' in (9) and (10) has a entity-like detonation, which is metonymic in the sense that its referent is defined in terms of its being an argument of the predicate 'to buy a house'. The presence of the adjacent noun *oni* constrains the interpretation of the semantically too general grammatical nominalization *tapas maroa* 'who bought a house' (Shibatani, this volume). Note that this analysis does not require to postulate a gap, since *tapas maroa* 'who bought a house' is a complete expression in an appositional relation with *oni* 'man'.

(9) Tapas maroa oni ma kahakoa.
[tapas maro-a] [oni] ma ka-hako-a
house buy-PERF.NMLZ:ABS man:ABS already go-DIM-PERF
'The man who bought a house is just gone.' (GC-EE-2014)

(10) Oni tapas maroa ma kahakoa.
[oni] [tapas maro-a] ma ka-hako-a
man:ABS house buy-PERF.NMLZ:ABS already go-DIM-PERF
'The man who bought a house is just gone.' (GC-EE-2014)

This kind of construction, in which a grammatical nominalization is coreferential with an external nominal constituent, occurs infrequently and only in certain specific conditions (see below). It is more natural for grammatical nominalizations in Iskonawa to exhibit coreferentiality relations with an overtly expressed internal argument as it is the case in the examples in (5)–(8) (the nominalization *Germanan hawe bea* 'the tortoise that German brought' refers to one of its internal arguments: *hawe* 'tortoise'), without any external nominal constituent. If we were to analyze nominalizations in (5)–(8) as relative clauses, they would be 'internally headed' and this type of relative clause is less prototypical than externally headed ones (Keenan 1985), which crucially are infrequent in Iskonawa. Furthermore, it is also very common to find examples like the ones in (2) and (3) in natural texts,

in which the referent of the grammatical nomination is not expressed either internally or externally, in what would correspond to a headless relative clause. Again, if we were to analyze those as so, we would need to conclude that non-prototypical relative clauses are more common than prototypical ones in Iskonawa texts. Thus, if we analyze grammatical nominalizations as relative clauses in Iskonawa, we would need to conclude that relative clauses in Iskonawa are highly unusual from a cross-linguistic perspective. A more straightforward approach to the Iskonawa data would be to analyze the examples discussed so far as entailing grammatical nominalizations, some of which may accomplish a relativization function. This perspective, which is in line with Shibatani's (2009, this volume) approach to nominalization, is the one assumed in this chapter.

Thus, the fact that grammatical nominalizations in Iskonawa, as in many other languages, accomplish a relativization function, which equals the function of relative clauses in the literature, does not mean that we should analyze these nominalizations as relative clauses. The main problem with such an analysis is that we would have to first look at these constructions as if they were relative clauses, and then assume that they are non-prototypical instances of this type of clause, since they only rarely show an overt external head. By contrast, if we assume that all these examples are just grammatical nominalizations in the sense proposed by Shibatani (2009), we are able to avoid the unnecessary complexity of the relative-clause analysis. Shibatani explains this in the following way:

> Grammatical nominalizations, especially those that show a clausal character, have often been considered a type of relativisation and are called "headless relatives" or "free relatives" as if they were derivatives of relative clauses. There is no basis for this other than the fact that they show formal resemblances to relative clauses […] and the skewed perspective many linguists have had about grammatical nominalizations, namely viewing them from the perspective of relative clauses.
>
> (Shibatani 2009: 187)

Shibatani's approach is very useful for understanding and analyzing the Iskonawa data, which suggest that we need to look at relativizing constructions from the perspective of nominalizations, and not at nominalizations from the perspective of relative clauses (see Zariquiey 2011: chapter 20, for a similar analysis regarding another Panoan language, Kakataibo; see also Valle & Zariquiey, this volume).

The prototypical relative-clause-like constructions are less common than non-prototypical ones in discourse because they are not relative clauses (they are grammatical nominalizations) and because external nominal constituents are obligatory only under some specific conditions. One crucial point regarding Iskonawa grammatical nominalizations in this relativizing function is that the distribution of "externally-headed" and "internally-headed" relativizing constructions, which

is at the core of the typology of relative clauses, is predictable based on the function of the coreferential argument in the nominalized construction. For those cases in which the coreferential argument is the A, S or P of the nominalized construction, both "externally-headed" and "internally-headed" constructions are acceptable (but the latter are strongly preferred in both elicitation and natural texts). In turn only "externally-headed" constructions are possible for those cases in which the coreferential argument of the grammatical nominalization is an oblique participant within the nominalized construction.

The following examples illustrate this. In (11), we find an elicited example, in which the grammatical nominalization *onin hawe bia* has two different possible interpretations, one based on coreferentiality with A: 'the man who captured the tortoise' and the other based on coreferentiality with P: 'the tortoise that the man captured'. Conversely, in (12) and (13), the coreferential participant is a locative adjunct from the perspective of the internal syntax of the grammatical nominalization and the preferred construction is the one that corresponds to an externally headed relativizing construction, that is, (12). In fact, (13) was considered as impossible by Iskonawa speakers and semi-speakers.

(11) [Onin hawe bia] ma kahakoa.
oni-n hawe bi-a ma ka-hako-a
man-ERG tortoise:ABS capture-PERF.NMLZ:ABS already go-DIM-PERF
'The man who captured the turtle is just gone.'
'The turtle that the man captured has just gone.' (GC-EE-2014)

(12) [German ohashina] tapas ma pohoa.
[German uha-shina] tapas ma poho-a
Germán:ABS sleep-PAST.NMLZ cottage already fall.down-PERF
'The cottage where Germán slept a few days ago fell down.' (NC-EE-2014)

(13) German tapasmena ohashina ma pohoa.
**[Germán tapas-mena uha-shina] ma poho-a*
Germán:ABS cottage-LOC sleep-PAST.NMLZ already fall.down-PERF
'The cottage where German slept a few days ago fell down.' (NC-EE-2014)

Thus, the different constructions illustrated here, which are reminiscent of internally-headed and externally-headed relative clauses, are just nominalizations with different internal properties, which directly correspond to their referential properties. If the metonymic properties of the grammatical nominalization point towards a non-core participant, it is simply not possible to include it as an overt constituent in the internal configuration of the grammatical nominalization. This situation triggers the inclusion of an external nominal constituent and this produces the externally-headed relative clause effect found in examples like (12). It would be

interesting to determine which are the functional motivations for this distribution. Notice that if the coreferential argument is a core argument of the grammatical nominalization, then, both internal and external codifications are available, but it is also possible to not express the coreferential participant overtly either inside or outside the nominalization.

4. Nominalization and switch-reference

4.1 The switch-reference function of grammatical nominalizations

Cross-linguistically, switch-reference constitutes a device for keeping track of core arguments between the clauses of complex sentences by means of verbal affixes that appear on dependent verbs in order to indicate whether the arguments (canonically, the subjects) of the two related clauses are coreferential or not. Panoan languages are well-known among South American languages for their complex switch-reference systems. In Panoan languages, most switch-reference markers distinguishes between S and A in the main clause, and indicate the time of the subordinate verb relative to the matrix predicate, which may or may not be the main predicate of the sentence. The encoding of all this information may produce highly sophisticated systems with large inventories of specialized markers, with both canonical (subject-based) and non-canonical (object to subject, subject to object, different objects and so on) switch-reference meanings. In Kakataibo, for instance, we find a switch-reference system composed of 21 suffixes (see Zariquiey 2011: chapter 21; Valle & Zariquiey, this volume).

Iskonawa differs radically from a Panoan language like Kakataibo, as I have been able to find only eight dedicated switch-reference markers, all of which are associated with the same-subject category. Other switch-reference categories (P > S/A, S/A > P, P > P, different subjects, different objects), which would have dedicated morphology in other Panoan languages, do not exhibit specialized morphology in Iskonawa and are simply encoded by means of grammatical nominalizations, similar to the ones illustrated previously in this chapter. This, as also discussed in Valle and Zariquiey (this volume) is crucial for the understanding of the development of switch-reference markers in Panoan.

In what follows, I present and discuss some examples of switch-reference constructions in Iskonawa, paying particular attention to the participation of grammatical nominalizations in their configuration. Before discussing those complex examples, let us start with examples of same-subject switch-reference, like the ones presented in (14) and (15). In (14), the main predicate is lexically transitive (*noi* 'love') and, therefore, the switch-reference marker for same subjects has to be

-hon 'S/A > A', which, in the example, gets a simultaneous interpretation. In turn, in (15) the matrix predicate is intransitive and the form of the switch-reference marker on the dependent verb is *-ah* 'S/A > S', which in this case gets a sequential interpretation (the dependent event is previous to the matrix event).

(14) Iso inahon isores noia.
iso *ina-hon* *iso-res* *noi-a*
spider.monkey raise-S/A > A spider.monkey-only love-PERF
'When (she$_i$) raised a spider monkey, she$_i$ truly loved it.' (JC-chachibai-2013)

(15) Piah niho icha poia.
pi-ah *niho* *icha* *poi-i*
eat-S/A > S forest.devil a.lot shit-IMPF
'After it$_i$ eats, the niho$_i$ (a type of forest demon) defecates a lot.'
(IC-niho-2013: 017)

The switch-reference suffixes in Examples (14) and (15) have correspondences in all the Panoan languages whose grammars have been thoroughly described: *-hon* and *-ah* are cognate with *-ʂon* and *-aʂ* in other Panoan languages like Shipibo-Konibo and Kakataibo, for instance, and in those languages these markers exhibit an equivalent function and distribution. The same can be said about the other same-subject switch-reference markers in Table 4: all of them find correspondences at least in some Panoan languages.

Table 4. Switch-reference markers in Iskonawa

Temporal relation	**Same subjects**		**Other categories**
(the dependent event is…)	**(transitive main predicate)**	**(intransitive main predicate)**	**(O > S/A, S/A > O, O > O, different subjects, different objects)**
Previous	*-hon*	*-ah*	*Grammatical*
Simultaneous	*-hon*	*-ah*	nominalizations
	-kin	*-i*	
Simultaneous durative	*-anan*		
Posterior	*-no*		

However, the situation varies when we look at other switch-reference categories, which, as I have already mentioned, are simply expressed by means of grammatical nominalizations in Iskonawa. Basically, Iskonawa uses the nominalizers *-a* and *-i*, which are also the perfective and imperfective markers of the language, for various switch-reference categories, which include different subjects, different objects, O > S/A, S/A > O and O > O, among other possibilities. Many of

the switch-reference constructions based on grammatical nominalizations are ambiguous and are open to different interpretations. This is true regarding the examples in (16) and (17), for instance. In (16) the grammatical nominalization *enan German mipia* may be interpreted as a switch-reference construction with the meaning 'after I hit German' and any switch-reference relation to the main predicate, excluding same subjects, is potentially possible: it may be the case that *German* (the P argument of the grammatical nominalization) or someone else (but never the subject) cried. A similar situation is found regarding the example in (17). The example in (16) can also be interpreted as including a grammatical nominalization in a relativization function and therefore it may also be translated as 'Germán, whom I hit, cried' or 'I, who hit Germán, cried' (but the latter interpretation was considered odd). Note that the relativization reading of (17) is not possible since the main verb is transitive and therefore we will need an ergative case marking (see (18)).

(16) Enan German mipia hohoa.
enan German mipi-a hoho-a
1SG.A Germán hit-NMLZ cry-PERF
'After I hit Germán, he/some one else cried.'
'Germán, whom I hit, cried.'
(?) 'I, who hit Germán, cried.' (NC-EE075–2014)

(17) Enan German mipia Maria kena.
enan German mipi-a Maria ken-a
1SG.A Germán hit-NMLZ Mary call-PERF
'After I hit Germán, he/someone else called Mary.' (NC-EE076–2014)

(18) Enan German mipiaton Maria kena.
enan German mipi-a-ton Maria ken-a
1SG.A Germán hit-NMLZ-ERG Mary call-PERF
'Germán, who I hit, called Mary.'
'I, who hit Germán, called Mary.' (NC-EE077–2014)

The examples presented in (16) and (17) reveal that, differently from what we know about other Panoan languages, Iskonawa does not have a rich switch-reference systems based on a large inventory of dedicated and highly specialized markers. On the contrary, although we find the same 'same subject' markers that we find in other Panoan languages, various switch-reference categories are expressed simply by means of grammatical nominalizations, which are open to ambiguity and are highly contextual. This is different from what we find in other Panoan languages, such as Kakataibo, in which different types of different subject switch-reference relations are expressed by means of highly specialized markers (see

Zariquiey 2011: Chs. 20; Valle & Zariquiey, this volume). However, one important point about this issue is that many of these different markers clearly came from nominalizations. Therefore, the situation in Iskonawa might be considered as representing a previous stage to the development of the particularly rich inventories of switch-reference markers found in languages like Kakataibo. In the next section, I discuss what is the relevance of the Iskonawa data for the understanding of the development of the more complex switch-reference systems found in other Panoan languages.

4.2 Nominalization and switch-reference in diachronic perspective

The Iskonawa data help us to understand two crucial points. The first one is that, as previously claimed by Valenzuela (2003: Ch. 20), the markers for 'same subjects' listed in Table 4 may be old. The evidence in favor of this is that Panoan languages with simple switch-reference systems, like Iskonawa, exhibit the same 'same subjects' markers found in those with richer systems: we find the same forms and the same typically Panoan distinction between transitive and intransitive. Basically, the Iskonawa data offer direct support to the idea that *-xon* and *-kin* 'S/A > A', and *-ax* and *-i* 'S/A > S' could be traced to an old stage of the development of the language family (either to Proto-Pano or at least to the linguistic ancestor of what Fleck 2013 calls the mainline branch of the family, which comprises all Panoan languages but the Northern or Mayoruna subgroup).

The second point is more interesting. In Iskonawa, 'different-subject' switch-reference categories are expressed by means of grammatical nominalizations. The nominalizers *-a* and *-i*, which are also the perfective and imperfective markers of the language, accomplish basically all the non-same-subjects' switch-reference categories that are expressed by means of dedicated markers in other languages, such as Kakataibo.[4] The important fact here is that, diachronically, it may be possible that at least some switch-reference markers for 'different subjects' found in the family also came from nominalized forms combined with a case marker. For instance, the Kakataibo marker *-këbëtan* 'simultaneous event, different subjects/objects, matrix transitive predicate' is likely to be the result of the combination of the general nominalizer *-kë* and the comitative = *bëtan*. In turn, the marker *-këxbi* 'simultaneous event, O > S' likely comes from the combination of the same general nominalizer *-kë*, the case marker = *x* 'S' and the adverbial enclitic = *bi* 'same, emphatic' (see also Valle and Zariquiey, this volume). It is still necessary to figure out more precise scenarios that trigger the combination of nominalizers and case

4. Note that the marker *-i* 'imperfective, imperfective nominalizer' is formally identical to the same-subject marker *-i* in Table 4.

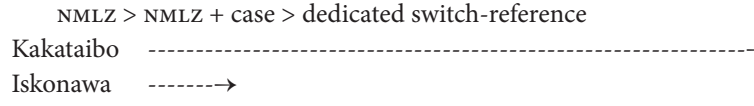
morphemes for encoding switch-reference, but specialization and disambiguation of otherwise ambiguous grammatical nominalizations may be among its potential functional motivations. Furthermore, this strategy is widespread among South American languages, as convincingly argued by Van Gijn (this volume).

Therefore, the pattern seems relatively clear: while 'same subjects' switch-reference markers seem old in Panoan and do not come from nominalizations (but from some other source, likely case markers; Valenzuela 2003: Ch. 20), 'different subjects' switch-reference markers are the result of a diachronic process whereby grammatical nominalizations like the ones found in Iskonawa became fused with case markers and other forms resulting in new markers with specialized and non-ambiguous 'different subject' meanings.

Following this argumentation, the Iskonawa situation may constitute a previous stage in relation to the complex switch-reference systems found in other Panoan languages. An old stage of Proto-Pano used grammatical nominalizations for different-subject switch-reference functions, but, at some point, these grammatical nominalizations became fused with other markers in order to produced the specialized forms that we find in other Panoan languages.

There is, however, one potential problem with this line of argumentation. Taking into consideration that Iskonawa is an obsolencent language, the simplicity of its switch-reference system might also be a consequence of the simplification process that often accompanies language obsolescence (Palosaari & Campbell 2011). Thus, Iskonawa might have had a more complex switch-reference system, similar to the one found in Kakataibo and other Panoan languages, and this system might have become simpler as a consequence of its obsolescence situation, producing a cycle like the one presented in (19), rather than the different stages of the process described in (20) for Iskonawa and Kakataibo.

(19) NMLZ > NMLZ + case > dedicated switch-reference > NMLZ

(20)


Thus, the diagram in (19) proposes a situation in which Iskonawa had at some point a rich 'Panoan-like' switch-reference system, which got lost as a consequence of language obsolescence. In turn, the diagram in (20) assumes that Iskonawa did not have this type of switch-reference system and that what is currently found corresponds to how the language was before the process of language obsolescence. Without previous records of the language it is not possible to decide between these two possible scenarios. In 1971, however, Eugene Loos recorded an Iskonawa word list, some sentences in paradigms and a transcribed and translated

text with 35 lines (see Zariquiey 2017). The important fact about this text is that it was recorded only 12 years after the Iskonawa people were contacted by the missionaries for the first time and, therefore, it is relatively unlike that at that time the language was already undergoing obsolescence. Interestingly, this text includes at least 8 instances of 'same subjects' switch-reference markers (as in Example (21)), but there is no evidence of dedicated 'non-same-subjects' markers at all in the text. There are, however, 4 clear cases in which a grammatical nominalization was used as a switch-reference device for different subjects. One example is offered in (22) (Loos' orthography has been adapted in both examples).

(21) Atsa bebahon pimakin.
[atsa beba-hon] pi-ma-kin
manioc cook-S/A > A eat-CAUS-REM.PAS.HAB
'Cooking manioc, (s)he used to feed (someone else).' (Loos, 11)

(22) Toa aka awen bene hoke.
[toa ak-a] awen bene ho-ke
son do-NMLZ 3SG.GEN husband arrive-PERF
'After (the tapir) made a son (be borne by her), her husband arrived.'
(Loos, 8)

The scarce evidence offered by Loos' 1971 text seems to point towards the hypothesis that what we currently find in Iskonawa in terms of its switch-reference system is not the result of a simplification process associated with its current obsolescence, but constitutes a old stage in the language. Although such an interpretation is tempting, we cannot know if the Iskonawa language was undergoing obsolescence before their speakers were contacted and started to live with the Shipibo-Konibo, or if, as the Iskonawa oral history suggests, their pre-contact society was the result of the co-existence of the remnants of multiple Panoan groups and, as a consequence, their language underwent some degree of pidginization. Therefore, Loos' data only indirectly supports the interpretation in (20). In any case, either innovative or conservative, the Iskonawa data constitutes a synchronic example of a pattern that is only a presupposed stage in the attempts to reconstruct the complex Panoan switch-reference system (Valle and Zariquiey; this volume).

The general pattern proposed in (20) is based on the fact that, being nominal elements, grammatical nominalizations can take case markers (see also Van Gijn, this volume). In modern Iskonawa, for instance, it is possible to find the case marker = *tian* 'temporal' in examples like the one in (17). This is illustrated in (23). The marker, whose use may be a recent loan from Shipibo-Konibo, is not obligatory and (17) and (23) seem to be synonymous.

(23) Enan German mipitian awen ewa kena.
[enan German mipi-a]=tian Maria ken-a
1SG.A Germán.ABS hit-PERF.NMLZ=TEMP Mary call-PERF
'After I hit Germán, he/someone else called Mary.' (NC-EE078–2014)

One final interesting phenomenon is that bare grammatical nominalizations may accomplish similar switch-reference functions in other Panoan languages. This is illustrated with the Kakataibo example in (24), where the grammatical nominalization *xu ʻikë* 'being a baby' can only be interpreted as having the P argument of the matrix clause as its subject. Examples like this one suggest that the development of the complex paradigms of 'different subjects' switch-reference markers might have started with the specialization of some bare grammatical nominalizations.

(24) Xu ʻikë kana ʻën Maria ʻunankën.
[xu ʻi-kë] kana ʻë=n Maria ʻunan-akë-n
small be-NMLZ NAR.1SG 1SG=A Maria.ABS know-REM.PAST-1/2p
'I met Maria when she (P) was a baby.'
* 'I met Maria when I (A) was a baby.'

5. Conclusions

The present paper has offered a first characterization of grammatical nominalizations in Iskonawa, an obsolescent Panoan language, paying particular attention to some of their most salient properties and functions. It has first been shown that grammatical nominalizations are highly isomorphic with independent (non-nominalized) clauses in terms of their internal morphosyntax, but it has also been shown here that the external syntax of nominalizations is highly nominal. A characterization of their relativization function has also been discussed, showing that both 'internally-headed' and 'externally-headed' relativizing constructions are found in the data and that their distribution is simply based on the semantic properties of grammatical nominalizations (basically, the type of argument they are coreferential with). Finally, a detailed discussion of switch-reference and its relations with nominalization has also been presented in this paper. While Iskonawa exhibits a set of dedicated 'same subjects' switch-reference markers, which are basically the same as the ones found in other Panoan languages, there are no 'different subjects' markers in the language and this function is accomplished by grammatical nominalizations. The interesting fact about this is that the specialized 'different subjects' markers found in other Panoan languages diachronically come from grammatical nominalizations combined with case markers and, therefore, the situation found in Iskonawa is likely to be an antecedent of them: basically

these markers might have developed from the use of grammatical nominalizations as highly ambiguous and poorly specialized switch reference constructions for non-same-subjects, exactly as we currently find in Iskonawa.

References

d'Ans, Marcel. 1973. Reclasificación de las lenguas pano y datos glotocronológicos para la etnohistoria de la Amazonía peruana. *Revista del Museo Nacional* § XXXIX: 349–369.

Comrie, Bernard & Thompson, Sandra A. 1985. Lexical nominalization. In *Language Typology and Syntactic Description, Vol. 3: Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 349–398. Cambridge: CUP.

Fleck, David W. 2013. Pano languages and linguistics. *Anthropological Papers of the American Museum of Natural History* 99.

Keenan, Edward. 1985. Relative clauses. In *Language Typology and Syntactic Description, Vol. 2: Complex Constructions*, Timothy Shopen (ed.), 141–170. Cambridge: CUP.

Loos, Eugene. 1999. Pano. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 227–250. Cambridge: CUP.

Malchukov, Andrei L. 2004. *Nominalization/verbalization. Constraining a Typology of Transcategorial Operations*. Munich: Lincom.

Noonan, Michael. 1985. Complementation. In *Language Typology and Syntactic Description, Vol. 2: Complex Constructions*, Timothy Shopen (ed.), 42–140. Cambridge: CUP.

Palosaari, Naomi & Campbell, Lyle. 2011. Structural aspects of language endangerment. In *The Cambridge Handbook of Endangered Languages*, Peter Austin & Julia Sallabank (eds), 100–119. Cambridge: CUP. https://doi.org/10.1017/CBO9780511975981.006

Payne, Thomas. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: CUP.

Shibatani, Masayoshi. 2009. Elements of complex structures, where recursion isn't: The case of relativization. In *Syntactic Complexity, Diachrony, Acquisition, Neuro-cognition, Evolution*, T. Givón & Masayoshi Shibatani (eds), 163–198. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.07ele

Shibatani, Masayoshi. What is nominalization? Towards the theoretical foundations of nominalization. In *Nominalization in Languages of the Americas* [Typological Studies in Language 124], Roberto Zariquiey, Masayoshi Shibatani & David W. Fleck (eds). Amsterdam: John Benjamins. (this volume)

Valenzuela, Pilar M. 2003. Transitivity in Shipibo-Konibo grammar. PhD dissertation, University of Oregon.

Valle, Daniel & Zariquiey, Roberto. Clausal nominalization in Kakataibo. In *Nominalization in Languages of the Americas* [Typological Studies in Language 124], Roberto Zariquiey, Masayoshi Shibatani & David W. Fleck (eds). Amsterdam: John Benjamins. (this volume)

van Gijn, Rik. Case markers as subordinators in South American indigenous languages. In *Nominalization in Languages of the Americas* [Typological Studies in Language 124], Roberto Zariquiey, Masayoshi Shibatani & David W. Fleck (eds). Amsterdam: John Benjamins. (this volume).

Zariquiey, Roberto. 2011. A grammar of Kakataibo-Kashibo. PhD dissertation, LaTrobe University.

Zariquiey, Roberto. 2015. *Bosquejo grammatical de la lengua iskonawa*. Lima & Boston MA: RCLL, CELACP & Latinoamericana Editores.
Zariquiey, Roberto. 2017. *Vocabulario iskonawa-castellano*. Lima & Boston MA: RCLL, CELACP & Latinoamericana Editores.

CHAPTER 16

Lexicalized nominalized clauses in Matses (Panoan)

David W. Fleck
American Museum of Natural History

Nominalization is ubiquitous in the Matses language. Many functions that are performed by relative and adverbial clauses in other languages are accomplished by nominalization in Matses. Verbs are nominalized by attaching one of 27 different suffixes to the verb stem in place of inflectional morphology, creating either a word that falls into the noun lexical category, or a multi-word nominalizations which I call here a "nominalized clause." Matses nominalized clauses, rather than being noun phrases built up around a de-verbalized noun, have essentially main-clause syntax, including main-clause case-marking frame. The present paper describes the internal and external syntax of Matses nominalized clauses in comparison to simple lexicalized nominalized words. Of particular interest is that nominalized clauses can become lexicalized.

1. Introduction

nominalization *n.* (the result of) the process whereby a non-nominal element is changed into an element that functions morpho-syntactically like nouns do in the language in question.
lexicalization *n.* the *diachronic* process whereby a complex word or series of words can no longer be understood based solely on the meaning of its components, thereby becoming a unit of the language's lexicon.
clause *n.* a unit of grammatical organization headed by a verb, whose constituents are morphosyntactically related to the verb in the same manner as in independent sentences in the language in question.

Nominalization in Matses serves two distinguishable functions: (1) to generate new nominal lexical items from verbs; and (2) to produce dependent clauses that stand in for arguments and serve relativization and other grammatical functions. Inspection of the internal syntax of multi-word nominalizations reveals that they have main-clause case-making frame and syntax. Because multi-word

https://doi.org/10.1075/tsl.124.16fle

nominalizations in Matses cannot be analyzed as noun phrases built up around a de-verbalized noun (e.g., by associating the arguments with it as genitives), I designate them "nominalized clauses" (Fleck 2003, 2008). Nominalized clauses can be contrasted with "(de)verbal nouns," single-word nominalizations that lack internal syntax and are, superficially at least, equivalent to the type of derivational nominalization that is common in European languages. Nominalizations can also be divided into lexicalized nominalizations and nonce nominalizations, the latter being *ad hoc* formations composed for the communicative purposes of the speech act at hand.

In my initial analysis of Matses nominalization (Fleck 2003), I made the observation that lexicalized nominalizations are in every respect syntactically identical to underived nouns in the Matses language, while a few aspects of the external syntax of nominalized clauses distinguish them from true nouns. In light of this, I posited a primary division between lexicalized nominalizations and nominalized clauses for the purposes of predicting their behavior within a sentence. But alas, I had neglected to inspect multi-word lexicalized nominalizations that came to light during the compilation of a Matses-Spanish dictionary (Fleck et al. 2012). What was revealed upon analysis of these multi-word lexicalized terms is that they have internal syntax identical to that of multi-word nonce nominalized clauses. Consequently, the morpho-syntactically relevant distinction to be made is actually between **lexicalized nominalizations** and **nonce nominalizations**, with the division between nominalized clauses and verbal nouns being purely a formal (i.e., word-count) distinction. Hence, the expression "lexicalized nominalized clause" is not self-contradicting after all.

This essay, however, is not just about vindicating apparent oxy-morons, and lexicalized nominalized clauses are not merely an anomaly to be pointed out and then dismissed. Rather, they will turn out to be the key for unmasking the true nature of nominalization in Matses. Before describing multi-word lexicalized nominalization in Section 5, first I will provide an overview of the types, morphology, and functions of nominalizations in Matses (Section 2) and present the basic features of Matses grammar (Section 3) that will be essential for analyzing the internal and external syntax of nonce nominalized clauses (Section 4). In the final discussion (Section 6) I will contrast Matses and English nominalization and consider whether they differ at their very core.

Matses is spoken by about 2000 speakers in Peru and about 1000 in Brazil, in the Amazonian rainforests along the Javari River and its tributaries. The Matses made first peaceful contact with the non-indigenous society in 1969, and all Matses speak Matses as their first language, many being essentially monolingual. Matses belongs to the Mayoruna branch of the Panoan family (Fleck 2013), and while this branch is quite divergent from the rest of the family, the analysis proposed here

may be profitably applied to other Panoan languages. In addition to its syntactic and functional properties, lexicalized nominalization in Matses is of interest in that it is a repository of obsolete verb roots and traditional knowledge and a source of linguistic jokes.

2. Overview of nominalization in Matses

Nominalization was described at length in my grammar of the Matses language (Fleck 2003: 292–321, 1011–1047) and in a subsequent article (Fleck 2008), which the reader may consult for further details and illustration. In this section I will describe the parameters of nominalization in Matses, contrast the form and function of action and participant nominalizations, and present the inventory of nominalizing suffixes.

2.1 Parameters of nominalization in Matses

In Matses, only verbs can be nominalized. All verbs, except the copula *ne* 'be' (but including the copula *ik* 'be, exist, Auxiliary'), can be nominalized. The existence of a lexicalized nominalization does not preclude the composition of a formally identical nonce nominalization.

All nominalization is accomplished through suffixation of verbs, by attaching nominalizing suffixes in the stead of finite inflection. Derivational (causative, emphatic, directional, etc.) suffixes can intervene between the verb root and the nominalizing suffix. There are a handful of lexemes that occur as both verbs and nouns without any overt marking, such as *isun*[1] 'urinate/urine' and *mado* 'son/to become the father of a male child for the first time,' but all these pairs involve nouns with very specific and unpredictable meanings (in fact, one cannot always be certain if the verb or the noun was the original meaning), and all these verbs can be nominalized through suffixation to obtain nouns with other (predictable) meanings. This and the fact that all verbs (save *ne* 'be') can be nominalized indicate that nominalization is a completely productive process in Matses. This productivity will be important for the final discussion on the roles of nominalization in Matses.

1. Orthography: *a*, *e*, *ë* (ɨ), *i*, *o*, *u*, *p*, *t*, *k* (= glottal stop syllable-finally), *b* (sometimes pronounced with some friction), *d* (= flap intervocalically), *m*, *n* (= velar nasal before *k,* bilabial nasal before *p,b*), *s*, *sh* (ʃ), *ṣh* (ʂ), *ts*, *ch* (tʃ), *ç̣h* (tʂ), *w, j*. Basic forms, rather than conditioned allomorphs, are used in examples.

2.2 Action nominalizations vs. participant nominalizations

In addition to the lexicalized vs. nonce distinction, a second important division can be made among nominalizations. Action/activity/process/event/state nominalizations and participant/argument nominalizations are two classes of nominalizations in Matses that are not only semantically distinct, but can also be defined on morho-syntactic bases.

Action nominalizations can be nominalized clauses or one-word nominalizations, but in Matses they are never lexicalized. Matses lacks abstract lexemes like *freedom* and *involvement*. Action nominalizations have a very limited distribution in that they occur only in the following syntactic positions:

i. as subject of a copular clause with an adjective as a copula complement:

(1) a. *[Mimbi kun opa=Ø kuessunne-**ak**]*$_{NMLZ}$ *iksa=mbo ik-e-k.*
2ERG 1GEN dog=ABS kill-**ACT.NMZR** bad=AUG be-NPAST-INDIC
'It is bad that you killed my dog.'
more literally, 'You killing my dog is bad.'

b. *Kuessunne-**ak** bëda=mbo ik-e-k.*
kill-**ACT.NMZR** good=AUG be-NPAST-INDIC
'It is good that he/she/they/it killed him/her/them/it.'

ii. as object of comparative and locative postpositions:

(2) *[Titado=Ø dadpen umbi chokoka-**ondak**]*$_{NMLZ}$*=no*
peach.palm=ABS many 1ERG plant-DIST.PAST.EXP.ACT.**NMZR**=LOC
tambis=n pe-ak-o-ṣh.
paca=ERG eat-REC.PAST.INFER-REC.PAST.EXP-3
'Pacas (large rodents) had fed where I planted many peach palms long ago.'

iii. as object of the verb *dan* 'mistake/assume incorrectly' (a complementation strategy, following the analysis in Fleck 2006c):

(3) *Debi=n [opa=n ṣhëkten=Ø bed-**ak**]*$_{NMLZ}$*=Ø*
Davy=ERG dog=ERG collared.peccary=ABS grab-**ACT.NMZR** = ABS
dan-o-ṣh.
mistake-REC.PAST.EXP-3
'Davy mistakenly thought that the dogs had captured the collared peccary.'

By contrast, participant nominalizations can be lexicalized and can occur in essentially any syntactic position in which a noun or noun phrase can occur. Specifically, participant nominalizations serve the following roles in the language:

i. Creation of new lexemes from verbs:

(4) a. *beska* 'sweep' + *-te* 'INST.NMZR'=*beskate* 'broom'

ii. Nonce composition of referential words:

b. *beska-te*
sweep-INST.NMZR
'what is/will be used to sweep' (e.g., a branch)

(5) *beska-nëdaid*
sweep-DIST.PAST.INFER.NMZR
'one who evidently swept long ago'
'what was evidently swept long ago'
'what was evidently used to sweep long ago'
'place evidently swept long ago'

iii. By filling a core or peripheral participant slot, nominalized clauses perform the following among other communicative functions:

- complex referential expressions (reminiscent of so-called headless relative clauses)

(6) *[Tied=Ø dëd-**kid**]$_{NMLZ}$=o=Ø cho cho-ambo*
swidden=ABS fell-**AGT.NMZR**=PL=ABS REDUP=PL come-not.yet
ik-e-k.
AUX-NPAST-INDIC
'The ones who will fell (the trees of) the swidden have not arrived yet.'

- predicate modification (an adverbial function)

(7) *[Aton mado=Ø buan-tiad-**kid**]$_{NMLZ}$=bi=Ø abentsëk=bi*
3GEN son=ABS take-ABIL-**AGT.NMZR**=EMPH=ABS alone=EMPH
nid-o-ṣh.
go-REC.PAST.EXP-3
'Rather than taking her son, she went alone.'
Literally: 'The one that could/should have taken her son went alone.'

- predication (through the use of a predicate nominal; note the switch-reference adverbial clause within the nominalization)[2]

2. A brief anecdote will hopefully help convince those readers who are skeptical about the possibility of nominalization performing a predicative function. A Matses schoolteacher wrote a reader about the natural history of 15 birds. In his first draft, only about half of the sentences were of the type in Example (8). But after several revisions, he had converted almost all of the sentences into participant nominalizations embedded in copular clauses. I objected, pointing

(8) *Piuṣh=Ø [tantan-kin akte=dapa=Ø pote-**kid**]$_{NMLZ}$=Ø*
tortoise=ABS swim-while:S/A>A river=large=ABS cross-**AGT.NMZR** = ABS
ne-e-k.
be-NPAST-INDIC
'The tortoise crosses large rivers swimming.'
Literally: 'The tortoise is one that crosses large rivers swimming.'

iv. By taking the place of a modifying noun, nominalized clauses perform the following among other communicative functions:

- relativization

(9) *[Mëkueste [nukin pachid=Ø pe-**kid**]$_{NMLZ}$]$_{NP}$=Ø kues-o-mbi.*
agouti 1+2GEN manioc=ABS eat-**AGT.NMZR**=ABS kill-REC.PAST.EXP-1A
'I killed the agouti that had been eating our manioc.'
Literally: 'I killed the agouti, the eater of our manioc.'

- attribution

(10) *[Daṣhkute [piu=mbo ik-**kid**]$_{NMLZ}$]$_{NP}$=Ø bun-e-bi.*
shirt red=AUG be-**AGT.NMZR** = ABS want-NPAST-1S
'I want the red shirt.'
Literally: 'I want the shirt (the one) that is red.'

While (9) and (10) are formally quite similar and could both be said to perform a relativization function, the construction in (10) is special in that this is the *only* way to modify a noun with an adjective, other than in lexicalized terms. The syntax of these relative-clause-like constructions and other nominalization types will be described further in sections 4 and 5, once the relevant basic features of Matses grammar have been laid out in Section 3.

2.3 Inventory of nominalizing suffixes

For convenience it is possible to classify the participant nominalizing suffixes into two categories based on whether they serve as a "case recoverability strategy" (i.e., identification of the event participant to which the nominalization refers, which is described in detail for Matses nominalizations in Fleck 2008):

---

out that in my recorded natural history accounts about half of the sentences were active (i.e., not nominalizations embedded in a copular clause), but he insisted that it sounded better his way. In the end, the book (Jiménez Huanán 2014) was published with 115 predicative nominalizations (of the type in Example 8), and only 5 active sentences. His own translations into Spanish were all active.

i. Differential participant nominalizers: specify or narrow down which participant of the event is the referent of the nominalization (e.g., the Agent, the Instrument, etc.), as in Examples (11a)–(c). These are in **bold** font in Table 1.
ii. Non-differential participant nominalizers: do not specify which participant is being referred to and therefore do not help identify the referent (11d). Rather, this group of nominalizing suffixes are specific with respect to tense and evidentiality (see Fleck 2007 for a description of Matses tense and evidentiality in nominalizations). These are underlined in Table 1.

Table 1. Paradigm of all 27[a] Matses nominalizing suffixes

	Remote Past		**Distant Past**		**Recent Past**		**Present, habitual or generic**	**Future**
	Experiential	**Inferentail**	**Experiential**	**Inferential**	**Experiential**	**Inferential**		
A	*-denned*	*-ampid*[c]	*-ondaid*	*-nëdaid*	*-boed*		***-kid***	
S						***-aid***		
O or APP[b]								***-te***[d]
Instrument								
Action	*-dennek*	*-ampik*[e]	*-ondak*	*-nëdak*	*-bok*	*-ak*		*-te*
Neg Action	*-nëdakma*				*-akma*		*-tema*	
Neg O/Inst	***-nëdakmaid***				***-akmaid***		***-temaid***	
Neg A/S							***-esa***	
Causer							***-anmës***	
Characterizer							***-sio***	

a This number does not include the following 4 very archaic nominalizing suffixes:*-bompid* (= *-denned*); *-bompik* (= *-dennek*); *-oṣhaid* (= *-ondaid*); *-oṣhak* (= *-ondak*), and includes the synonymous forms in footnotes c-e.
b APP = "affected peripheral participant," that is, a non-core participant (i.e., not A, S or O) that was significantly affected by the event, such as an arrow broken in the process of shooting an animal, or ground that was charred where something was burned.
c Where *-ampid* occurs, the synonymous form *-nëdampid* also occurs.
d Where *-te* occurs as a participant nominalizer, the less generic form *-tekid* also occurs.
e Where *-ampik* occurs, the synonymous form *-nëdampik* also occurs.

(11) a. *kues-kid*
kill-AGT.NMZR
'one who kills'
'one who always kills'
'one who evidently killed (him/her/it/them)'.
'one who is killing (him/her/it/them)'
'one who will kill (him/her/it/them)'

b. *kues-an-kid*
kill-ANTIPASS-AGT.NMZR
'killer, assassin'
c. *kues-temaid*
kill-NEG.O/INST.NMZR
'what should not be killed'
'what should not be used to kill'
d. *kues-boed*
kill-REC.PAST:EXP:NMZR
'one who recently killed (him/her/it/them)'
'one whom (he/she/it/they) recently killed'
'weapon that (he/she/it/they) recently used to kill (him/her/it/them)'
'wound (resulting from him/her/it/them killing him/her/it/them)'
(in all cases entails that the speaker witnessed the killing event)

For the present paper, the relevance of this distinction is that only nominalizations with differential participant nominalizers end up becoming lexicalized, specifically, *-kid*, *-aid*, *-te*, *-anmës*, and *-sio*.[3] This is not necessarily due only to the lack of a case-recoverability function of the non-differential nominalizers, since it is also pertinent that these suffixes *always* specify specific tense and evidential values. The key feature seems to be that *-kid*, *-aid*, *-te*, *-anmës*, and *-sio* can refer to participants of generic or habitual events and states.[4] It is also noteworthy that differential participant nominalizers with negative meanings do not form terms that become lexicalized.

The unusually large size of the inventory of Matses nominalizing suffixes (31 counting the archaic ones) and the observation that inflectional markers have similar forms (e.g., *-denne* is the remote past experiential finite inflectional suffix) make it tempting to segment at least some of the nominalizers. Positing the forms *-ed* ~ *-id* 'Non-differential Participant Nominalizer' *-ak* ~ *-k* 'Action Nominalizer' and *-ma* 'Negative' to segment these nominalizers would leave only a few irregular and suppletive forms. However, this segmentation is discouraged by:

3. The suffix *-sio* is primarily a noun modifier with the meaning 'big' or 'dear'. It is used infrequently as a nominalizer to indicate that a person is characterized by performing an action excessively, and lexicalized nominalizations with *-sio* are almost exclusively nicknames; eg., *mua* 'lie' + *-sio* = *muasio* 'big liar'.

4. As an alternative and equally viable analysis, one could split up the suffixes *-kid*, *-aid*, and *-te* to differentiate their use to refer generic/habitual events vs. tense-evidentiality-specific events.

i. the lack of productivity of these three proposed suffixes with other possible inflectional suffixes (such as *-tsia* 'Future Conditional' or *-nui* 'Nonpast Uncertainty';
ii. the fact that *-ak* has a distinct specific meaning when occurring alone as a nominalizer (see Table 1);
iii. and the cognate forms in the closely-related language Matis suggest that there was never a form *-ed*/*-id*, but that rather this is a reduction of the nominalizer *-kid*, in which case the same objection in (ii) would apply.

In any case, the generalization holds whether these are segmented or not: nominalizations that code specific tense and evidential information are less likely to become noun lexemes.

3. Basic features of Matses grammar essential for analyzing nominalizations

The Matses language is predominantly dependent-marking, with possessor marking in possessive noun phrases, and strict ergative-absolutive case marking on independent arguments as the principal means of identifying grammatical relations; but it does have nominative-accusative subject person agreement (third person vs. first or second person) with some finite verbal inflections (12).

(12) a. A O
Tumi=n *mibi* *is-o-ṣh.*
man's.name=ERG 2ABS see-REC.PAST:EXP-3
'Tumi saw you.'

b. O A
Tumi=Ø *mimbi* *is-o-k.*
man's.name=ABS 2ERG see-REC.PAST:EXP-1/2
'You saw Tumi.'

c. S oblique
Nid-o-ṣh *Tumi=Ø* *aton* *ṣhubu=no.*
go-REC.PAST:EXP-3 man's.name=ABS 3GEN house=LOC
'Tumi went to his house.'

d. S oblique
Mibi *nid-o-k* *Tumi=bëd.*
2ABS go-REC.PAST:EXP-1/2 man's.name=COM
'You went with Tumi.'

Absolutive (S/O) case is unmarked (coded as "=Ø" in the examples); ergative (A) case is marked with the phrase-level enclitic = *n*, identically to instrumental and

genitive cases; all other noun phrases are obliques (optional, peripheral participants) and overtly marked as such by either phonologically bound or free postpositions (e.g., *=bëd* 'Comitative,' = *no* 'Locative,' *ëkëduk* 'inside') (12c-d). Matses has 28 prefixes that designate body-parts and semantic extensions of these (Fleck 2006b). When attached to a verb, these prefixes can function as locative postpositions, specifying a spatial relation between an extra participant and the absolutive argument of the clause:

(13) S extra participant
Kuesban=Ø ṣhubu *an-diad-e-k.*
bat=ABS house inside-hang-NPAST-INDIC
'Bats hang inside houses.'

The status of this "extra participant" is ambiguous in that it is unmarked for case (like S and O arguments), but its position is fixed (directly before the prefix) and argument tracking in switch-reference clauses and other aspects of Matses grammar do not treat it as a core argument; it is as the extra participant is the postpositional object of the prefix. Note that prefixation does not always involve an extra participant, in which case the prefix refers to a part of, or a location associated with, the S or O argument.

No other grammatical categories besides case are obligatorily marked on nouns; for example, number is marked optionally (on either nouns or verbs) and there is no gender distinction. Personal pronouns exhibit case-specific forms, though the first-person plural inclusive pronoun, the archaic second-person plural pronouns, and the third-person zero-pronoun do not distinguish A vs. O (Table 2). While full nouns in the ergative and genitive (and instrumental) case are identical, as can be seen in Table 2 ergative and genitive pronouns have distinct forms; thus personal pronouns will be essential for determining the case of arguments in nominalized clauses.

Table 2. Matses personal pronoun paradigm

	S/O	A	Genitive
1 (Singular or 1 + 3)	*ubi*	*umbi*	*kun*
2 (Singular or Plural)	*mibi*	*mimbi*	*min*
4 (Singular or Plural)	*abi*	*ambi*	*aton* [a]
1 + 2		*nuki*	*nukin*
2 Plural (archaic)		*mitso*	*mitson*
2 Plural (archaic)		*miki*	–
3 (Singular or Plural)		*Ø*	*aton* [a]

a There is no third vs. fourth person distinction for genitive pronouns.

The covert third-person pronoun (indicated as "Ø" in the examples) might seem like a potential source of confusion, but because overt first- and second-person arguments are required in almost all clause types, are case-specific, and there are no ambitransitive verbs, the simple absence of one or more expected core arguments usually unequivocally reveals the presence of third-person covert pronouns in a clause (14).

(14) Ø Ø *çhuşhka-o-şh.*
3ERG 3ABS reprimand-REC.PAST:EXP-3
'He/She/They reprimanded him/her/them/it.'

Fourth-person (i.e., third-person coreferential) pronouns (Fleck 2008) are used only in dependent clauses and will be exemplified in Section 4.1.

All verbs are strictly classified according to lexical transitivity classes (Table 3), and their valence can only be altered via overt derivational morphology (causative, applicative, reflexive/anticausative/passive, antipassive and reciprocal suffixes).

Table 3. Transitivity classes of verbs in Matses

Transitivity Class Subtype (Valence)	Core Functions and Marking	Number of Roots	Examples
Intransitive:			
(Simple) intransitive (1)	S=Ø	>400	run, cry, fall, die
Extended intransitive (2)	S=Ø, E=Ø [a]	ca. 4	want, forget, have
Transitive:			
(Mono)transitive (2)	A=*n*, O=Ø	>400	kill, see, eat, know
Ditransitive (3)	A=*n*, O=Ø, O=Ø	ca. 8	give, take, tell

a E=non-subject extended intransitive argument

Table 4. Differences between main-clause and nominalized-clause internal syntax

	main clauses	nominalized clauses
constituent order:	any order	verb final
fourth-person suffixes:	cannot occur	obligatory if coreference with a higher clause

Verb-final constituent order appears to be the most neutral, but ordering of constituents within a main clause is free of syntactic restrictions (as shown in Example 12), while word order within phrases is comparatively rigid.

There is almost no inter-clausal coordination in the language; instead, the predominant means of clause combination in Matses involves using dependent clauses in noun, adjective and adverb syntactic slots. By "dependent" I mean here simply that these clauses cannot compose a discourse-ready sentence without being

incorporated into a main clause.[5] An interesting feature of Matses grammar is that almost all dependent clauses are formed by attaching class-changing suffixes to verbs, namely via nominalization, adjectivalization and adverbialization. The infinitive-like desiderative complement clause is the only type of dependent clause that cannot be synchronically identified as formed by means of a class-changing suffix. Thus, there are only 4 (formally- and morphosyntactically-defined) types of dependent clauses in Matses:

Adverbialized clauses (including switch-reference clauses)
Adjectivalized clauses
Nominalized clauses
Complement clause (evidently historically a nominalized clause; see Fleck 2006c)

When the first three of these clause types are incorporated into the adverb, adjective, or noun slots of a higher clause, they take the same morphology and are otherwise treated grammatically the same manner as single words of their corresponding class, with a few qualifications that will be elaborated on for nominalized clauses in the following section.

4. Grammar of nominalized clauses

4.1 Internal syntax

The internal syntax of all the different types of dependent clauses in Matses is essentially the same, and differs from that of main clauses in only the following three ways:

i. the verb must be clause-final (otherwise constituent order is free);
ii. coreferential arguments are optionally or obligatorily absent in some dependent clause types; and
iii. fourth-person pronouns (in their coreferential function) do not occur in main clauses.

Considering that verb-final order is the pragmatically neutral and most common constituent order in main clauses, the verb-final requirement in dependent clauses entails only that some pragmatically motivated word orders with arguments

5. In previous descriptions I have called these "subordinate clauses," with the same intended meaning. Here I attempt to use a more neutral term. By referring to nominalized clauses as dependent or subordinate, I do not mean subordinate to a head noun, but to a higher clause.

following the verb are not possible. Therefore, one cannot distinguish a dependent clause based solely on its constituent order.

In *participant* nominalizations the referent of the nominalization cannot occur overtly within the nominalized clause, neither as a full noun or as a pronoun. In other words, there is always a so-called "gap" in participant nominalizations. However, since order is free (for any element other than the verb) and the referent of the nominalization can be an oblique, one cannot identify the location of this "gap"; and since arguments of the nominalization can be present but covert (if third person), one cannot identify the referent solely by the "gap."

The covert third person pronoun can be anaphoric (he/she/it/they) or indistinct/generic (i.e., one, someone, something, etc.; Fleck 2006a: 562). Considering this, one could posit an analysis wherein the absent referent of the nominalization (the "gap") is a third-person covert pronoun within the clause. But this is a moot point and of no descriptive utility. Nevertheless, my literal translations of Matses participant nominalizations include the word "one", but this is in consequence to the structure of the English language, rather than a reflection of my analysis.

Meanwhile, in action nominalizations, where the referent is not one of the participants, any or all participants can be included in the clause, as in (15)

(15) *[Debi=n ënden bawen=Ø anseme-ak]*$_{NMLZ}$*=Ø*
Davy=ERG early large.catfish=ABS fish.with.hook-ACT.NMZR=ABS
bëda=mbo ik-e-k.
good=AUG be-NPAST-INDIC
'It is good that Davy fished a *bawen* catfish early.'

Fourth person (i.e., third-person co-referential) pronouns occur obligatorily in any nominalization where a core argument (A, S, or O) in the nominalized clause is co-referential with a core argument (or a genitive noun that is part of the argument) in a higher clause, as in (16).

(16) *[**Ambi** pinchuk=n tuska-bued]*$_{NMLZ}$*=Ø*
4ERG thorn=INST pierce-REC.PAST.EXP.NMZR=ABS
dë-chik-ban-tsëk-kin ankun-ban-e-k chido=n.
tip-pull.out-ITER-DIM-while:S/A>A string-ITER-NPAST-INDIC woman=ERG
'The women$_i$ string **the ones (i.e., seeds) that they$_i$ perforated with thorns**, pulling the end (of the string) out and through (to make a necklace).'

Determining the co-reference relations of fourth-person pronouns is rather complex and has already been described in detail in Fleck (2008). Here I will only point out that in cases like the one in Example (16), *ambi* reveals that the A argument of the nominalized clause is in the ergative case, rather than the genitive case, as the third-person (coreferential or non-coreferential) pronoun is *aton* (see Table 2). All

such cases where pronouns occur within nominalized clauses show that nominalized clauses (and all dependent clause types) have the same case-marking frame as main clauses, as can be observed in the following example.

Table 5. Differences between participant nominalizations and action nominalizations.

participant nominalizations	**action nominalizations**
internal syntax ("gap")	
the participant that is the	all participants can occur
referent of the nominalization	overtly within the
cannot appear in the clause	nominalization
external syntax	
can occur in essentially all noun	very limited number of noun
slots	slots
can serve a relativization function	no relativization function
lexicalization	
some types can become lexicalized	never lexicalized

(17) *Ad=en neshka-aṣh, "Mitsana*
like.that=MANR:TR pluck-after:S/A>S 2GEN: 3POSS
*nëid=Ø ne-enda-k, mado, [**umbi***
this.one=ABS be-NPAST:PRMS-INDIC son **1ERG**
*se-bued]$_{NMLZ}$=Ø. **[Mimbi***
shoot.with.arrow-REC.PAST.EXP.NMZR=ABS **2ERG**
se-aid]$_{NMLZ}$=bi=k=Ø, umbi pe-e-k"
shoot.with.arrow-PAT.NMZR=EMPH=SEPAR=ABS 1ERG eat-NPAST-INDIC
ke-onda-ṣh.
say-DIST.PAST.EXP-3
'After plucking them (wood-quails) like that, he said, "This one is yours, son, the one that I have shot. And the one that you have shot I will eat."

It should be emphasized that participants of nominalized events are never associated with the verb of a nominalization through genitive marking or other strategies not used in main clauses.

4.2 External syntax

Nominalized clauses are case marked identically to regular noun phrases. Because nominalized clauses are always verb final, case marking (18) or postpositional enclitics (19) are attached to the verb, following the nominalizing suffix (through some noun-phrase enclitics can intervene, such as the plural marker):

(18) *[Wesnid=Ø pe-tsëk-**kid**]$_{NMLZ}$ = **n** onkete=Ø*
curassow = ABS eat-DIM-**AGT.NMZR=GEN** language = ABS
tantia-ben-ondaṣh ne-e-bi.
know-INCHO:TR-after:DIST.PAST.EXP:S/A>S do-NPAST-1S
'I do it (i.e., speak Spanish) now, after I long ago learned the speech of lowly curassow eaters (i.e., rural Amazonians).' [i.e., explaining why he doesn't speak the standard dialect of Spanish]

(19) *[Umbi is-**ondak**]$_{NMLZ}$=**no** dadpen ik-onda-ṣh*
1ERG see-**DIST.PAST.EXP:ACT.NMZR=LOC** many be-DIST.PAST.EXP-3
kauṣhchued=Ø.
tetra=ABS
'Where I looked long ago, there were many tetras (type of small fish).'

While action nominalizations, like the one in (19), have a very limited distribution (see Section 2.2), a participant nominalization can occur in any slot in which a noun can occur; that is, in a core argument (A in 20, and O in 21a) or peripheral participant slot, or as part of a noun phrase as a genitive (as in Example 18) or non-genitive modifying a second nominal (Section 4.3). The only exception is as the possessee of a genitive phrase (21b), unless the nominalization has become lexicalized (21c).

(20) *[Kopiok ka-bued]$_{NMLZ}$=n dadpen=kio*
left call-REC.PAST.EXP.NMZR=ERG many=AUG
shik-se-bud-ne-ak.
chest-pierce-DUR-DISTR-NARR.PAST
'The one they called Lefty had shot many (woolly monkeys) in the chest.'

(21) a. *Tumi=n [Dësi=n sika-aid]$_{NMLZ}$=Ø*
man's.name=ERG woman's.name=ERG strain-PAT.NMZR=ABS
chiwin-o-ṣh.
spill-REC.PAST.EXP-3
'Tumi spilled what Dësi had strained.'

b. **Tumi=n* ***kun*** *[Dësi=n sika-aid]$_{NMLZ}$=Ø*
man's.name=ERG **1GEN** woman's.name=ERG strain-PAT.NMZR=ABS
chiwin-o-ṣh.
spill-REC.PAST.EXP-3
'Tumi spilled my what Dësi had strained.'

c. *Tumi=n* ***kun*** *sikaid=Ø chiwin-o-ṣh.*
man's.name=ERG **1GEN** strained.beverage = ABS spill-REC.PAST.EXP-3
'Tumi spilled my strained beverage.'

This distinction of whether a nominalization can be possessed or not is not based on whether the nominalization contains more than one word or whether the

subject of the nominalization is present or not, as might be assumed upon comparing (21b) and (21c). Indeed, (21c) could not be interpreted with *sikaid* as a nonce formation, with a meaning 'Tumi spilled my what had been strained.' The following examples illustrate more clearly how the same form is treated differently with respect to possessibility when in one case it is a nonce formation (22a) and in the other a lexicalized label (22b).

(22) a. **aton kues-te*
3GEN hit-INST.NMZR
'his$_i$ thing that he$_{i/j}$ will use to hit'
b. *aton kueste*
3GEN stick/club
'his stick/club'
(e.g., a length of bamboo, a paddle)

While the ability to be possessed is the only strict syntactic distinction between lexicalized and nonce nominalization, it is an important one, considering that in Matses only noun phrases can be possessed by genitives.

In addition to case marking, nominalized clauses, lexicalized or nonce, can take any other morphology that noun phrases do, including plural (Example 6), emphatic (Examples 7 and 17), augmentative, diminutive, and other enclitics. Additionally, they can be followed by any of the grammatical particles that follow noun phrases, such as *pado* 'deceased' *chedo* 'et cetera' or *penkio* 'Negative', which will be illustrated in Section 5.2.

4.3 Relativization

Just as a noun root can modify another noun root in Matses (23), a participant nominalization can modify a second nominal (24 & 25).[6]

(23) a. *bëdi bitsi*
jaguar skin
'jaguar hide(s)'
b. *bëdi=n bitsi*
jaguar=GEN skin
'jaguar's/jaguars' hide(s)'

6. Note that Matses does not have a compounding process that involves phonological fusion of two noun stems or any special grammar that distinguishes noun-noun compounds from noun-noun NPs. Any of the noun phrase types in (23) and (24) can be used as nonce expressions and all can become lexicalized.

(24) a. ? ?*[pambid=Ø pe-kid]*$_{NMLZ}$ *bitsi*
meat=ABS eat-AGT.NMZR skin
'carnivore hide(s)'

b. *[pambid=Ø pe-kid]*$_{NMLZ}$*=n* *bitsi*
meat=ABS eat-AGT.NMZR=GEN skin
'carnivore's/carnivores' hide(s)'

3. *[tëstuk ma-we-kid]*$_{NMLZ}$ *bitsi*
epiphyte head-lie-AGT.NMZR skin
'margay hide(s)' (a margay is a small arboreal feline whose skin is valuable in the illegal fur trade; its name refers to its habit of lying on tree limbs under the cover of epiphytes (arboreal plants) to ambush its prey)

(25) *[[Kun mado [Ducho ka-boed]*$_{NMLZ}$*]*$_{NP}$*=n tita]*$_{NP}$*=Ø*
1GEN son Lucho call-REC.PAST:EXP:NMZR=GEN mother=ABS
"'Kuen-enda!' ka-onda-ṣh-i" ke-onda-ṣh.
run.off-NEG.IMPER tell-DIST.PAST.EXP-3-1O say-DIST.PAST.EXP-3
'The mother of my son, the one they call Lucho, said "They told me 'Don't run off!'".
Alternate translation: 'The mother of my son who is called Lucho…'

The context of the example in (25) is that the narrator has many sons, several of whom, including this one, are named Dunu, but they have more recently acquired Spanish names, which narrow down the reference. While noun phrases like those in (23a) and (25) have parallel structures, there is a semantic distinction to be made. While in (23) one entity (jaguar) modifies a second, distinct entity (hide), in (25) both nominals refer to the same entity (Lucho). I have not yet found text examples or overheard cases where a nominalized clause modifies a second nominal without the two referring to the same entity, other than when the nominalized clause is lexicalized (24c). Elicited examples like (24a) are disfavored and generally corrected with genitive phrases, like (24b). As with the English translations, non-genitive phrases like (23a) are used to refer to a type of hide, with the term jaguar taking on a generic reference, while ones like those in (23b) tend to be interpreted as the hide of (a) particular jaguar(s). Note that "carnivore" is a nonce formation, as it is not relevant in Matses classification of animals, and that if replaced with a lexicalized nominalization, as in (24c), it becomes completely acceptable. While the possessibility restriction explained in Section 4.2 seems to be a grammatically motivated pattern, the disinclination toward non-coreferential modification by nonce nominalized clauses may be a semantic one; that is, their tendency to refer to specific referents as opposed to generic categories.

In the cases where both the modifying and the modified nominals refer to the same entity, there is much more freedom with respect to: i) the relative ordering of the nominals, (ii) whether they are both case-marked, and (iii) whether they occur separate from each other. While this applies to any pair of nouns, noun phrases or nominalizations, this syntactic freedom is exploited to the fullest degree by nominalized clauses. Much can and has been said about the positional possibilities of nominalized clauses in this relativization function (Fleck 2003, 2008). Here I will limit myself to a few basic observations:

– A nominalized clause and the nominal(s) it modifies can be treated as a single noun phrase, with a single case marker indicating their grammatical function. While this is ambiguous with absolutives (being zero-marked), this is clear with ergative arguments (26), genitives (25) and obliques.

(26) *[Tsusio, kun şhëni=sio pado, Kuibusio pado,*
old.man 1GEN older.namesake=large deceased man's.nickname deceased
*[aton chido dadpen ik-kid]$_{NMLZ}$]$_{NP}$=**n** Ø "Utsi*
3GEN woman many have-AGT.NMZR = **ERG** 3ABS REDUP=DISTR
utsi=en kuëdën-me-ta!" ka-onda-şh.
other=MANR:TR sing-CAUS-IMPER tell-DIST.PAST.EXP-3
'The old man, my late big older namesake, the late Bearded One, the one who had many wives, told them (his wives) "Have some of them sing!"'.

– Especially when either the nominalization or the modified noun phrase is long, there is the option of marking case on both (or more) nominals:

(27) *Adnubik, [[kun şhëni pado]$_{NP}$=n opa]$_{NP}$=**n** [[piktsëk*
meanwhile 1GEN older.namesake deceased=GEN dog=**ERG** small
*ik-kid]$_{NMLZ}$ daëd ik-tsëk-kid]$_{NMLZ}$=**n** aid=**n**=kio*
be-AGT.NMZR two be-DIM-AGT.NMZR = **ERG** that.one=**ERG** = EMPH
[bëdi=n taë]$_{NP}$=Ø bedan-onda-şh.
jaguar=GEN footprint=ABS follow-DIST.PAST.EXP-3
'Meanwhile, my late older namesake's dogs, ones that were two little ones, those very ones, followed the jaguar's spoor.'

– Nominalized clauses can occur disjunct from the modified noun:

(28) *Poshto=n **inkuente=Ø** ik-e-k [[kuëte*
woolly.monkey=GEN **tail**=ABS exist-NPAST-INDIC tree
kuidi]$_{NP}$=Ø widënua-te]$_{NMLZ}$=Ø.
branch=ABS hold-INST.NMZR=ABS
'The woolly monkey has a prehensile tail.'
Literally: 'The woolly monkey's tail exists, one that is used to hold onto tree branches.'

– When the nominalized clause and the modified noun are disjunct, both must be case-marked:

(29) *[Ad-kid]*$_{NMLZ}$*=**n**=bi* *kachina=Ø* *pe-e-k*
be.like.that-AGT.NMZR = **ERG** = EMPH chicken=ABS eat-NPAST-INDIC
*mapiokos=**n**.*
common.opossum=**ERG**
'The one that is like that, the common opossum, eats chickens.'

– Two nominalized clauses can modify a single noun phrase, sometimes with the two nominalizations being treated as a single noun phrase, and sometimes not.

(30) *[Kun dawës pado]*$_{NP}$*=n* *[[Dashe=Ø*
1GEN older.male.cross.cousin deceased=ERG man's.name=ABS
kuëmëd-kid]$_{NMLZ}$*=bi* *[Tsiakketsëkkid=Ø*
be.named-AGT.NMZR=EMPH man's.nickname=ABS
ka-ondaid]$_{NMLZ}$*]*$_{NP}$*=n* *Ø* *shik-kues-kio-onda-şh.*
call-DIST.PAST.EXP.NMZR=ERG3 ABS chest-hit-EMPH-DIST.PAST.EXP-3
'My late older male cross cousin, the one who was named Dashe but whom they used to call Tsiakketsëkkid, shot him (a non-Indian) squarely in the chest.'

– Two nominalized clauses commonly refer to a single entity without an overt non-nominalized modified nominal. I have no examples of a phrase being formed in such cases, though absolutive arguments are formally ambiguous in this respect:

(31) *[tantia-kid]*$_{NMLZ}$*=n=kio* *[ënden më-kiad-kid]*$_{NMLZ}$*=n=kio*
know-AGT.NMZR=ERG=EMPH long.ago hand-learn-AGT.NMZR=ERG=EMPH
bëda=mbo=en *isan-Ø* *chokueshka-e-k.*
good=ADVZR=MANR:TR palm.species=ABS mash-NPAST-INDIC
'Ones who know well, ones who learned long ago, mash *isan* palm fruits well.' (refers to women)

(32) *[Kuidi=Ø cho-kid]*$_{NMLZ}$*=Ø* *[bëda=mbo ik-kid]*$_{NMLZ}$*=Ø*
branch=ABS have-AGT.NMZR=ABS good=AUG be-AGT.NMZR=ABS
is-şhun *këwëte=Ø* *te-e-k* *Matses=n.*
see-after:S/A>A hook=ABS cut-NPAST-INDIC Matses=ERG
'After finding a branched one, a good one, the Matses cut it into a hook.' (refers to a sapling)

As argued in Fleck (2003, 2008), Matses does not have canonical relative clauses, but rather nominalized clauses can be considered to have a "relativization

function" in addition to their other functions. The nominalizations in the present section could be described as being in an apposed relation to the modified noun and/or the other nominalized clause.

The relevance of this relativization function for the present paper is that lexicalized nominalizations can consist of either a simple nominalization or a nominalization modifying a noun (or a more complex nominal). Both types will be illustrated in Section 5. It should be noted that when I speak of multi-word lexicalized nominalizations, a modified noun is not counted as part of the nominalization, but it can form part of the lexeme.

Although I said that I would not go into a detailed discussion here of the different types of relative-clause-like nominalizations, one type deserves closer attention. In Matses, adjectives cannot directly modify a nominal, except as part of a lexicalized phrase. As illustrated in Example (10), in nonce formations adjectives must occur in a nominalized clause in order to attribute the quality they denote to a nominal. Adjectives modify nominals directly only when they form part of a lexicalized label, like those in (33), and as such are similar to English compounds like *bluebird*.

(33) a. *kana* 'macaw' + *piu* 'red' = *kana piu* 'scarlet macaw'
b. *matses* 'person' + *uşhu* 'white'
= *matses uşhu* 'person of European descent'
c. *kuëte* 'tree' + *pachi* 'soft' = *kuëte pachi* 'giant *Ceiba* tree'
d. *isese* 'liver' + *padish* 'weak' = *isese padish* 'lung'
e. *bakuë* 'baby' + *çhëşhë* 'black' = *Bakuë Çhëşhë* 'Black Baby (nickname)'

If one wished to refer to a black species of macaw from another country or a macaw that was painted black, etc., the only means of expressing it as a phrase is the structure in (34).

(34) *kana [çhëşhë=mbo ik-**kid**]*$_{NMLZ}$
macaw black=AUG be-**AGT.NMZR**
'black macaw' Literally: 'macaw (one) that is black'

If one uses a expression with the structure in (33) when it is not a lexicalized label, it will be interpreted as an attempt to coin a new term or as a joke. When a party of sunburnt tourists arrived in a boat at a Matses village, one man made the following exclamation, which was a source of much merriment among the rest of the onlookers:

(35) *[Matses piu]*$_{NP}$*=Ø cho cho-ak!*
people red=ABS REDUP=DISTR come-REC.PAST.INFER
'Red people have come!'

The humor stemmed not just from the reference to the unfortunate state of the visitors' skin, but in that the remark was phrased as if it were a legitimate label for a race of humans, as was later explained to me, who had laughed heartily along with the Matses despite having missed the main point of the joke. While this facetious expression never caught on, during my 20 years among the Matses I have seen several noun-adjective phrase coinages gain wide currency, especially nicknames and names for Western manufactured items. Temporary vs. permanent states or other semantic considerations are irrelevant for the selection of the attributive constructions in (33) versus (34); rather the structure in (33) is limited to lexemes, coinages, and jokes. With nominalizations, on the other hand, the structure is the same for lexemes and *ad hoc* phrases, but the grammar is sensitive to lexicalized status of a nominalization.

4.4 Summary of grammatical differences between types of nominalizations

Table 7 emphasizes that while the term "nominalized clause" may be useful for characterizing the internal syntax of multi-word nominalizations in Matses and their status as dependent clauses, they differ from one-word nominalizations only with respect to the number of words they contain. In fact, considering that third-person pronouns are covert in Matses, one could propose that one-word nonce nominalizations (like the one in Example 11) are equivalent to a "one-word nominalized clause." One might assume, as I did in my dissertation, that one-word nominalizations would differ in that nominalized clauses are not "lexicalizable," but this is not so, as will be illustrated in the following section.

5. Lexicalized nominalized clauses

5.1 Internal grammar

All the examples used to illustrate the present section are the only or the principal name for animals, plants, ailments, or (historical or mythical) ethnic groups, so there is no doubt that these are lexemes in the Matses language. As lexemes, they are not synchronically parseable, since the meaning of the lexeme is not a predicTable product of the meanings of the component words and morphemes. However, since the examples I have selected have not undergone reduction or any other type of phonological alteration, the same forms could just as well be produced as nonce nominalizations with a meaning that I will include below the lexical gloss as the "nonce meaning".

Examples with the different nominalizing suffixes will be analyzed in turn, since different nominalizers illustrate different aspects of the internal syntax. For example, the following three lexemes are all formed with the suffix *-kid* 'Agent Nominalizer,' which selects the S or A of the verb, and therefore can illustrate the case-marking of the O argument (36) or of obliques (37 and 38), but not of the subject. Note that the lexeme in (36) is a noun modified by a nominalization.

(36) modified noun O transitive verb
chiki *[şhëkëd=**Ø** pe-kid]*$_{NMLZ}$
hawk lizard=ABS eat-AGT.NMZR
lexical gloss: 'species of hawk'
nonce meaning: 'lizard-eater hawk'.

(37) oblique transitive verb
*di=**n*** *tsi-mak-kid*
palm.fiber=INST butt-wrap-AGT.NMZR
lexical gloss: 'historical Indian tribe'
nonce meaning: 'one(s) who wrap(s) the notch-end (of their arrows) with *Astrocaryum* palm fiber twine'

(38) oblique transitive verb
*danto=**n*** *se-an-kid*
knee=LOC pierce-ANTIPASS-AGT.NMZR
lexical gloss: 'mythical tribe whose members stabbed people in the knees'
nonce meaning: 'one(s) that pierce(s) on the knee' or 'on-the-knee piercer(s)'

(39) extra participant transitive verb
akte *an-che-kid*
water inside-eat.unchewed-AGT.NMZR
lexical gloss: 'agami heron'.
nonce meaning: 'one that eats in the water'

As can be observed in (36–38), O arguments, instruments, and locatives are case-marked in the same manner as in nonce nominalized clauses (and main clauses). Likewise, in (39) it can be seen that extra participants (the "objects" of body-part/locative prefixes presented in Section 3) are unmarked, as in any other clause type (cf. Example 13). See Example (24c) for another prefixed nominalized verb, and note that in (37) the prefix occurs without an extra participant, as is also common in any clause type.

Multi-word nominalizations formed with the instrumental nominalizing suffix *-te* are more informative in that the referent of the nominalization being an oblique, any of the core arguments may appear overtly in the lexeme:

(40) O transitive verb
*ṣhawi=**Ø*** *annoṣhka-te*
blowgun=**ABS** sand.inside-INST.NMZR
lexical gloss: 'vine snake' (type of very thin snake)
nonce meaning: 'one for sanding the inside of blowguns'

(41) A transitive verb
*pesa=**n*** *madad-te*
toucanet=**ERG/GEN** brood-INST.NMZR
lexical gloss: 'crabwood tree (*Carapa guianensis*; *andiroba* in Spanish; a species of timber tree)'
nonce meaning: 'one used by toucanets to brood (their eggs)'
nonce meaning if=*n* is interpreted as the genitive maker: 'toucanet's brooder'

(42) S intransitive verb
*achu=**Ø*** *kuëdën-te*
howler.monkey=ABS sing-INST.NMZR
lexical gloss: 'large tree type used to make canoes'.
nonce meaning: 'one for howler monkeys to sing'

The following pair of examples are with the nominalizer *-aid*, which selects as its referent an O, an S, or any oblique that is considered to have been affected by the event:

(43) A transitive verb
*majan=**n*** *pe-aid*
spirit=ERG/GEN eat-PAT.NMZR
lexical gloss: 'jock itch'
nonce meaning: 'what was bitten by a spirit'
nonce meaning if=*n* is interpreted as the genitive maker: 'spirit's bite'

(44) modified nominal A transitive verb
[ṣhëkten neste]$_{NP}$ *[**ambi** pe-aid]*$_{NMLZ}$
collared.peccary medicinal.plant **4ERG** bite-PAT.NMZR
lexical gloss: 'tree in the Leguminosae family'
nonce meaning: 'collared peccary$_i$ medicine that it$_i$ eats' i.e., one of several types of medicinal plants used when spirits of collared peccaries (boar-like mammals) cause an illness; this particular plant species being distinguished from others in its class by the fact that collared peccaries eat it

While Example (43) has the same dual possible interpretation as (41), the lexeme in (44) resolves the issue in that the genitive interpretation is not possible because ergative and genitive pronouns have different forms: the third-person (coreferen-

tial or non-coreferential) genitive pronoun *aton* would be expected if subjects of nominalizations were in the genitive case.

The final example of this section illustrates the use of the cause nominalizer *-anmës*:

(45) O (causee) intransitive verb
ëşhë=Ø *nën-anmës*
eye=ABS hurt-CAUSER.NMZR
lexical gloss: 'conjunctivitis'
nonce meaning: 'one that makes one's eyes hurt'
alternative nonce meaning: hurter of the eye, eye hurter (cf. *ëşhë dauë* 'eye medicine')

The suffix *-anmës* is particularly interesting in that it increases the valence of a verb and selects the introduced argument, the causer, as its referent (see Fleck 2001 for more examples and discussion of this suffix). Causative dependent and main clauses, formed with the causitivizer *-me*, likewise exhibit an increase in valence, as described in Fleck (2002).

Although most multi-word lexicalized nominalizations involve only two words, they can nevertheless be rather complex. For example, in (46) the O is itself a nominalization, and the higher nominalized verb is a verbalized adjective. Lexicalized nominalizations can even involve reduplicated verbs (47) or quotations (48). Despite their complexity, the internal syntax of these examples is precisely the same as that of a main clause or nonce nominalized clause.

(46) O adjective → transitive verb
pe-te=Ø *piu-wa-te*
eat-INST.NMZR=ABS red-make:TR.VBZR-INST.NMZR
lexical gloss: 'spice plant in the ginger family'
nonce meaning: 'one for making food red'

(47) O transitive verb
nidaid=Ø *poşh* *poşhka-kid*
ground = ABS REDUP=DISTR perforate-AGT.NMZR
lexical gloss: 'species of beetle'
nonce meaning: 'one that perforates the ground (making many holes)'

(48) quoted complement transitive verb
sededen *ke-kid*
sunbittern.vocalization say-AGT.NMZR
lexical gloss: 'sunbittern (shorebird)'
nonce meaning: 'one that says "sededen"'

In conclusion, lexicalized multi-word nominalizations do not have any special internal grammar different from that of nonce nominalized clauses (or main clauses). This is the "key" to understanding Matses nominalization I alluded to in the introduction, and which will be elaborated on in the final discussion.

5.2 External grammar

As summarized in Table 6, there are tree differences between nonce and lexicalized nominalizations. The first one, that only the nominalizers *-kid*, *-te*, *-aid*, *-anmës* and *-sio* occur in lexicalized nominalizations, is just as true for multi-word lexicalized nominalization, with one possible, where a nominalization with the nominalizer *-ampid* could be argued to have obtained lexical status:

Table 6. Differences between grammar of lexicalized and nonce (participant) nominalizations.

nonce nominalizations	lexicalized nominalizations
formed with any of the 27	formed with only 5 of the
nominalizing suffixes	nominalizing suffixes
cannot be possessed	can be possessed
disfavored for modifying nouns	can readily modify nouns,
that are not co-referential	coreferential or not

Table 7. Differences between external grammar of nominalized clauses and nonce (one-word) verbal nouns.

none

Table 8. Inventory of lexicalized nominalizations in a Matses dictionary

Number of words	Conservative count[a]	Unrestrained count[b]
1 word	307	276
2 words	110	132
3 words	3	12
TOTAL lexicalized nominalizations:	420	420

a Counts as one word: (i) reduplicated forms (like Example 47); (ii) copulas contracted with the preceding word, or (iii) onomatopoetic forms followed the quotative verb *ke* 'say' (similar to Example 48) when it is ambiguous whether they form single phonological word.
b Counts any instance of (i)–(iii) as two words.

(49) *Matses=n chui-ampid*
Matses=ERG tell-REM.PAST.INFER.NMZR
possible lexical gloss: 'Matses mythology'
nonce meaning: 'what was (evidently) told long ago by the Matses'

The second difference, that only lexicalized nominalizations may be possessed, is shown to hold for multi-word lexicalized nominalization in the following examples. Multi-word lexicalized terms, whether nominalizations (50) or genitive noun phrases (51) can be possessed, while nonce multi-word nominalizations cannot (52), despite having the same structure as lexicalized nominalization.

(50) *Kun [pesa=n madad-te]$_{NMLZ}$ ne-e-k.*
1GEN crabwood.tree be-NPAST-INDIC
'It is my crabwood tree.' (e.g., claiming it to harvest it for sale; see Example 42 for the nonce gloss of the name of this tree species)

(51) *Kun [dide=n këku]$_{NP}$ ne-e-k.*
1GEN night.monkey=GEN cowtree be-NPAST-INDIC
'It is my *naranjo podrido* tree/fruit (*Couma macrocarpa*; lit. night monkey's cowtree).'

(52) *Kun [chankuëşh=n madad-te]$_{NMLZ}$ ne-e-k.*
1GEN toucan=GEN/ERG brood-INST.NMZR be-NPAST-INDIC
* 'It is my [toucan's brooding place/nest].'
* 'It is my [one (tree/place/nest) for brooding by toucans].'
* 'It is my toucan's (i.e., a pet) brooding place/nest.'

The fact that (52) is not possessible suggests that it has not acquired lexeme status.

The third difference between nonce and lexicalized nominalizations, the disinclination toward the modification of non-coreferential nouns, was shown to apply to multi-word nominalization in the beginning of Section 4.3 (example 24c). Here follows an additional example where a lexicalized name of a palm tree modifies a second, non-coreferential noun:

(53) *[[Juka nis-te]$_{NMLZ}$ bakuë]$_{NP}$=Ø kuesban=n pe-kid.*
manioc grate-INST.NMZR fruit = ABS bat-ERG eat-HAB
'Bats eat stilt palm [lit. manioc grater] fruits.' (the spiny roots of this palm are used to grate manioc)

The final example illustrates how a multi-word nominalization (the nickname of a captured woman who passed on conjunctivitis to the Matses), just like the noun phrase in the same example, can be followed by a noun phrase particle (particles in bold).

(54) *[Kun tita* ***amano]***$_{NP}$*=Ø [[Ëşhë Nën-anmës]*$_{NMLZ}$
1GEN parallel.aunt **non.blood**=ABS eye hurt-CAUSER.NMZR
pado]$_{NP}$*=Ø [ën-kin=bi bed-aid]*$_{NMLZ}$*=Ø*
deceased = ABS finish-while:S/A > A=EMPH grab-PAT.NMZR=ABS
ne-e-k.
be-NPAST-INDIC
'My non-blood-related parallel aunt, the late Eye Illness Causer, was captured last.'

5.3 Inventory of lexicalized nominalizations

As defined in the introduction and discussed in Section 6.2, lexicalization (as I use the term in this paper) is a *diachronic* process. Consequently, at any particular point in time there will be terms that have not completed the process and many lexicalized terms may persist for some time with ambiguous lexical status. Nevertheless, fully lexicalized items can be identified through several means prior to applying syntactic tests. I consider the following to be useful indicators of lexical status of nominalizations in Matses:

i. Irregular phonological reduction, fusion or alteration.
ii. Semantic shift, whereby the meanings of the component morphemes do not add up to the meaning of the nominalization.
iii. The nominalization is the only or the principal name for an entity.

The multi-word nominalizations exemplified in the preceding sections are only a small sample of the actual number in the Matses lexicon. Of the 2268 noun lexemes catalogued in a Matses-Spanish dictionary (Fleck et al. 2012), 665 (29%) are or contain nominalizations. When nominalizations that occur more than once in the lexicon (e.g., as modifiers of more than one nominal, or as the same nominalization modified by adjectives or size-specifying enclitics) are counted only once, the result is that there are 420 different nominalizations in the lexicon. Table 8 shows how this last number is broken down by the number of words in the nominalization. Nominals modified by a nominalization, like *chiki* 'hawk' in Example (36), were not counted as part of the nominalization.

It should be noted that 40 of the one-word nominalizations had some feature that did not allow them to have nonce interpretations: 20 contained verb roots absent from the modern language; and 20 contained phonological reductions or alterations. Only one multi-word nominalization could no longer have a nonce interpretation, *Abi Doduaid* 'Lobo Creek', which seems to contain a reduplication of the hypothetical verb *do* 'ascend' (possibly the source of the suffix *-do* 'upward').

All of the multi-word lexicalized nominalizations, even this stream name, follow the clausal syntactic properties described in the Section 5.1.

Another pattern that will be of interest in Section 6 is that although nonce nominalizations may contain many derivational verbal suffixes before the nominalizing suffixes, lexicalized nominalized verbs tend to be morphologically much simpler: only 34 of the 420 nominalized lexemes in Fleck et al. (2012) contain one suffix before the nominalizer, and none more. Specifically, these suffixes are *-an* 'Antipassive' (25 instances), *-tsëk* 'Diminutive' (4), *-me* 'Causative' (3), *-nan* 'Reciprocal' (1) and *-ad* 'Reflexive' (1).

6. Discussion

6.1 A comparison with English nominalization

There are three relevant differences between English and Matses nominalization:

i. Nominalization (in the traditional sense) in English is not only notoriously irregular, but its productivity is highly constrained, while Matses nominalization is completely productive.
ii. Multi-word English nominalizations have phrasal syntax, while in Matses they have clausal syntax.
iii. English has a dedicated clause type for relativization in addition to having nominalizations that can also perform this communicative function; meanwhile, relativization in Matses is accomplished purely by nominalization, and this is not its only or principal function.

I propose that (i) and (iii) are not coincidental: one would expect wider productivity if nominalization were the only periphrastic means of modifying nominals in English.

For the present discussion I will focus on participant nominalizations and take as an example the agentive nominalizer *-er*, which is considered among the most productive nominalizers in English. Consider the following Merriam-Webster definitions of ***loser***, illustrated with my own sample sentences:

1. someone or something that loses a game, contest, etc.
   e.g., The loser (of the card game) has to wash the dishes. [refers to a participant of a specific event]
   e.g., Losers (of contests) always have an excuse. [generically refers to a participant of any event]

2. a person who is not successful, attractive, etc.
   e.g., Edward is a loser. [refers to a type of person, who may never have competed]

The second meaning of *loser* is a good example of a lexicalized nominalization, but of greater interest to the present discussion is the first meaning. If it is transparently analyzable, why should the first meaning of *loser* be included in a dictionary? The implication is that the author considers it to have lexeme status. In a more fine-grained treatment of lexicalization, it could be said that meaning 1 of *loser* is conventionalized, but has not undergone enough phonetic or semantic changes to be considered fully lexicalized. If we deem nominalizations like *loser* to be (partial) lexemes, it would entail that nonce nominalizations with English *-er* would be essentially limited to new formations. Yet even new nonce formations with *-er*[7] are typically spoken with some hesitation, often written with quotation marks, may be interpreted as the speaker or writer attempting to coin a new term, and tend to be judged as speech errors if uttered by a child or non-native speaker. They are not real words, not yet at least. My interpretation is that nominalization in English (as traditionally defined) is productive only for coining new terms which must be conventionalized before being acceptable.[8]

Even if one takes a less extreme position with respect to the productivity of English nominalization, the situation is nevertheless clearly quite different in languages like Matses where nominalization is as productive and unconstrained as verbal inflection. Meanings equivalent to the first meaning of *loser* do not belong in a Matses dictionary, since nonce nominalizations do not need to be conventionalized to be recognized as legitimate words (or clauses). Indeed, considering that there are 27 different productive nominalizing suffixes, that multiple derivational suffixes (from an inventory of at least 40 different derivational suffixes) can be included preceding the nominalizing suffix, and one of the 28 body-part prefixes can

7. Nonce usages of *-er* would include nominalization of verbs that are typically not considered to be suffixable with *-er* (e.g., *dier*, *disappearer*, *meeter*); of verbs for which nominalizations with *-er* exist but may not be known to the speaker (e.g., *stayer* to mean 'one who stays' rather than 'horse with endurance'); or of verbs for which nominalizations with *-er* are well-known, but have a meaning different from the intended meaning (e.g., *stinker* to mean 'someone who stinks' rather than 'something of poor quality', or 'contemptible person' or *toddler* to refer to an adult who toddles as opposed to a child in a certain age category).

8. I mean acceptable in a formal since, since nominalization in English can be used creatively or facetiously, analogously to the Matses example in (35). I propose that it is the unusualness of using nominalization for reference rather than coining that makes such utterances interesting or humorous. It would be interesting to study whether such non-standard usages ever end up becoming lexemes.

also be attached, to include all the possible combinations as entries would result in the heaviest dictionary yet published, not to mention multi-word nominalizations.

With respect to the second difference listed above, it is noteworthy that a synchronically unanalyzable nominalization can head a multi-word nominalization in English (e.g., *Bob is a plumber of mansions*). Conversely, Matses nominalizations that have become synchronically unsegmenTable due to phonetic change or because the verb no longer exists in the language cannot head nominalized clauses. For example, it can be inferred that the noun *mekte*, an archaic term for 'paddle,' was formed by a historical verb *mek* 'to paddle' plus the instrumental nominalizer *-te*, yet *mekte* cannot be used to head a nominalized clause. This seems to be applicable as well to transparent lexicalized nominalizations, but it is harder to show because they can always be simultaneously interpreted as nonce nominalizations. What this difference suggests is that English multi-word nominalizations are formed by taking a word from the lexicon and associating any participants with it through the use of genitives and other means available for forming noun phrases (although a psycholinguistic study would be required to propose this with any certainty). Meanwhile, in Matses it is as if a pre-formed clause is nominalized as a unit, rather than being built up with clausal syntax around a deverbal noun.

6.2 Lexicalization paths

One of the most common phrases that I have come across in language-specific and general descriptions of nominalization is that "nominalization creates new noun lexemes." Yet seldom do the authors elaborate on just how the derived forms enter the lexicon. Obviously a single utterance of a nonce nominalization does not automatically gain entry into the lexicon of the speech community.

The discussion in the preceding section might have led the reader to ask: do nominalizations in English and Matses become lexemes by different means? My answer is "Probably." Since Matses does not have a long written tradition and comparison with other Panoan languages does not reveal much, a general description of lexicalization of nominalizations in Matses is necessarily speculative. Nevertheless, since this phenomenon has been studied for English, based on the differences between the two languages it is possible to assess whether the lexicalization paths found in English are viable for Matses.

In the typical lexicalization path for English nominalizations (and compounds), a concept first comes to be relevant in a speech community, which creates an atmosphere receptive to a new coinage, then after the coined nominalization gains acceptance and currency it develops independently semantically and phonologically as (or after, depending on how lexicalization is defined) it becomes part of the speech community's lexicon. Because in Matses nonce derivational

nominalization is not restricted to coinages, at least some Matses nominalizations may have a longer road to lexicalization, starting out as simply nonce nominalizations that describe a concept that becomes important later.

Leaving aside this speculation, the relevant point is that in Matses regular nonce nominalizations and nonce nominalizations made with intentions to coin have precisely the same grammar. The finding that multi-word lexicalized nominalizations have the same internal grammar as nonce nominalized clauses is the keyhole that allows us a peek at the lexicalization path of Matses nominalizations. Meanwhile, the observation made at the end of Section 4.3 that noun-adjective phrases are a coining strategy that differs from the construction used for nonce attribution reveals that there are different lexicalization paths in other parts of the grammar.

One caveat is that the only nominalizations that end up becoming lexicalized, as mentioned in Section 2.3, are *-kid*,*-aid*, *-te -anmës*, and *-sio.* However, these do not form construction types different from participant nominalizations formed with the other nominalizers, and their privilege of forming nominalizations that can be lexicalized is not a grammatical property, but a semantic one; specifically their ability to code habitual tense and generic actions. Meanwhile, the rest of the nominalizers code specific tense, aspect and/or evidentiality, and are therefore less likely to lexicalize into nouns that become independent of the verbal semantics and the specific event for which they are composed. This trend is reiterated with respect to the complexity of nominalizations: lexicalized nominalizations generally contain fewer words and fewer suffixes, lending themselves more easily to a generic event interpretation that eventually becomes independent of the verbal semantics. In light of this, could it be a coincidence that the antipassive, which backgrounds the O argument, is the most common suffix among those found in lexicalized nominalizations?

"Lexicalized nominalized clause" may sound like an oxy-moron because many see the lexicon as a place where clauses are outlawed. Although "lexicalized nominalized clause" is an accurate formal characterization, I would concede that from a communicative perspective multi-word nominalizations do not necessarily function as clauses. This is because lexicalization of nominalizations in Matses involves a process whereby the derived word or clause comes to refer directly to an entity (like a typical noun does), rather than through the event or state coded by the verb and its arguments. In other words, a nominalized clause is no longer interpreted primarily by deciphering it as a clause, but as a unitary sign. Nevertheless, speakers who are not familiar yet with the lexeme can gain a clue to its meaning by interpreting it as a clause, and it is clear that even those who handle it primarily as a unitary sign can simultaneously recognize it as a composite unit. One line of evidence for the latter claim is that a source of humor for the Matses is to translate

lexicalized nominalized clauses into Spanish literally. I once heard much laughter generated by the literal translation of *maçho pe-kid* 'species of army ant' as *mordedor de viejas* ('biter of old women').

Morpheme gloss abbreviations:

1	First Person
2	Second Person
3	Third Person
4	Fourth person
A	more agentive argument of transitive verb
ABIL	Abilitative
ABS	Absolutive
ACT	Action
ADVZR	Adverbializer
AGT	Agent
ANTIPASS	Antipassive
AUG	Augmentative
AUX	Auxiliary
CAUS	Causative
COM	Comitative1
DIM	Diminutive
DIST.PAST	Distant Past
DISTR	Distributive
DUR	Durative
EMPH	Emphatic
ERG	Ergative
EXP	Experiential (evidentiality)
GEN	Genitive
HAB	Habitual
IMPER	Imperative
INCHO	Inchoative
INDIC	Indicative
INFER	Inferential (evidentiality)
INST	Instrumental
ITER	Iterative
LOC	Locative
MANR	Manner
NARR.PAST	Narrative Past
NEG	Negative
NMZR	Nominalizer
NPAST	Non-past
O	less agentive argument of transitive verb
S	subject of intransitive verb
PAT	Patient
PL	Plural
POSS	Possessee
PRMS	Permission
REC.PAST	Recent Past
REDUP	Reduplication
SEPAR	Separate
TR	Transitive
VBZR	Verbalizer

References

Fleck, David W. 2001. Culture-specific notions of causation in Matses grammar. *Journal de la Société des Américanistes* 87: 177–196.

Fleck, David W. 2002. Causation in Matses (Panoan, Amazonian Peru). In *The Grammar of Causation and Interpersonal Manipulation* [Typological Studies in Language 48], Masayoshi Shibatani (ed.), 373–415. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.48.15fle

Fleck, David W. 2003. A grammar of Matses. PhD dissertation, Rice University.

Fleck, David W. 2006a. Antipassive in Matses. *Studies in Language* 30(3): 541–573. https://doi.org/10.1075/sl.30.3.03fle

Fleck, David W. 2006b. Body-part prefixes in Matses: Derivation or noun incorporation? *International Journal of American Linguistics* 72(1): 59–96. https://doi.org/10.1086/505279

Fleck, David W. 2006c. Complement clause type and complementation strategies in Matses. In *Complementation: A Cross-linguistic Typology*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 224–244. Oxford: OUP.

Fleck, David W. 2007. Evidentiality and double tense in Matses. *Language* 83(3): 589–614. https://doi.org/10.1353/lan.2007.0113

Fleck, David W. 2008. Coreferential fourth-person pronouns in Matses. *International Journal of American Linguistics* 74(3): 279–311. https://doi.org/10.1086/590084

Fleck, David W. 2013. Panoan languages and linguistics. *Anthropological Papers of the American Museum of Natural History* 99.

Fleck, David W., Uaquí Bëso, Fernando Shoque & Jiménez Huanán, Daniel Manquid. 2012. *Diccionario Matsés-Castellano, con Índice Alfabético Castellano-Matsés e Índice Semántico Castellano-Matsés*. Iquitos: Tierra Nueva Editores. <http://repositorio.pucp.edu.pe/index/handle/123456789/49614?show=full>

Jiménez Huanán, Daniel Manquid. 2014. *Cun Nidaidën Icquid Podo Choquido: Aves de Mi Tierra*. Lima: UNICEF. <http://acateamazon.org/wp-content/uploads/2015/06/Las-aves-de-mi-tierra-.pdf>

CHAPTER 17

Nominalization and its pervasiveness in Xavante

Adriana M. Estevam
Independent Scholar

The aim of this presentation is to show the ubiquitous character of nominalizations in Xavante, a Jê language spoken in central Brazil by approximately 15,000 speakers. After a brief presentation of background information, we will give a synchronic description of lexical and grammatical nominalizations. The section devoted to grammatical nominalization points out the hybrid nature of constituents we are inclined to call 'nominalized clauses': it offers an analysis of their internal (clausal) structure and their syntactic (nominal) behaviour. A final section presents a diachronic hypothesis involving several contexts where nominalizations can be argued to have operated at the clause level in a previous state of the language.

Introduction

Spoken in the Mato Grosso state of Brazil by approximately 15,000 speakers, the Xavante language is a member of the central group of languages of the Jê family. It is an agglutinative, head-marking, head-final language. It shows nominalizations at the lexical level, but it also displays means for allowing a clause to function as a nominal constituent, i.e., it also allows grammatical nominalization (Shibatani 2009). In addition to these clearly synchronic nominalizing processes, the configurations of two types of independent clauses suggest that they arose from nominalizations. Taken together, these properties confer a ubiquitous character to nominalizations in Xavante. The purpose of this paper is to show this pervasiveness by adopting two perspectives, one synchronic, and one diachronic. In other words, our objective is twofold: to offer a synchronic description of nominalization, and to suggest a diachronic hypothesis that explains how finite constructions may have arisen from nominal predications. By describing how this linguistic operation is exploited at different levels in the Xavante language, we hope to offer a contribution to the studies of nominalization, showing by the same token how this pervasiveness suggests different ways of apprehending nominalization.

https://doi.org/10.1075/tsl.124.17est

The paper is organized as follows. Section 1 provides background information on the Xavante language. Section 2 offers a synchronic description of lexical nominalization (2.1) and grammatical nominalization (2.2). Section 3 suggests a diachronic hypothesis whereby verbal constructions involving stative verbs (3.1), negation (3.2), and aorist aspect (3.3) resulted from the reanalysis of nominal (or non-finite verbal) constructions.

1. Background information

This section describes some general features of Xavante grammar relevant for the discussion of nominalization in this language. This information concerns two parts of speech: nouns, treated in Section 1.1, and verbs, described in Section 1.2. Each section describes their morphosyntactic characterization and their internal subclasses. Section 1.3 shows how nouns and verbs share some properties, suggesting a certain parallelism between the two categories.

1.1 Nouns

1.1.1 *Noun classes*

Nouns in Xavante are divided into two subclasses, according to whether they require the overt expression of a certain participant associated with the noun; we will refer to these subclasses by the widely used terms "alienable" and "inalienable".

Inalienable nouns require the expression of an associated participant; either morphologically, by a pronominal prefix (1a), or syntactically, by a noun phrase immediately preceding the inalienable noun (1b).

(1) a. *a-paihi*
2-arm
'your arm'
b. *baʔõtõ paihi*
girl arm
'the girl's arm'

Alienable nouns are morphosyntactically independent: they do not require the expression of an associated participant. In order to associate a participant with an alienable noun, this relation can only be established syntactically, since these nouns do not inflect. As example (2) below shows, the associated participant is introduced by the morpheme *te*, which has the appearance of a genitive marker. In fact, it differs from a genitive marker only with respect to its word class: instead of being a postposition (or, more generally, a grammatical morpheme), it is a noun,

but it is functionally equivalent to a postposition. Following Queixalós (2005), we will refer to this noun as a 'generic relational noun'. Being generic, it co-refers with any alienable noun; being inalienable, it can establish a relation between two nominal entities.[1] In (2), the alienable noun *aibö* 'man' is accompanied by the expression of a first person associated participant introduced syntactically by means of the inalienable noun *-te* 'possession of'.

(2) *Ãhãta ĩĩ-**te** aibö=hã awaʔawi=hã wa=wa-höibaprédub ni*
DEM 1SG-GRN man=DEMC NOW=DEMC 1NOM=1PL.ABS-get.old INDF
wasiré=hã.
together=DEMC
'My husband (lit. 'this man of mine'), now, we are getting old together'.

The next section shows the morphological process functionally equivalent to the use of the generic relational noun *-te* by which alienable nouns may be bound to an associated participant.

1.1.2 *Noun stems*

Noun stems may be simple or complex. For our present purposes, it is only necessary to describe one type of complex noun stem here. These result from a derivational process marked by the prefix *siP-*,[2] as in (3). This morphological process applies to the class of alienable nouns and derives inalienable nouns. In the example below, the nominal base *aibö* 'man' is alienable; as such, it does not require the expression of an associated participant. After derivation with the *siP-* prefix, the resulting noun *siaibö* 'man of' is inalienable and requires the expression, either morphologically or syntactically, of an associated participant: in (3), this participant is morphologically marked by the second person prefix *a-*. We deliberately leave out the gloss of the *siP-* prefix here, in order to comment on it in Section 1.3.3.

(3) *A-**si**-aibö te=za te wa-zawi=zaʔra.*
2-SI-man N1NOM=PROSP [3ERG]AUX 3PL.ABS-like=PL
'Your men should like us.'

1. It would take too long to expose the syntactic properties of the generic relational noun here. These show that *te* is in the process of changing from being the head of a noun phrase (where the alienable noun is an adjunct) to a 'genitive' postposition adjoined to the alienable noun, head of the phrase.

2. The capital *P* symbolizes a bilabial consonant with no audible release which is dropped before a vowel, voiced before a voiced stop and nasalized before a nasal consonant.

1.2 Verbs

1.2.1 *Verb classes*

Two main verb classes may be distinguished according to their valency: intransitives take one argument, which is co-indexed by an absolutive prefix (4); transitives take two arguments, the patient-like one co-indexed by an absolutive prefix, and the agent-like one co-indexed by ergative markers (5).

(4) *Ãme wa=wa-höimana ni.*
here 1NOM=1PL.ABS-live INDF
'We live here.'

(5) *E te=za ĩ-wa-ʔmadö?*
INTER.P N1NOM=PROSP 2ERG-1PL.ABS-see
'Will you see us?'

Some differences exist in this morphological marking between unmarked aspect (4 and 5) and aorist aspect (6 and 7), specifically with transitive verbs. However, these differences do not consist of a morphological split: whichever the aspect, subjects of intransitives are co-indexed by a series of absolutive prefixes, while arguments of a transitive verb are morphologically coded by ergative-absolutive markers. The ergative indexes are reduced to a binary opposition between second person versus all others in the unmarked aspect, whereas they distinguish all persons in the aorist aspect, albeit separately from the verb (on the *te* auxiliary).

(6) *Rowẽ na wa=za oto wa-höimana.*
happiness INSTR 1NOM=PROSP from.this.point 1PL.ABS-live
'We will live happily from now on.'

(7) *Wa=za wa-te ai-ʔmadöʔö.*
1NOM=PROSP 1PL.ERG-AUX 2ABS-see
'We will watch over you.'

Observe that subjects of intransitive and transitive verbs are also coded by a small set of two nominative clitics, contrasting only first person (*wa*) with all other persons (*te*).

Among intransitive verbs, a particular subclass consists of a group of (what we call) *stative* verbs. These are formally distinct from other intransitive verbs in that (i) they cannot predicate independently (they must co-occur with an auxiliary), and (ii) their subject is marked only by the absolutive prefix series, the use of the nominative clitics being ungrammatical with these verbs. Semantically, these verbs express stative predications, as illustrated in example (8), and will be discussed further in Section 3.1.

(8) *E ai-wasutu di?*
INTER.P 2ABS-tire AUX
'Are you tired?'

1.2.2 *Verb stems*

Verb stems may be simple or complex. Only two types of complex stems will be illustrated here, as only these two are relevant for the subsequent discussion. The first type consists of stems derived by the prefix *roP-*. In terms of valency, this morpheme functions as an antipassive: it blocks the absolutive co-indexing of the patient-like argument of transitive verbs. Semantically, it sometimes modifies the base in an unpredictable manner. In example (9a), the patient argument of the transitive verb *zabu(i)* 'visit' is specified by a pronominal prefix, whereas in the antipassive version of the same verb in (9b), the patient is expressed in a postpositional phrase, with no pronominal marking on the verb.

(9) a. *Oto da-zabui õ di.*
from.this.point 3H.ABS-visit NEG PART
'(They) didn't visit him after that.'
b. *Pé a-no zô **rob**-zabu=aba!*
EXH 2-elder.brother for ANTP-visit=COL
'Go visit your elder brother! (lit. 'Go visiting, to your elder brother.')'

We mention this type of complex verb stem because they are found in deverbal nominalizations described in § 2.1.1.

The second type of complex verb stems – of greater interest for our purposes – are characterized by the pair of formatives[3] *aP-* ~ *siP-*. These segments, found in the initial position of certain verb stems, are analyzable at most as fossilized morphemes. Although they may be bound to independently existing verb stems, they occur mostly with roots unable to function as stems. They do not have a semantic or grammatical function; instead, the characteristic of the *aP-* ~ *siP-* formatives is that they alternate according to the aspectual distinction previously mentioned: *aP-* stems are used in the unmarked aspect, as in (10), whereas *siP-* stems are used in the aorist aspect, as in (10).

(10) a. *Wa-ʔrata ma wa=za oto **a**pi=zaʔra ni.*
1PL-grandmother DAT 1NOM=PROSP from.this.point cook=PL INDF
'We will now cook for our grandmother.'

3. Following Bauer (1983: 16–17), we use this term because 'all morphs are formatives, but not all formatives are morphs, so that the use of the term 'formative' can avoid the issue of whether a particular element of a word realizes a morpheme or not', since 'a formative is defined as a distributional segment of a word-form independent of whether or not it is also a morph.'

b. *Ãne za wa-ʔrata ma=hã ʔre wa-ñipi=zaʔra.*
like.this PROSP 1PL-grandmother DAT=DEMC PVB 1PL.ABS-cook=PL
'This is how we will cook for our grandmother.'

1.3 Parallels between nouns and verbs

This section discusses the possible parallels observable between nouns and verbs. These parallels arise when comparing the use of certain clitics and particles (1.3.1), access to predicate function (1.3.2), and valency (1.3.3).

1.3.1 *Clitics and particles*

Some grammatical markers (such as the plural, collective and diminutive morphemes) are used similarly on verbs and nouns.[4] The next example illustrates the use of the collective marker *waʔwa*: in (11a) it is used directly following the inalienable noun *usu* 'group of' to quantify over its associated participant marked by the third person (honorific) prefix *da-*, in (11b) it follows the verb *ñoʔre* 'sing' to quantify over the subject participant.

(11) a. *Da-usu=waʔwa waihuʔu õ di.*
3H-group=COL [3ABS]be.known NEG PART
'I don't know their group (lit. 'their group is not known [to me]').'
b. *Abareʔu te=tãma da-ñoʔre=re=waʔwa.*
group.name N1NOM=3+DAT 3H.ABS-sing=DIM=COL
'The *Abareʔu* are singing for us.'

The case of the particle *mono* is less straightforward, because the use of this morpheme with verbs manifests some semantic extension. Nevertheless, a semantic connexion can arguably be established to show that the same morpheme may head a noun phrase, as in (12a), or a verb phrase, as in (12b): the distributive value of *mono* found in a nominal context may maintain this meaning with verbs, or express a secondary motion ('walking') or aspectual meaning (iterative, progressive or imperfective) because the distribution is spatially conceptualized in the first case, and temporally in the second.

4. We consider these markers to be clitics instead of affixes based on the criterion of 'freedom of host selection' (see Haspelmath & Sims 2010: 198) and therefore write them separately to emphasize their grammatical independence. Other criteria for treating these markers as clitics are freedom of position (for the diminutive) and phonological erosion (for the collective).

(12) a. *Duréi wa-te aihiʔrata mreme=**mono**=hã te*
long.time 1PL-GRN elder word=DISTR=DEMC [1SG.ERG]AUX
a-wapari.
[3ABS]COL-listen
'I have been listening to the words of our elders for a long time.'

b. *Te=mreme=**mono**.*
N1NOM=[3ABS]talk=IMPERF
'He is talking.'

1.3.2 *Predicate function*

Verbs are prototypically used in predicate function; in order to function as arguments or adjuncts, they must be nominalized by the processes described in Section 2. Nouns typically head phrases acting as arguments or complements of a postposition, but are not limited to those functions; noun phrases may also be used as predicates.[5] In such cases, the referent of the subject is included in the class of referents designated by the nominal predicate, as in example (13).

(13) *Õhã warazu.*
3PR non.indian
'He is a "white man".'

This situation, whereby noun phrases can function as predicates without a copula, is reminiscent of omnipredicative languages (Launey 1986), where all categories are prototypically predicates. In these languages, because any expression is fundamentally predicative, a particular morpheme is necessary to indicate that a constituent is a referring expression. According to Queixalós (2006), this 'referrer' can be lost when omnipredicativity is lost, a situation that could have happened in Xavante. Moreover, contrary to typically omnipredicative languages, nominal predicates in Xavante do not establish an existential predication; if the actual state of the language did evolve from an 'omnipredicative' stage, the possibility for nouns to express an existential predication was also lost along the way.

1.3.3 *Valency*

Section 1.1.1 described noun classes in terms of alienable and inalienable nouns. This distinction can be accounted for in terms of valency, as proposed by Queixalós (2005): alienable nouns have a valency of one (they take one argument when they are in predicate function), whereas inalienable nouns have a valency of two (they

5. Changing the temporal value of the sentence does not reveal a copula: for instance, 'he was a white man' would be *Õhã waraza-ʔrata*, where a nominal compound including the verb stem *ʔrata* 'be old' is used to express the past.

take two arguments when they are in predicate function, one external and one internal to the noun phrase they head). This account is not only congruent with the fact that nouns can act as predicates, but also with the fact that a noun may change classes. Recall from Section 1.1.2 that inalienable nouns can be derived from alienable nouns by the prefix *siP-*: in terms of valency, this morpheme increases the valency of alienable nouns, since it allows for the expression of a new nominal complement. Consequently, in order to emphasize the parallelism between nominal and verbal predicates, we propose to call the morpheme *siP-* a 'nominal applicative', adopting the term coined by Ribeiro (2002).

Recall from § 1.1.1 that establishing a relation between an alienable noun and an associated participant is also possible syntactically, by means of the inalienable noun *-te* 'thing of, possession of'. Although the generic relational noun *te* is not directly involved in nominalization processes, it will be an important argument for our hypothesis concerning grammatical nominalizations, presented in Section 3. This section will also show why it is possible to hypothesise that the nominal applicative *siP-* has become part of the stem of a subclass of intransitive verbs.

These parallels should not obscure the fact that nouns and verbs are distinct classes of the language, as shown by their morphosyntactic properties. Among these, as the next section shows, is the possibility of nominalizing a verbal element.

2. Nominalizations

Nominalizations in Xavante can apply to the lexical level, as shown in 2.1, or to the clause level, as discussed in 2.2.

2.1 Lexical nominalizations

This section presents the three types of lexical nominalizations found in the language. All are derivations resulting in deverbal nouns: these refer to an agent participant (2.1.1), a non-agent participant (2.1.2), or to the action/state denoted by the verb they derive from (2.1.3).

2.1.1 *'Agent' deverbal nouns*

'Agent' deverbal nouns result from a productive derivational process marked by the suffix *-ʔwa* applied to transitive and intransitive verb stems. The resulting noun belongs to the class of inalienable nouns: it must be preceded by a nominal complement or be inflected for it, except when the slot for the absolutive argument of the verbal stem is occupied by the morpheme *roP-* (see § 1.2.2). Some examples are given in the second column of (14a), where the first column presents verb

stems from which the corresponding agent nouns are derived. The first hyphen shows that, when inflected, the verb stem is preceded by an absolutive prefix, and the noun by a person prefix or a nominal complement.

(14) a.

-madöʔö	'watch'	*-madöʔö-**ʔwa***	'person in charge, chief'
-poto	'create'	*-poto-**ʔwa***	'creator'
romñoré	'study'	*romñoré-**ʔwa***	'student'
-sipi	'(to) cook'	*-sipi-**ʔwa***	'(a) cook'
-wede	'medicate, treat'	*-wede-**ʔwa***	'doctor'

The examples in (14b), taken from a text, illustrate the productivity of agent noun derivation. The short extract given here shows how the speaker produced three synonymous agent nouns (one after the other) in order to express his thought. The first and third of these nouns require the expression of an argument; in the first case, this argument is expressed morphologically by the first person prefix *wa-*, in the last, it is expressed syntactically by the noun phrase *waʔrãi* 'our heads'.[6]

(14) b. *wa-ʔrãʔõtõ-**ʔwa**, wa-ma rowaihuʔu-**ʔwa**, wa-ʔrãi waihuʔu-**ʔwa***
1PL-unite-NMZR 1PL-DAT teach-NMZR 1PL-head teach-NMZR
'those who united us (lit. 'our uniters'), our teachers (lit. 'to us teachers'), our mentors (lit. 'teachers of our heads')'

2.1.2 *'Non-agent' deverbal nouns*

Similarly, the suffix *-zé* derives nouns from transitive and intransitive verbs. The resulting noun also requires a complement (expressed either morphologically by a prefix, or syntactically by a noun phrase), except when the stem contains the *roP-* prefix (see § 1.2.2). The semantics of nominalizations formed with this suffix, however, is less specific than the result of the *-ʔwa* nominalization. As the examples in (15a) show, the nouns derived by the *-zé* suffix can refer either to an action, a location, or even an instrument. To illustrate each of these meanings, a contextual occurrence of a deverbal noun expressing an action is provided in (15b), a location in (15c), and an instrument in (15d).

(15) a.

-höimana	'live, exist'	*-höimana-**zé***	'life, existence'
rowasuʔu	'tell'	*rowasuʔu-**zé***	'communication'
romhuri	'work'	*romhuri-**zé***	'work/working room'
-sipi	'cook'	*-sipi-**zé***	'cooking/kitchen'
-wede	'medicate, treat'	*-wede-**zé***	'medication, medicine'

6. These might look like clause nominalizations, but Section 2.2.1.1 shows how clause nominalizations have other properties not observed here (namely, a particular nominalizer *ĩ-* and ergative marking of the agent-like participant).

b. *Pizawaipo zô si, da-siʔwapé-**zé**.*
trophy for only 3GNC-compete-NMZR
'The competition (is) only for the trophy.'

c. *Romhuri-**zéb** ʔre te=za taré ãma ro-waza.*
work-NMZR INES N1.NOM=PROSP for.nothing [3]PVB ANTP-heap
'They are going to heap them up for nothing in the working room.'

d. *Da-wede-**zé**=hã e momo te ʔsawari=zaʔra?*
3GNC-treat-NMZR=DEMC INTER.P place N1.NOM [3ABS]toss=PL
'Where will they toss the medicine?'

In both *-zé* and *-ʔwa* nominalizations, when the noun is derived from a transitive verb, the complement it requires corresponds to the patient(-like) argument of the transitive verb.

2.1.3 *'Action' deverbal nouns*

Finally, a third nominalizing process, involving the prefix *simi- ~ ñimi-*, derives inalienable nouns from transitive verb stems. Unlike the preceding derivations, the resulting deverbal nouns can all inflect for an argument (independently of the shape of the stem) and the argument they require refers to the agent-like participant of the corresponding transitive verb. They can refer to the action/state denoted by the verb stem or to its result. Some examples are given in (16), in citation form in (a), contextualized in (b).

(16) a. *-zawi* 'love, like' *-**ñimi**-zawi* 'love, friendship'
-poto 'create' *-**ñimi**-poto* 'creation'
rowairébé 'decide' *-**ñimi**-rowairébé* 'decision'
rowaihuʔu 'teach' *-**ñimi**-rowaihuʔu* 'teaching'

b. *Ãhãna=hã da-**ñimi**-roti na=hã*
nowadays=DEMC 3GNC-NMZR-counsel INSTR=DEMC
wa-zapaʔa=zaʔra=mono õ di.
1ABS-consider=PL=ITR NEG PART
'Nowadays we don't take into account their advice.'

Interestingly, the last two derivations are not mutually exclusive and may result in some redundancy, specifically, when the *-zé* nominalization results in a noun denoting an activity. As Example (17) below illustrates, the deverbal noun *-ñimi-romhuri* 'work of' in (a) – resulting from the prefixal nominalization – can serve as a base for the derivation with the *-zé* suffix, illustrated in (b). These examples were taken from the same text, depicting the various activities Xavante women consider part of their work. Despite their morphological difference, both nouns *-ñimiromhuri* and *-ñimiromhurizé* denote the activity of 'working' or the result of this activity.

(17) a. *Ãne uburé-pese wa-ñimi-romhuri=hã, piʔõ.*
so all-completely 1PL-NMZR-work=DEMC woman
'So all this is our work, the women.'
b. *Uburé wa-ñimi-romhuri-zé=hã.*
all 1PL-NMZR-work-NMZR=DEMC
'All this is our work.'

Most frequently, however, the *-zé* marker serves a disambiguating function, as does the *zé* marker used in grammatical nominalizations (see § 2.2.1.2).

2.2 Grammatical nominalization

This section looks at nominalizations from the level of the sentence. Section 2.2.1 deals with internal and external properties of clause nominalization, while Section 2.2.2 depicts some aspects of clause nominalization and discourse.

2.2.1 *Clause nominalization*

Clause nominalization in Xavante is a productive process that allows a clause to function as a nominal constituent. It is marked by the verbal prefix *ĩ-*.[7] Any verb can take this morpheme and head a constituent functionally equivalent to a noun phrase in the sense that this constituent has access to the syntactic functions that characterize noun phrases. Although such a constituent has the external properties of a noun phrase, its internal properties are those of a clause, albeit some restrictions apply in comparison to independent clauses, because it is headed by a non-finite verb. The next two sections review the internal (clausal) properties and external (nominal) properties of this type of nominalization.

7. As the editors kindly pointed out to us, the existence of a formal difference between grammatical and lexical nominalization is uncommon in the languages surveyed in this volume. However, morphemes similar to lexical nominalizers do show up in grammatical nominalizations (see § 2.2.1.2), although in a disambiguating function. Moreover, this difference is not so surprising if one follows Cristofaro's (this volume) view that a better understanding of nominalization can be gained by adopting a diachronic approach. By examining the structure of Xavante, it is possible to hypothesize that the markers for grammatical and lexical nominalizations have two different origins: the latter were probably nouns at an earlier stage of the language (see § 2.2.1.2 for more details), whereas the former could have developed from a demonstrative pronoun. Two properties of the *ĩ-* marker suggest this possible origin, one being its position before the nominalized constituent, the other being its form. Although possible cognates for a demonstrative pronoun are not found in other Jê languages, many have a third person possessive prefix *i-*, which could be the reflex of a former demonstrative pronoun. (The grammaticalization of demonstrative pronouns into possessive markers is well attested according to Diessel (1999: 128).)

2.2.1.1 *Internal properties.* The constituent signaled by brackets in (18) is an example of a nominalized clause. It is headed by the verb *sebre* 'cook', marked by the inflectional nominalizing prefix *ĩ-*. Because the verb is transitive, its subject is morphologically expressed by the inflected auxiliary *-te*, used with transitive verbs in all dependent clauses, with negation and in independent clauses to express aorist aspect. Its object is syntactically expressed by the noun phrase *bö* 'annatto'. (Further examples will show that the object argument can also be marked morphologically on the verb by an absolutive prefix.) In this example, the nominalized clause is the complement of the postposition *zô* 'for, after'.

(18) *[Bö wa-te* ***ĩ-sebre*** *]$_{NMLZ}$ zô te=wei mo.*
annatto 1PL.ERG-AUX NMZR-[3ABS]cook for N1NOM=hither [3ABS]go
'He came after the annatto we cooked.'

The independent counterpart of this nominalized clause is given in (19) for comparison. Recall that, in an independent affirmative clause, a verb can express non-aorist or aorist aspect, as shown in (19a) and (19b) respectively. In an independent negative clause, illustrated in (19c), the verbal morphology is the same as that used in the aorist construction.

(19) a. *Bö wa=sebre ni.*
annatto 1NOM=[3ABS]cook INDF
'We cooked the annatto (now).'
b. *Bö wa=wa-te sebre.*
annatto 1NOM=1PL.ERG-AUX [3ABS]cook
'We cooked the annatto (always).'
c. *Bö wa-te sebre õ di.*
annatto 1PL.ERG-AUX [3ABS]cook NEG PART
'We did not cook the annatto.'

As the following tests will show, nominalized clauses share the same morphosyntactic properties as negative propositions: the verb is non-finite. The important difference between the two structures is that, while negative independent clauses are headed by the morpheme *di*, nominalized clauses are headed by the non-finite verb marked by the prefix *ĩ-*. Observe in (20) that the nominalized clause illustrated above is incompatible with the first person nominative clitic *wa* used in independent clauses (with both aspects) and with the marker *ni*, also used in independent clauses (but only in the non-aorist aspect).

(20) a. * *bö* ***wa****=wa-te ĩ-sebre zô te=wei mo*
annatto 1NOM=1PL.ERG-AUX NMZR[3ABS]cook for N1NOM=hither [3ABS]go

b. **bö wa-te ĩ-sebre **ni** zô te=wei*
annatto 1PL.ERG-AUX NMZR[3ABS]cook INDF for N1NOM=hither
mo
[3ABS]go

Although the verb marked with the (grammatically) nominalizing prefix cannot be accompanied by all the markers typically found in independent clauses, this does not entail that it is lexically nominalized and that it should be analyzed as the head of a noun phrase; it merely shows that the verb is non-finite. The internal structure of the constituent headed by such a verb form is that of a clause, not of a noun phrase: the 'patientive' argument is not expressed as would be expressed a nominal argument – by a prefix (*ĩ-* or *ti-*) or by a noun phrase immediately preceding the head noun – and the morpheme *-te* is not the generic relational noun.

Although the nominalized clause suffers some restrictions because of its non-finiteness, it can nevertheless be accompanied by morphemes found in independent clauses, such as the habitual preverb[8] *ʔre* or the imperfective particle *mono*, as illustrated in (21). Compare the occurrence of these morphemes inside the nominalized clause in (21a) with their occurrence in independent clauses in (21b) and (21c). Example (21a) also shows that the object of the verb is indexed by an absolutive prefix, in this case, the first person plural *wa-*.

(21) a. *Õhõ tô, Norõsuʔrã **ʔre** [**ʔre***
DEM indeed Couto.Magalhães INESS HAB
*ĩ-**wa**-rob=zaʔra=**mono**]*$_{NMLZ}$
NMZR-1PL.ABS-live=PL=IMPERF
'(It is) in that one indeed, in Couto Magalhães, (where) we lived.'

b. *Te=mreme=**mono**.*
N1NOM=[3ABS]talk=IMPERF
'He is talking.'

c. *E taré we **ʔre** ai-mo?*
INTER.P gratuitously hither HAB 2ABS-go
'You come here just like that?'

The morphemes provided in the last example to illustrate the internal properties of nominalized clauses are internal to the verb phrase. The following data show that modal particles, such as the irrealis *éré* and the prospective *za*, can also occur

8. We use the term 'preverb' following Craig & Hale (1988) to refer to grammatical morphemes occupying a preverbal position and clearly related to postpositions, both formally and semantically. Typically, these markers serve the purpose of indicating a relation between the verb and a noun phrase, but they can also acquire an aspectual function.

inside nominalized clauses. These particles can occur inside or outside the verb phrase in independent clauses but are not found inside a noun phrase.

(22) a. *[Ĩsépu wa-te* ***éré*** *ĩ-a-zéptö]*$_{NMLZ}$ *ʔre*
sick.one 1PL.ERG-AUX IRR NMZR[3ABS]COL-cure HAB
nomro=mono õ di.
[3ABS]stay=IMPERF NEG PART
'The sick we were going to cure did not stay.'

b. *Te=éré ti-ña [****za*** *te ĩ-wazari]*$_{NMLZ}$
N1NOM=IRR 3ABS-say PROSP [3SG.ERG]AUX NMZR[3ABS]mix
na=hã.
INSTR=DEMC
'He said he would mix it.'

c. *Ma ʔmai-wĩrĩ ni tô* ***éré****.*
PERF [3ABS]SG-kill INDF indeed IRR
'Someone nearly killed him.'

d. *Höiwahö wa=****za*** *oto wa-waʔutu=zaʔra=ni aʔö.*
afternoon 1NOM=PROSP from.this.point 1PL.ABS-rest=PL=INDF little
'In the afternoon we will rest a little.'

There are no restrictions concerning the syntactic expression of arguments inside the nominalized clause. The subject and object of the verb with the *ĩ-* prefix can either be omitted or occur inside the nominalized clause, just as they can occur or not in an independent clause. In (18), the object of the non-finite verb heading the nominalized clause occurs inside it. In (22b), however, there is no syntactic expression of the 'patientive' argument of the verb heading the nominalized clause.

The following examples show that the subject can be syntactically expressed or not inside the nominalized clause. In (23a), the third person pronoun *õ* is the subject of the verb *wairébé* 'go out' and occurs inside the nominalized clause. In (23b), the participant *tebe* 'fish', introduced in the first sentence, is not repeated inside the nominalized clause as the subject of the verb *döʔö* 'die'.

(23) a. *Ãhãna=hã maparané aʔamo oto [õ=norĩ=hã*
today=DEMC two month from.this.point 3PR=COLL=DEMC
ĩ*-wairébé]*$_{NMLZ}$ *na=hã.*
NMZR[3ABS]go.out INSTR=DEMC
'Today it has been two months since they have moved.'

b. *Ãma ʔre awẽ=mono wamhã, tebe te=za döʔö.*
[3]PVB HAB [3ABS]dawn=IMPERF SUB fish N1NOM=PROSP [3ABS]die
'When it dawns, the fish are dead.'
*Mararé za piʔõ sô sisaʔre, [****ĩ****-döʔö]*$_{NMLZ}$ *zô.*
early PROSP woman [3]for [3ABS]rush NMZR-[3ABS]die for
'The women will go for them early, for the ones that are dead.'

Adjuncts can also occur inside a nominalized clause, as the following example illustrates: the nominalized clause, shown in brackets, contains the adjunct *höiwa u* 'to the sky'.

(24) *Taha Parinaiʔa ãté [höiwa u ĩ-da-watobro]*$_{NMLZ}$
AN.PR PROP.N DUB sky DIST.LOC NMZR-3HON.ABS-go.out
'That Parinai'a is, it seems, the one who went to the sky.'

2.2.1.2 *External properties.* Nominalized clauses have access to all the syntactic functions that characterize a noun phrase in Xavante: they can act as subject, object, complement of a postposition or predicate. The following examples illustrate a nominalized clause, signaled by brackets,[9] in each of these nominal functions. In (25a), the nominalized clause is the subject of the clause; in (25b), it is the object; in (25c), it is the predicate. A nominalized clause functioning as complement of a postposition is given in (25d).

(25) a. *[Bö wa-te wasiwi ĩ-sebre]*$_{NMLZ}$*=hã wẽ na*
annatto 1PL.ERG-AUX 1COLL.ERG NMZR-[3ABS]cook=DEMC good INSTR
ma te=watobro.
PFC N1NOM=[3ABS]come.out
'The annatto we cooked is good (lit. 'came out well').'

b. *[Mimi da-te ĩ-ubumro]*$_{NMLZ}$ *wa=wa-te*
firewood 3GNC.ERG-AUX NMZR-[3ABS]gather 1NOM=1PL.ERG-AUX
waibu.
[3ABS]take
'We would take the firewood they had gathered.'

c. *Ãhã wede-ñorõ=hã [marã ʔrep si ʔre*
DEM plant-fiber=DEMC bush INESS only HAB
ĩ-höimana=zaʔra=mono]$_{NMLZ}$
NMZR[3ABS]exist=PL=IMPERF
'These fibers are the ones found only in the bush.'

The following example, which shows a nominalized clause functioning as complement of the postposition *da* 'to', illustrates the use of a nominalized clause parallel to the use of a noun phrase. The same postposition, required by the complex

9. We have no formal evidence for arguing weather the nominalized clauses include or not participants expressed as noun phrases. An interpretation of the nouns *bö* 'annatto' in (25a) or *mimi* 'firewood' in (25b) as being external to the nominalized clauses (which would serve a relativization function) runs the risk of being ethnocentric. Considering that the language shows no signs of adjectives – see Estevam (2015) –, and consequently, no use of adjectival phrases modifying nouns, we chose to include the noun phrases inside the brackets: indeed, we see no reason to believe that nominalized clauses could have an attributive function.

predicate *uware di* 'is weak' to mark its complement, is used after the noun phrase *romhuri* 'work' and after two nominalized clauses.[10]

(25) d. *Duré ihöibaté=hã uware di romhuri* ***da****=hã, uware*
also young.man=DEMC weak AUX work to=DEMC weak
di [te ĩ-rosaʔrata=zaʔra=mono]$_{NMLZ}$ ***da****=hã, [sisõwa*
AUX [3ERG]AUX NMZR-think=PL=IMPERF to=DEMC future
rowaʔõno=hã te ĩ-ʔmanharĩ=zaʔra=mono]$_{NMLZ}$ ***da****=hã.*
plan=DEMC [3ERG]AUX NMZR[3ABS]make=PL=IMPERF TO=DEMC
'Also, young men are (too) weak for work, weak for thinking, to make plans for the future.'

In terms of reference, a nominalized clause can refer to either participant encoded as a core argument of the verb, as well as a peripheral argument, a circumstance, and even to the event or state denoted by the verb. Examples (23b) and (24) above illustrate nominalized clauses referring to the argument of the intransitive verb heading the clause. In examples (25a) and (25b) above, the nominalized clauses refer to the patient argument of the transitive verb heading the nominalized clause. An instance of a nominalized clause referring to the agent argument of a transitive verb is given below in (26). The nominalized clause in (27) refers to the participant who experiences the state expressed by the verb heading the nominalized clause.

(26) *Wahã, ĩĩ-prédub õ ñerẽ, ʔwatébrémi wasihö=zaʔra=mono wamhã,*
1PR 1SG.ABS-grow NEG SUB boy [3ABS]fight=PL=IMPERF SUB
[te da-wi ĩ-romhöri-uʔösi=mono]$_{NMLZ}$
[1SG.ERG]AUX 3GNC-MAL NMZR-separate-PERM=IMPERF
'As for me, even though I wasn't big, when boys would fight, I would always be the one separating them.'

(27) *Taha wa ĩ-ʔra=hã [asa ti-mama te*
AN.PR because 3-son=DEMC in.return 3CORF-father [3SG.ERG]AUX
ĩ-sawi-pe]$_{NMLZ}$
NMZR[3ABS]love-INTSF
'That is why the son (is the one who) loves back his father very much.'

A nominalized clause referring to a circumstance was seen in (21a), a case of location nominalization. Likewise, a nominalized clause may refer to the action denoted by the verb, without any overt indication to mark such an interpretation.

10. There is also a subordinator *da*, which marks subordinate clauses, homonymous to the postposition in question, and a verb *romhuri* 'to work', identical in shape to the noun in question. However, the constituent *romhuri da* is not a subordinate clause because the verb *romhuri* 'to work' is transitive: in a subordinate clause, it would be in the aorist form *te romhuri* (inflected for third person). See, for instance, Example (28).

This is illustrated below in example (28a). The data given in (28b) shows that the same nominalization can also refer to the result of the action denoted by the verb.

(28) a. *Warazu=hã [uburé ti-ñimi-romhuri-zéb na=hã*
white.man=DEMC all 3CORF-NMZR-work-NMZR INSTR=DEMC
te ĩ-romhuri=zaʔra=mono]$_{NMLZ}$=hã wapu di.
[3SG.ERG]AUX NMZR-work=PL=IMPERF=DEMC [3ABS]light/easy AUX
'The white men, all the work they do with their machine is easy.'

b. *Ĩhöʔa ñib-ʔri=hã wa-ñimi-romhuri, [wa-te*
priest N.APL-house=DEMC 1PL-NMZR-work 1PL.ERG-AUX
ĩ-romhuri=zaʔra=mono].
NMZR-work=PL=IMPERF
'The house of the priests is our work, the work we have done.'

As suggested by the preceding examples, nominalized clauses can assume a variety of syntactic and semantic functions, with no specific marking other than the nominalizing prefix *ĩ-*. The interpretation of nominalized clauses is thus strongly context-dependent. As Shibatani (this volume) points out, grammatical nominalizations evoke a variety of non-uniform concepts and, consequently, are used in situations where the context points to the most relevant denotation. Of course, cases of ambiguity may arise. Two morphemes may be used when the nominalized clause is judged potentially ambiguous: the first, *ʔwa*, explicitly signals that the nominal constituent refers to an agent, as example (29) illustrates; the second, *zé*, marks the nominalization as referring to a location (30a) or an instrument (30b).

(29) *Tamé wa-ñoʔa=hã [õ nasi ĩĩ-ma ĩ-ñarĩ] ʔwa ma*
then 1PL-front=DEMC 3PR DUR 1SG-DAT NMZR[3ABS]tell NMZR PERF
waze.
[3ABS]send.away
'Then, they fired in front of us the one who was telling (that) to me.'

(30) a. *[Ã höiwi ʔre ĩ-sina=mono] zé.*
DEM airplane HAB NMZR[3ABS]land=IMPERF NMZR
'(It is) a place for this airplane to land.'

b. *[Piʔõ poʔreõʔwa duré da-te ãma ʔre*
woman disobedient also 3H.ERG-AUX [3]PVB HAB
ĩ-wasisi=zaʔra=mono] zé.
NMZR[3ABS]tie=PL=IMPERF NMZR
'(It is) also for tying disobedient women.'

Although the morphemes *ʔwa* and *zé* are glossed 'nominalizers' for convenience, we do not analyze them as nominalizers *per se*, because the nominalizing function is assumed by the *ĩ-* prefix. Rather, they specify the exact interpretation the

nominalized clause should have. In this respect, they differ from their homonyms, the *-ʔwa* and *-zé* suffixes used to produce lexical nominalizations (§ 1.1). Clearly, these functional morphemes must have developed from a common origin, possibly inalienable nouns (requiring the formal expression of a core argument), absent from the language today.[11] Interestingly, the result of this grammaticalization reflects the fact that nominalizations can operate both at the lexical and the grammatical level. However, the nominalizing function of these grammatical morphemes is only assumed by the suffixes, used to create new words denoting entities that are new and relevant enough to be inserted in the lexicon. The grammatical words *ʔwa* and *zé* did not assume a nominalizing function presumably because clause-level nominalizations arise to denote concepts necessitated by a given context and useful only in such a context. They only serve a disambiguating purpose because nominalized clauses are marked by a specific morpheme and their interpretation can be contextually inferred.

The importance of context for understanding the use of clause nominalization is even clearer when nominalization serves discursive purposes. These are presented in the following section.

2.2.2 *Nominalization and discourse*

Nominalization can serve the purpose of focussing or de-focussing attention on certain aspects of information. The first case occurs with insubordination, depicted in Section 2.2.2.1. The second case can account for an apparent case of refinitization, explained in Section 2.2.2.2.

2.2.2.1 *Insubordination.* The previous section showed that nominalized clauses share an important number of external properties with noun phrases. However, there is one property of grammatical nominalizations that distinguishes them from noun phrases: insubordination. Unlike noun phrases, nominalized clauses can function as independent clauses. Should nominalized clauses in such a use be interpreted as predicates functioning with an elided subject (whose referent is recoverable from context)? This would be consistent with the fact that noun phrases are usually dropped once their referent is active in the minds of the speech act participants. However, the idea that nominalized clauses are used in predicate function with an elided subject cannot be supported for a number of reasons.

11. In Mẽbengokre, a Northern Jê language, the corresponding morphemes *dʒʌ* and *dʒwɔ̃j* are argued by Salanova (2008: 115) to be 'semantically impoverished versions of the nouns 'container' and 'owner', respectively'.

Firstly, although it might be possible to defend such a hypothesis when the covert subject would refer to a participant, it would be impossible to sustain it when the nominalized clause refers to a situation. Such cases arise easily, as in example (31).

(31) *Taha wam=hã [ĩ-da-wasutu-uptabi]*$_{NMLZ}$ *duré [ĩ-ʔrubu]*$_{NMLZ}$
AN.PR during=DEMC NMZR-3GNC-tire-be.real also NMZR-be.thirsty
'At that moment people are very tired and thirsty.'

In other words, nominalized clauses can be used independently to express a thetic judgement[12] (Kuroda 1972), or to mark sentence focus (Lambrecht 1994: 222): in (31), the nominalized clause is linked to the preceding sentences by a discourse connector, but there is no topic and the whole structure conveys new information.

Secondly, nominalized clauses used independently very often have a pragmatic effect of emphasis or confirmation. This effect is enhanced by intensifiers such as *wẽ* 'well' or *uptabi* 'really' that frequently occur inside the nominalized clause. (The latter occurs in examples (31) and (32b).) In this respect, nominalized clauses are frequently used as an answer to a question, to confirm the assertion emphatically. The following example illustrates a typical situation where speaker A asks speaker B to assert that an entity X has a property Y. Speaker A forms a question with the *di* auxiliary (32a), while speaker B answers using a nominalized clause (32b). example (33) shows that the use of intensifying morphemes, although frequent, is not mandatory.

(32) a. *E wẽ di?*
INTR.P [3ABS]be.good/pretty AUX
'Is it good/pretty?'
b. *[Ĩ-wẽ-uptabi]*$_{NMLZ}$ *!*
NMZR-[3ABS]be.good/pretty-be.real
'It's really good/pretty!'

(33) *E hö di? [Ĩ-hö]*$_{NMLZ}$ *!*
INTR.P [3ABS]be.cold AUX NMZR-[3ABS]be.cold
'Is it cold? It's cold, all right!'

12. To make explicit the relation between nominalization and thetic judgement in Xavante is an objective we would like to pursue in another article. We will only offer here a few references where this relation has been mentioned, but a full survey of the literature would surely reveal many examples: for Austronesian languages and Modern Arabic, see Sasse (1987); for Yukagir, see Maslova (1997); for examples concerning Amerindian languages specifically, see Carlin (2011) for the Cariban language Trio and Queixalós (2016) for Sikuani.

In such cases, the nominalized clause cannot be interpreted as the complement of a covert assertive predicate: sentences such as '(I'll say) that it is really good!', '(you bet) that it is cold!' would require a nominalized clause marked by a postposition (or, alternatively, a subordinate clause).[13] Because the insubordinate use of a nominalized clause marks the assertion of a fact by the speaker, it can carry strong emotional marks.

As Givón (2016: 284) points out, insubordination and re-finitization can be argued to be two distinct processes. Indeed, the case of insubordination we have just seen involves nominalized clauses used as independent clauses without any sign of re-finitization: although they function as independent clauses, they retain all their properties of grammatical nominalization and do not allow the use of markers found in finite clauses, such as the nominative clitics *wa* (for first person) and *te* (for all other persons). In contrast, the next section presents a possible case of insubordination and re-finitization.

2.2.2.2 *Re-finitization.* Inflectional morphology of transitive verbs shows a resemblance with nominalized clauses which suggests that one of its values is the result of insubordination, namely, when the agent is second person. The pair of examples in (34) show that there are no overt markers of a second person agent in nominalized clauses, compared to other persons (marked by the *-te* morpheme). In example (35), the same verb *rẽme* 'to abandon' is inflected for a second person agent in an independent clause: the agent is marked by the prefix *ĩ-*.[14] In other words, finite transitive clauses with a second person agent are probably the result of reanalyzed nominalized clauses. The reason for using nominalized clauses to express an event with a second person agent could have been pragmatic: factors such as politeness could prevent speakers from explicitly saying that the hearer is the agent responsible for the event described. A sentence such as (35) could have meant 'the house is the one abandoned (by you)' and, consequently, the nominalizer *ĩ-* was reanalyzed as a second person agent prefix.

(34) a. *ʔri te ĩ-rẽme te=za.*
house [1SG.ERG]AUX NMZR-[3ABS]abandon N1.NOM=[3ABS]stand
'The house I abandoned is standing (over there).'

13. The only verb of speech that takes a direct object (animate or inanimate) is *wasuʔu* 'tell'; the possibility of using this particular verb in a main clause with a nominalized clause as object remains to be verified with speakers. Nevertheless, because nominalized clauses never occur in our entire corpus as complements of an overt verb of speech, we are inclined to believe that insubordination has reached the stage of 'constructionalization' proposed by Evans (2007: 374).

14. In a number of verbs, a second person agent is also marked by the apocope of the last syllable of the verb stem in final position.

b. *ʔri ĩ-rẽme te=za.*
house NMNZ-[3ABS]abandon N1.NOM=[3ABS]stand
'The house you abandoned is standing (over there).'

(35) *ʔri ĩ-rẽme.*
house 2ERG-[3ABS]abandon
'You abandoned the house.'

This hypothetical insubordination is congruent with the particular status of the second person prefix *ĩ-*: it is the only overt ergative person marker prefixed on transitive verb stems. (Other persons acting as the agent participant of transitive verbs are marked elsewhere in the verb phrase.) If our hypothesis is correct, this shift of grammatical nominalization towards independent clause resulted in complete re-finitization, as the type of clause illustrated in (35) can be accompanied by the full array of markers found in finite clauses (such as, for instance, the nominative clitic *te* used for non first person subject).

This case of insubordination and re-finitization is probably not the only process in Xavante whereby a non-finite dependent clause has been reanalyzed into a finite independent clause, as we will try to show in the next section. Indeed, after this overview of nominalization in Xavante from a synchronic perspective, we would like to propose in the following section a diachronic hypothesis whereby nominal structures were reanalysed into verbal structures, thereby explaining some of the nominal properties of certain verbal configurations.

3. Diachronic hypothesis

The general idea that a class of finite independent clauses may have evolved from nominalized clauses is not new in the realm of Jê linguistics: the case has been examined and argued for by Alves (2010) to explain alignment phenomena in Timbira. Alves reconstructs a nominalization strategy used for subordinate clauses in Proto-Northern Jê, reanalyzed in northern Jê languages as main finite clauses: see, for instance, Gildea & Alves (to appear) and references therein.

In the same fashion, three different structures involving verbal predicates in Xavante independent clauses can be argued to have originated from a nominal predication. The first one concerns stative verbs (3.1); the second involves negative polarity (3.2); the third is linked to aspectual determination (3.3).[15]

15. Because the evolution of this putative nominalized structure had different results in the various Jê languages, we were only able to find one case similar to our hypothesis for Xavante in related languages: the one concerning negative clauses. According to Alves (2010: 469–470),

3.1 Stative verbs

The structure of predications involving stative verbs suggests that these are the result of nominal predications having diachronically evolved into verbal finite clauses. The initial stage of this evolution would have been an existential predication, marked by the copula *di*. This morpheme is analyzed as an existential copula by McLeod & Mitchell (2003: 74), who worked with the Xavante in the Culuene area in the 1970s and collected data such as (36).

(36) *Aibö* ***di***.
man EXIS.COP
'There are men.'

Today, in the region of São Marcos, where we collected our data, the preceding sentence is not grammatical. (More generally, the morpheme *di* cannot follow a noun phrase.) The equivalent meaning of (36) would be expressed as in (37). In other words, the former existential copula *di* has been replaced by the morpheme *iré*.

(37) *Aibö iré=hã*.
man EXIS.COP=DEMC
'There are men.'

According to the examples in Sousa Filho (2007), the predicative morpheme *-di ~ -ti ~ -ki* in Xerente (the sister language of Xavante in the central group of Jê languages) can be combined with a noun to express an existential predication. Two of the examples provided by Sousa Filho (2007: 215–216) are reproduced below: (38a) illustrates the existential predication with the animate noun *wapsã* 'dog', (38b), with the inanimate noun *kuba* 'canoe'. In both cases, the noun is marked by the predicative suffixe *-di*.

(38) a. *Tonmẽ wapsã-di*.
here dog-PRED
'There is / are (a) dog(s) here.'
b. *Kɨ-mba kuba-di*.
river-INES canoe-PRED
'There is / are (a) canoe(s) in the river.'

---

negative clauses in Proto-Northern Jê were formed from a nominalized clause followed by a negative existential verb, of which the nominalized clause is the argument. The verb in the originally nominalized clause has been reanalyzed in Northern Jê languages as the main verb, and the negative existential verb has become a negation marker. No equivalent analysis is proposed for stative verbs nor aorist aspect, but it is assumed for finite clauses expressing 'recent past' in Timbira.

In the Xavante dialect we examined, although the *di* marker is not an existential copula, it is used in two other contexts. One reveals the primary function of the *di* morpheme: as examples (39b) and (39c) illustrate, it is an auxiliary for verbs that can receive a stative aspectual determination, thus referring to a quality or to a non-intrinsic property of the subject. Assuming that *di* was an existential copula that could be preceded by a noun phrase, it is possible that, in a previous state of the language, nouns were used in existential predications in order to express states, qualities, or non-intrinsic properties of nominal referents.[16]

The examples below show another sign pointing in this direction when we turn to the shape of verb stems. The verbs employed in (39) are representative of two subclasses of intransitive verbs: *tete* 'to be hard/unripe' in (39c) is representative of intransitive verbs which only have one stem, whereas another subclass behaves like the intransitive verb meaning 'to be ripe', which has the two stems *apo* ~ *sipo* shown in (39a) and (39b) respectively. When the latter class of verbs is combined with the auxiliary *di*, the stem used is the one beginning with the *siP-* formative. Again, considering that *di* could have been a copula allowing a noun phrase to function in an existential predication, the use of the stem with the initial *siP-* formative suggests that this verb could have originally been an inalienable noun derived by the prefix *siP-* from an alienable noun. In other words, these structures suggest that existential predications involving nouns – which left traces in the resulting verbal morphology – were reanalyzed into verbal predications: 'it is ripe' could literally have been 'there is ripeness of (a fruit)'.

(39) a. *Uwaiʔrewawẽ e ma tô* ***apo?***
Avocado INTER.P PERF FCT ripen
'Has the avocado ripened?'

b. *E* ***sipo*** *di?*
INTER.P [3ABS]ripen AUX
'Is it ripe?'

c. *Zahadu tete di.*
still [3ABS]hard AUX
'It is still unripe.'

Recall from Section 1.2.2 that some verbs present a pair of complex stems: one beginning with the formative *aP-*, the other with *siP-*. It could be that the two suppletive forms of the same verbal base result from a reanalysis of two distinct

16. Incidentally, there are no adjectives in Xavante. Typological studies show that languages with little or no adjectives categorize 'adjectival notions' either nominally or verbally. If our hypothesis is correct, it suggests that languages can opt for both categories not only synchronically but also diachronically. It might be one of the factors motivating reanalysis of nominal (or non-finite verbal) structures into verbal (or finite) structures.

derivational processes. While the first can be assimilated to an anticausative prefix, blocking the expression of an agentive argument from a verbal base,[17] the second could have been the nominal applicative applied to a nominal base (homonymous with the verbal stem).

With these observations in mind, we can now turn to the second structure that possibly evolved from an anterior nominal configuration, which also involves the *di* morpheme: negative sentences.

3.2 Negation

In negated sentences, the use of the *di* marker is compulsory;[18] this is illustrated by the examples (40) – with an intransitive verb – and (41) – with a transitive verb. The data in (40) is also representative of the fact that the subclass of intransitive verbs presented earlier, i.e. with two stems – showing the initial *aP-* ~ *siP-* alternation – selects the stem beginning with the *siP-* formative in negated clauses. In other words, negative predicates have properties similar to those found in stative predicates, arguably because both inherited traits from a common nominal predicative structure.

(40) a. *Wa=tô api.*
1NOM=FCT cook
'I did the cooking.'

b. *Ĩĩ-**ñipi** õ **di** tô.*
1SG.ABS-cook NEG PART FCT
'I didn't do the cooking.'

(41) a. *Asaro wa=wami.*
rice 1NOM=[3ABS]sieve
'I sieved the rice.'

b. *Asaro te **wamirĩ** õ **di**.*
rice [1SG.ERG]AUX [3ABS]sieve NEG PART
'I didn't sieve the rice.'

Further evidence that the configuration of negative sentences in Xavante is diachronically derived from the structure of a nominal existential predication comes from the morphosyntactic properties of negated transitive verbs, illustrated in (41b). In negative sentences, transitive verbs have to be accompanied by the

17. A cognate for this morpheme is found in Apinajé (Oliveira 2005: 131).

18. The same occurs in Xerente, a closely-realated Jê language: according to Souza Filho (2007: 140), negation in Xerente is marked by the combination of the negation adverb *kõ* and the predicative morpheme *-di*.

marker *te*. At first glance, this morpheme might seem to be the generic relational noun *-te* 'thing of, possession of', mentioned at the end of Section 1.1.1. Recall that the function of this particular inalienable noun is to establish a relationship between the referent of an alienable noun and another participant, differing from a genitive adposition only by its morphosyntactic category. In other words, the structure of negative sentences could be paraphrased as 'there is no X-doing of Y', where 'X' and 'Y' would be noun phrases expressing a patient and an agent, respectively. Thus, two markers involved in negated transitive verbs suggest a nominal origin for negative sentences: the morpheme *di* (which points to an existential predication) and the morpheme *te* (which points to a nominal constituent). But the nominal appearance of negative sentences ends here. Syntactically, negative sentences do not bear the structure of a noun phrase: in a complex noun phrase, the complement of an inalienable noun heading the phrase has to occupy the position immediately preceding it. Moreover, the morphological properties of the marker *te* and of the generic relational noun show that these are two distinct morphemes.[19]

If negative sentences cannot be interpreted as negated nominal predications in a strict sense, they could be argued, however, to be composed of a complex predicate consisting of the auxiliary *di* and a non-finite verb. Recall from 2.1.1 that the auxiliary *di* is necessary with verbs expressing a state. However, the interpretation of *di* as an auxiliary expressing a state in a negated predication is questionable, not only for semantic reasons, but also because of its distributional and functional properties. The auxiliary *di* used with non-finite verbs expressing a state has a limited distribution, whereas the marker *di* is compulsory with the negation morpheme *õ* in independent clauses for any verb. Also, the morpheme *di* disappears in subordinate clauses, whether accompanying a stative or negated verb. In other words, in negative sentences the marker *di* seems only to function as the head of the sentence, without contributing any specific semantic content.

Finally, one last observation points to a 'nominal' characterization of negation (providing 'nominal' is taken in a larger sense, including non-finite clauses): negated verbs are incompatible with morphemes occurring in typically finite clauses. These markers include person prefixes/clitics and some aspectual markers. The following examples show the agrammaticality of the non-first person nominative clitic *te* and of the perfect particle *ma* with negated verbs.

(42) a. *Ma te=abʔru.*
PFC N1NOM=get.mad
'He got mad.'

19. Their inflections, given in appendix, are clearly different.

b. *Sibʔrui õ di.*
[3ABS]get.mad NEG PART
'He didn't get mad.'

c. * ***ma te=****sibʔrui* *õ di*
PFC N1NOM=[3ABS]get.mad NEG PART

This is congruent with the fact that negation is marked the same way with nominal predicates. In (43), the negation of the nominal predicate *baʔõtõ* 'girl' is marked by the negative particle *õ* followed by the morpheme *di*.

(43) *Õhã baʔõtõ õ* ***di*** *oto, õhã oto piʔõ.*
3PR girl NEG PART from.this.point 3PR now woman
'She is not a girl anymore, she is a woman.'

In short, negative sentences show a non-finite structure, composed of an abstract particle *di* heading the clause and a non-finite verb unable to function with the complete array of markers used in finite clauses. The fact that the marker *di* is still necessarily used as the head of a negative non-finite clause despite semantic loss suggests that negative verbal predication might have been nominal originally.

A further step of this hypothetic diachronic change from nominal to verbal predication would be the disappearance of the *di* marker and the reanalysis of such a non-finite structure into finite verbal predication. This can be observed with verb forms expressing a particular aspect, presented in the next section.

3.3 Aspect

In independent affirmative clauses, the verbal paradigms reveal a binary aspectual opposition, illustrated in the following examples. The 'unmarked' or 'underspecified' aspect is lexically determined: its interpretation depends on the inherent semantic characteristics of the verb. It is opposed to the 'aorist' aspect, formally marked, which is employed when referring to the distant past, an iterative action, a frequent or habitual situation, a general characterization; that is, whenever no connection is made between the situation described by the verb and the moment of speech.

The sentences in (44) illustrate this aspectual contrast with the intransitive verb *api* ~ *ñipi* 'to do the cooking', occurring in (a) with the 'unmarked' aspect and in (b) with the aorist. Notice that this intransitive verb belongs to the subclass of verbs showing the initial *aP-* ~ *siP-* alternation, and that the *siP-* stem is used in

the aorist. Also, the first person nominative clitic *wa* can be used with the aorist, but not the marker *ni*.[20]

(44) a. *Wa-ʔrata ma wa=za oto **api**=zaʔra*
1PL-grand.mother DAT 1NOM=PROSP from.this.point do.the.cooking=PL
ni.
INDF
'We will now do the cooking for our grandmother.'

b. *Ãne za wa-ʔrata ma=hã ʔre*
this.way PROSP 1PL-grand.mother DAT=DEMC HAB
*wa-**ñipi**=zaʔra.*
1PL.ABS-do.the.cooking=PL
'This is how we will usually do the cooking for our grandmother.'

c. * *ãne za wa-ʔrata ma=hã ʔre*
this.way PROSP 1PL-grand.mother DAT=DEMC HAB
*wa-ñipi=zaʔra **ni***
1PL.ABS-do.the.cooking=PL INDF

Compare the finite aorist verb in (45a) with its negative counterpart, in (45b). Recall that a negated verb is non-finite: it is dependent on the head of the sentence *di* and is incompatible with morphemes found in finite clauses, such as the first person nominative clitic *wa* (45c). The aorist, on the other hand, is finite: it is the head of the clause and can be used with the morpheme *wa*, as in (45a). The fact that some intransitive verbs exhibit the same *siP-* stem in the aorist aspect and with negation suggests that these have a common origin. This is to say that in the aorist, we see, yet again, a structure reminiscent of a nominal configuration, where a functional head occupying the last position of the sentence has been dropped, giving way to the reinterpretation of a nominal or non-finite configuration into a finite verbal structure.

(45) a. ***Wa**=ʔre wa-ñipi=zaʔra.*
1NOM=HAB 1PL.ABS-do.the.cooking=PL
'We usually do the cooking.'

b. *Wa-**ñipi**=zaʔra õ di.*
1PL.ABS-do.the.cooking=PL NEG PART
'We don't do the cooking.'

c. * ***wa**=wa-ñipi=zaʔra õ di*
1NOM=1PL.ABS-do.the.cooking=PL NEG PART

20. This morpheme, glossed 'indefinite', is used when the subject is an indefinite, generic or honorific third person, an honorific second person, or a non-singular first person.

In (46), the contrast between non-aorist and aorist aspect is illustrated with the transitive verb *ʔmaña(rĩ)* 'to make'. Here, the aorist verb must follow the morpheme *-te*, which we analyze as an auxiliary specific to transitive verbs in the aorist aspect. Although it could be argued to be an ergative pronoun, we do not adopt this analysis because of the presence of the pronominal clitic *wa*.

(46) a. *Uhi wa=za ʔmaña.*
beans 1NOM=PROSP [3ABS]make
'I am going to make beans.'
b. *Verdura wa=da-me te ʔre a-ʔmañarĩ.*
vegetables 1NOM=3GNC-COMIT [1SG.ERG]AUX HAB [3ABS]COL-make
'I used to make the vegetables with them (the other girls).'

Observe also that the verb stem loses its final syllable in the non-aorist aspect when it occupies the last position of the verb phrase. The fact that this syllable remains in the aorist aspect suggests the trace of a morpheme occupying the last position in the structure and heading the construction. However, the disappearance of this functional head and the possibility of combining the predicate with morphemes such as the perfect particle *ma* show that the predicate is a finite verb phrase. In (47), the perfect particle *ma* occurs with the verb *waihuʔu* in the non-aorist aspect in (a) and with the aorist aspect in (b).

(47) a. ***Ma** ti-wi waihuʔu=zaʔra.*
PFC 3-ABL [3ABS]learn=PL
'They learned it from them.'
b. ***Ma** oto da-te waihuʔu, romhuri=hã.*
PFC from.this.point 3GNC.ERG-AUX [3ABS]learn work=DEMC
'They learned it, the work.'

In other words, when a finite verb is used in the aorist aspect, the clause shows traces of a nominal structure. These include the trace of a morpheme occupying the last position and heading the clause (as shown by the impossibility of using a morpheme in final position, such as the indefinite subject marker *ni*, as well as the mandatory use of the 'long' verb stem), the *siP-* stem selected by the subclass of intransitive verbs characterized by the *aP-* ~ *siP-* initial alternation, the *-te* morpheme used with transitive verbs.

To conclude this section, which has focused on nominal structures from a diachronic perspective – and which has, in a sense, been more about verbalization than nominalization – we would like to propose that these structures are pertinent to the question of nominalizations from a synchronic perspective. In this respect, independent clauses with negative polarity and aorist aspect represent cases of

what could be called 'semantic' nominalization.[21] More specifically, verbs in these clauses are not marked by a specific nominalizer, but they lose some grammatical and semantic features typically associated with verbs: finiteness and temporal instability in the case of negation (verbs in negative sentences are non-finite and treated as components of a stative predication), internal temporal structure in the case of aorist aspect (aorist verbs lack aspectual determination in the sense of internal temporal structure).

This suggests that nominalizations can be seen to operate synchronically in Xavante at all levels: not only is nominalization possible with lexemes and entire clauses, it also applies to certain types of predications. In other words, the data examined here suggest that nominalization is a linguistic operation that can cut across layers of linguistic organization, just as would be expected from a cognitive linguistics approach. In this perspective, it seems to us that Langacker (1991)'s analysis of nominalization in English applies perfectly to the facts in Xavante: according to Langacker (1991: 33), nominalization can occur at the stem level (which characterizes a 'process type'), at the clause level (which characterizes a 'grounded instance of the process type') and at an intermediate level (which characterizes an 'ungrounded instance of the type').

Conclusion

This article has attempted to depict the pervasiveness in Xavante of nominalization, which operates in a number of ways at different levels of the language. This allows a panoramic observation of different facets of nominalization and suggests different perspectives for its analysis. A common trait which seems to emerge from a synchronic and diachronic point of view is that nominalization allows reference to a situation: synchronically, reference to a situation can be obtained through clause nominalization and 'semantically' nominalized predicates; diachronically, reference to a situation was obtained using referring expressions in an existential predication.

21. Or 'metaphorical', as Salanova (2013: 7, footnote 8) puts it. Nominalization manifested in verbal inflectional morphology (through the contrast between finite and non-finite verbal forms) is a central question in Jê languages. Oliveira (2005: 190) considers that 'nonfinite verb forms are the most nominal form of verbs' in Apinajé. Alves (2004: 155) hypothesizes that past tense verbal forms in Apãniekrá are originally non-finite forms used in nominalized constructions. More generally, Salanova (2007: 3–4) argues, based on Mẽbengokre, 'that nominalization underlies many aspectually-conditioned splits described in the literature, as well as being at the core of the perfect construction in languages such as French and Italian.'

Lexical nominalizations create new lexemes derived from verbal and nominal bases. Nominalizations based on verb stems result in alienable and inalienable nouns denoting a participant, an action/state or result. In the case of inalienable nouns, the argument they require refers either to the 'patient' or to the 'agent' argument of the corresponding verbal base, depending on the derivation. Nominalizations based on noun stems are more restricted: they result only in inalienable nouns requiring a 'possessor' argument.

The semantic range of grammatical nominalizations is the same as lexical nominalizations: they also denote concepts such as participants and actions. However, they serve a very different purpose: whereas lexical nominalizations are motivated by the need to create and store in the lexicon new lexemes denoting new realities, grammatical nominalizations are used in a context-dependent fashion.

Verbal predicates can be considered 'semantically' nominalized when they lose their internal temporal structure. This occurs with negation, expressed by a non-finite structure formally similar to stative predicates, and with aorist aspect, bearing many formal properties similar to negation. Theses constructions are characterized by morphosyntactic properties which suggest that they were originally nominal existential predications. This points to the idea that, because nominalizations denote temporally stable concepts, they can serve an ontological function: they can be used in an assertion to specify what is (or isn't) in a specific context. They allow speakers to present events as facts.

Interestingly, there is no nominalizing marker proper to verbal negative or aorist predicates; the nominal character of these constructions is apparent in their structural properties, resulting from a reanalysis of an existential predication. On the other hand, clause nominalizations are clearly marked by a specific prefix. This morpheme, which allows clauses to function syntactically like nominal constituents, could have derived historically from a demonstrative which would point to a referent and thus allow the whole expression to refer. In the absence of a specific object, this demonstrative could have pointed to a place (demonstratives in Xavante can also have a locative interpretation[22]) and, in an abstract sense, to a situation. This would account for the insubordinate use of nominalized clauses and its different readings: events are presented as facts in thetic judgements, confirmation or emphatic expressions.

Finally, because nominalizations allow speakers to present events as facts, their use can be motivated by pragmatic factors, such as politeness. Referring directly to some agent participants, if considered offensive, could be avoided by using a nominal structure. Through insubordination, a unique ergative prefix arose from a clause-nominalizing morpheme to mark second person agents on transitive verbs.

22. The same is true in indigenous languages of North America (Mithun 1999: 132).

Consequently, the question that emerges from the picture of nominalization in Xavante is: what account can provide a unified explanation for all the different uses of nominalization in this language? Semantically, nominalization results in linguistic units that denote an entity-like concept (Shibatani, this volume). Formally, this process may be derivational or inflectional. Functionally, it serves an ontological purpose. We would like to propose that it is this ontological function of nominalization which allows this process to be maximally exploited in Xavante. Whether describing a participant or a fact, nominal constructions may be used either to refer, to predicate or to assert (depending on the context in which they are used) because their ontological function presupposes a 'domain of existence', and this domain may be limited to a referent, a clause, a sentence, or a portion of discourse. In other words, the interpretation of nominalizations – as existence of a referent or the occurrence of a fact – depends on anchoring their denotation in a specific context, which provides a support for their ontological function.

Abbreviations

1	first person
2	second person
3	third person
ABS	absolutive
AN.PR	anaphoric pronoun
ANTP	antipassive
AUX	auxiliary
COL	collective
COMIT	comitative
CORF	coreferential
DAT	dative
DEM	demonstrative
DEMC	demarcative
DEOBJ	deobjectal
DIST.LOC	distant locative
DISTR	distributive
DUB	dubitative
DUR	durative
ERG	ergative
EXH	exhortative
EXIS.COP	existential copula
FCT	factual
HAB	habitual
IMPERF	imperfective
INDF	indefinite
INES	inessive
INSTR	instrumental
INTER.P	interrogative particle
INTSF	intensifier
IRR	irrealis
MAL	malefactive
NMZR	nominalizer
N.APL	nominal applicative
N1	non first person
NEG	negation
NOM	nominative
PART	particle
PERM	permansive
PFC	perfect
PL	plural
PR	pronoun
PROP.N	proper noun
PROSP	prospective
PVB	preverb

GNC	generic	SG	singular
GRN	generic relational noun	SUB	subordinator
H	honorific	VOC	vocative

References

Alves, Flávia de Castro. 2004. O Timbira falado pelos Canela Apãniekrá: Uma contribuição aos estudos da morfossintaxe de uma língua Jê. PhD dissertation, UNICAMP.

Alves, Flávia de Castro. 2010. Evolution of alignment in Timbira. *International Journal of American Linguistics* 76(4): 439–475. https://doi.org/10.1086/658054

Bauer, Laurie. 1983. *English Word-formation*. Cambridge: CUP. https://doi.org/10.1017/CBO9781139165846

Carlin, Eithne B. 2011. Theticity in Trio (Cariban). *International Journal of American Linguistics* 77(1): 1–31. https://doi.org/10.1086/657326

Craig, Colette G. & Hale, Ken. 1988. Relational preverbs in some languages of the Americas: Typological and historical perspectives. *Language* 64: 312–343. https://doi.org/10.2307/415436

Diessel, Holger. 1999. *Demonstratives: Form, Function, and Grammaticalization* [Typological Studies in Langage 42]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.42

Estevam, Adriana Machado. 2015. Expressando conceitos de qualidade em Xavante: Adjetivos ou verbos? *Moara* 43: 141–165.

Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.

Gildea, Spike & Alves, Flávia de Castro. To appear. Reconstructing the source of nominative-absolutive alignment in two Amazonian language families. In *Reconstructing Syntax: Cognates and Directionality* [Brill Series in Historical Linguistics], Eugenio Luján, Jóhanna Barðdal & Spike Gildea (eds). Leiden: Brill.

Givón, T. 2016. Nominalization and re-finitization. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada-Fernández (eds), 271–296. Amsterdam: John Benjamins.

Haspelmath, Martin & Sims, Andrea D. 2010. *Understanding Morphology*. London: Hodder Education.

Kuroda, Sige-Yuki. 1972. The categorical and the thetic judgment. Evidence from Japanese syntax. *Foundations of Language* 9(2): 153–185.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Discourse Referents*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511620607

Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar*, Vol. 2: *Descriptive Application*. Stanford CA: Stanford University Press.

Launey, Michel. 1986. Catégories et opérations dans la grammaire nahuatl. Thèse d'État, Université Paris IV.

Maslova, Elena. 1997. Yukagir focus in a typological perspective. *Journal of Pragmatics* 27: 457–475. https://doi.org/10.1016/S0378-2166(96)00038-0

McLeod, Ruth & Mitchell, Valerie. 2003. *Aspectos da língua xavante*. Cuiabá: Publicações da SIL.

Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: CUP.

Oliveira, Christiane C. 2005. The language of the Apinajé people of Central Brazil. PhD dissertation, University of Oregon.

Queixalós, Francesc. 2005. Posse em katukína e valência dos nomes. In *Novos estudos sobre línguas indígenas*, Aryon Rodrigues & Ana Suelly A. C. Cabral (eds), 177–202. Brasília: Editora Universidade de Brasília.

Queixalós, Francesc. 2006. The primacy and fate of predicativity in Tupi-Guarani. In *Lexical Categories and Root Classes in Amerindian Languages*, Ximena Lois & Valentina Vapnarsky (eds), 249–287. Bern: Peter Lang.

Queixalós, Francesc. 2016. The role of nominalization in theticity. A Sikuani contribution. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada-Fernández (eds), 205–244. Amsterdam: John Benjamins.

Ribeiro, Eduardo Rivail. 2002. O marcador de posse alienável em Karirí: Um morfema Macro-Jê revisitado. *LIAMES* 2: 31–48.

Salanova, Andrés. 2007. Nominalizations and aspect. PhD dissertation, MIT.

Salanova, Andrés. 2008. Uma análise unificada das construções ergativas em Mẽbengokre. *Amérindia* 32: 109–134.

Salanova, Andrés. 2013. Event nominals in Mẽbengokre and the theory of nominalizations. *Canadian Journal of Linguistics* 58(2): 267–292. https://doi.org/10.1017/S0008413100003030

Sasse, Hans-Jürgen. 1987. The thetic/categorical distinction revisited. *Linguistics* 25: 511–580. https://doi.org/10.1515/ling.1987.25.3.511

Souza Filho, Sinval Martins de. 2007. Aspectos morfossintáticos da língua Akwẽ-Xerente (Jê). PhD dissertation, Universidade Federal de Goiás.

Shibatani, Masayoshi. 2009. Elements of complex structures, where recursion isn't: The case of relativization. In T. Givón and Masayoshi Shibatani (eds), *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution*, 163–198. Amsterdam: John Benjamins.

Appendix

Table 1. Inflectional paradigm of the *-te* auxiliary

			Number	
			Singular	**Plural**
Person	1		ø-te	wa-te
	2	unmarked	mo	
		honorific	aa-te	
	3	unmarked	ø-te	te-te ~ ø-te
		honorific	da-te	
		generic	da-te	

Table 2. Inflectional paradigm of the generic relational noun *-te*

			Number		
			Singular	**Collective plural**	**Discrete plural**
Person	1		ĩĩ-te	wa-te	wa-te zaʔra
	2	unmarked	a-te	a-te waʔwa	a-te zaʔra waʔwa
		honorific	aa-te	aa-te waʔwa	aa-te zaʔra waʔwa
	3	unmarked	ĩ-te	ĩ-te zaʔra	
		honorific/generic	da-te	da-te waʔwa	da-te zaʔra waʔwa
		coreferential	ti-te	ti-te zaʔra	

CHAPTER 18

Innovation in nominalization in Tupí-Guaraní languages

A comparative analysis of Tupinambá, Apyãwa and Nheengatú

Aline da Cruz and Walkíria Neiva Praça
Universidade Federal de Goiás / Universidade de Brasília

This paper focuses on nominalizations within Tupí-Guaraní, a sub-group of the Tupí linguistic family. For this purpose, we first analyze two very conservative Tupí-Guaraní languages, Tupinambá and Apyãwa (Tapirapé), and then compare them with a very innovative Tupí-Guaraní language, Nheengatú (língua geral). By doing so, we show that the nominalizers are correlated to more general typological properties of the sub-group, such as the omnipredicative pattern. More specifically, we address the historical development of these languages showing that the loss of omnipredicative properties led to the restructuration of the forms and the functions of nominalization.

1. Introduction

The Tupí-Guaraní languages, spoken in South America, compose an uncontroversial grouping within the Tupian family. Nominalization is a pervasive feature of Tupí-Guaraní languages, and a number of nominalizers have been reconstructed to Proto-Tupí-Guaraní. Documents from the 16th century register eight different kinds of nominalizations in Tupinambá, the first language described in Brazil. A more recently described language, Apyãwa, traditionally known as Tapirapé, possesses the majority of the same type of nominalizations found in Tupinambá, which facilitates the analysis of the old patterns of the language. In contrast, Nheengatú, a language that descends from Tupinambá, has innovated the forms and the functions of nominalizations. In this paper we investigate the historical development of the nominalization patterns in these three languages. We find that while nominalization is a pervasive construction in Tupinambá and Apyãwa, Nheengatú has lost many of the nominalizers and restricted their productivity. We

https://doi.org/10.1075/tsl.124.18cru

propose that the differences in the uses of nominalizations are associated with a deep typological change, from the conservative omnipredicative pattern to an innovative non-omnipredicative pattern.

The term omnipredicativity was coined by Launey (1986, 1994, 2004) to describe languages in which most lexical entries can function as predicates and arguments are subordinate predicates designating an entity (or rather describing the notional value of the term). Other theoretical approaches, such as generative theory, also deal with the topic, but in this approach the term used is nonconfigurationality. The term nonconfigurationality was coined by Hale (1983) to describe the same type of language analyzed by Launey (1994), but focuses on the constituency and the hierarchy of syntactic functions rather than on the projection of lexical units in the syntactic positions of argument and predicate. Most of the properties observed by Launey to define omnipredicativity were also found in Tupinambá (Queixalós 2006) and in Apyãwa (Praça 2007), but they have been lost in Nheengatú (Cruz 2011).

This paper is organized as follows. Section 2 briefly presents the three languages used in this work. Section 3 gives the necessary morphosyntactic background, by presenting the main characteristics of omnipredicative languages. Section 4 describes the main features of nominalization in Tupinambá, followed by the description of nominalization in Apyãwa in Section 5. Section 6 introduces the types of nominalizations in Nheengatú. After describing the main features of nominalization found in Tupinambá, Apyãwa and Nheengatú, a discussion on the restructuration of the forms and functions of nominalization in the historical development of Nheengatú is presented in Section 5. The paper concludes with a summary of the historical development of nominalization in these Tupí-Guaraní languages, from a conservative to an innovative pattern.

2. Languages

In this section, we provide background information on the three Tupí-Guaraní languages investigated in this study: Tupinambá, Apyãwa and Nheengatú. Tupinambá was the first language with whiTch the Portuguese colonizers had contact. It was documented in the 16th and 17th centuries. Historical sources written on and in this language include poems and theatre pieces by Anchieta (1977)[1]; grammars by Anchieta (1990 [1595]) and by Figueira (1880 [1621]); a catechism by Araújo (1618); and vocabularies. Thanks to such documents, it is possible to analyze the

1. According to Leite (2004), the theatre pieces were written by Anchieta between 1561 and 1567.

extinct Tupinambá language. Previous analyses of nominalization in Tupinambá have been carried out by Rodrigues (1953, 2010) and Lemos Barbosa (1956).

According to the early historical documentation, Tupinambá was widely spread in the coast of Brazil. For this reason, the language was used for inter-ethnic communication, not only between indigenous groups, but also between indigenous groups and the colonizers. In 1616, with the foundation of a fort that latter become the city of Belém, the colonizers decided to bring Tupinambá speakers to participate in the colonization of the Amazon region (Bessa Freire 2004). As Tupinambá was spread deeper into the Amazon region, it underwent changes and became known as *língua geral Brasílica*. Nowadays, Brasílica is known by the remaining speakers as Nheengatú (*nheen* 'language', *katu* 'good', the good language). It is spoken by approximately 8,000 speakers in the Upper Rio Negro by people of Arawak descent; more specifically, the Baré from the Upper Rio Negro, the Baniwa from the Lower Rio Içana, and the Warekena from Rio Xié. In the course of the 19th and 20th centuries, these groups have switched from their traditional Arawakan languages and adopted Nheengatú as their mother tongue. More recently, some of these groups are passing through another language shift, since some communities in these areas have been switching from Nheengatú to Brazilian Portuguese. Even though Nheengatú is considered a co-official language in the municipality of São Gabriel da Cachoeira in Brazil, there are few studies about its current stage of development. Aspects of currently spoken Nheengatú morphosyntax was studied by Moore, Facundes and Pires (1993), Cruz (2011, 2014, 2015), and Moore (2014). Nominalizations are only briefly described in Cruz (2011: 246–255).

While the Tupinambá had contact with Portuguese colonizers since the 16th century, the Apyãwa established contact with the non-indigenous society only in the 20th century. This situation may have allowed Apyãwa to retain a more conservative morphology with respect to Proto-Tupí-Guarani, as reconstructed by Jensen (1999).

Apyãwa is spoken by approximately 950 speakers who live in two indigenous territories, the Tapirapé/Karajá territory and the Urubu Branco (Tãpi'itãwa) territory, in the northeastern part of the state of Mato Grosso, in Brazil. Although several studies have already analyzed aspects of Apyãwa morphosyntax (for instance, Almeida et al. (1983), Leite (1990) and Praça (2007)), nominalization was studied only by Praça (2007), and not in depth.

3. The omnipredicative pattern in conservative Tupí-Guaraní languages

According to Lemos Barbosa (1956), one of the most remarkable features of Tupinambá is that "the distinction between nouns and verbs is not clear, because

all nouns can become predicates, and all infinitive verbs are true nouns. They seem to have two statuses: verbal and nominal".[2] The observation made by Lemos Barbosa (1956) about Tupinambá is supported by the analyses of various other Tupí-Guaraní languages. Queixalós (2006) hypothesized that members of this family are descendants of a language in which all lexical entries were predicative (a state of 'omnipredicativity'; cf. Launey (1994)). The omnipredicative properties were maintained in Tupinambá, and, according to Praça (2007), they also remain in Apyãwa. In contrast, Nheengatú has lost most of these characteristics, and can no longer be considered an omnipredicative language. In the remainder of this section, we discuss the omnipredicative properties of Tupinambá and Apyãwa.

In Tupinambá and Apyãwa, nouns and verbs can occur as predicates and arguments, without any derivational process. The predicative function is not morphologically marked, thus, the absence of marking indicates the predicative function, both in nouns, as illustrated in (1) from Apyãwa and (2) from Tupinambá, and in verbs, as in (3) from Apyãwa and (4) from Tupinambá.

(1) *marare-ø* ***i-memyr***
cow-RF 3.NA-offspring
'The cow has an offspring.'
(lit: (There is) her offspring in relation to the cow.)

(2) *pajé-ø* ***i-posáŋ***
shaman-RF 3.NA-medicine
'The shaman has medicine.' (Rodrigues 2010: 111; our glossing)

(3) *ka'i-ø* ***a-xe'eg*** *a-ka-wo* *'ywyrã-ø r-e*
monkey-RF 3.A-speak 3CO-be-CVB tree-RF LK-POS
'The monkey is speaking in the tree.'

(4) ***a-ɨur***
1SG.A-come
'I came.' (Gregório 1980 apud Navarro (1998: 75; our glossing))

In both Tupinambá and Apyãwa, the argumental function is marked by the suffix *-a*, which has the allomorph *-ø*.[3] When inflected with the suffix *-a*, nouns and verbs function as arguments. Notice that the nouns *memyr-a* in (5) from Apyãwa, and *mosáŋ-a* in (6) from Tupinambá are suffixed with *-a*, which indicates that

2. In the original: "*a distinção verbo-nome não é nítida, pois todo nome pode tornar-se predicativo, e todo verbo no infinitivo é um verdadeiro nome. Os mesmos parecem ter dois 'status': o verbal e o nominal.*" (Lemos Barbosa 1956: 393).

3. In both Tupinambá and Apyãwa, the allomorph *-a* occurs after consonants, while *-ø* occurs after vowels.

they function as arguments. The suffix *-a* is also attached to verbs to function as arguments, as illustrated in (7) from Apyãwa and in (8) from Tupinambá. It is important to observe that the suffix *-a* does not promote lexical class change, and therefore is not a nominalizer, since it occurs with both nouns and verbs. According to Queixalós (2001, 2006), the suffix *-a* turns originally predicative stems into referential units, which, as consequence, can function as arguments, whereas the absence of the suffix would indicate a predicate – a pattern which the author compares with the uses of the morpheme *ang* in Tagalog and its cognates in other Austronesian languages (cf. Lemaréchal (1989, 1991)).

(5) ***i-memyr-a*** *a-xaj'a*
3.NA-offspring-RF 3.A-cry
'Her offspring cried.'

(6) *pajé* ***m-osáng-a*** *o-j-kuáβ*
shaman INDF-medicine-RF 3.A-3.NA-know
'The shaman knows medicine.' (Rodrigues 2001:110; our glossing)

(7) ***xe=ø-xe'eg-a*** *mĩ i-ãrõãrõ*
1SG.NA=LK-speak-RF HAB 3.NA-be.beautiful
'My speech is always beautiful.'

(8) *sjé=ø-maʔenwár* ***ne=r-úr-a*** *r-esé*
1SG.NA=LK-remember 2SG.NA-LK-come-RF LK-of
'my remembering of your arrival'[4]
(Figueira 1687: 157–158 *apud* Rodrigues 1996; our glossing)

According to Rodrigues and Cabral (2002), the suffix *-a* can be reconstructed to Proto-Tupí-Guaraní.[5] Even though cognates of the suffix are found in all languages of the Tupí-Guaraní subfamily, its nature is a matter of debate. Rodrigues (1953), Lemos Barbosa (1956) and Almeida, Irmãzinhas de Jesus and Paula (1983) analyzed the suffix *-a* as a 'nominalizer'. However, as illustrated in examples (5) to (8) above, this suffix can co-occur with both nouns and verbs, since all lexical entries can take *-a* to function as an argument. Another hypothesized analysis in the literature considers the suffix *-a* to be a case marker. In such view, the suffix has been called a 'nominal index' (Rodrigues (1953), Lemos Barbosa (1956)); a 'nominal case' (Jensen 1989); an 'argumentive case' (Rodrigues (1996), (2001)), Praça (1999); a 'nuclear case' (Seki (2000), Borges (2006)); onomastic case (Adelaar

4. Translated by Rodrigues (1996) as "*bem me lembro da vinda de você*" [I do remember your arrival].

5. As there is not enough information about other Tupí languages, a morphological reconstruction of Proto-Tupí has not yet been carried out.

1997). The analysis adopted in this paper is the one proposed by Queixalós (2001, 2006), which considers that the suffix should not be seen as neither a case marker nor a nominalizer, but rather as indicating the function of referrer.

In many Tupí-Guaraní languages, the suffix *-a* has disappeared completely, or it has fused with the root, as a fossilized form with no morphological value (cf. Section 6). In contrast, in Apyãwa and Tupinambá, the occurrence of this suffix is very productive. This productivity may be intrinsically related to the strong omnipredicativity found in these languages. Thus, since the main lexical entries, nouns and verbs, are generated in the lexicon as predicates, they require the presence of *-a* to be able to function as an argument. As shown in (7) and (8) above, intransitive verbs inflected with the referrer *-a* can occur as arguments. Intransitive verbs may also be nominalized, as exemplified in (9). Notice that both *xe'eg* 'speak' and its nominalized form *xe'eg-ãw* 'speech' must be inflected with the suffix *-a* to occur as arguments. In contrast, transitive verbs can only occur as arguments when nominalized, as shown in (10).

(9) *ie ã-enow* *ne=ø-xe'eg-**ãw**-a*
I 1SG.A-heard 2SG.NA=LK-speak-NMLZ-RF
'I heard your act of speaking.'

(10) *tãxão-ø ø-xokã-**ãw**-a* *i-kãto*
pig-RF LK-kill-NMLZ-RF 3.NA-good
'The killing of the pigs was good.'

Even derived nominal forms (i.e., those subjected to deverbal morphological processes), such as nominalizations, are inherently predicates. For instance, in (11a), the nominalized form *xe=r-o'y-pepa-kyxi-ãw* occurs as an existential predicate, which could be translated as 'my scissors (exist)'. The nominalized form takes the referrer to institute argument, as illustrated in (11b).

(11) a. *xe=r-o'y-pepa-kyxi-ãw*
1SG.NA=LK-arrow-wing-cut-NMLZ
'I have scissors.' (lit: My arrow-wing-cutter (exists).)
b. *e-m-or* *ne=r-o'y-pepa-kyxi-ãw-ã*
2SG.IMP-CAUS-come 1SG.NA=LK-arrow-wing-cut-NMLZ-RF
'Lend me your scissors.'

The contrast between the pair of examples in (11) is particularly interesting because it shows that nominalization in an omnipredicative language is not just a process that turns a verb into a noun in order to allow it to occur as argument. In fact, this is not necessary, since intransitive verbs can occur as arguments with the same morphology required by nouns. In other words, if nouns and verbs can occur as arguments when inflected by suffix *-a*, what would be the use of

nominalizations in these languages? The answer should take into account the semantic implications of creating a new lexical item: nominalization establishes a semantic link with the source verb and adds semantic features in order to create nouns to designate events, agents, results, patients, instruments and so on. For instance, in (12) the nominalized verb *ʾo* 'ingest', inflected with the suffix *-a*, occurs as argument. In contrast, in (13) the verb *ʾo* 'ingest' is nominalized creating a reference to an entity that expresses the 'act of seed ingesting'. Notice also that in (13) the nominalized verb occurs as a predicate at the syntactical level. It is noteworthy that the nominalized verb is combined with the negation circumfix *n*=…*-i*, which, other than in nominalizations, is used exclusively with independent predicates.

(12) *ãpi-ø a-kwããw tãtã-ø* ***ø-ʾo-pã-ãw-a***
mother-RF 3.A-know banana-RF LK-ingest-COMP-NMLZ-RF
'The mother knew about the eating of all the bananas.'

(13) *aʾỹj-a mĩ* ***n=i-ʾow-ãw-i***
seed-RF HAB NEG=3SG.NA-ingest-NMLZ-NEG
'There is (exists) no seed ingesting (referring to the murici seed).'

Although nouns and verbs can occur as predicates, and inflected by the suffix *-a*, they can also occur as arguments, these two word classes can be distinguished by their morphological properties. In Apyãwa, only verbs can be combined with the nominalizer *-ãw* (cf. Section 5). Thus, the combination of a noun with a nominalizer is completely ungrammatical, as illustrated by the grammatical test in (15) with Apyãwa speakers:

(15) * *i-memyr-ãw-a*
intended meaning 'the event of having children'

4. Nominalization in Tupinambá

Rodrigues (1953, 2010), based on the analysis of 16th century documents, lists eight nominalizers in Tupinambá, each with discrete functions. Six of these nominalizers derive a verb into a nominal; one derives a postpositional phrase or an adverbial phrase into a nominal; and one is used for creating nominals from predicates. The core function of each nominalizer will be briefly described in the remainder of this section. Due to the limited information available on this extinct language, it is not possible to give detailed information on the properties of each of these nominalizers.

4.1 Deverbal nominalization

In Tupinambá seven different suffixes can derive argument-denoting nominals. Based on the semantic features of the nominals that they create, they are classified as agentive nominalizer (*-ár*), event and instrument nominalizer (*-aβ*), theme/patient nominalizer (*-pɨr*) and resultative nominalizer (*emi-*). Nominalizers can also have aspectual and modal semantic features, such as the habitual agentive nominalizer (*-βor*) and the suffix *-swér*, which indicates an agent who enjoys doing the action denoted by the basic verb.

The suffix *-ár* (allomorph *-sár*) derives a noun from an action verb, yielding a nominal meaning 'the one who performs the action denoted by the verb', such as in (16) below. As other nouns, the nominalized form takes a prefix from the NA (non-agent-like argument) set, as exemplified in (17).

(16) *moyáŋ-ár*
make-NMLZ
'author' (Rodrigues 2010)

(17) *xe-juka-sár-a*
1SG.NA-kill-NMLZ-RF
'the killer of me' (Anchieta 1990 [1595]: 31)

The two other agentive nominalizers *-βór* and *-swér* differ from *-ár* in that they bear aspectual and modal properties. The suffix *-βór* indicates an agent that often makes the action denoted by the verb: 'the one who often does the action denoted by the verb', as shown in (18). Rodrigues (2010) calls *-βór* an 'habitual agentive nominalizer'. According to Jensen (1989: 116), the suffix *-βór* registered in Tupinambá has a cognate in Old Guarani. The form was recorded only in documents from the colonial period. No description of currently spoken Tupí-Guaraní languages indicates the existence of cognates of *-βór*.

(18) *kayẽ-βór*
run.away-NMLZ
'the fugitive' (Rodrigues 2010)

A third agentive nominalizer in Tupinambá is *-swér*, which indicates that the agent likes to do the action denoted by the verb. In (19) the intransitive verb *pukáˀ* 'laugh' was combined to the suffix *-swér*, creating a noun whose meaning is the 'one who likes to laugh'. In the pair of examples in (20a), the intransitive verb *ñeéng* 'speak' is combined with a set A (agent-like argument) prefix when it occurs in an independent clause, as in (20b), while its nominalized form requires a set NA prefix.

(19) *puká-swér*
laugh-NMLZ
'a joyful person' (Rodrigues 2010)

(20) a. *a-ñeéng*
1SG.A-speak
'I spoke.' (Anchieta 1990 [1595]: 51v)
b. *xe-ñeéng-ixuér*
1SG.A-speak-NMLZ
'I am talkative.' [6] (Anchieta 1990 [1595]: 51v)

There are two more participant nominalizers in Tupinambá, *emi-* and *-ɨpɨr*, which are used to indicate the theme or patient. The prefix *emi-* was analyzed by Rodrigues (1953) as 'nominalization of object', while the suffix *-pɨr* was analyzed as 'nominalization of patient'. In this paper, we maintain the analysis of *-pɨr* as 'patient nominalizer', but we analyze *emi-* as a 'resultative nominalizer', using the term defined by Comrie and Thompson (2007: 340). According to these authors, a resultative nominalizer "forms nouns designating the result, or the typical or 'cognate' object of an action". This is the function of the nominalizer *emi-* in Tupinambá, as illustrated in (21). The suffix *-pɨr* creates a noun with a more patientive meaning; that is, 'the thing/person that suffered the action denoted by the verb', as given in (22) and (23) below.

(21) *ajẽté kó né r-apé-0 a'é né r-emi-ekár-a*
in.reality DEM 2SG.NA LK-way-RF DEM 2SG.NA LK-NMLZ-look-RF
'In reality, this is your way, this is the thing that you were looking for.'
(Anchieta *apud* Rodrigues et al. 2006: 24)

(22) *i-yuká-pɨr*
3.NA-kill-NMLZ
'the one who is killed' (Rodrigues 2010)

(23) *o-j-posánóng-ɨpe jané ø-jár-a a'é i-namí-monók-ɨpɨr-a*
3.A-3.NA-cure-Q 1PL.NA LK-lord-RF DEM 3.NA-ear-cut-NMLZ-RF
'Did our Lord cure the one who had his ear cut?' (Araújo 1618)

The suffix *-áβ* is a more generic nominalizer that creates nouns for expressing event, instrument or location. As an instrument nominalizer, *-áβ* is suffixed to a verbal root to create a noun, meaning the 'thing used to perform the action denoted by the verb', as illustrated in (24). This suffix can also indicate a locative nominal; that is, a noun meaning the 'place where the action denoted by the verb

6. In Anchieta's translation, '*eu sou falador, tenho inclinação a falar*'.

occurred', as illustrated (25). The nominalizer *-áβ* can also indicate a 'time in which the action denoted by the verb occurred', as illustrated in (26). Rodrigues (2010) calls this suffix an 'instrumentive nominalizer', however, this term highlights only one of the functions of the morpheme and obscures its locative semantics and its function of denoting events in general. In this work, the suffix *-áβ* is labelled 'event nominalizer'.

(24) *enúβ-áβ*
hear-NMLZ
'an instrument for listening' (Rodrigues 2010)

(25) *moyáŋ-áβ*
make-NMLZ
'place where something is made' (Rodrigues 2010)

(26) *kaʔ-ú-aβ*
cauim-drink-NMLZ
'occasion on which people drink cauim' (Rodrigues 2010)

4.2 Nominalization of other categories

In Tupinambá nouns can be created from postpositional phrases. For this purpose, the suffix *-swár* is combined with postpositional phrases, as illustrated in (27) and (28). The form is used to express an entity from a specific special or temporal origin.

(27) *pó-pe-swár*
hand-LOC-NMLZ
'the one that is in the hand' (Rodrigues 2010)

(28) *kaʔá-βo-swár*
forest-LOC:DIFUSE-NMLZ
'the one who lives in the forest' (Rodrigues 2010)

Tupinambá also has a grammatical nominalization that takes predicates as its basis. The suffix *-mba'e* is attached to transitive or intransitive predicates. The nominalized form maintains some finite features. Despite being a nominalization, it takes set A person-marking prefixes, which are otherwise restricted to independent clauses. For instance, the set A prefix that occurs with a transitive verb occurs in the independent clause in (29a) likewise occurs in its nominalized form in (29b). The nominalization with *-mba'e* contrasts with the agentive nominalization in its finite properties: while grammatical nominalizations occur with set A prefixes, which otherwise occur with verbs, the agentive nominalization takes the prefix of set NA, which are used with nouns, stative verbs and postpositions. This can

be observed by comparing example (29b) with example (17), reproduced here as (29c). The grammatical nominalization can also occur with intransitive predicates, as in the pair of examples in (30). In *Arte de Gramatica da língua mais usada na costa do Brasil*, Anchieta does not provide examples of grammatical nominalization that takes nominal predicates as its base .[7]

(29) a. *o-juká*
3.NA-kill
'He/They killed.' (Anchieta 1990 [1595]: 17v)

b. *o-juká-bae ixé*
3.NA-kill-NMLZ I
'I am the one who kills.' (Anchieta 1990 [1595]: 30v)

c. *xe-juka-sár-a*
1SG.NA-kill-NMLZ-RF
'the killer of me' (Anchieta 1990 [1595]: 31)

(30) a. *o-só*
3.A-go
'He/They went.'

b. *ixe o-só-βaʔé*
I 3.A-go-NMLZ
'I am the one who went.' (Anchieta 1990 [1595]: 30v)

4.3 Summary of nominalization in Tupinambá

In Sections 4.1 and 4.2, we have presented the nominalizers found in Tupinambá by Rodrigues (1953). Table 1 provides a list of all nominalizers registered in the

Table 1. Tupinambá's nominalizers

	Form	Type of nominalizer	Examples
Verbal to Nominal	*-ár*	agentive	(16), (17)
	-βor	habitual agentive	(18)
	-swér	propensive agentive	(19), (20)
	emi-	resultative	(21)
	-pir	theme/patient	(22), (23)
	-aβ	event, instrumentive, location	(24)–(26)
Other Non-nominal to nominal	*-swár*	circumstantial	(27), (28)
	-mbaʔé	grammatical	(29b), (30b)

7. More investigation in other Tupinambá documents is necessary to fill this gap. As will be shown in Sections 5.2 and 6.2.2, grammatical nominalizations of nominal predicates occur in both Apyãwa and Nheengatú.

language. The table indicates the form and the basic semantics, and provides reference to the main examples used in this paper that illustrate the use of each affix.

5. Nominalization in Apyãwa

In this section we describe the nominalizations in Apyãwa. When compared to Proto-Tupí-Guaraní, Apyãwa displays very conservative morphology; for instance, it preserves the suffix *-a*, the morphemes for increasing and decreasing valence, the majority of nominalizers, the majority of particles, among other properties. For this reason, the analysis of Apyãwa allows the investigation of the conservative patterns of nominalization in Tupí-Guaraní languages.

Nominalization in Apyãwa is a general, productive, and regular process that creates a type of nominal that can function as an argument or as a predicate. There are four types of verb-based nominalizations and two types of nominalizations derived from other word classes. The verb-based nominalizations constitute relational nouns[8]; in other words, they always have a nominal complement, which is expressed with person-indexing prefixes of set NA (*xe-* '1SG'; *xane-* '1INCL'; *are-* '1EXCL'; *ne-* '2SG'; *pe-* '2PL'; *i-* ~ *ø-* ~ *t-* ~*h-* '3') or with noun phrases. The two other types of nominalizations are the circumstantial nominalization and the predicate nominalization. Circumstantial nominalizations take an adverbial phrase as their bases, whereas predicate nominalizations take as their bases intransitive predicates.

5.1 Deverbal nominalization

The verb-based nominalizations in Apyãwa are formed by attaching nominalizing affixes to verbal bases. There are four types of nominalizations in the language. Three of them are formed with suffixes: *-ãw* (*-ãw* ~ *-tãw*) 'event, instrument and location nominalizer', *-ãr* (*-ãr* ~ *-tãr*) 'agentive nominalizer', and *-pyr* (*-ipyr* ~ *-pyr*) 'patient nominalizer'. One nominalization is formed with a prefix: *emi-* 'resultative nominalizer'. Generally, the derived noun maintains reference to the patient, which is expressed as its obligatory nominal complement.

8. The concept of 'relational nouns' is syntactically defined: in a nominal predicate, relational nouns select two arguments (as in examples (1) and (2)), whereas non-relational nouns select only one argument. The concept of relational *vs.* non-relational nouns is closely related to the concept of bivalent vs. monovalent nouns proposed by Queixalós (2005 and 2010). The relational vs. non-relational opposition is the syntactic counterpart of the opposition between inalienably possessed nouns vs. alienably possessed nouns, whose definition is purely semantic.

The suffix *-ãw* is attached to transitive and intransitive verbal roots, forming nouns that denote events, instruments and locations. The nouns formed with *-ãw* always have an obligatory nominal complement; in other words, they are relative nouns. The suffix *-ãw* can occur with any type of verb: stative intransitive verbs, as in (31), or active intransitive verbs, as in (32), and transitive verbs, as in (33) and (34).

(31) *wãkiri ne=ø-kywe-ãw-a* ***i-ãi~ãiw***
Walkíria 2SG.NA=LK-be.skinny-NMLZ-RF 3.A-be.ugly~RED
'Walkiria, you getting skinnier is very ugly.'

(32) *koxãwiri-ø a-ixãk* ***i-xãj'a-ãw-a***
Koxãwiri-RF 3.A-see 3.NA-cry-NMLZ-RF
'Koxãwiri saw her crying.' (lit: 'Koxãwiri saw the crying of hers.')

(33) *ã-ixãk rãka* ***i-nopỹ-ãw-a***
1SG.A-see R.PST 3.NA-beat-NMLZ-RF
'I saw his getting beaten up.'

(34) *t-yro-paej-tãw-a*
3.NA-clothes-wash-NMLZ-RF
'wash basin'

The agentive nominalizer *-ãr* derives nouns from transitive verbal roots, indicating the one who performs the action denoted by the verb. The derived noun entails reference to a patient, which is its obligatory nominal complement, through the person prefixes from set NA, as illustrated in (35), or through noun phrases, as in (36).

(35) ***t-yro-paej-tãr-a*** *a-xe'eg a-ka-wo 'y-pe*
3.NA-clothes-wash-NMLZ-RF 3.A-talk 3.CO-be-CVB water-LOC
'The "laundry women" are talking at the river.'

(36) *xãrio-ø miãr-a* ***ø-kotok-ãr-a***
Xãrio-RF deer-RF LK-kill-NMLZ-RF
'Xãrio is a killer of deer.'

The 'patient nominalizer' *-ipyr* occurs suffixed to transitive verbal bases, indicating that the entity has suffered or suffers an action denoted by the verb. The derived noun is inflected only with the allomorph *i-* of the third person prefix from set NA, which indicates the patient. In this construction, the agent of the action is not expressed, as illustrated in (37), (38) and (39) below. The reference to the patient can be recovered through a lexical item, as illustrated in (39).

(37) ***i-nopỹ-pyr-a*** *a-xa'ja a-ka-wo*
3.NA-beat-NMLZ-RF 3.A-cry 3.CO-be-CVB
'The one who was beaten is crying.'

(38) *are rãka ara-'o* *kwaxi-ø* ***i-kotok-pyr-a***
we R.PST 1EXCL.A-ingest coati-RF 3.NA-poke-NMLZ-RF
'We ate the coati that was speared.'

(39) *marãxe'i-ø* ***i-moon-ipyr-a***
Marãxe'i-RF 3.NA-paint-NMLZ-RF
'Marãxe'i was painted.' (Lit: 'Marãxe'i was the painted one.')

The nominalizing prefix *emi-* derives nouns from transitive verbal bases. The result from the derivation is a noun that maintains the same valence, so that both participants of the event continue to be expressed. In this kind of construction, the agent is expressed as a genitive phrase, indicating the possessor. The reference to the patient is done by the nominalizing prefix *emi-*, as illustrated in (40), (41) and (42) below. It is noteworthy that, whereas the 'patient nominalizer' *-ipyr* demostes the agent, the 'resultative nominalizer' requires its expression as the genitive possessor.

(40) ***xe=r-emi-py-kwer-a*** *konomĩ-ø epe a-aka*
1SG.NA=LK-NMLZ-wrap-NPST-RF boy-RF DEM 3.A-be
'The boy I wrapped is there.'

(41) *xere=ø-ypy-ø* *agỹ-ø* ***r-emi-ãpa-ø*** *xawie*
1INCL.NA=LK-first-RF PL-RF LK-NMLZ-make-RF POS
'[This basket is] equal to the ones made by our elders.'

(42) *ãw'i'i ãpĩ-ø* *i-maãpyk-i* ***t-emi-'o-ø***
little mother-RF 3.NA-cook-2SG.A 3.NA-NMLZ-ingest-RF
'The mother cooked a little of what is for eating.'

5.2 Nominalization from other categories

The other types of nominalizations are the circumstantial nominalization and the grammatical nominalization. The base for circumstantial nominalization is adverbial expressions (postpositional phrases and adverbs in general). The suffix *-wãr* occurs only with adverbial expressions, forming names of entities characterized by the circumstance associated with it. In (43), (44) and (45), the suffix *-wãr* occurs directly attached to postpositions, while in (46) it occurs with an adverb.

(43) *ãpĩ-ø a-ãpa-'i ma-ãkyg-a* ***r-opi-wãr-a***
mother-RF 3.A-make-ATT hand-finger-RF LK-PERL-NMLZ-RF
'The mother carefully made the ring.'
(Lit: 'The mother carefully made the one which is on the hand's finger.')

(44) *ere-ma-tarak ke kwe takypy-ø* ***r-e-wãr-a***
2SG.A-CAUS-tear DUB FUT stake-RF LK-POS-NMLZ-RF
'You might tear that which is on the stakes.'

(45) ***karoka~roka-mõ-wãr-a*** *pa ke mĩ a-pa~par rõ'õ*
afternoon~RED-LOC-NMLZ-RF INFER DUB HAB 3.A-go.out~RED N.ASS
'It seems that the afternoon group is leaving.'

(46) ***ãxe'i-wãr-a*** *a-pãw*
yesterday-NMLZ-RF 3.A-finish
'The one from yesterday is finished.'

The suffix *-ama'e* is attached to any type of intransitive predicate: active and stative intransitive verbal predicates and nominal predicates. The form cannot be suffixed to transitive predicates. The nominalized form maintains some finite features. Despite being nominalized, it takes the set A person markers, which are commonly used in independent clauses, as illustrated in (47) and (48). In the case of nominalization with a stative verb as its base, as in (49), or with a nominal predicate as its base, as in (50) and (51), the nominalized form is inflected with prefixes from set NA. The resulting noun from this type of nominalization occurs only with the participants of the third person, indicating that the entity is characterized as an experiencer or as an attribute expressed by the base.

(47) *a-ixãk akoma'e-ø* ***a-yj-ama'e-ø***
3.A-see man-RF 3.A-run-NMLZ-RF
'He saw the man who ran.'

(48) *ã-ow parãxi-ø* ***a-kãxym-ama'e-kwer-a***
1SG.A-find pencil-RF 3.A-disappear-NMLZ-NPST-RF
'I found the pencil which had disappeared.'

(49) *akoma'e-ø* ***i-eew-ama'e-ø*** *n=a-ãpa-j ka-ø*
man-RF 3.NA-lazy-NMLZ-RF NEG=3.A-make-NEG field-RF
'The man who is lazy does not have a cultivated field.'

(50) *konomĩ-ø* ***ø-ewek-ama'e-ø*** *were-ka ewa-'i-ø*
boy-RF 3.NA-belly-NMLZ-RF 3.CC-be worm-ATT-RF
'the boy whose belly has worms'

(51) ***i-pepa-e'ym-ama'e-ø*** *a-manõ*
3.NA-wing-NEG-NMLZ-RF 3.A-die
'the one who does not have a wing died'
(Referring to a chicken attacked by a dog and nursed by children)

It is noteworthy that grammatical nominalization differs from other types of nominalization by the fact that the process does not only transform verbs into nouns, as illustrated in (47), (48) and (49), but also can take a nominal predicate as its base; that is, it can nominalize a nominal predicate, as illustrated in (50) and (51). Due to this last property, grammatical nominalization can also be interpreted as subordination. In the specialized literature, *-ama'e* has received various interpretations: 'relative noun', 'predicate nominalizer' and 'relative nominalization', by Rodrigues (1953, 2001); 'subject nominalization' by Jensen (1998: 542); and as 'relative agent' by Almeida, Irmãzinhas de Jesus and Paula (1983: 32).

Traditionally, all deverbal nominalizations are considered derivational processes, because they produce a change in word-class. However, the regularity and productivity of nominalizations in Apyãwa may let us think that in this language nominalizations are better analyzed as 'word-class changing inflection', as defined by Haspelmath (1995).

5.3 Summary of the nominalizations in Apyãwa

In Section 5 we have presented the nominalizers found in Apyãwa. Our analysis is summarized in Table 2. The first column presents the morpheme form, followed by a classification of its base in the second column. The third indicates if the nominalizer can occur to transitive and/or intransitive verbs. Then, the fourth column indicates the basic semantics. The reader is referred in the last column to examples that illustrate each nominalizer.

Table 2. Nominalizers in Apyãwa

Morpheme	Base	Verb type	Type of nominalizer	Examples
-ár	verb	trans.	agentive	(35)–(36)
-ãw	verb	trans., intrans.	event, instrumentive, location	(31)–(34)
-pyr	verb	trans.	passive	(37)–(39)
emi-	verb	trans.	resultative	(40)–(42)
-wãr	adverbials	N/A	circumstantial	(46)–(45)
-ama'e	predicate	trans.	grammatical	(47)–(51)

Notice that Apyãwa has cognates of the majority of Tupinambá nominalizers, except for the propensive agentive nominalizer and the habitual agentive nominalizer, as can be seen by comparing Table 1 and Table 2.

6. Nominalization in Nheengatú

In this section we describe nominalization in Nheengatú, a language that descends from Tupinambá and has been used for interethnic communication since the 16th century, as explained in Section 2. As result of this long and intense contact with other languages, a large number of innovation has taken place in Nheengatú, including in the patterns of nominalization. This section investigates the nominalizers of 21st century Nheengatú, as spoken along the Upper Rio Negro, by the Baré, Baniwa and Warekena people. We first give the necessary morphosyntactic background (6.1), and then analyze the forms and functions of nominalization (6.2).

6.1 Morphosyntactic background of Nheengatú

In Section 3, we saw that one of the most remarkable characteristics of Tupinambá and Apyãwa is the fact that in these languages both nouns and verbs can occur as predicate; and, with the suffix *-a* 'referrer', they can also occur as arguments. Queixalós (2006: 24) suggested that this property should be seen as a modern evidence of a period of omnipredicativity, in which there was a superclass of predicates. As emphasized by Launey (1994), the lack of differentiation is only in the ability to occur as predicates and arguments, as nouns and verbs have distinct morphological properties.

In the development from Tupinambá to Nheengatú, the language has lost its omnipredicative properties. In consequence, the characteristics that differentiate nouns and verbs have pervaded the syntax: verbs can no longer occur as argument, a function that is now restricted to the class of nouns. Additionally, the suffix *-a* 'referrer' has lost its morphosyntactic functions. The form has been fossilized as part of a large number of roots. For instance, the Tupinambá noun *jaguar* 'jaguar' could be inflected with the suffix *-a*, resulting in *jaguar-a*. In contrast, in Nheengatú the cognate is *jaguara* 'dog', in which the vowel *-a* has become part of the root, and therefore can no longer be interpreted as a referrer.[9]

9. According to Queixalós (2006: 21), in Paraguayan Guarani the form *-a* has also lost its morphological function. These two languages had in common a history of intense contact with other languages.

The loss of the referrer and the restriction of the argument function to nouns have had impacts on the functions of nominalizations. While in Apyãwa and Tupinambá both nouns and verbs require the referrer to occur as arguments; in Nheengatú, nouns do not need any morphological device to occur as arguments. Moreover, in this language, verbs must be nominalized to occur as arguments. This means that in Nheengatú the ability to occur as argument is the primary function of nouns and the predicative function is primary for verbs. In the pairs of examples below, verbs occur as predicates in (52a) and (53a), and, nominalized, they occur as arguments in (52b) and (53b).

(52) a. *maita **u-yupiru** Namuĩ*
how 3SG.A-begin Anamoim
'How did [the community of] Anamoim begin?'

b. ***u-yupiru-sa** kua Namuĩ pe-kua=rã kuri*
3SG.A-begin-NMLZ DEM Anamoim 2PL.A-know=SUB FUT
'The beginning [founding] of Anamoim, you are going to know (now).'

(53) a. *cawa-miri **puranga***
wasp-DIM beautiful
'The little wasp was beautiful.'

b. *a-nheẽ=ntu **i-puranga-sa***
1SG.A-tell=RESTR 3SG.NA-beautiful-NMLZ
'I only say what is beautiful.' [Lit: 'I only say the beauty.']

6.2 Nominalization in Nheengatú

In Sections 4 and 5, we saw that Tupinambá had eight nominalizers, while Apyãwa has six. Nheengatú has reduced this system to just four nominalizers: *-sa(wa)*, *-sa(ra)*, *-wara*, and *waa*. This section contains a description and an analysis of the constructions identified for productive nominalization in Nheengatú. We first describe deverbal nominalizations in 6.2.1, and then nominalizations from other categories in 6.2.2.

6.2.1 *Deverbal nominalizations*

Deverbal nominalizations in Nheengatú are divided into participant nominalizations and event nominalizations. Syntactically, both kinds of nominalizations assume nominal functions, such as head of nominal phrases in argument position and as nominal predicate. Morphologically, event nominalizations and participant nominalizations differ with respect to the number of non-finite features they display.

6.2.1.1 *Participant nominalization.* According to Foong, Grunow-Harsta and Wrona (2011: 3), participant nominalizations refer to first order ontological entities (e.g. people, objects, locations). Consequently, they assume semantic roles such as agents, patients and locatives. In Nheengatú only one nominalization refers to participants: the agentive nominalization *-sa(ra)*. Agentive nominalization occurs in prototypical nominal functions. They can occur not only as core arguments, as illustrated in (54) and (55), but also as nominal predicates, as illustrated in (56).

(54) ***u-puraki-sa(ra)*** *tau-sika*
3CO-work-NMLZ 3PL.A-arrive
'The workers arrived.'

(55) *tau-senui* ***u-yu-mbue-sa(ra)****-ita*
3PL.A-call 3CO-R/R-teach-NMLZ-PL
'They called the students.'

(56) *ae* ***yane-u-mbue-sa(ra)***
he 1PL.NA-3CO-teach-NMLZ
'He is our teacher'.

Participant nominalizations in Nheengatú displays what Givon (2001: 25) considers prototypical adjustments that create noun phrases from finite verbal phrases. The nominalized verb may occur with person prefixes of set NA to indicate the nominal complement, as illustrated in (56) above. Even though they take these prefixes, they also occur with the prefix *u-*, which is obligatory and cannot be changed for any other marker. The form *u-* seems to be a bleached form without pronominal function, at least if we consider it only in a synchronic analysis. In the diachronic perspective, we can hypothesize that the frozen mark *u-* comes from the prefix *o-* that was used in Tupinambá to indicate coreference between the subject of the verb and the determiner of any other noun in the sentence.

Participant nominalizations are lexically restricted in terms of the kind of verbs that can be the bases for the process: only transitive, as shown in (55) and (56) above, and active intransitive verbs, as in (54) above, can be nominalized with the suffix *-sa(ra)* 'agentive nominalizer'. Stative verbs cannot co-occur with the suffix *-sa(ra)*, unless derived with the prefix *mu-* (with nasalization as its allomorph) 'causative'. What actually happens is that when combined with *mu-*, stative verbs become transitive verbs, and thus can occur with *-sa(ra)*. In (57), a grammatical test shows that the non-inflectional stative verb *pinima* 'be colourful' cannot be combined with the suffix *-sa(ra)*. The verb *pinima* has to be derived with the prefix *mu-* 'causative', and then the transitive verb *mpinima* 'to paint' can be nominalized, as illustrated in (58) below.

(57) **pinima-sara*
be.colourful-NMLZ
intended meaning: 'the one who paints'

(58) ***u-m-pinima-sa(ra)*** *u-sika*
3CO-CAUS-be.colourful-NMLZ 3SG.I-arrive
'The painter arrived.'

The agentive nominalizer is used in conservative uses of the language, such as written texts and formal speech. In everyday speech, the form is used mainly by elderly speakers. The younger generation tends to avoid words formed with agentive nominalization; instead they tend to use loanwords from Portuguese. The form *-sara* is very often reduced to *-sa*. This reduced form is homophonous with the suffix *-sa* 'event nominalizer'.

6.2.1.2 *Event nominalizations.* Event nominalization is used to create secondary ontological entities, such as nouns of actions and nouns of states. The form *-sa* 'event nominalizer' can be combined with any kind of verb to create nouns. In (59) the transitive verb *kua* 'to know' is nominalized, forming the noun *takuasa* 'their knowledge'. In (60) the active intransitive verb *vivei* 'to live' is nominalized creating the noun *yaviveisa* 'our way of living'. Stative verbs can also be nominalized, as in examples (61) and (62) below. In (61) the non-inflectional stative intransitive verb *puku* 'be long' is nominalized to create *pukusa* 'length'. In (62) the inflectional descriptive intransitive verb *suri* 'be happy' is nominalized, creating *tarurisa* 'their happiness'.

(59) *ta-su* *ta-pita* ***ta-kua-sa*** *irũ*
3PL.A-go 3PL.A-stay 3PL.A-know-NMLZ COM
'They stay with their knowledge.'

(60) *yane-kultura* *yane-kostume* ***ya-vivei-sa*** *nhaã nunka re-pudei* *re-perdei*
1PL.NA-culture 1PL.NA-custom 1PL.A-live-NMLZ DEM never 2SG.A-can 2SG.A-loose
'Our culture, our traditions, our way of living, these, you can never loose.'

(61) *u-yaxiu* *pituna* ***puku-sa***
3SG.A-cry night be.long-NMLZ
'She cried the whole night.'
(Lit: 'She cried by the length of the night.')

(62) *ta-maã=ntu* *ta-iku* ***ta-ruri-sa*** *rupi*
3PL.A-see=RESTR 3PL.A-be 3PL.NA-be.happy PERL
'They are looking through their happiness.'

Event nominalizations exhibit a mix of both nominal and verbal features. They maintain inflectional properties of the verbal base, such as the pronominal agreement. More specifically, the prefixes of set A are retained without any formal adjustment. For instance, consider the verb *manduai* 'to remember / to think' in (63) and its nominalized counterpart in (64). Both the verb and its nominalized form are inflected with set A prefixes.

(63) ***ya-manduai*** *s-ese*
1PL.A-remember 3SG.NA-POS
'We remember that.' (Lit: "We remember about this.")

(64) *u-mu-aiwa* ***ya-manduai-sa***
3SG.A-CAUS-be.spoiled 1PL.A-remember-NMLZ
'It spoils our thoughts.'

On the syntactic level, action and state nouns behave similarly to simple nouns. To achieve a better understanding of the morphosyntactic properties of nominalizations, we shall compare the action and state nouns to sentences expressing approximately the same information. We will also compare action and state nouns to non-derived nouns (Comrie and Thompson (2007: 344). Examples (65) and (66) allow the comparison between action nouns and sentences. In (65), the transitive verb *kua* 'know' selects two arguments: a subject, *yaneramunha ambira* 'our grandparents', and an object, manifested by the demonstrative *ae*. The prefix *ta-* '3rd person plural (set A)' indicates the agreement between the subject, *yane ramunha ambira* 'our grandparent', and the verb *kua* 'know'. In (66) there are two grammatical nominalizations based on the verb *kua* 'know': *yakuasa* 'our knowlegde' and *takuasa* 'their knowledge'. Note that both the verb *kua* and its nominalized form *kuasa* express the "experiencer" by a set A prefix. For the comparison between simple nouns and nominalized nouns we present example (66). Simple nouns, like *yane-ramunha* 'our grandfather', express their possessors with a set NA prefix, whereas nominalized verbs, such as *ya-kua-sa* and *ta-kua-sa,* must occur with a set A prefix. The asymmetry between simple nouns and nominalized verbs indicates that event nominalizations retain verbal patterns.

(65) *yane-ramunha* *ambira* ***ta-kua*** *ae*
1PL.NA-grand.parent dead 3PL.A-know DEM
'Our grandparents knew that.'

(66) ***ya-kua-sa*** *maye* *yane-ramunha* *ambira* ***ta-kua-sa***
1PL.A-know-NMLZ be.like 1PL.A-grandfather dead 3PL.A-know-NMLZ
'Our knowledge is like the knowledge of our late grandfathers.'

6.2.2 *Nominalization from other categories*

In this section, we analyze Nheengatú nominalizations derived from non-verbal categories: the circumstantial nominalization (formed with *-wara*) and the predicate nominalization (formed with *waa*). In Sections 4 and 5 we saw that in Apyãwa and Tupinambá, nouns can be created from postpositional phrases and adverbs. The same occur withthe suffix *-wara* 'circumstantial nominalizer' from Nheengatú, as illustrated in (67) with a postpositional phrase, and in (68) with an adverbial phrase:

(67) ***kua-rupi-wara*** *te se-ramunha*
DEM-PERL-NMLZ FOC 1SG.NA-grandfather
'My grandfather WAS FROM HERE!'

(68) *yane-anama-ta* ***ike-wara****-ita*
1PL.NA-family-PL here-NMLZ-PL
'Our family is from here.'

Furthermore, the suffix *-wara* can occur with any expression used to indicate the location from where an entity has originated. Locations can be expressed not only by postpositional phrases and adverbs, but also by noun phrases that indicate place nouns. Thus, *-wara* can occur with noun phrases that indicate place, as shown in (69) below.

(69) *kua Werekena-ita* ***São José-wara****, a-sui-wara*
DEM Warekena-PL São José-NMLZ DEM-ABLAT-NMLZ
'These Warekena (people) are from São José, from there.'

Note that the possibility of suffixing *-wara* directly to nouns is an innovation of Nheengatú. As shown in Section 4 and 5, the cognate forms *-swár* from Tupinambá and *-wãr* from Apyãwa can be suffixed only to postpositional phrases and adverbial phrases. In example (70) from Apyãwa, place nouns occur inside a postpositional phrase, and, only then, can be suffixed by *-wãr*. In the pair of examples in (71), we applied a grammatical test to speakers of Apyãwa in order to verify the possibility of combining *-wãr* directly to the place noun. Speakers strongly indicate that the construction formed in (71a) is ungrammatical, and recommend the construction in (71b) instead. As for Tupinambá, it is not possible to perform grammatical tests; however, all the instances of *-swár* presented by Rodrigues (2010: 16), show it suffixed to postpositions.

(70) *maj-a rõ'õ rãkã'e a-o'o orokotãj-pe-wãr-a*
snake-RF N.ASS NARP 3.A-bite Orokotãj-LOC-NMLZ-RF
'The snake bit the one who is from Orokotãwa.'

(71) a. * *Wakiri-ø* *Brasilia-wār-a*
Walkíria-RF Brasília-NMLZ-RF
b. *wakiri-ø* *Brasília-pe-wār-a*
Walkíria-RF Brasília-LOC-NMLZ-RF
'Walkíria is the one from Brasília.'

We can hypothesize that the fact that *-wara* in Nheengatú can be suffixed to place nouns is a change induced by linguistic contact. In Portuguese, some suffixes can be combined with place nouns in order to create nouns that denote the one who comes from a specific place. For instance, *-ense* is suffixed to the place noun *Santa Catarina* for creating the noun *catarinense* 'the one who comes from Santa Catarina'.

Nominalized nouns created by the suffix *-wara* behave similarly to simple nouns. At the syntactic level, they occur in prototypical nominal functions. They can occur as core arguments, such as subject in (72) or object in (73); and also as nominal predicates, as in (67), (68) and (69) above. Additionally, at the morphological level, nouns derived by *-wara,* as any other noun, can be suffixed with a plural marker, as illustrated in (72) and (73).

(72) *aiwã u-pita* ***kui-wara-ita***
CONJ 3SG.A-stay NOW-NMLZ-PL
'Then, the ones from now stayed.'

(73) *ti=u-m-ba* ***isana-wara-ita***
NEG=3SG.A-CAUS-finish Içana-NMLZ-PL
'It didn't finish the ones from Içana.'

On the semantic level, the suffix *-wara* creates nouns that indicate an entity characterized by its spatial origin. By metaphoric extension, the reference to time can also be expressed by the nominalizer *-wara*. In (74) the word *kuxima-wara* means 'an entity that comes from the past', such as elders and old music.

(74) ***kuxima-wara-ita*** *tau-sika* *tau-partiri*
formerly-NMLZ-PL 3PL.A-arrive 3PL.A-share
'The elders used to arrive and share.'

Nheengatú also has a grammatical nominalization. The particle *waa* (and its clitic allomorph *=wa*), which is a cognate of the form *-mbaé* from Tupinambá and *-ama'é* from Apyãwa, occurs with any predicate to create a nominalized form that can occur in a typical noun position as a core argument. In (75) a nominalized noun occurs as subject; in (76) it occurs as object; and in (77) it occurs as the complement of an existential construction. Additionally, nominalized nouns can occur as complement of a genitive phrase, as illustrated in (78) and (79). As any other nominal, nominals created by the nominalizer *waa* can also create a nominal predicate, as illustrated in (80).

(75) ***nhaã yawaka pe iku waa*** *u-sendu-pa yande*
DEM heaven LOC be NMLZ 3SG.A-listen-COMP we
'That one who is in heaven listens to everything from us.'

(76) *ape paa u-maã* ***tau-pui u-iku waa*** *nhaã kaverna-wasu upe*
CONJ REP 3SG.A-see 3PL.A-jump 3SG.A-be NMLZ DEM cave-AUM LOC
'Then, (they say that) he saw the ones who were jumping inside the cave.'

(77) *aikue yepe puranga waa*
EXIST INDF be.beautiful NMLZ
'There was one who was beautiful.'

(78) *ta-sendu até mairame paa nhaã galu u-nheengai waa ora*
3PL.A-listen until when REP DEM rooster 3SG.A-sing NMLZ hour
'(They say that) they listened until the hour at which the rooster crowed'.
[Lit: '(They say that) they listened until the hour of the rooster to crow.']

(79) *aikue yepe taina,* ***tuyu waa*** *raira*
EXIST INDF child old.man NMLZ son
'There was a child, who was the son of the one who was old.'

(80) *aitenhaã taina* ***taira ta-yapi waa***
DEM child son 3PL.A-threw NMLZ
'That child was the son who they threw away.'

Grammatical nominalizations with *waa* have morphological properties of nouns, such as combination with a plural marker. According to Cruz (2015), the grammaticalization of the form *-ita* 'plural' into a suffix is a recent development of Nheengatú and it may be related a deep typological change that the language is suffering due to intense linguistic contact. In (81) and (82) the nominalized nouns are inflected with the plural suffix. Notice that when the particle *waa* receives the plural marker, it loses phonological material and it occurs as a clitic, whose host is the last word of the predicate, which is been nominalized.

(81) *uwiara* ***u-manu=wa-ita*** *ara*
today 3SG.A-die=NMLZ-PL day
'Today is the death day.'

(82) *ai=te paa nhaã pedasu itá-ita maxi posu upe=wa-ita*
3SG=FOC REP DEM piece stone-PL leper well LOC=NMLZ-PL
'(They say that) he becomes that stone, which is in the well of the lepers.'
(Lit: '(They say that) he is those pieces of stone, which are in the well of the lepers.')

The particle *waa* (and its clitic allomorph) creates a grammatical nominalization in which some finite properties are maintained. The nominalized form retains the set A person prefixes , as shown in (83) below. Aspectual markers and tense markers also remain in the nominalized form. In the examples below, the grammatical nominalizations maintain the clitics that indicate aspect: the clitic *wã* for perfective in (83) and the clitic *re* for imperfective in (84). In (85) the particle *kuri* indicates future tense.

(83) *ti=será inde re-manu=wã waa*
NEG=Q you 2SG.A-die=PFV NMLZ
'Was it not you who already died?'

(84) *primeiru paa u-pisika nhaã puranga=re waa*
first REP 3SG.A-take DEM be.beautiful=IPFV NMLZ
'They say that he took that one who was beautiful.'

(85) *aitenhaã kuri puxuera waa u-yapi kuri ui kua kiti*
DEM FUT be.ugly NMLZ 3SG.A-throw FUT cassava.flour DEM ALAT
'The one who is ugly is going to throw flour here.'

Based on the finite properties of nominalizations formed with the particle *waa*, Cruz (2011: 510) analyzed the constructions formed with *waa* as relative clauses. In that publication, the morpheme *waa* was analyzed as having developed into a relativized, even though it is cognate with nominalizers in other Tupí-Guaraní languages. In the present paper we analyze *waa* as a grammatical nominalizer because it creates arguments and cannot establish an assertion. Furthermore, the nominal output of the grammatical nominalization has properties of nouns, such as the possibility of occurrence as an argument and combination with plural markers.

6.2.3 *Summary of nominalizations in Nheengatú*

Our analysis is summarized in Table 3.

Table 3. Nominalizers in Nheengatú

Morphem	Base	verb type	Type of nominalizer	Examples
-ár	verb	trans., intrans.	agentive	(54)–(58)
-aw	verb	trans., intrans.	event	(59)–(66)
-wara	adverbial	n.a	circumstantial	(73)–(74)
waa	predicate	trans., intrans.	grammatical	(75)–(85)

7. Comparison of nominalizations in Apyãwa, Tupinambá and Nheengatú

In this paper, we have compared the nominalizers of three Tupí-Guaraní languages: Tupinambá, Apyãwa, and Nheengatú. The analysis confirms that Tupinambá and Apyãwa are very conservative in relation to Proto-Tupí-Guaraní, while Nheengatú displays a number of innovations. Tupinambá is the most conservative: it displays eight nominalizers. Apyãwa is also conservative: it maintains the majority of the nominalizers, except the habitual agentive nominalizer and the propensive agentive nominalizer. In this language, the agentive nominalizer *-ár* has extended its function to all cases of agentive nominalizations. Apyãwa maintains resultative and patient nominalizers separate. In contrast, Nheengatú has lost half of the nominalizers. This language reduced all participant nominalizers to one form (*-sara*), which continues to contrast with the event nominalizer (*-sa*). Even though Nheengatú has reduced the number of verbal to nominal nominalizations, the nominalizers which create nouns from other categories (i.e. from non-verbal forms) remains without functional modification: *-wara* as 'circumstantial nominalizer', and *waa* as 'grammatical nominalizer'. The comparison of the forms of the nominalizers in these three languages is summarized in Table 4.

The main difference between the innovative Nheengatú language, and the conservative Tupinambá and Apyãwa languages is in the functions they display and their typological implications. More specifically, we hypothesize that the difference between the function of nominalization in Tupinambá and Apyãwa, in contrast to the function of nominalization in Nheengatú, is related to the omnipredicative properties of the conservative languages that have been lost in the innovative one. On the one hand, Tupinambá and Apyawã are omnipredicative. This means that all lexical entries are primarily predicates. In these conservative languages the suffix *-a* 'referrer' must be combined with a lexical entry for it to be able to serve a referential function, and, thus, to occur as argument. On the other hand, Nheengatú has lost most of its omnipredicative properties. In this innovative language, the separation between nouns and verbs is easier to recognize: nouns are the unique class of words that can occur as arguments, while verbs are the class of words with predicate function. The fact that nouns can also occur as predicates is a vestige of the old pattern of the language.

According to Queixalós (2006), omnipredicativity in Tupí-Guaraní languages are in different stages of changes towards non-omnipredicative patterns. As we saw in Section 3, Apyãwa and Tupinambá have a high degree of omnipredicativity: in these languages, both nouns and verbs can occur as predicates and they can occur as arguments when inflected by a referrer. The ability of both nouns and verbs to occur as predicates has been retained in Nheengatú. However, verbs must be suffixed with a nominalizer to occur as arguments.

Table 4. Comparison of Nominalizers in Tupinambá, Apyãwa and Nheengatú

Form	Basic semantics	Tupinambá	Apyãwa	Nheengatú
**-ár*	agentive	Transitive verbs, intransitive verbs	Transitive verbs	Transitive verbs, intransitive verbs
**-βór*	habitual agentive	?		
**-cwér*	propensive agentive	?		
**emi-*	resultative	Transitive verbs	Transitive verbs	
**-pɨr*	patient	Transitive verbs	Transitive verbs	
**-áβ*	event	Transitive verbs, intransitive verbs, stative intransitive verbs	Transitive verbs, intransitive verbs, stative intransitive verbs	Transitive verbs, intransitive verbs, stative intransitive verbs
**-cwár*	circumstantial	Postpositional phrase, adverbial phrase	Postpositional phrase, adverbial phrase	Postpositional phrase, adverbial phrase, noun phrase
**-βaʒé*	predicate	Transitive verbs, intransitive verbs, stative intransitive verbs, Nominal predicates (?)	Intransitive verbs, stative intransitive verbs, Nominal predicates	Transitive verbs, intransitive verbs, stative intransitive verbs, Nominal predicates

In an omnipredicative language, such as Tupinambá and Apyãwa, the function of nominalizers is (a) to create new lexical entries; and (b) to introduce semantic details into the newly created forms; they may indicate participants, events, circumstances, and even aspectual and modal properties. However, it is not their function to make it possible for a lexical entry to occur as argument, since this function is performed by another morpheme: the referrer. As pointed by Queixalós (2006), the referrer indicates that a lexical element (both nominal or verbal) no longer performs the predicate function. Typologically, this kind of morpheme is also found in languages that are not genetically linked to the Tupí-Guaraní linguistic subfamily, such as Austronesian languages. In short, omnipredicative languages need a referrer to make any word (verbal or nominal) an argument. In contrast, in non-omnipredicative languages, such as Nheengatú, nominalizers have a triple function: (a) they create new lexical entries; (b) they introduce semantic details into the created forms, and (c) they allow non-nominal entries to occur as arguments, due to the fact that in this type of language argumenthood is exclusively restricted to nominals.

8. Final remarks

The analysis presented in this paper illustrates how an omnipredicative pattern may develop into non-omnipredicative pattern. According to Queixalós (2006), an initial state is characterized by the primarily predicative nature of all (or almost all) lexical entries. In this stage argumenthood is not in the lexicon, but is instead provided by an inflectional morpheme that is combined with any lexical item that occurs as argument. Note that even in the omnipredicative stage, nouns and verbs may be kept apart in their morphological properties. In this paper two Tupí-Guaraní languages are presented as examples of an omnipredicative stage, Tupinambá and Apyãwa, while one language of the same family is present as an example of a non-omnipredicative language, Nheengatú. The comparison allows one to observe that an omnipredicative language can develop into a non-omnipredicative language (in this case the change was probably catalyzed by contact with non-omnipredicative languages, such as Brazilian Portuguese). As the language become non-omnipredicative, the difference between nouns and verbs becomes more pronounced: the sharing morphology between nouns and verbs becomes less evident; and the referrer loses its functions and disappears either through phonetic bleaching or fusion with the lexical stem.

With respect to nominalization, we can observe a deep change in its forms and functions. In omnipredicative languages, as illustrated by Tupinambá and Apyãwa, the functions of (a) creating new lexical entry and (b) introducing semantic details, is kept apart from the function of (c) enabling a non-nominal to occur as an argument. While (a) and (b) are performed by a number of nominalizers; the function (c) is performed by a referrer. In non-omnipredicative languages, such as Nheengatú and Indo-European languages, the functions (a), (b) and (c) are performed by the nominalizers. At this point, we suggest an epistemological digression: since the majority of theories on linguistics were postulated in relation to non-omnipredicative languages, our view of nominalization was that its functions were necessarily (a), (b) and (c). However, the analysis of omnipredicative languages indicates that typologically we must make a separation between these three functions.

List of standard abbreviations

1	first person	INDF	indefinite
2	second person	LK	linker
3	third person	LOC	locative
ALL	allative	NA	set NA (non-agent like argument)

ATT	attenuative
AUM	augmentative
CAUS	causative
CC	commitative causative
CO	correference
COM	commitative
COMP	completive
CONJ	conjunction
CVB	converb
DIM	diminutive
DUB	dubitative
EXIST	existential
FOC	focus
FUT	imminent future
A	set A (agent-like argument)
IMP	imperative
INFER	inferential
IPFV	imperfective
INCL	inclusive
NARP	non-attested remote past
N.ASS	non-assertive
NEG	negation
NMLZ	nominalizer
NPST	nominal past
PERL	perlative
PFV	perfective
PL	plural
POS	postposition
Q	question marker
RF	referrer
REP	reportative
RESTR	restrictive
R.PST	recent past
R/R	reflexive / reciprocal
S.D	spatial demonstrative
SG	singular
SUB	subordinator

References

Adelaar, Willem. 1997. O sufixo onomático e a preservação das consoantes do fim de raiz em Guarani paraguaio. *ABRALIN: Boletim da Associação Brasileira de Lingüística* 20: 11–19. Maceió: Universidade Federal de Alagoas.

Almeida, A. Irmãzinhas de Jesus & de Paula, Luiz Gouvêa. 1983. *A língua Tapirapé*. Rio de Janeiro: Xerox do Brasil.

Anchieta, José. 1990[1595]. *Artes de gramática da língua mais usada na Costa do Brasil*, facsimile from the first edition. São Paulo: Loyolla.

Anchieta, José. 1977. *Teatro de Anchieta. Originais acompanhados de tradução verificada, introdução e notas pelo Pe.* Armando Cardoso S. J., Vol. III. São Paulo: Edições Loyolla.

Araújo, Antonio. 1618. *Catecismo na língua brasílica*. Lisbon: P. Crasbeeck.

Bessa Freire, José Ribamar. 2004. *Rio Babel: A história das línguas na Amazônia*. Rio de Janeiro: Atlântica.

Borges, Monica Veloso. 2006. Aspectos fonologicos e morfossintaticos da lingua ava-canoeiro (Tupí-Guarani). PhD dissertation, UNICAMP.

Comrie, Bernard & Thompson, Sandra A. 2007. Lexical nominalization. In *Language Typology and Syntactic Description*, Vol. 3: *Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 334–381. Cambridge: CUP. https://doi.org/10.1017/CBO9780511618437.006

Cruz, Aline. 2011. *Fonologia e gramática do Nheengatú: A língua geral falada pelos povos Baré, Warekena e Baniwa*. Utrecht: LOT. <http://www.lotpublications.nl/publish/issues/Cruz/index.html>

Cruz, Aline. 2014. Reduplication in Nheengatú. In *Reduplication in Indigenous Languages of South America*, Gale Goodwin Gómez & Hein van der Voort (eds), 114–140. Leiden: Brill. https://doi.org/10.1163/9789004272415_006

Cruz, Aline. 2015. The rise of number agreement in Nheengatú. *Revista do Museu Goeldi* 10(2). <http://ref.scielo.org/ysgxqk>

Figueira, Luís. 1878[1687]. *Arte de grammatica da língua brasilica*. Lisboa: Miguel Deslandes.

Figueira, Luís. 1880[1621]. *Arte de gramatica da lingua brasilica*. Rio de Janeiro: Lombaerts e C.

Gregório, José. 1980. *Contribuição Indígena ao Brasil*. Belo Horizonte: União Brasileira de Educação e Ensino.

Givón, Talmy. 2001. *Syntax. An Introduction* 2 Vols. Amsterdam: John Benjamins.

Hale, Kenneth. 1983. Walpiri and the grammar of nonconfigurational languages. *Natural Language and Linguistic Theory* 1: 5–47. https://doi.org/10.1007/BF00210374

Haspelmath, Martin. 1995. Word-class changing inflection. In *Yearbook of Morphology*, Geert Booij & Jaap van Marle (eds), 43–66. Dordrecht: Kluwer.

Jensen, Cheryl. 1989. *O desenvolvimento histórico da língua Wayampi*. Campinas: Editora da UNICAMP.

Jensen, Cheryl. 1998. Comparative Tupi-Guarani morphosyntax. In *Handbook of Amazonian Languages*, Vol. 4, Desmond C. Derbyshire & Geoffrey K. Pullum (eds), 480–618. Berlin/New York: Mouton de Gruyter.

Jensen, Cheryl. 1999. Tupí-Guaraní. In *The Amazonian Languages*, Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds), 125–163. Cambridge: CUP.

Launey, M. 1994. *Une grammaire omniprédicative. Essai sur la morphosyntaxe du nahuatl*. Paris: CNRS Editions.

Launey, Michel. 1986. Catégories et opérations dans la grammaire Nahuatl. PhD dissertation, Université de Paris-Sorbonne Paris 4.

Launey, Michel. 2004. The features of omnipredicativity in Classical Nahuatl. *STUF – Language Typology and Universals* 57(1): 49–69.

Leite, Serafim S. J. 2004. *História da companhia de Jesus no Brasil*. São Paulo: Edições Loyolla.

Leite, Yonne. 1990. Para uma tipologia ativa do Tapirapé. Os clíticos referenciais de pessoa. *Estudos em Línguas Indígenas Cadernos de Estudos Linguísticos* 18: 37–56.

Lemaréchal, Alain. 1989. *Les parties du discours. Sémantique et syntaxe*. Paris: Presses Universitaries de France (PUF).

Lemaréchal, Alain. 1991. *Problèmes de sémantique et de syntaxe en palau*. Paris: CNRS.

Lemos Barbosa, Antônio. 1956. *Curso de Tupi Antigo*. Rio de Janeiro: Livraria São José.

Moore, Denny. 2014. Historical development of Nheengatu (Língua Geral Amazônica). In *Iberian Imperialism and Language Evolution in Latin America*, Salikoko S. Mufwene (ed.), 108–142. Chicago IL: University of Chicago Press. https://doi.org/10.7208/chicago/9780226125671.003.0004

Moore, Denny, Facundes, Sidney & Pires, Nádia. 1993. Nheengatu (LGA), it's history, and effects of language contact. *Proceedings of the Meeting of the Society for the Study of the Indigenous Languages of the Americas* 2(4): 93 – 118.

Navarro, Eduardo de Almeida. 1998. *Método moderno de tupi antigo: A língua do Brasil dos primeiros séculos*. São Paulo: Editora Vozes.

Praça, Walkíria. 2007. Morfossintaxe da língua Tapirapé. PhD dissertation, University of Brasilia.

Praça, Walkíria Neiva. 1999. Nomes como predicados na língua Tapirapé. MA thesis, University of Brasilia.

Queixalós, Francesc. 2005. Posse em Katukína e valência dos nomes. In *Novos estudos sobre línguas indígenas brasileiras*, Aryon Dall'igna Rodrigues & Ana Suelly Arruda Câmara Cabral (eds), 177–102. Brasília: Universidade de Brasília.

Queixalós, Francesc. 2010. Grammatical relations in Katukina-Kanamari. In *Ergativity in Amazonia* [Typological Studies in Language 89], Spike Gildea & Francesc Queixalós (eds), 235–284. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.89.10que

Queixalós, Francesc. 2006. The primacy and fate of predicativity in Tupi-Guarani. In *Lexical Categories and Root Classes in Amerindian Languages*, Valentina Vapnarsky & Ximena Louis (eds), 249–288. Bern: Peter Lang.

Queixalós, Francesc. 2001. *Des noms et des verbes en tupi-guarani: État de la question*. Munich: Lincom.

Rodrigues, Aryon Dall'igna, Cabral, Ana Suelly Arruda Câmara, & Carretta Corrêa-da-Silva, Beatriz. 2006. Evidências linguísticas para a reconstrução de um nominalizador de objeto **-mi- em Proto-Tupi. *Estudos da Língua(gem)* 4(2): 21–39.

Rodrigues, Aryon Dall'igna & Cabral, Ana Suelly Arruda Câmara. 2002. *Línguas Indígenas Brasileiras*, Vol. 1: *Fonologia, gramática e história Revendo a classificação interna da família Tupi-Guarani* [Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da ANPOLL], Aryon Dall'igna Rodrigues & Ana Suelly Arruda Câmara Cabral (eds), 327–337. Belém: Editora Universitária Pará.

Rodrigues, Aryon Dall'igna. 1996. Argumento e predicado em Tupinambá. *Boletim da ABRALIN* 19: 57–66.

Rodrigues, Aryon Dall'igna. 1953. Morfologia do verbo Tupi. *Letras* 1: 121–152.

Rodrigues, Aryon Dall'igna. 2010. Estrutura do Tupinambá. In *Língua e culturas Tupi*, Aryon Dall'igna Rodrigues & Ana Suelly Arruda Câmara Cabral, 11–42. Brasília & Campinas: LALI & Nimuendajú.

Rodrigues, Aryon Dall'igna. 2001. Sobre a natureza do caso argumentativo. In *Des noms et des verbes en tupi-guarani: État de la question*, Francesc Queixalós (ed.), 103–114. Munich: Lincom.

Seki, Lucy. 2000. *Gramática do Kamaiurá – Língua Tupi-Guarani do Alto Xingu*. Campinas & São Paulo: Editora da Unicamp & Imprensa Oficial.

Yap, Foon Ha, Karen, Grunow-Harsta, Karen & Wrona, Janick. 2011. Nominalization strategies in Asian languages. In *Nominalization in Asian Languages: Diachronic and Typological Perspectives* [Typological Studies in Language 96], Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona eds, 1–57. Amsterdam: John Benjamins.

Subject index

Recent scholarship has confirmed earlier observations that nominalization plays a crucial role in the formation of complex constructions in the world's languages. Grammatical nominalizations are one of the most salient and widespread features of languages of the Americas, yet they have not been approached as foundational grammatical structures for constructions such as relative clauses and complement clauses. This is due to an imbalance in past scholarship, which has tended to focus on these constructions at the expense of the nominalization structures underlying them. The papers in this collection treat grammatical nominalizations in their own right, and as a starting point for the investigation of their uses in complex grammatical structures. A representative sample of Amerindian languages, with focus on South America, examines properties of grammatical nominalizations such as their multiple functions, their internal and external syntax, and their diachronic development. Among the far-reaching theoretical conclusions reached by the studies in this volume is that the various types of relative clauses recognized in the typological literature are actually no more than epiphenomena arising from the different uses of grammatical nominalizations.

ISBN 978 90 272 0244 4

John Benjamins Publishing Company