Typological Studies in Language 126

Diverse Scenarios of Syntactic Complexity

edited by Albert Álvarez González,
Zarina Estrada-Fernández
and Claudine Chamoreau

John Benjamins Publishing Company

Diverse Scenarios of Syntactic Complexity

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 126

Diverse Scenarios of Syntactic Complexity
Edited by Albert Álvarez González, Zarina Estrada-Fernández and Claudine Chamoreau

Diverse Scenarios of Syntactic Complexity

Edited by

Albert Álvarez González
University of Sonora

Zarina Estrada-Fernández
University of Sonora

Claudine Chamoreau
CNRS (SEDYL/CELIA - CEMCA)

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/tsl.126

Cataloging-in-Publication Data available from Library of Congress.

ISBN 978 90 272 0343 4 (HB)
ISBN 978 90 272 6230 1 (E-BOOK)

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

Table of contents

Introduction

The notion of syntactic complexity

Generally speaking, something is complex when it consists of many interrelated parts, and complexity can be viewed as a matter of the number of parts and interrelations present in a network of connections, in a system. In general, complexity has been characterized as the number and variety of elements and the elaborateness of their interrelational structure (Rescher 1998: 1; Hübler 2007: 10). An increase in complexity thus corresponds, at the most general level, to the increase in hierarchical organization, that is, an increase in the number of hierarchical levels within a system (Givón 2009: 4).

In the case of languages, this means that syntactic complexity increases as linguistic elements are hierarchically grouped into phrases, clauses, and sentences, implying that "more structural units/rules/representations mean more complexity" (Hawkins 2004: 252). This idea that complexity is associated with the level of hierarchical organization within a system can be clearly seen, for instance, in the following operationalizations of syntactic complexity based on:

- the number of rules that operate in the syntax of a language; the more, the more complex (cf. Kortmann & Szmrecsanyi 2012: 9),
- the degree of clausal embedding in a structure or the upper limit allowed by the grammar of a language; the more embedded clauses, the more complex (Karlsson 2009: 192),
- and the number of phrasal nodes that a syntactic unit (e.g., a phrase or a clause) dominates; the more nodes, the more complex (Szmrecsanyi 2004: 1033).

In all cases, syntactic complexity refers to hierarchical relationships of inclusion and subsumption between syntactic elements.

The study of syntactic complexity proposed in this book is concerned with the different ways of combining basic clauses to form more complex structures. It is thus focused on the phenomenon of clause combinations, which results in a multi-clause construction that is part of the different construction types of

https://doi.org/10.1075/tsl.126.int

interclausal connectivity usually found in the world's languages, such as serial verbs, complement clauses, relative clauses, adverbial clauses, clause chains, and coordination. These different clause-combining strategies exhibit different degrees of syntactic integration, the earlier ones showing relatively "tight" syntactic integration between two verbs or clauses, and the later ones "looser" syntactic integration. These strategies are thus usually arranged between one extreme representing a single clause, and the other extreme representing two grammatically distinct clauses, illustrating a continuum of interclausal integration from most integrated to least integrated (Payne 2006: 288).

The type of syntactic complexity that concerns us most in this book involves the embedding of clauses inside other clauses, resulting in complex clauses represented by the combination of an independent clause and one or more dependent clauses. It is commonly assumed that an independent clause is one that is fully inflected and capable of being used in discourse on its own, and a dependent clause is one that depends on some other proposition for at least part of its inflectional information. The usual suspects of clause dependency are found in an integrated intonational contour, relational government in the encoding of arguments and reduced finiteness (Givón 2001: 327).

Presentation of the book

This book represents a logical continuation of two previous books (Comrie & Estrada 2012; Chamoreau & Estrada 2016), also published in the same *Typological Studies in Language* Series and also with a focus on the study of syntactic complexity. While the first one was focused on *Relative Clauses in Languages of the Americas* (TSL 102), the second one was devoted to two important features often observed in relation to complex clauses: *Finiteness and Nominalization* (TSL 113). This third book adopts a more general perspective, extending its domain within the study of syntactic complexity in order to propose a wide range of topics associated with clause combination, and trying to illustrate *Diverse Scenarios of Syntactic Complexity.*

This book was born from the papers presented at the annual *Seminar on Syntactic Complexity* held each November at the University of Sonora in Hermosillo (Sonora, Mexico) from 2015 to 2017, and organized by the research group "Estudios lingüístico-tipológicos y etnoculturales en lenguas indígenas y minoritarias" ("Typological and Ethnocultural Studies on Minority and Indigenous Languages", CA81) from the same university. The main objective of these workshops was to bring together linguists working on syntactic complexity understood as the combination of clauses resulting in complex sentences, in order to study this phenomenon

in a diversity of languages and from a diversity of theoretical perspectives. This workshop and now the book presented here represent a clear indication of the growing interest in understanding how syntactic complexity works and evolves in the world's languages, how syntactic complexity varies across languages, how it is influenced by language contact, how it is acquired.

These two-day workshops allowed for the discussion of several questions about different phenomena related to syntactic complexity in different aspects of (morpho)syntax and from different theoretical perspectives, including both functional and formal/minimalist approaches. The different topics covered in these events included different types of clause combining strategies such as relative clauses, complement clauses, adverbial clauses, serialization, and clausal nominalizations, but also the switch reference systems involved in clause chains, the role of insubordination and the influence of language contact in the development of syntactic complexity as well as the acquisition of complex clauses in child language and the grammaticalization processes leading to this syntactic complexity. Derived from these events, we are pleased to present now the papers collected in this volume, which all investigate issues that are related to syntactic complexity, from both the theoretical and the typological perspectives. This book thus addresses a variety of questions on syntactic complexity from different backgrounds and viewpoints, and based on the study of different languages.

All the works here collected aim at covering a phenomenon related to clause combining in a particular language. The diversity of topics present in this selection of papers is a good illustration of the varied aspects involved in the construction and development of syntactic complexity. For instance, in terms of subordination, the reader will find studies that concentrate on a particular type of subordination like complement clauses (Rojas, Åfarli & Subbarao, Messineo), adverbial clauses (Steele, Guerrero) or cause/reason adverbial clauses (Álvarez), but also papers in which the three major types of subordination (relative, complement and adverbial clauses) are included in the analysis (Estrada, Chamoreau). Even though subordination is better represented, coordination is also part of the topics discussed in this book (Chamoreau, Álvarez). The evolutionary dimension of interclausal connectivity is very well represented since almost all the papers deal with the development of syntactic complexity (the only exception is Steele), either concerning the origins of complex clauses and of clause-combining markers, or regarding the grammaticalization processes leading to these constructions. Since ontogeny is an important developmental domain for understanding the genesis of syntactic complexity, the acquisition of complex clauses in child language is also addressed in this book (Rojas), offering the possibility to see that ontogeny (language acquisition) and diachrony (historical change) can function as parallel developments in the construction of syntactic complexity. In that regard, Rojas and Álvarez, from an

ontogenetic and a diachronic perspective, respectively, show that insubordinated clauses can represent a developmental stage prior to the genesis of complex clauses. The switch-reference systems involved in the combination of clauses are discussed in several works (Davis & Comrie, Guerrero, Álvarez), while some others address how syntactic complexity is shaped and influenced by clausal nominalizations (Estrada), serialization (Messineo), or language contact (Davis & Comrie, Åfarli & Subbarao). All these different topics contribute to cover the "Diverse Scenarios of Syntactic Complexity" promised in the title of this book.

But the diversity is not only present in the topics. The value of this volume is also due to the fact that different perspectives and different languages are proposed in order to explore the phenomenon under analysis. Although most of the papers are more or less explicitly functionalist, two papers are situated within formalist frameworks (Åfarli & Subbarao, and Steele) and another within the framework of Construction Grammar (Rojas). Åfarli & Subbarao argue for the benefits of using a specific subtype of Minimalism in order to explain grammatical mixing in a specific case of long-term language contact, Steele uses a version of Generalized Phrase Structure Grammar (GPSG) to characterize adverbial clauses in Luiseño, and Rojas adopts a constructional approach to complex sentence formation in early language acquisition. Regarding the languages studied in this book, although the Americas are the best represented with studies on languages from Mexico such as Yaqui (Guerrero, and Álvarez), Pima Bajo (Estrada), Purepecha (Chamoreau), and Spanish (Rojas), from Southern California (USA) with Luiseño (Steele), and from Argentina with Toba (Messineo), two other papers are focused on Dakkhini Hindi-Urdu (Åfarli & Subbarao) from Southern India and on Kobon and Haruai (Davies & Comrie) from Highland New Guinea. As such, this book gives the opportunity for readers to expand both their typological and their theoretical knowledge about syntactic complexity in a variety of languages.

Overview of the papers

In the opening paper *Switch-reference in Kobon and Haruai: Areal influences within Highland New Guinea,* John Davies and Bernard Comrie propose an interesting description of the switch-reference systems in Kobon and Haruai, two neighboring but unrelated non-Austronesian languages spoken in the Highlands of Papua New Guinea, in the southwest corner of Madang Province. Although the switch-reference morphologies of the two languages are very different, the authors show that both languages exhibit some striking similarities on a number of structural switch-reference patterns, especially in tracking the referent of the grammatical subject in experiencer constructions and in dealing with overlapping reference.

Since these similarities are going against more widespread cross-linguistic patterns, they propose that the convergence of patterns cannot be plausibly attributable to common ancestry and are therefore most plausibly attributed to language contact, suggesting a long-term contact situation between both languages.

The article *Models of grammar and the outcomes of long-term language contact: Language mixing in Dakkhini* presented by Tor A. Åfarli & Karumuri V. Subbarao, focuses on some grammatical mixing outcomes in Dakkhini Hindi-Urdu (Indo-Aryan), a contact language spoken in the southern states of India, resulting from the long-term language contact between Hindi/Urdu (Indo-Aryan) and Telugu (Dravidian), in which Telugu is assumed to be the matrix language (thus providing the grammatical structure of the language) and Hindi/Urdu the embedded language (thus providing the morphemes). Based on the analysis of embedded questions and complement clauses in the three languages involved, the authors demonstrate that a theory of language mixing such as the Matrix Language Frame Model (Myers-Scotton 1993, 2002) is not adequate to account for Dakkhini-type contact outcomes. They argue that this model was developed to model "online" code-switching situations where the speakers are bilingual, as opposed to long-term contact situations of the Dakkhini type, in which speakers are not necessarily bilingual. The authors thus propose a generative exoskeletal model (labeled "Exoskeletal Frame Model"), which is better suited to explain sustained long-term mixing outcomes (as exemplified by Dakkhini), but can also explain short-term bilingual mixing outcomes as found in "online" code-switching, and even non-mixing outcomes. The key element of this general model of grammar is the assumption of a distinction between an underlying abstract syntactic skeleton and the instantiation of that skeleton by functional and lexical exponents. In the case of Dakkhini, the authors argue that the syntactic skeleton is mainly provided by Telugu, whereas both the functional and lexical exponents are mostly provided by Hindi/Urdu.

In the next paper, *Constructional grounding in emerging complexity: Early* COMP-*que constructions in Spanish acquisition,* Cecilia Rojas-Nieto adopts a constructional grounding view to emerging complexity which assumes that components of complex constructions may be used as free clauses before they are integrated into a complex frame. The author studies a case of complex sentence formation in Spanish early acquisition, proposing the existence of a grounding relation between main clauses with a complement-taking verb and free, insubordinated clause constructions marked by the complementizer *que*. The author provides developmental evidence that supports a constructivist and constructional approach to emerging complexity, in which each constructional piece is learned by itself and used independently before becoming jointly assembled in complex constructions, showing that children have access to and may learn these independent frames from parental models.

In *The predicates of Luiseño clausal adjuncts*, Susan Steele offers a formal treatment of the essential properties of adverbial clauses ('clausal adjuncts' according to Steele's denomination) in Luiseño, a Uto-Aztecan language of southern California, comparing the morphological properties of the words functioning as predicates in each of the three subordinate clause types (relative, completive and adverbial clauses). Her study considers the notional properties (person, number, and temporality) as well as the syntactic potential (final) of a word, all included in its articulated informational structure (the set of attributes and their values containing the notional and combinatory information associated with a word form). Based on a non-morphemic approach, in which no one-to-one match between the values in an informational structure and the formatives that comprise the word is assumed, the author proposes that adverbial clauses in Luiseño are informationally different in ensuring access to their temporal properties, since the temporality of the adverbial clause must be determined relative to that of the clause to which it is adjoined, and they are morphologically different in that this characteristic coexists with properties of form otherwise associated with predicates whose temporal properties are not accessible. The independence or dependence of the adverbial clause's subject relative to its embedding domain is thus explained by the fact that if the subject of a clausal adjunct is not given internal to the clause, it is determined relative to that of the clause to which it is adjoined.

The next paper is devoted to *Adverbial subordinators in Yaqui*. In this paper, Lilián Guerrero explores the complexity of the clause linkage markers introducing adverbial clauses in Yaqui, a Uto-Aztecan language spoken in northwestern Mexico. She identifies two different kinds of adverbial connectives in this language: (i) clause-final markers, which can be bound morphemes without lexical content or postpositions with specific meanings, and (ii) clause-initial markers, which are free particles with lexical content. Considering the expression of the adverbial clause subjects, the author shows that clause-final markers are associated with more integrated adverbial clauses than the ones introduced by clause-initial markers, which adopt the structural features of coordination instead of subordination. Based on the contrasts observed among the adverbial clauses taking final subordinators and those introduced by initial markers, the author suggests that the latter may be new in the Yaqui grammar. Concerning the bound morphemes (*-kai* and *-o*), they can introduce several adverbial clauses including manner, purpose, temporal, concessive and conditional meanings, and seem to pattern like a switch-reference system, since they present a consistent, though not obligatory distribution (*-kai* tends to be used when the main and dependent subjects are the same, *-o* tends to occur when the subjects are different).

The paper *Grammaticalization of the linking devices with ka in Purepecha* by Claudine Chamoreau is interested in the linking properties associated with the

marker (-)*ka* in Purepecha, a language isolate spoken in the highlands of the State of Michoacan, in western Mexico, from both synchronic and diachronic perspectives. In synchrony, the different uses, forms, and positions of the linking devices marked by *ka* are presented. The author shows that both free and bound forms of this marker can be used at the initial position of the linked unit for conveying different coordinating functions between phrases and between clauses, whereas in subordinating functions, two markers *-ka* are involved: one that is suffixed to the dependent clause initial element in order to form relative, complement or adverbial subordinators, and another that is obligatory suffixed to the final verb indicating the dependent subjunctive mood. In diachrony, the author proposes two different grammaticalization pathways: one deriving the subordinating suffix *-ka* used in initial subordinators from the free coordinator *ka* and another deriving the subjunctive suffix *-ka* from the assertive mood marker *-ka* used for speech-act participants.

In *Syntactic nominalizations in Pima Bajo: Diachronic diversity,* Zarina Estrada-Fernández analyses the syntactic nominalizations involving the use of three different suffixes (*-dam*, *-kɨg*, and *-ka*) in Pima Bajo, a Uto-Aztecan language from northwestern Mexico, spoken in the central part of the states of Chihuahua and Sonora. She shows that the nominalization constructions marked by these markers, which can function as complement, relative and adverbial clauses, have a mixed status in which nominal and verbal properties are combined, illustrating that syntactic nominalizations correspond to a scalar and gradual phenomenon along a continuum. Considering the presence or absence of five different features in the syntactic nominalizations marked by *-dam*, *-kɨg*, and *-ka*, the author proposes a hierarchical organization of these syntactic nominalizations from more to less nominalized constructions, associated with different degrees of nominalization. Regarding the diachronic dimension involved in these nominalizations, the author explores the origin of the nominalizing suffixes *-ka* and *-kɨg,* and proposes that the suffix *-ka* comes from a stative suffix associated with the posture verb *kaat* 'to be lying' and that the suffix *-kɨg* has its diachronic origin in the combination of the stative suffix *-ka* and the emphatic demonstrative *hɨgai*.

In her study on *Syntactic complexity and grammaticalization in Toba language (Guaycuruan)*, Cristina Messineo analyses the grammaticalization paths originated from two different types of multi-verb constructions in this language spoken in the province of Chaco (Argentina): (i) complement clauses with modal and phasal verbs, and (ii) serial verb constructions with motion verbs. In the first case, the process of grammaticalization involves a change of the grammatical status of the phasal or modal verb, in which the phasal verbs develop into aspectual auxiliaries (inceptive or resultative) or into words with a prepositional function. The modal verbs evolve into different kinds of modal auxiliaries (epistemic, deontic, immediate future) or into an aspectual marker (ingressive). In the second case, serial

verb constructions with motion verbs illustrate the grammaticalization paths from directional verbs to progressive aspect auxiliaries, from a locative verb to a durative aspect auxiliary, and ultimately constitute the origin of directional and locative suffixes in the language.

The last contribution, *From discourse to syntax: The use of the discourse marker* bwe *in the creation of interclausal connectives in Yaqui* by Albert Álvarez González shows how Yaqui has recruited the element *bwe,* a discourse connective of discontinuity (that is, a spoken discourse marker that introduces a topic-shift), in order to participate in the creation of two new interclausal connectives that also correspond to thematic reorientation devices: the cause/reason adverbial connective *bwe'ituk* and the adversative connective *bweta.* Despite the two markers having the same thematic reorientation function, the author points out the differences between their evolutionary paths, by comparing the old adversative and cause/reason adverbial clauses and the new constructions marked by the *bwe-* connectives. In the case of *bweta,* the recruitment of *bwe* is very recent. Here, the marker has an increase in subjectivity and expressively reinforces the original adversative meaning conveyed by the old conjunction *t(ep)a,* without any further syntactic rearrangements and structural differences between the new and the old constructions. In contrast, the genesis of *bwe'ituk* is less recent and its formation implies an evolutionary process of explicitness-driven maturation associated with more overt complexity and less intersubjective/pragmatic meaning. In addition, there is also a process of insubordination, a syntactic rearrangement regarding the encoding of the adverbial clause subject (from zero in cases of same subjects or accusative marking for different subjects to nominative marking in both cases), and another rearrangement of the position of the connective within the adverbial clause (from clause-final to clause-initial position). The first syntactic rearrangement has caused the loss of the old switch-reference system and a change from more to less clause integration, whereas the second syntactic modification is clearly in accordance with the *bwe* function as a discourse topic-shifter, which always occupies the chain-initial position.

References

Chamoreau, Claudine & Estrada-Fernández, Zarina (eds). 2016. *Finiteness and Nominalization* [Typological Studies in Language 113]. Amsterdam: John Benjamins.

Comrie, Bernard & Estrada-Fernández, 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

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. https://doi.org/10.1075/z.146

Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199252695.001.0001

Hübler, Alfred W. 2007. Understanding complex systems. *Complexity* 12(5): 9–11. https://doi.org/10.1002/cplx.20178

Karlsson, Fred. 2009. Origin and maintenance of clausal embedding complexity. In *Language Complexity as an Evolving Variable*, Geoffrey Sampson, David Gil & Peter Trudgill (eds), 192–202. Oxford: OUP.

Kortmann, Bernd & Szmrecsanyi, Benedikt. 2012. Introduction: Linguistic complexity second language acquisition, indigenization, contact. In *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*, Bernd Kortmann & Benedikt Szmrecsanyi (eds), 6–34. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110229226

Myers-Scotton, Carol. 1993. *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: OUP.

Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198299530.001.0001

Payne, Thomas E. 2006. *Exploring Language Structure*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511806483

Rescher, Nicholas. 1998. *Complexity: A Philosophical Overview*. New Brunswick NJ: Transaction.

Szmrecsanyi, Benedikt. 2004. On operationalizing syntactic complexity. In *Le poids des mots. Proceedings of the 7th International Conference on Textual Data Statistical Analysis Louvain-la-Neuve, March 10–12, 2004*, Vol. 2, Gérard Purnelle, Cédrick Fairon & Anne Dister (eds), 1032–1039. Louvain-la-Neuve: Presses Universitaires de Louvain.

PART I

Syntactic complexity and language contact

Switch-reference in Kobon and Haruai

Areal influences within Highland New Guinea

John Davies* and Bernard Comrie**, ***
* SIL International / ** Max Planck Institute for Evolutionary Anthropology / *** University of California Santa Barbara

Kobon and Haruai are two neighboring unrelated Papuan languages with similar syntactic typology. Within the overall space of cross-linguistic variation, their switch-reference systems are remarkably similar, though not identical, for instance strictly tracking the referent of the grammatical subject. Kobon uses same-subject marking when the referent of the controller is properly included in that of the target, and also in the inverse configuration when both noun phrases are of the same grammatical person; different-subject marking is only used when the referent of the target is properly included in that of the controller and the two noun phrases are of different persons. Haruai has essentially the same system, but allows more variation to express subjective assessment of degrees of coreferentiality.

Keywords: switch-reference, Papuan languages, experiencer construction, overlapping reference, language contact

1. Introduction

Kobon [kpw] and Haruai [tmd] are two neighboring but not demonstrably genealogically related non-Austronesian languages spoken in the Highlands of Papua New Guinea, in the southwest corner of Madang Province.[1] Haruai has borrowed

1. An earlier version of this article was presented at the workshop "Switch-reference: State of the Art and Where to Go from Here?" within the Annual Meeting of the Societas Linguistica Europaea, Split, Croatia, in September 2013, and as part of an external evaluation at the Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany in November 2013. We are grateful to all those who participated in the ensuing discussions, and also to anonymous readers. Both authors share joint responsibility for the article, but fieldwork on Kobon was carried out exclusively by John Davies [JD], on Haruai by both Bernard Comrie [BC] and JD. Since the kinds

https://doi.org/10.1075/tsl.126.01dav

a substantial amount of vocabulary from Kobon, no doubt in part in response to word taboo (Comrie 1988, 2000). The morphologies of the two languages are very different. They share striking similarities in overall syntactic typology. Before the establishment of government control in the area from the mid-1960s, there is no direct evidence of the nature of contact between Kobon and Haruai. Genetic studies (Bhatia et al. 1989; Jenkins 1987) suggest that the Haruai are more recent arrivals to the area – though by no means recent in absolute terms – and this presumably initiated a long period of chronologically varying degrees of contact between the two communities and languages; at least at present, it is not possible to say anything more detailed than this.

2. Switch-reference in Kobon and Haruai: The basics

Both Kobon and Haruai have switch-reference systems, whereby the verb of a medial clause is marked according to whether its subject is the same as (coreferential with) or different from that of the final clause on which it is dependent.[2] We therefore refer to the subject of the final clause as the controller of switch-reference, the subject of the medial clause as its target. In both languages, the use of same-subject marking is frequent. In Haruai, the use of different-subject marking is also frequent; in Kobon, it competes with use of the final verb form – this was true even in traditional Kobon as first encountered by JD in the 1970s but has become increasingly more so over the years.

---

of examples relevant to this article are often rare in available corpora, we have supplemented textual examples (identified as such) with elicited material.

2. When three or more clauses are joined by means of switch-reference, the antepenultimate and preceding clauses may be treated as dependent either on the immediately following clause ("local" switch-reference) or as dependent on the final clause ("global" switch reference) (Comrie 1998: 427–431). Example (24), the only one in this article to have more than two clauses linked by switch-reference, illustrates local switch-reference. The clause on which another clause is dependent can be referred to as that clause's matrix clause.

Kobon (Davies 1981: 185)[3]

(1) *Yad kaj pak-em, ram ud ar-nab-in.*
I pig kill-ss.1sg house take go-fut-1sg
'I will kill a pig and take it to the house.'

(2) *Yad kaj pak-nö, ne ram ud ar-nab-ön.*
I pig kill-ds.1sg you house take go-fut-2sg
'I will kill a pig and you will take it to the house.'

Haruai

(3) *An hön-a pal-ön, rag ram-a du-öl-a.*
we pig-foc kill-ss carry house-foc go-prs(1pl)-decl
'We killed a pig and took it to the house.'

(4) *Naŋ kwöi-a nöl-man, nuŋ rag ram-a du-a.*
you sweet_potato-foc give-ds.2sg he carry house-foc go(prs.3sg)-decl
'You gave [him] a sweet potato and he took it to the house.'

2.1 Switch-reference morphology in Kobon and Haruai

Both Kobon and Haruai index the person-number of the subject, and only the subject, in verb forms, with further details as given below.

Kobon: The paradigms are given in Table 1 for same subject (ss), different subject (ds), and, for comparison with final verb forms, remote past (rempst) and perfect (prf). The ss and ds forms do not distinguish tense-aspect-mood (TAM). Note that basically all paradigms make the same 8-way person-number distinction, except that ds merges 2sg and 3sg, and ss merges 3sg and 3pl.

3. In the literature on switch-reference, the terms "medial" and "final" are often restricted to clauses showing a relation of switch-reference, but this is not the practice in the broader field, e.g. in discussions of Turkish or Japanese, where "final" means any clause that could appear as an independent clause, "medial" any clause whose structure shows that it is not final; we use the terms here in this broader sense. We prefer to avoid the terminology "(non)finite", which brings with it its own problems, as discussed for instance in the contributions to Nikolaeva (2007). For one approach to finiteness in Haruai, see Comrie (2016).

Table 1. Kobon verb morphology (selected)

		SS	DS	REMPST	PRF
SG	1	*-em*	*-nö*	*-nö*	*-bin*
	2	*-mön/-ön*	*-ö*	*-na*	*-ban*
	3	*-öm*	*-ö*	*-a*	*-öp*
DU	1	*-ul*	*-lo*	*-lo*	*-bul*
	2/3	*-mil*	*-lö*	*-lö*	*-bil*
PL	1	*-un*	*-no*	*-no*	*-bun*
	2	*-mim*	*-be~ -pe*	*-be~ -pe*	*-bim*
	3	*-öm*	*-lö*	*-la*	*-bal*

Haruai: Final verb forms make a 5-way person-number distinction, as illustrated in Table 2. Different person-numbers sharing the -Ø ending are distinguished by different allomorphs of the tense suffix; the future (FUT) paradigm of *ab* 'throw' is given as an example, with the tense morpheme *-(ö)n*.

Table 2. Haruai finite verb morphology (selected)

		Finite	FUT
SG	1	-Ø	*ab-n* (phonetically: *abɨn*)
	2	*-ö*	*ab-n-ö* (phonetically: *abɨnö*)
	3	-Ø	*ab-ön*
PL	1	*-ŋ*/-Ø	*ab-n-ŋ* (phonetically: *abɨniŋ*)
	2/3	palatalization	*ab-öñ*

SS has a single realization, as in Table 3.

Table 3. Haruai same-subject verb morphology

SS	*-ön*

DS has different realizations according to whether the time reference is non-future or future. In the non-future, as shown in Table 4, a 3-way person-number distinction is made, with merger of 3SG, 1PL, and 2/3PL.

Table 4. Haruai non-future different-subject verb morphology

DS.NFUT	SG	1	*-mön*
		2	*-man*
	Other		*-m*, *-mn* (the latter phonetically: *-mɨn*)

In the future, DS distinguishes singular from plural only, as in Table 5.

Table 5. Haruai future different-subject verb morphology

DS.FUT	SG	*-aŋ*
	PL	*-ŋ*

With a first person singular subject, *-mön* can also be used, alongside *-aŋ*, with future time reference. To simplify glosses, for the non-future DS is used, for the future DS.FUT; in the non-future, person-number is indicated for the 1SG and 2SG, not for the form that merges all other combinations.

3. Switch-reference in Kobon and Haruai: What is tracked?

In both Kobon and Haruai, the grammatical subject of a clause can be identified by means of the person-number indexing of the verb of that clause.

Roberts (1997: 161–174) notes that in many languages of Papua New Guinea with switch-reference, where there is a difference between the grammatical subject and the highest participant in terms of agentivity or topicality, it is the participant highest in agentivity/topicality whose reference is tracked, although he does not give detailed statistics. Reesink (1983: 235) suggests that this is the norm. Both Kobon and Haruai differ from the pattern described by Roberts and Reesink in that they strictly track the grammatical subject in all constructions. In work on switch-reference in Papuan languages, including Roberts and Reesink, the exceptions most often cited to strict subject control of switch-reference are experiencer constructions, to which we turn in this section. The experiencer construction in Kobon and Haruai is the one shown in (5) (for Kobon) and (7) (for Haruai). We refer to the noun phrase in the position of 'we' in (5) as the Experiencer, the noun phrase in the position of 'thing' in (5) as the Stimulus. Turning then to the behavior of experiencer constructions in switch-reference: In Kobon Example (6), only a final verb form is possible (neither SS nor, perhaps surprisingly, DS). In Haruai Example (8), the DS marker is used (and a final verb form is also possible), while the SS marker is impossible. Thus, both languages share the impossibility of using the SS marker in this configuration, though they differ in the details of the actual marking found, Kobon allowing only a final verb form, Haruai preferring DS marking.

Kobon (Davies 1981: 100, 101)

(5) *Hon nan g-öp.*
we thing do-PRF.3SG
'We are ill (lit. Thing does us).'

(6) *Aig nan g-öp, ar-ag-nab.*
Aig thing do-PRF.3SG go-NEG-FUT.3SG
'Aig is ill and will not go.'

Haruai

(7) *Nɨ kiö pɨl-a.*
I hunger shoot(PRS.3SG)-DECL
'I am hungry (lit. Hunger shoots me).'

(8) *Hömlö römal rag ur nɨm-mɨd-mɨn, kiö pɨl-öŋ-a.*
banana cut carry cook eat-IPFV-DS hunger shoot-PST(3SG)-DECL
'[He] cut, brought, cooked, and ate some bananas, and felt hungry.' [textual example]

Another difference between Kobon and Haruai is that Kobon can use the SS marker (as well as DS and a finite verb) where the two Stimuli can be judged to be coreferential, as in (9) – whatever causes him to be hungry also causes me to be hungry; Haruai does not allow this, as exemplified in (10).

Kobon

(9) *Nɨpe kiö l-öp/l-öm/l-ö, yad abe kiö l-öp.*
he hunger put-PRF.3SG/-SS.3SG/-DS.3SG I also hunger put-PRF.3SG
'He is hungry and I am hungry too.'

Haruai

(10) *Nuŋ kiö pɨl-mɨn/*pɨl-ön, nɨ aipam kiö pɨl-a.*
he hunger shoot-DS/*-SS I also hunger shoot(PRS.3SG)-DECL
'He is hungry and I am hungry too.'

4. Switch-reference in Kobon and Haruai: Overlapping reference

Where the referent of one subject is properly included in the referent of the other, different languages of Papua New Guinea with switch-reference systems have different solutions to whether such combinations are considered SS or DS. Table 6 is adapted slightly from Roberts (1997: 158). In this table, *Target* is the person-number of the subject of the medial clause (the one that is marked SS or DS), *Controller* is the person-number of the final clause. In all examples in section 4, overlapping reference is presupposed.

Table 6. Switch reference and overlapping reference in selected Papuan languages

Target	Controller	Alamblak Angave Irumu	Amele Ono Suena	Nend Waskia	Kewa	Kobon	Usan
1PL	1SG	SS	SS	SS	SS/DS	SS	SS
1PL	2/3SG	SS	SS	DS	DS	SS	SS/DS
1SG	1PL	SS	DS	SS	SS/DS	SS	SS
2/3SG	1PL	SS	DS	DS	DS	DS	DS

While Alamblak, Angave, and Irumu consistently use SS for all configurations, the following factors intervene in the other languages:[4]

a. Direction of inclusion: In all cases where relevant, inclusion of the controller in the target facilitates use of SS; this is the only relevant factor for Amele, Ono, and Suena, and interacts with factor (b) in Kobon and Usan.
b. Grammatical person: Having both subjects of the same grammatical person facilitates use of SS; this is the only relevant factor in Nend, Waskia, and Kewa, and interacts with factor (a) in Kobon and Usan.

4.1 Overlapping reference in Kobon

Kobon, as described by Davies (1981: 190–199) and as confirmed by JD in work with native speakers in June 2013, has more fully the system as set out in Table 7.

Table 7. Switch reference and overlapping reference in Kobon

Target	Controller	
1PL	1SG	SS
1PL	2/3SG	SS/F
2PL	2SG	SS
2PL	3SG	SS/F
3PL	3SG	SS
1SG	1PL	SS/F
2/3SG	1PL	DS/F
2SG	2PL	SS/F
3SG	2PL	DS/F
3SG	3PL	SS/F

4. We leave out of account whatever may be responsible for the choice between alternants in the entry "SS/DS"; to the extent that this is known, reference should be made to Roberts (1997) and the original references cited there.

The occurrence of switch-reference forms is thus governed by the following principles:

a. If the controller is included in the target and both are of the same grammatical person, only SS is allowed, as in (11).

(11) *Uri nöp hol bɨ mɨhau ram ap-ul, yad nöp*
now indeed we.DU man two house come-SS.1DU I you.ACC
pak-pin.
hit-PRF.1SG
'Just now we two came to the house and I hit you.'

b. If the controller is included in the target and they are of different grammatical persons, SS is allowed, and so is a final verb form, as in (12).

(12) *Uri nöp hol bɨ mɨhau ram ap-ul/au-bul, ne yɨp*
now indeed we.DU man two house come-SS.1DU/-PRF.1DU you I.ACC
pak-pan.
hit-PRF.2SG
'Just now we two came to the house and you hit me.'

c. If the target is included in the controller and both are of the same grammatical person, SS is allowed, and so is a final verb form, as in (13).

(13) *Yad kaj pak-em/pak-nab-in, hon ram ud ar-nab-un.*
I pig kill-SS.1SG/-FUT-1SG we house take go-FUT-1PL
'I will kill a pig and we will take it it to the house.'

d. If the target is included in the controller and they are of different grammatical persons, either DS or a final verb form is allowed, as in (14).

(14) *Nɨpe hainö kaj pak-ö/pak-nab, hon ram ud ar-nab-un.*
he later pig kill-DS.3SG/-FUT-3SG we house take go-FUT-1PL
'Later he will kill a pig and we will take it to the house.'

4.2 Overlapping reference in Haruai

Comrie (1998) describes Haruai as having the same system with respect to occurrence of SS and DS as Kobon. More detailed work by JD and BC in April–May 2013 sets out a more nuanced picture, with the following generalizations emerging primarily from carefully controlled elicited data.

a. If the controller is included in the target, only SS is allowed, irrespective of the grammatical persons of the noun phrases involved, provided that no third person participant is included in the subject of the target clause without also being included in the subject of the controller clause.

(15) *An nöbö hogw-a hön-a pal-ön, nɨ rag ram-a du-m-a.*
we man two-FOC pig-FOC kill-SS I carry house-FOC go-PST(1SG)-DECL
'We two (i.e. you and I) killed the pig and I took it to the house.'

(16) *An nöbö hogw-a hön-a pal-ön, nagö rag ram-a du-m-ö.*
we man two-FOC pig-FOC kill-SS you carry house-FOC go-PST-2SG
'We two (i.e. you and I) killed the pig and you took it to the house.'

(17) *An nöbö hogw-a hön-a pal-ön, nuŋ rag ram-a du-öŋ-a.*
we man two-FOC pig-FOC kill-SS he carry house-FOC go-PST(3SG)-DECL
'We two (i.e. he and I) killed the pig and he took it to the house.'

In (15) and (16), the subject pronoun 'we' of the target clause is interpreted as 'you and I', so no third person participant is involved. In (17), the subject pronoun of the target clause is interpreted as including a third person participant, i.e. as 'he and I', but that third person participant is also included in (in fact, is identical to) the subject of the controller clause. (See (19) for the case where the third person participant is not included in the subject of the controller clause.)

a′. However, this can be overridden if there are strong reasons to regard the two noun phrases as non-coreferential, e.g. if the joint participants in the first clause are in a taboo relationship, such as son-in-law and mother-in-law, with the result that the events described in the first clause are not viewed as a joint action. Two older speakers allowed only DS in (18); younger speakers were less categorical.

(18) *Möxɨj nɨ yöŋö rɨn-a yax rag hö hör yörakw*
mother_i_1 I with wood-FOC cut carry come separately there
nɨg-mɨn, nuŋ rag ram-a du-a.
put-DS she carry house-FOC go(PRS.3SG)-DECL
'My mother-in-law and I collected and made separate piles of firewood there and she carried it to the house.'

a″. If the situation is basically as under (a) but the subject of the target clause includes a third person participant that is not included in the subject of the controller clause, then either SS or DS is allowed; the choice depends on the extent to which the speaker considers the two referent sets to be the same or different.

(19) *An nöbö ad-a hön-a pal-ön/pal-mɨn, nagö rag ram-a*
we man group-FOC pig-FOC kill-SS/-DS you carry house-FOC
du-m-ö.
go-PST-2SG
'We men (incl. at least one third party) killed the pig and you took it to the house.'

Note that the crucial semantic distinction between (16) and (19) is that 'we' in (16) is interpreted to include speaker and addressee alone, while in (19) it includes at least one third-person referent (minimally 'he and I', in fact plus others given the use of *nöbö ad* 'group of men').

b. If the target is included in the controller, and the two noun phrases are of different grammatical persons, then only DS is allowed.

(20) *Nagö hön-a pal-aŋ, an nöbö ad-a rag ram-a*
you pig-FOC kill-DS.FUT.SG we man group-FOC carry house-FOC
di-n-iŋ-a.
go-FUT-1PL-DECL
'You will kill a pig and we men will take it to the house.'

c. If the target is included in the controller, and the two noun phrases are of the same grammatical person, then at least in elicitation, there is a significant difference depending on whether or not one of the subjects is third person.

c′. If none of the participants is third person, then only SS is used, i.e. for the configuration 'I … we …' where 'we' includes only the speaker and addressee.

(21) *Nɨ hön-a pal-ön, an rag ram-a di-n-iŋ-a.*
I pig-FOC kill-SS we carry house-FOC go-FUT-1PL-DECL
'I will kill the pig and we (incl. addressee(s), but not third parties) will take it to the house.'

c″. If one of the participants is third person, then the situation is much more complex, sometimes with variation allowed in elicitation, sometimes with textual examples contradicting elicitation. Under elicitation usually DS was given, though textual examples of SS in the case of two third person subjects are attested.

(22) *Naŋ hön-a pal-aŋ, ñiŋ rag ram-a di-öñ-a.*
you pig-FOC kill-DS.FUT.SG you.PL carry house-FOC go-FUT(2PL)-DECL
'You will kill the pig and you-all (incl. at least one third party) will carry it to the house.'

(One practical problem in working with constructed examples like (22) is that Haruai *ñiŋ* is both 'you.PL' and 'they', and there is no distinction in verb morphology between 2PL and 3PL.)

(23) *Nuŋ hön-a pal-aŋ, ñiŋ rag ram-a di-öñ-a.*
he pig-FOC kill-DS.FUT.SG they carry house-FOC go-FUT(3PL)-DECL
'He will kill the pig and they will carry it to the house.'

(24) *Möhöd hön yogw rag du nöböhöd-a il öröu nɨg-ɨm,*
o_sister pig bag carry go o_brother-FOC near set put.DS
mö_pɨg-a hön yogw rag du nöbö_pɨg-a il öröu nɨg-ön, hön
y_sister-FOC pig bag carry go y_brother-FOC near set put-SS pig
bli rɨg ur-öi-arɨm.
some stone cook-PRS(3PL)-PTCL
'The older sister took [her] bag of pork and put it down near the older brother, the younger sister took [her] bag of pork and put it down near the younger brother, they cooked some pork on [hot] stones.' [textual example]

Example (24) includes three clauses. In the shift from the first clause to the second, the subject shifts from the older sister to the younger sister, a clear instance of non-coreference, and the DS marker is used. In the shift from the second to the third clause, the subject shifts from the younger sister to the combination of the two sisters (minimally; the brothers might also be included), and this is treated as SS.

Note that where Haruai differs from Kobon, it is through use of the DS marker where Kobon would have a SS marker. This may be connected with a general feature of Haruai, namely the possibility (though rarely exploited in practice) of using a DS marker even with fully coreferential subjects in order to indicate different events, e.g. a major break in time or location (Comrie 1998: 431–432). Kobon does not show this phenomenon. (Neither Kobon nor Haruai, incidentally, shows the inverse phenomenon of using SS with non-coreferential subjects in order to tie things together as a single event.)

Table 8 summarizes the differences between Kobon and Haruai; the information in Table 8 is slightly simplified to avoid any loss of overview, and the text above should be referred to for precise details.

Table 8. Summary of similarities and differences between Kobon and Haruai switch-reference

	Kobon	**Haruai**
morphology	SS distinguishes person-number	SS does not distinguish person-number
	DS distinguishes person-number	DS distinguishes person-number, but with much syncretism
	DS insensitive to tense of matrix clause	DS sensitive to tense of matrix clause
tracked element	tracks grammatical subject	tracks grammatical subject
	can treat Stimuli as coreferential	cannot treat Stimuli as coreferential
	with coreferential Experiencers, only Final verb form allowed	with coreferential Experiencers, allows DS

(*continued*)

Table 8. (*continued*)

	Kobon	Haruai
overlapping reference	if controller is included in target, uses SS (also Final verb if they are of different grammatical persons)	if controller is included in target, uses SS (but DS possible to express distance between participants, or if a third party is involved)
	if target is included in controller and they are of the same grammatical person, SS (or Final verb) is used	if target is included in controller and they are of the same grammatical person, SS is used (but DS is possible if one of the participants is third person)
	if target is included in controller and they are of different grammatical persons, DS (or Final verb) is used	if target is included in controller and they are of different grammatical persons, DS is used

5. Conclusion and prospects

Haruai and Kobon have radically different switch-reference morphologies, but on a number of structural switch-reference patterns they are remarkably close to one another, though not identical. In particular, both languages track the grammatical subject in experiencer constructions, and either have or approach a system for dealing with overlapping reference whereby SS is used when either the controller is included in the target or both noun phrases are of the same grammatical person. In cases of overlapping reference, Haruai follows a pattern basically similar to that found in Kobon, but sometimes differing from Kobon in allowing more variation to express subjective assessment of the degree of coreferentiality between the two noun phrases. These striking similarities, sometimes going against more widespread cross-linguistic patterns as in the case of experiencer constructions, are not plausibly attributable to common ancestry and are therefore most plausibly attributed to language contact.

It would be interesting to compare Kobon and Haruai with other languages of the area. However, this part of Papua New Guinea is still in some respects a linguistic last frontier, and we lack comparably detailed work on the other languages of the area. This includes Haruai's only proven linguistic relative, Hagahai-Pinai. For Kobon's closest linguistic relative, Kalam, there may well be enough material available, thanks to the work of Ralph Bulmer, Ian Saem Majnep, Andrew Pawley, Lyle Scholz, and others, but the relevant analysis remains to be done. The Kalamic languages and the South Adelbert languages form a genealogical grouping within the Madang branch of the Trans-New Guinea family, but despite recent pioneering

work on the Sogeram sub-branch of South Adelbert by Daniels (2015), not enough is yet known of the details of switch-reference to permit adequate comparison. This applies even more so to languages of the Ramu branch of the Lower Sepik-Ramu family. Comparison of Kobon and Haruai with the broader neighborhood must therefore remain a task for future research.

Abbreviations

ACC	accusative
DECL	declarative
DS	different subject
DU	dual
F	final
FOC	focus
FUT	future
IPFV	imperfective
NEG	negation
NFUT	non-future
PL	plural
PRF	perfect
PRS	present (in Haruai: present/recent past)
PST	past
PTCL	particle
REMPST	remote past
SG	singular
SS	same subject

References

Bhatia, Kuldeep, Carol Jenkins, Madhuri Prasad, George Koki & Julie Lomange. 1989. Immunogenetic studies of two recently contacted populations from Papua New Guinea. *Human Biology* 61: 45–64.

Comrie, Bernard. 1988. Haruai verb structure and language classification in the Upper Yuat. *Language and Linguistics in Melanesia* 17(1–2): 140–160.

Comrie, Bernard. 1998. Switch reference in Haruai: Grammar and discourse. In *Productivity and Creativity: Studies in General and Descriptive Linguistics in Honor of E.M. Uhlenbeck* [Trends in Linguistics: Studies and Monographs 116], Marc Janse (ed.), 421–432. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110806946.421

Comrie, Bernard. 2000. Language contact, lexical borrowing, and semantic fields. In *Languages in Contact* [Studies in Slavic and General Linguistics 28], Dicky Gilbers, John Nerbonne & Jos Schaeken (eds), 73–86. Amsterdam: Rodopi.

Comrie, Bernard. 2016. Finiteness in Haruai. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada-Fernández (eds), 71–82. Amsterdam: John Benjamins.

Daniels, Don. 2015. A Reconstruction of Proto-Sogeram: Phonology, Lexicon, and Morpho-syntax. PhD dissertation, University of California, Santa Barbara.

Davies, John. 1981. *Kobon* [Lingua Descriptive Studies 3]. Amsterdam: North-Holland.

Jenkins, Carol L. 1987. Medical anthropology in the Western Schrader Range, Papua New Guinea. *National Geographic Research* 3: 412–430.

Nikolaeva, Irina (ed.). 2007. *Finiteness: Theoretical and Empirical Foundations*. Oxford: OUP.

Reesink, Ger P. 1983. Switch reference and topicality hierarchies. *Studies in Language* 7(2): 215–246. https://doi.org/10.1075/sl.7.2.03ree

Roberts, John R. 1997. Switch-reference in Papua New Guinea: a preliminary survey. In *Papers in Papuan Linguistics No 3* [Pacific Linguistics A-87], Andrew Pawley (ed.), 101–241. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Models of grammar and the outcomes of long-term language contact

Language mixing in Dakkhini

Tor A. Åfarli and Karumuri V. Subbarao
Norwegian University of Science and Technology NTNU
University of Delhi & University of Hyderabad (formerly)

In this paper, we lay the groundwork for an explanation of the language mixing outcomes witnessed in the grammar of contemporary Dakkhini, which is the result of long-term diachronic contact between Hindi/Urdu and Telugu. We argue that an Exoskeletal Frame Model is well suited to account both for Dakkhini-type contact outcomes, and for "online" code-switching outcomes. The key element of our model is the assumption of a distinction between an underlying abstract syntactic skeleton and the instantiation of that skeleton by functional and lexical exponents. In the case of Dakkhini, we argue that the syntactic skeleton is mainly provided by Telugu, whereas both the functional and lexical exponents by and large are provided by Hindi/Urdu.

Keywords: code-switching, Dakkhini, exoskeletal frame, language contact, language mixing

1. Introduction

There is a variety of language contact situations in the South Asian subcontinent. For example, Dakkhini Hindi-Urdu (Indo-Aryan) spoken in the southern states of India, Konkani (Indo-Aryan) and BhalavaLi Bhasha (Indo-Aryan) spoken in the Kannada region, SaurasTra (Indo-Aryan) spoken in the Tamil speaking region, and some Tibeto-Burman languages spoken in Nepal, Tripura, Assam, and Nagaland which have been in constant contact with some Indo-Aryan languages. However, in this article we restrict our attention to the outcomes of some of the changes that took place in the syntax of Dakkhini Hindi-Urdu (Dakkhini, hereafter), which are the mixing outcomes of sustained long-term contact between Hindi/Urdu and Telugu.

https://doi.org/10.1075/tsl.126.02afa

Our ultimate goal is to try to analyze and explain the grammatical mixing outcomes of the long-term language contact that resulted in contemporary Dakkhini. In so doing, we ask if existing theories of language mixing are adequate for that purpose. We examine one such theory in some depth, namely the Matrix Language Frame Model (MLFM) (Myers-Scotton 1993, 2002), and we argue that it falls short of the task, but simultaneously we notice that the MLFM is strictly speaking not intended to account for long-term contact outcomes of the Dakkhini type, mainly because it is designed just to account for "online" mixing outcomes where the speakers are bilingual, often referred to as code-switching (CS). Still, we find this limitation to be a serious shortcoming, because it implies that the MLFM is not general enough as a model of grammar. In fact, a model of grammar should strive to cover all linguistic phenomena that are generated by an I-language, and should not be specially designed to deal with just a subset. Therefore, we conclude that there is an explanatory lacuna with regard to the explanation of grammatical mixing outcomes of sustained long-term contact of the Dakkhini type. In languages of the Dakkhini type, the speakers are not necessarily engaged in "online" CS.

In contrast to the MLFM, we argue that a generative exoskeletal model that we – for the purpose of this article – label an Exoskeletal Frame Model (EFM) seems to be well equipped to account both for sustained long-term mixing outcomes (as exemplified by Dakkhini), and for short-term bilingual mixing outcomes as found in "online" CS. Moreover, since the type of model that we propose belongs to a group of exoskeletal models of grammar that are originally motivated on monolingual grounds (Marantz 1997; Borer 2003, 2005a, b; Åfarli 2007; Lohndal 2014), we suggest that the EFM is a general model that can account for both long- and short-term mixing outcomes, as well as non-mixing outcomes.

The article is organized as follows. In Section 2, we consider existing theories of language mixing outcomes with a focus on the MLFM. In Section 3, we present some basic facts about and data from Dakkhini (3.1 and 3.2), and we discuss the shortcomings of the MLFM in dealing with the Dakkhini data (3.3). Then, in Section 4, we present the alternative EFM, and in Section 5 we try to explain how the EFM can deal with Dakkhini data, using different types of subordination in Dakkhini as a case study. Section 6 concludes the paper.

2. Theories of language mixing outcomes: The MLFM

Existing theories of language mixing outcomes typically focus on different types of CS outcomes, be it single-word CS, or CS of larger chunks, see e.g. Muysken (2000) for a typology. CS may be characterized as a bilingual state where two (or

more) languages are "activated" in an individual speaker simultaneously to produce "online" mixing outcomes.

There exist several prominent mixing/CS theories or approaches to the analysis of mixing/CS, for instance Pieter Muysken's as discussed in e.g. Muysken (2000, 2008, 2012) or Jeff MacSwan's Minimalist approach (MacSwan 1999, 2000, 2005, 2009, 2013). However, here we want to focus on Carol Myers-Scotton's MLFM (Myers-Scotton 1993, 2002; Jake, Myers-Scotton & Gross 2002) as an example of an existing theory of mixing/CS that initially may seem promising for an analysis of Dakkhini, but which, we will argue, still falls short of the task. MacSwan's Minimalist CS theory is not discussed here, but a criticism of his theory is found in Åfarli (2015a), see also Chan (2007).

The following quotation gives the gist of the MLFM (González-Vilbazo & López 2011: 846–847):

> The leading idea of the MLFM is that the participating languages in a code-switching event are utilized asymmetrically. One language is the *matrix language*, providing the grammatical skeleton of the clause and determining restrictions on, e.g. word order, agreement, etc. The other language is the *embedded language*, which provides phrases whose insertion is acceptable to the extent that it does not violate restrictions of the matrix language.

This is illustrated in the following example from Hindi – English CS (from Myers-Scotton 1993: 107–108):

(1) *Some Englishmen* ne[1] *tribal girls* ko phusalaa liyaa.
some Englishmen NOM tribal girls ACC seduce did
'Some Englishmen seduced the tribal girls.'

Here Hindi provides the grammatical skeleton or syntactic frame (e.g., as seen in the SOV word order). Thus, Hindi provides the core inflectional morphemes (e.g., the case and tense markings), but lexical content phrases (in *italics*) are code-switched as whole DP chunks from English. In the terms of the MLFM, Hindi is the host or Matrix language (ML), whereas English is the guest or Embedded language (EL). Notice that the English DP chunks have English number inflection, which means that these DP chunks are so-called EL islands, in MLFM's terminology.

Another example is provided by English – Norwegian CS as found in the immigrant variety American Norwegian (c. 1850–1950) (Haugen 1953; Åfarli 2015b).

---

1. The postposition *ne* in Hindi-Urdu is an ergative case marker which Myers-Scotton (1993) labels as a nominative case marker. It may be noted that Hindi-Urdu is a split ergative language (see Subbarao 2012 for details).

(2) Så *play*-de dom *game*-r.
then play-PAST they game-PL
'Then, they played games.'

Here Norwegian provides the syntactic frame, as seen e.g. from the V2 structure of (2) (English is not a V2 language, but Norwegian is). Inflectional affixes are Norwegian, even on the content morpheme stems code-switched from English (in *italics* in (2)). In American Norwegian, this pattern is remarkably regular: stems may be English, but inflections are Norwegian almost without exception. Again, in MLFM terms, Norwegian is the host or ML, whereas English is the guest or EL.

The mixing patterns in (1) and (2) are both examples of insertional CS in the typology of Muysken (2000), but in (1) whole DP chunks with micro-level nominal inflections (number inflections) belonging to the EL are inserted in the ML structure accommodating ML macro-level case frames, while in (2) only lexemes/stems are inserted, with all inflectional affixes belonging to the ML. Notice that (2) provides a particularly clear illustration of one of the basic principles of the MLFM, namely the System Morpheme Principle, which dictates, among other things, that inflections must come from the ML, a principle that we will see later is problematic in the context of Dakkhini.

The mixing patterns in (1) and (2) are found in different types of mixing/CS situations in different geographical and historical locations around the globe. Therefore, they cannot be seen as patterns resulting from special local circumstances, but rather, they must be seen as patterns resulting from basic properties of language mixing or CS as such, see e.g. Myers-Scotton (1993, 2002), Kamwangamalu (1997), and Åfarli (2015a: 13–18) for examples from different language mixing cases. See also the literature on "mixed languages", e.g. Bakker & Muysken (1995), Bakker (2003), Velupillai (2015: Chapter 3).

The MLFM seems to capture the mixing patterns in (1) and (2) quite well. Consider the following characterization of the MLFM, quoted from Bentahila (1995: 135–6):

> [T]he M[atrix] L[anguage] must set the frame, which involves specifying the required order of the morphemes and providing all syntactically relevant system morphemes; the E[mbedded] L[anguage] will contribute only content morphemes which are set into a basically M[atrix] L[anguage] structure.

The example in (2) exemplifies the predicted pattern perfectly. As for (1), the English DP chunks are EL islands that are inserted in superior DP structures in the ML structure, so the relevant system morphemes that are provided by the Matrix Language (Hindi) are the case inflections only.

Notwithstanding the success of the MLFM in dealing with language mixing as shown in (1) and (2), in the next section we will see some evidence that, for purely empirical reasons, the MLFM does not seem to be a good tool for the analysis of language mixing as exhibited in Dakkhini. We will first give a very brief presentation of Dakkhini and consider some Dakkhini data (Sections 3.1 and 3.2). Then, we will discuss why the MLFM is not an adequate model for the analysis of Dakkhini (Section 3.3). In Section 4, we will also see that the MLFM, in our view, has some basic theoretical shortcomings that make it less adequate as a model of grammar.

3. Dakkhini

3.1 Dakkhini: Basic facts

Dakkhini is a contact language that (today) has features both from Hindi/Urdu (Indo-Aryan) and from Telugu (Dravidian). It is a variety of some Indo-Aryan languages transplanted in the 14th Century and it has a literary tradition, culminating in the 17th Century, after which it ceased to be a literary language (Khan 1974). Today, its grammatical structure is heavily influenced by Telugu, while the lexical and functional items are mainly taken from Hindi/Urdu, see especially Arora (2004) for a book length study, see also Mohiddin (1980), Mustafa (2000), and Subbarao (2012).

Dakkhini is a verb-final language with the unmarked word order SOV. As for the grammatical mixing aspects of Dakkhini, Telugu basically provides the grammatical structure of the language (its "soul"), thus apparently constituting the ML in MLFM terms, whereas Hindi/Urdu provides the morphemes (its "body"), apparently constituting the EL. The rationale for this characterization is that Dakkhini, even though both its lexical and functional morphology is generally taken from Hindi/Urdu, still has the word order and employs almost all the grammatical constructions found in Telugu, such as Backward Control, the Double Dative Construction, participial relative clauses etc., see Arora & Subbarao (1989), Arora (2004), Arora & Subbarao (2004), Subbarao & Arora (1988), Subbarao & Arora (2009), and Subbarao (2012).

We will see in Section 3.3 that the tendency that Hindi/Urdu provides all the morphemes, both functional and lexical, is a problem for an analysis of Dakkhini in MFLM terms.

3.2 Dakkhini: Some examples

The role of Telugu as the ML in Dakkhini is seen in the following examples with embedded questions (=S2 in the examples), adapted from Arora (2004: 12).

(3) Hindi/Urdu (Indo-Aryan): *ki* as Initial Complementizer (IC)
Mujhe kyā patā [$_{S2}$ *ki* rām kab āyega]?
I + DAT what known IC Ram when will-come
'How do I know when Ram will come?'

(4) Dakkhini: *ki* as Final Complementizer (FC)
[$_{S2}$ rām kab ātā ē *ki*] mere ku kyā mālum?
Ram when comes FC I + DAT what known
'How do I know when Ram will come?'

(5) Telugu (Dravidian): *o:* as FC
[$_{S2}$ rāmuDu yeppuDu ostāD *-o:*] nā.ku yēmi telusu?
Ram when comes FC I + DAT what known
'How do I know when Ram will come?'

As can be seen from these examples, Dakkhini and Telugu share the same structure, which is different from that of Hindi/Urdu, even though Hindi/Urdu and Dakkhini share almost all morphemes. Notice in particular that Hindi/Urdu and Dakkhini employ the same embedded complementizer (*ki*), which is different from the one employed in Telugu (*-o:*).

This indicates that function words (like complementizers) in Dakkhini are taken from the EL (i.e. Hindi/Urdu) and are syntactically reanalyzed resulting in a change in position from an IC to an FC to suit the Telugu ML pattern.

Corresponding to the data in (3)–(5), consider *that*-clauses in the three languages under discussion, exemplified in (6)–(8), again adapted from Arora (2004: 12).

(6) Hindi/Urdu: *ki* as IC
Mujhe nahi: patā [$_{S2}$ *ki* si:tā gã:v cali: gayi: hai].
I + DAT NEG known IC Sita village has gone is
'I did not know that Sita has gone to the village.'

(7) Dakkhini: *bol ke* as FC
[$_{S2}$ si:tā gã:v ku cale gayi: *bol ke*] mere ku mālum nai:.
Sita village DAT went away FC I + DAT known not
'I did not know that Sita had gone to the village.'

(8) Telugu: *ani* as FC
[$_{S2}$ sīta u:ri ki wellindi *-ani*] nāku teliyadu.
Sita village DAT went FC I + DAT not known
'I did not know that Sita had gone to the village.'

Hindi/Urdu has the IC *-ki*. In *that*-clauses, as expected, Telugu employs an FC, which is the complementizer *-ani*, literally meaning 'having said'. Dakkhini, surprisingly, does not employ the Hindi/Urdu *ki* in clause final position, nor does it use Telugu *-ani*. Instead, Dakkhini uses an innovation, *bol ke*, a FC, which, like Telugu *-ani*, literally means 'having said'. Apparently, the complementizer *bol ke* is a Hindi/Urdu form calqued from the Telugu complementizer *-ani,* and like its Telugu counterpart it is a FC. The complementizer *bol ke* in Dakkhini performs a variety of functions such as a reason marker, purposive marker, labelling marker in naming, conditional marker etc., like the quotative complementizer does in Telugu and other Dravidian languages.

Interestingly, in (6)–(8) we see that Dakkhini avoids borrowing the Telugu form, but instead Dakkhini uses a Hindi/Urdu form, even though that form is not the form actually used in the corresponding construction in Hindi/Urdu. However, the basic meaning of the complementizer, its FC position, and the functions that it performs are identical in Dakkhini and Telugu. We take this to indicate that not only do Dakkhini and Telugu employ the same underlying skeletal ML frame, but that common underlying frame also has certain abstract syntactico-semantic properties that are reflected in both languages. We will suggest an analysis in EFM terms incorporating this assumption in Section 5.

3.3 Dakkhini and the MLFM

The reason why the MLFM seems initially to be a promising tool for the analysis of Dakkhini is that Dakkhini exhibits an asymmetry between the languages involved (Hindi/Urdu and Telugu) that is reminiscent of the asymmetry between an EL and ML. From the point of view of the MLFM, the problem with Dakkhini is that, as we saw in the previous section, not only some of the lexical content morphemes, but basically all morphemes, both lexical and functional/grammatical, come from the Hindi/Urdu lexicon. This contradicts MLFM's System Morpheme Principle (Myers-Scotton 1993: 83), which dictates that the functional system morphemes must come from the ML, i.e. Telugu in the case of Dakkhini. Actually, what would be expected according to the MLFM is that, if Dakkhini were a "well-behaved" mixing or CS variety, it would employ the Telugu complementizer in (4) and (7), even though the lexical content morphemes are from Hindi/Urdu. As we have seen, that prediction is not borne out.

But does the MLFM really make wrong predictions regarding Dakkhini? The MLFM is designed as a theory of CS, but contemporary Dakkhini is not an "online" CS variety, as witnessed by the fact that its speakers are not necessarily bilingual. In urban areas, most of the speakers use only Dakkhini and only some of them are bilingual. In non-urban areas, the situation is reversed; most of the speakers (up to maximum 75–80%) are bilingual since they studied the Telugu medium in schools, and only very few are monolingual Dakkhini speakers. Thus, in its contemporary incarnation Dakkhini should rather be seen as a monolanguage with a bilingual history that still has left a very visible stamp on the language. Although Dakkhini possibly were a CS variety at some stage of its history, it is now an independent language with features both from Hindi/Urdu and Telugu. Therefore, one could claim that the MLFM does not primarily make wrong predictions as regards Dakkhini. Rather, the MLFM is irrelevant for Dakkhini, and therefore makes no predictions about that language.

However, we want to maintain that this is not a reasonable approach to the problem of the possible relevance of the MLFM to Dakkhini. There are two possibilities as regards CS and the history of Dakkhini. First, if Dakkhini were a CS variety at some point in its history (and we really do not know enough about the history of this language to be sure), the MLFM should be relevant for the analysis by virtue of that fact alone, since then Dakkhini could be seen as some kind of a "fossilized" outcome of CS. Second, even if there is no CS stage proper in the history of Dakkhini, one could still hypothesize that the MLFM could be relevant for Dakkhini, just because Dakkhini seems to exhibit something that we can descriptively recognize as an ML – EL type asymmetry (with the exception that functional system morphemes generally follow the EL, not the ML). Or, to look at the matter from a historical angle, the MLFM may potentially be relevant simply because the language mixing witnessed in Dakkhini must be the outcome of some sort of close contact between the languages involved (i.e. their speakers), even though that contact is not properly characterized as being of a CS variety proper.

In any case, we have argued that the MLFM makes the wrong empirical predictions as regards Dakkhini, and thus we wish to claim that this shows that the MLFM is not general enough since it accommodates only certain types of contact/mixing outcomes. Thus, the problem can be formulated as follows. In Dakkhini there appears to be some kind of ML/EL type asymmetry that calls for explanation, but that asymmetry is not explained by the ML/EL theory par excellence, namely the MLFM. See also the discussion in Boussafara-Omar (2003) of a possibly related problem for the MFLM.

If the MLFM cannot handle the Dakkhini mixing data, what can? In the remainder of this paper, we will argue that an exoskeletal model that we label the *Exoskeletal Frame Model*, i.e. the EFM, has the required generality and is better

suited as an explanatory model both for the Dakkhini mixing data and for "online" CS data, and we suggest that the EFM should be seen as a general model of grammar that is appropriate for monolingual data, as well.

4. An exoskeletal frame model: EFM

4.1 The EFM is a generative competence model

We have demonstrated in the preceding section that the MLFM cannot deal adequately with the Dakkhini data. In addition to this empirical shortcoming concerning Dakkhini type data, we want to point out that from our point of view the MLFM has a serious theoretical shortcoming, as well. The problem is that the MLFM is not a general theory of grammar; instead it is designed as a theory about certain bilingual mixing phenomena. In other words, it is not a so-called null theory (Mahootian 1993). Moreover, the MLFM cannot be said to be a model that deals with speakers' grammatical competence, i.e. their I-languages; thus, it is not properly a generative UG-based theory either (MacSwan 2005, 2009, 2013; González-Vilbazo & López 2011; Åfarli 2015a).

In our view, what is needed is a generative competence model that is a null theory and whose ultimate goal therefore is the explanation of all kinds of bilingual and monolingual phenomena of grammar. The main motivation for this is of course that, because of theoretical parsimony, we do not want to postulate a new theory or model for every class of linguistic phenomena that we encounter. We want a theory/model whose ultimate ambition is to encompass all kinds of grammatical phenomena. We suggest that the EFM (or exoskeletal approaches in general) is a good candidate for being the cornerstone of such a general theory. The EFM is a generative neo-constructional or exoskeletal theory (Borer 2003). Exoskeletal theories can be seen as theories within the Minimalist Program (Chomsky 1995), and therefore they are generative competence theories, but nevertheless they have been developed in opposition to one particular assumption within mainstream Minimalism, namely the lexicalist assumption that syntactic structure is projected from the properties of lexical items. Instead, all exoskeletal analyses share the basic assumption that syntactic structure is generated independently of lexical items (both functional items and lexical content items). Empirically, the most important advantage that exoskeletal approaches have as compared to mainstream lexicalist Minimalism is that they are able to account very neatly for lexical creativity and argument structure flexibility, which is ubiquitous in languages. Notice also, and this is particularly important to us, that exoskeletal theories are motivated on monolingual grounds. Language mixing phenomena have been brought under the

purview of exoskeletal approaches only very recently, see e.g. Grimstad, Lohndal & Åfarli (2014).

We will not go into the monolingual motivation for exoskeletal theories here, as there is already a rich literature dealing with that, see e.g. van Hout (1996), Borer (2003, 2005a, b, 2013), Åfarli (2007), Ramchand (2008), Lohndal (2012, 2014), and Nygård (2013), as well as literature within the Distributed Morphology framework (Halle & Marantz 1993; Marantz 1997, 2013; Harley & Noyer 1999; Embick & Noyer 2007).[2] Instead, we want to focus on how an exoskeletal approach, notably the EFM, can deal with language mixing. In 4.3 we will show that the EFM can deal with "online" mixing data of the CS type, and in Section 5 we will see that it can also accommodate mixing data of the Dakkhini type. Since the EFM is developed on the basis of monolingual data in the first place, we hypothesize that it is a generative null theory in the favored sense, and that it is equally well designed to accommodate both different types of mixing phenomena, as well as monolingual phenomena. However, before we discuss how the EFM deals with different sorts of language mixing, we will, in 4.2 below, take a closer look at the structure and properties of the EFM itself. Notice that we are able to just sketch some of the main features of the EFM here; for more detailed discussion, see the literature referred to.

4.2 The structure and properties of the EFM

Importantly, in exoskeletal theories in general and in the EFM in particular, syntactic structures, in incarnations as syntactic frames or templates with certain abstract syntactico-semantic properties, are generated independently of the concrete lexical and functional items, which are inserted into those structures or frames as exponents at a later stage in the derivation. As mentioned above, this is in contrast to mainstream Minimalist theories (Chomsky 1995; MacSwan 2009, 2013) where syntactic structures are generated based on features that are inherent in the lexical and functional elements that comprise the clause.

A quote from Marantz (2013: 153) about neo-constructional or exoskeletal approaches may illuminate the shift. Marantz claims that current developments in linguistic theory

---

2. In this article, the EFM is presented as belonging to a family of exoskeletal approaches. Our analysis could also presumably be implemented in the more well-known Distributed Morphology framework (Halle & Marantz 1993; Marantz 1997, 2013), since that framework has many of the crucial properties of the EFM and other exoskeletal analyses; the EFM could even be seen as a variant of Distributed Morphology. However, such an implementation is not attempted in this paper. See Alexiadou et al. (2015) for an analysis of language mixing phenomena in a more pronounced Distributed Morphology approach.

> have shifted discussion away from verb classes and verb-centered argument structure to the detailed analysis of the way the structure is used to convey meaning in language, with verbs being integrated into the structure/meaning relations by contributing semantic content, mainly associated with their roots, to subparts of a structured meaning representation.

Marantz' notion of "structured meaning representations" should be taken to correspond to verb frames in van Hout (1996), or templates in Borer (2005a, b), or frames in Åfarli (2007).

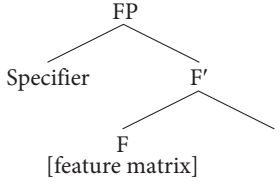
Even though the MLFM and the EFM are fundamentally different as regards their theoretical underpinnings (as discussed in 4.1), the EFM agrees with the MLFM regarding two important assumptions that have to do with the general framework of the analysis, namely (i) that frames, corresponding to MLFM's ML frames, are generated independently of lexical items, and (ii) that lexical insertion, i.e. insertion of exponents, takes place late in the derivation (Late Lexical Insertion).

As compared to mainstream Minimalist approaches, the role of features is toned down in the EFM. Still, we assume that the abstract syntactico-semantic frames contain feature matrices, but then only for the functional projections. Actually, the syntactic frame is taken to be a structural scaffolding for the sentence, consisting of a backbone of abstract functional heads and their projections. Each functional head is composed of functional feature matrices. Functional exponents are inserted in the abstract functional heads, and the functional exponents are designated to strictly correspond to the feature matrix in question. Thus, the functional exponents (functional morphemes, system morphemes) cannot be easily taken from another language than that of the frame, since the feature matrices of the functional backbone typically vary from language to language.

On the other hand, specifier positions and adjunct positions are much more liable to insertion of items generated on the basis of material from other languages, since those positions are not composed of (possibly complex) feature matrices that must match with relevant exponents, but at most must obey certain selection restrictions, for instance that arguments must be nominal.

The relevant structure is exemplified in (9).

(9)



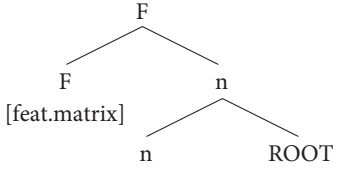
A nice example of a language that inserts whole argument DPs from another language is Michif, a mixed language based on Cree and French (Bakker & Papen

1997). In Michif, DPs are as a rule French (both the noun itself and determiners etc.), whereas the verbal system is Cree. Thus, in Michif, the sentential or verbal frame is Cree with verbal morphology being dictated by the Cree frame, whereas French DPs are inserted in the "free" specifier positions. Notice that *internal* to the DP, the frame is French, thus dictating French functional morphology inside the DP. In MLFM's terms, this amounts to claiming that Cree is the ML, with DPs as EL islands. In ELF's terms, as already indicated, Cree provides the sentential or verbal exoskeletal frame, whereas DPs generated from French nominal frames are inserted into the argument positions of that verbal frame.[3]

We have claimed that the sentential frame, according to the ELF, only comprises functional projections. What about lexical elements, i.e. lexical content item stems? We assume that lexical stems are created from vastly underspecified roots which acquire word class category only after being embedded under a functional categorizer, e.g. n creating noun stems, or v creating verb stems, see e.g. Marantz (1997); Grimstad, Lohndal & Åfarli (2014); Alexiadou et al. (2015). Inflectional functional projections then come on the top of the categorizer projection.

Thus, e.g. a noun has the following structure:

(10)



From a language mixing view, notice crucially that whereas the functional head in F strictly restricts which exponents are relevant, effectively restricting the possible exponent to the designated one that match just that feature matrix in the language, the elements matching the n stem are much more numerous. In fact, any stem exponent bearing a nominal specification is a possible exponent for n, irrespective of language. In other words, any [$_{n}$ n + ROOT] stem structure from any language may provide the starting point for a derivation in any other language. Thus, any noun or verbal stem from any language may in principle be mixed into any other language, but the stems still must be inflected by the exponents of the frame language, see

3. Notice that mixing of the Michif type, where DPs are taken from French (including French functional morphology) and verbs and verb inflectional morphology are from Cree, is predicted to be exceptional for reasons of acquisition burden, since two sets of functional morphology must be mastered. Nevertheless, if the sociolinguistic setting is conductive (i.e. that there are at some stage in the development of the language speakers fluent in both languages involved) such mixing is still possible (as indeed shown by the existence of Michif).

e.g. Grimstad, Lohndal & Åfarli (2014), Grimstad et al. (2014), Åfarli & Jin (2014), Åfarli (2015a, 2015b), Alexiadou et al. (2015), Riksem (2018).

In summary, even though both functional items and contentful lexical stems are inserted into the frame, the restrictions that they have to obey are radically different. The functional exponents of a language belong to a closed pool of items that must strictly match the features of the functional positions where they are inserted. As a main pattern, therefore, functional exponents cannot be made-up or mixed from another language. In contrast, the lexical stems are not subject to any matching requirements apart from word class category. They are just inserted as modifiers of the structure and may therefore be picked from any language, or even be created anew on the spot, as in made-up words. A corresponding free-for-all applies to specifier and adjunct positions, even though the *internal* structure of arguments or adjuncts will contain functional elements that must obey strict restrictions in accordance with the language of the frame of the argument or adjunct.

Notice that, in the EFM, morpheme order, cf. MLFM's "Morpheme Order Principle" (Myers-Scotton 1993: 83), is determined by the independently generated frame (i.e. the ML frame). Thus, morpheme order is given once the frame is generated, and there is no need for an independent and separate "Morpheme Order Principle".

Notice also that MLFM's "System Morpheme Principle" (Myers-Scotton 1993: 83) is not really an independent principle given the EFM analysis. Rather, its effect follows from the matching requirement of functional exponents, i.e., since functional exponents must match the feature matrices of the frame, and since the functional exponents of a language comprise a closed class, it follows naturally that the functional exponents will belong the same language as the frame. No independent and separate principle is needed.

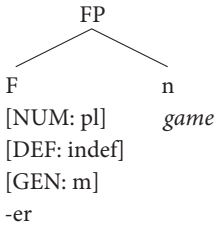
4.3 How the EFM deals with CS type mixing data

Let's see how the EFM system works for mixing phenomena of the CS type. Consider the American Norwegian mixing phenomena illustrated in (2), repeated here:

(2) Så *play*-de dom *game*-r.
then play-PAST they game-PL
'Then, they played games.'

The plural DP *gamer* consists of a contentful English stem and a Norwegian plural affix. It has the partial structure shown in (11) (Notice that the internal structure of the n-stem is not shown here, as it is irrelevant for the point being made).

(11)



The typical exponent for the functional matrix [pl, indef, m] in Norwegian is *-er*, so that particular affix is licensed here, and it is naturally taken from the Norwegian pool of functional exponents, since the frame and its functional feature matrix is generated by the Norwegian I-language. On the other hand, there are no feature matching requirements applying to the lexical stem position apart from a nominal requirement, and any item that the speaker has in his/her mental lexicon may be inserted.[4] In this example, the speaker picks an English nominal stem and inserts it into his/her American Norwegian frame. Therefore, after movement of the lexical n stem to the functional position, we get an English lexical stem that receives Norwegian inflection, which is the typical mixing pattern in American Norwegian (see Åfarli 2015a, 2015b; Grimstad et al. 2014). A comparable argumentation applies to the mixed verb *playde*, where the verbal stem is English and the inflectional tense suffix *-de* is one of the Norwegian past tense suffixes.

Consider now the Hindi – English CS example in (1), repeated here, which illustrates argument DP mixing.

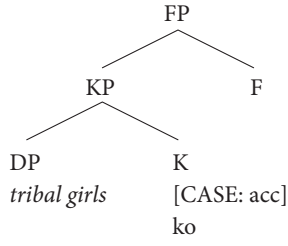
(1) *Some Englishmen* ne *tribal girls* ko phusalaa liyaa.
some Englishmen NOM tribal girls ACC seduce did
'Some Englishmen seduced the tribal girls.'

Here, the English DPs are EL islands in Myers-Scotton's (1993) terms. For instance, the direct object *tribal girls ko* contains an English DP with an English adjective and noun, and the noun has an English plural affix. However, the direct object also contains the Hindi accusative marker *ko*. Therefore, this example is only partially parallel to the Michief example mentioned earlier. In that example, the whole

4. One might wonder why a speaker may want to access the lexical resources of an "alien" language in this way. We believe the answer is simply that the other language may contain more apt lexical items for given purposes (especially in cases where the lexicon of a heritage language like American Norwegian is attrited), or that the speaker wants to spice up his/her utterances for some reason. Surely, using linguistic resources from "alien" languages seems to be very common in all languages. Since there are no formal prohibitions against inserting "alien" items in the open position in the structural frame, it is predicted that such items will be employed whenever the speaker finds it convenient or appropriate.

argument DP, which was French, were substituted into an argument position in a superior Cree frame. In (1), however, the English argument DP is substituted into a Hindi case frame. A possible structure is sketched in (12).

(12)



Notice that the specifier position of KP requires no stricter selection or matching than any other specifier position does, so it is expected that the DP position in [Spec, KP] is in principle open for mixing like any other argument position is.

To round off this section, notice that the EFM analysis that we have sketched above amounts to a type of relexification analysis (Velupillai 2015: 177 ff.), so our proposal may be seen as an implementation of a relexification analysis in an exoskeletal framework. Also, thus far the EFM analysis of data like (1) and (2) is not very different from the MLFM analysis of such data, when considered empirically, even though we have claimed that the EFM has an important theoretical edge over the MLFM because the EFM but not the MLFM is a generative null theory. Thus, one might think that the EFM is equally ill-equipped as the MLFM to deal with the Dakkhini data. We will see in the next section that the EFM is nevertheless a more promising tool for analyzing Dakkhini than the MLFM is, because where the MLFM postulates absolute principles, the EFM can exploit its feature matrix – exponent relation to open up for a natural analysis of the Dakkhini data.

5. An exoskeletal EFM analysis of Dakkhini

5.1 Dakkhini and the EFM

How is the EFM able to deal with the Dakkhini data where lexical content items and functional morphology are both as a rule taken from the same language, which is not the language that provides the syntactic frame? After all, the default empirical predictions of the EFM are the same as of the MLFM, namely that the language of the frame also provides the functional morphology.

The crux of the difference between the MLFM and the EFM is the following. The MLFM postulates an absolute principle (the System Morpheme Principle) that stipulates that the language of the frame provides the inflectional morphology. In

contrast, in the EFM this effect follows from the contingent specification of functional feature matrices and the contingent specification of insertion requirements of exponents, along with a general matching requirement between functional matrices and their functional exponents. Crucially, as we will see, that matching requirement may have different outcomes depending on the constitution of the functional matrix and/or the constitution of the feature requirements of the functional exponent.

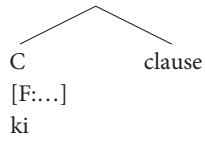
Specifically, if there is a change in the functional matching between a given feature matrix and the corresponding functional exponent for that matrix, that could be due to two different causes (or a combination of the two). First, it could be due to a change in the underlying feature matrix, or, second, it could be due to a change in the insertion restrictions connected to the given functional exponents. We assume that the insertion restrictions of exponents or, in the case of Dakkhini, the exponents themselves are more susceptible to change than the feature matrices belonging to the frame. Specifically, we hypothesize that in sustained long-term contact, there is a tendency that functional exponents for underlying feature matrices may be reconstituted so that existing exponents may receive new insertion criteria or, as seems to be the case in Dakkhini, new exponents with new insertion criteria may be introduced into the language. Actually, Dakkhini seems to have entered a stage where functional exponents from what was originally the guest language (Hindi/Urdu) are constituted as exponents for functional feature matrices in Telugu structural frames. Also, sometimes the new functional exponents employed are plain innovations. Thus, in the case of Dakkhini, the guest language, i.e. Hindi/Urdu, not only provides contentful lexical items, but (most of the) functional exponents as well. We will now show how the Dakkhini data discussed in Section 3.2 can be analysed in this light.

The reconstitution suggested above is assumed to be an effect of sustained long-term contact, and its structural outcome is the development of functional exponents that do not belong to the frame language. We hypothesize that the existence of such non-frame functional exponents in language mixing may be taken as evidence for sustained long-term contact.

5.2 Dakkhini and the complementizer *ki*

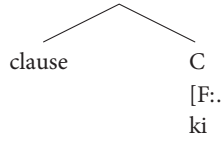
In the embedded questions in Hindi/Urdu, Dakkhini, and Telugu, recall (3)–(5), the complementizer *ki* occurs in clause-initial position in Hindi/Urdu, see (13). Notice that here and further below the exact functional feature matrix of C is not made explicit, since it is not important for the point being made.

(13)

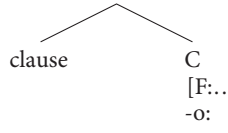


In Dakkhini, the "same" complementizer is clause final, see (14), like the corresponding Telugu complementizer *-o:* is in (15).

(14)



(15)



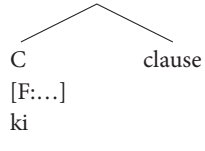
Given the overall structural similarities between Dakkhini and Telugu, there are strong indications that Dakkhini employs Telugu exoskeletal frames (see also Arora 2004), and by hypothesis Telugu functional feature matrices, as well. What has happened, given our analysis, is that Hindi/Urdu *ki* is reconstituted as an FC to match the Telugu functional matrix that is relevant for the complementizer position.

Notice, incidentally, that the reconstitution in question may be quite minor in this particular case, since the feature matrix for the Telugu C-position in embedded questions and the feature matrix for the corresponding Hindi/Urdu position may be identical. Therefore, it is rather a question of choice between the Telugu complementizer and the Hindi/Urdu complementizer. Thus, the choice of the Hindi/Urdu complementizer may be facilitated by the general strength of the Hindi/Urdu lexicon in the case of Dakkhini, e.g. due to the sustained long-term contact.

5.3 Dakkhini and *bol ke*

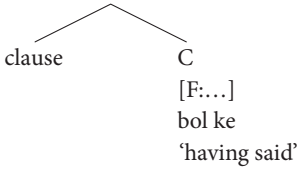
Consider now the next set of Dakkhini data discussed in Section 3.2, namely the *that*-clauses. Again, in Hindi/Urdu, the complementizer *ki* is in initial position, see (16).

(16)

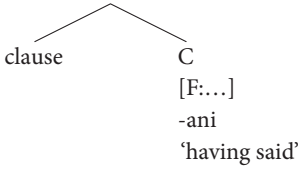


However, Dakkhini employs the "new" complementizer *bol ke* in this case, and *bol ke* is in clause final position, as shown in (17). The corresponding Telugu complementizer is *-ani*, which is also in clause final position, as shown in (18).

(17)



(18)



Again, given the overall structural similarities between Dakkhini and Telugu, we are led to assume that Dakkhini employs Telugu exoskeletal frames, and by hypothesis also Telugu functional feature matrices. However, in this case, it is not a reconstituted Hindi/Urdu complementizer that is used. Rather, a new exponent is created, made from Hindi/Urdu material, to match the Telugu functional matrix in the C-position. Thus, *bol ke* is an innovation of a functional exponent created specifically in Dakkhini.

Interestingly, *bol ke* is a calque in Dakkhini, and it is calqued from Telugu. Both Telugu *-ani* and Dakkhini *bol ke* (in contrast to Hindi/Urdu *ki*) have meanings that can be paraphrased as "having said" (cf. Section 3.2). This suggests that there are common abstract feature properties associated with the C-position of the shared Telugu/Dakkhini matrix frame.[5] Notice crucially that this strongly indicates that the underlying feature matrix of Dakkhini is of the Telugu type, whereas the pool of functional exponents does not easily accept Telugu items. Rather than accepting Telugu functional exponents, Dakkhini either reconstitutes and uses the Hindi/Urdu ones, or creates innovations.

---

5. Notice that the complementizer *bol ke* in Dakkhini, like *-ani* in Telugu, is grammaticalized from lexical content material. The development of lexical content items into functional system items is of course very common in language. Still, what must be noted in this case is that the core of the lexical meaning seems to be carried over into the functional item. This is indicated by the fact that both *bol ke* and *-ani* mean "having said". Thus, in our terms it seems like the abstract feature matrix of the complementizer position in Dakkhini contains the "having said" content in some way (in addition to the purely functional complementizer features), thus facilitating the creation of the *bol ke* calque.

5.4 Wider implications of our analysis

It is in the nature of languages that they vary as to functional categories and morphology. In some languages, verbs have complex inflectional morphology, whereas in others inflectional morphology on the verb is simple. Likewise, the inflection of nouns varies greatly among languages. For a given language (as defined by its functional frame), that language makes available just a very restricted pool of functional exponents, where each member has very specific instructions as to the functional feature matrix that it instantiates. Also, given this variation between languages, acquisition of functional morphology is burdensome, and for that reason alone the speaker tends to stick to the functional exponents of the language of the frame. This makes sense from the point of view of acquisition since in first language acquisition a given frame was acquired simultaneously with a designated set of exponents for that frame. Therefore, it is predicted that in the default case, the functional morphology will belong to the same language as the frame.

Still, the EFM does not formulate this correlation as an absolute principle, because it is in principle possible that the insertion requirements of the exponents are reconstituted, and that the functional morphology is taken from a language different from that of the frame. We have argued that Dakkhini is an instance of this, and we have further argued that this is because the language mixing witnessed in Dakkhini is of a long-term sustained type, where functional exponents of a guest language (as compared to the language of the frame) as well as innovations based on the guest language have been co-opted (along with the lexical stems of that guest language) and reconstituted to fulfill the role of functional morphology. We predict that under the conditions of long-tern language contact, mixing varieties similar to that of Dakkhini should be found.

Our analysis amounts to a typology of language contact mixing outcomes, where short-term "online" mixing of the CS type will typically exhibit the patterns discussed in Section 4.3, whereas long-term contact tend to result in mixing of the Dakkhini type.

6. Conclusion

While the EFM is a restrictive theory, we have tried to show that it has the required theoretical flexibility to explain the Dakkhini data, and in particular, to explain the fact that functional morphology is not necessarily provided by the ML in all cases of language mixing. The key mechanism that allows for the required flexibility is the distinction between an underlying functional feature matrix and the instantiation of that matrix by a functional exponent. Normally, the instantiation of a functional

feature matrix by a functional exponent is by its nature restricted (features must match), as opposed to the insertion of lexical stems or whole phrases, which is free (as explained in Section 4). In cases of long-term sustained contact, however, the restrictions on the insertion of functional exponents may be reconstituted. We have tried to show this to be the case in Dakkhini, where both lexical content stems and functional exponents are either innovations or taken from Hindi/Urdu.

We have also argued that Dakkhini may not be a CS language, but in fact a monolanguage in the sense that its speakers are not necessarily bilingual. Still, a linguistic theory must be sufficiently general to accommodate this kind of mixing languages. We have suggested that the EFM that we have tried to develop is a key component in such a general theory, and that such a theory hopefully can accommodate not only the standard "monolanguages" like Hindi/Urdu, Telugu, English, Norwegian, French, Cree etc., but also "online" CS mixing like in American Norwegian or in Hindi–English mixing outcomes, as well as sustained long-term non-CS mixing, as in Dakkhini.

Regrettably, we have been able to investigate only a tiny sample of Dakkhini mixing in this article. We hope to be able to cast our empirical net wider in future research, both in Dakkhini and other cases of mixing, in order to see if our hypotheses still stand or need revision, or must be outright discarded.

Acknowledgements

First of all, thanks a lot to three anonymous reviewers for very helpful remarks to an earlier version of this paper. Also, many thanks to the members of the EXOGRAM research group at NTNU (Ragnhild Eik, Maren Berg Grimstad, Terje Lohndal, and Brita Ramsevik Riksem) for useful suggestions. Earlier versions of this paper have been presented at events at the University of Southern California (FASAL 3), the Deccan College in Pune (LCI-2013), and the Ambedkar Open University in Hyderabad (special thanks to Prof. Mohammed S. Hayat for organizing the Deccani seminar in October 2012 where we had the fortunate opportunity to present our paper). Thanks to the audiences at each event for useful comments and suggestions.

Abbreviations

ACC	accusative	IC	initial complementizer
CS	codeswitching	ML	matrix language
DAT	dative	MLFM	matrix language frame model
DEF	definite	NEG	negation
EFM	exoskeletal frame model	NOM	nominative
EL	embedded language	NUM	number
F	feature	PAST	past tense
FC	final complementizer	PL	plural
GEN	gender		

References

Åfarli, Tor A. 2007. Do verbs have argument structure? In *Argument Structure* [Linguistik Aktuell/Linguistics Today 108], Eric Reuland, Tanmoy Bhattacharya & Giorgos Spathas (eds), 1–16. Amsterdam: John Benjamins. https://doi.org/10.1075/la.108.04afa

Åfarli, Tor A. 2015a. A syntactic model for the analysis of language mixing phenomena: American Norwegian and beyond. In *Moribund Germanic Heritage Languages in North America*, B. Richard Page & Michael T. Putnam (eds), 12–33. Leiden: Brill. https://doi.org/10.1163/9789004290211_003

Åfarli, Tor A. 2015b. Hybrid verb forms in American Norwegian and the analysis of the syntactic relation between the verb and its tense. In *Germanic Heritage Languages in North America: Acquisition, Attrition and Change* [Studies in Language Variation 18], Janne Bondi Johannessen & Joseph C. Salmons (eds), 161–177. Amsterdam: John Benjamins. https://doi.org/10.1075/silv.18.07afa

Åfarli, Tor A. & Jin, Fufen. 2014. Syntactic frames and single-word code-switching: A case study of Norwegian–Mandarin Chinese. In *The Sociolinguistics of Grammar* [Studies in Laguage Companion Series 154], Tor A. Åfarli & Brit Mæhlum (eds), 153–170. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.154.08afa

Alexiadou, Artemis, Lohndal, Terje, Åfarli, Tor A. & Grimstad, Maren Berg. 2015. Language mixing: A distributed morphology approach. In *NELS 45: Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistic Society*, Vol. 1, Thuy Bui & Deniz Özyildiz (eds), 25–38. Scotts Valley CA: Create Space Independent Publishing Platform.

Arora, Harbir. 2004. *Syntactic Convergence: The Case of Dakkhini Hindi-Urdu*. Delhi: University of Delhi, Publication Division.

Arora, Harbir & Subbarao, Karumuri V. 1989. Convergence and syntactic reanalysis: The case of *so* in Dakkhini. *Studies in the Linguistic Sciences* 19(1): 1–18.

Arora, Harbir, & Subbarao, Karumuri V. 2004. Syntactic change and convergence. In *Non-Nominative Subjects*, Vol. 2 [Typological Studies in Language 61], Peri Bhaskararao & Karumuri V. Subbarao (eds), 25–48. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.60.04aro

Bakker, Peter. 2003. Mixed languages as autonomous systems. In *The Mixed Language Debate. Theoretical and Empirical Advances*, Yaron Matras & Peter Bakker (eds), 107–150. Berlin: De Gruyter.

Bakker, Peter & Muysken, Pieter. 1995. Mixed languages and language intertwining. In *Pidgin and Creoles. An Introduction* [Creole Language Library 15], Jacques Arends, Pieter Muysken & Norval Smith (eds), 41–52. Amsterdam: John Benjamins.

Bakker, Peter & Papen, Robert A. 1997. Michif: A mixed language based on Cree and French. In *Contact Languages. A Wider Perspective* [Creole Language Library 17], Sarah G. Thomason (ed.), 295–363. Amsterdam: John Benjamins. https://doi.org/10.1075/cll.17.12bak

Bentahila, Abdelali. 1995. Review of Duelling languages: Grammatical structure in code switching by Carol Myers-Scotton. *Language* 71: 135–140. https://doi.org/10.2307/415966

Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. In *The Nature of Explanation in Linguistic Theory*, John Moore & Maria Polinsky (eds), 31–67. Stanford CA: CSLI.

Borer, Hagit. 2005a. *Structuring Sense: In Name Only*, Vol. I. Oxford: OUP.

Borer, Hagit. 2005b. *Structuring Sense: The Normal Course of Events*, Vol. II. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199263929.001.0001

Borer, Hagit. 2013. *Structuring Sense: Taking Form*, Vol. III. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199263936.001.0001

Boussafara-Omar, Naima. 2003. Revisiting Arabic diglossic switching in light of the MLF model and its sub-models: The 4-M model and the Abstract Level model. *Bilingualism: Language and Cognition* 6: 33–46. https://doi.org/10.1017/S1366728903001032

Chan, Brian Hok-Shing. 2007. Code-switching, word order and the lexical/functional category distinction. *Lingua* 118: 777–809. https://doi.org/10.1016/j.lingua.2007.05.004

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA: The MIT Press.

Embick, David & Noyer, Rolf. 2007. Distributed morphology and the syntax – morphology interface. In *The Oxford Handbook of Linguistic Interfaces*, Gillian Ramchand & Charles Reiss (eds), 289–324. Oxford: OUP.

González-Vilbazo, Kay & López, Luis. 2011. Some properties of light verbs in code-switching. *Lingua* 121: 832–850. https://doi.org/10.1016/j.lingua.2010.11.011

Grimstad, Maren B., Lohndal, Terje & Åfarli, Tor A. 2014. Language mixing and Exoskeletal Theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* 41: 213–237. https://doi.org/10.7557/12.3413

Grimstad, Maren B., Riksem, Brita Ramsevik, Åfarli, Tor A. & Lohndal, Terje. 2014. The inadequacy of feature-based lexicalist theories: A case-study of American Norwegian. Paper presented at The Fifth Annual Workshop on Immigrant Languages in the Americas, UCLA, Los Angeles, *17–19 October*.

Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20. Essays in Linguistics in Honor of Sylvian Bromberger*, Kenneth Hale & S. Jay Keyser (eds), 111–176. Cambridge MA: The MIT Press.

Harley, Heidi & Noyer, Rolf. 1999. Distributed morphology. *Glot International* 4: 3–9.

Haugen, Einar. 1953. *The Norwegian Language in America. A Study in Bilingual Behavior,* Vol. I & II. Philadelphia PA: University of Pennsylvania Press. https://doi.org/10.9783/9781512820522

van Hout, Angeliek. 1996. Event Semantics and Verb Frame Alternations. PhD dissertation [TILDIL Dissertation Series], Tilburg University.

Jake, Janice, Myers-Scotton, Carol & Gross, Steven. 2002. Making a minimalist approach to codeswitching work: Adding the Matrix Language. *Bilingualism: Language and Cognition* 5: 69–91. https://doi.org/10.1017/S1366728902000147

Kamwangamalu, Nkonko M. 1997. Language contact, code-switching, and I-languages: Evidence from Africa. *South African Journal of Linguistics* 15: 45–51. https://doi.org/10.1080/10118063.1997.9724105

Khan, Masud Husain. 1974. Dakhni-Urdu. In *History of Medieval Deccan* (*1295–1724*), Vol. II, Haroon K. Sherwani & P. M. Joshi (eds), 17–34. Hyderabad: The Government of Andhra Pradesh.

Lohndal, Terje. 2012. Without specifiers. PhD dissertation, University of Maryland, College Park.

Lohndal, Terje. 2014. *Phrase Structure and Argument Structure: A Case Study of the Syntax – Semantics Interface*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199677115.001.0001

MacSwan, Jeff. 1999. *A Minimalist Approach to Intra-sentential Code-switching.* New York NY: Garland.

MacSwan, Jeff. 2000. The architecture of the bilingual faculty: Evidence from intrasentential code switching. *Bilingualism: Language and Cognition* 3: 37–54. https://doi.org/10.1017/S1366728900000122

MacSwan, Jeff. 2005. Codeswitching and generative grammar: A critique of the MLF model and some remarks on 'modified minimalism'. *Bilingualism: Language and Cognition* 8: 1–22. https://doi.org/10.1017/S1366728904002068

MacSwan, Jeff. 2009. Generative approaches to code-switching. In *The Cambridge Handbook of Linguistic Code-Switching*, Barbara E. Bullock & Almeida Jacqueline Toribio (eds), 323–350. Malden MA: Wiley-Blackwell. https://doi.org/10.1017/CBO9780511576331.019

MacSwan, Jeff. 2013. Code-switching and grammatical theory. In *The Handbook of Bilingualism and Multilingualism*, Tej K. Bhatia & William C. Ritchie (eds), 309–335. Cambridge: CUP.

Mahootian, Shahrzad. 1993. A Null Theory of Code-switching. PhD dissertation, Northwestern University.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Pennsylvania Working Papers in Linguistics* 4.2., Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds), 201–225. Philadelphia PA: University of Pennsylvania, Penn Linguistics Club.

Marantz, Alec. 2013. Verbal argument structure: Events and participants. *Lingua* 130: 152–168. https://doi.org/10.1016/j.lingua.2012.10.012

Mohiddin, S. Khader. 1980. *Dakhani Urdu*. Annamalainagar: Annamalai University.

Mustafa, Khateeb S. 2000. *A Descriptive Grammar of Dhakkini*. New Delhi: Munshiram Manoharlal Publishers.

Muysken, Pieter. 2000. *Bilingual Speech. A Typology of Code-Mixing*. Cambridge: CUP.

Muysken, Pieter. 2008. *Functional Categories.* Cambridge: CUP. https://doi.org/10.1017/CBO9780511755026

Muysken, Pieter. 2012. Root/affix asymmetries in contact and transfer: Case studies from the Andes. *International Journal of Bilingualism* 16: 22–36. https://doi.org/10.1177/1367006911403211

Myers-Scotton, Carol. 1993. *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: OUP.

Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780198299530.001.0001

Nygård, Mari. 2013. Discourse ellipses in spontaneously spoken Norwegian: Clausal architecture and licensing conditions. PhD dissertation, Norwegian University of Science and Technology NTNU, Trondheim.

Ramchand, Gillian Catriona. 2008. *Verb Meaning and the Lexicon*. Cambridge: CUP.

Riksem, Brita Ramsevik. 2018. Language mixing in American Norwegian noun phrases. PhD dissertation, Norwegian University of Science and Technology NTNU, Trondheim.

Subbarao, Karumuri V. 2012. *South Asian Languages. A Syntactic Typology*. Cambridge: CUP. https://doi.org/10.1017/CBO9781139003575

Subbarao, Karumuri V. & Arora, Harbir. 1988. On extreme convergence: The case of Dakkhini Hindi-Urdu. In *Language Change and Language Variation*, A. Mukherjee (ed.), 105–122. Osmania University, Hyderabad. (Also in: *Indian Linguistics* 49, 1988).

Subbarao, Karumuri V. & Arora, Harbir. 2009. The conjunctive participle in Dakkhini Hindi-Urdu: Making the best of both worlds. *Indian Linguistics* 70: 359–386.

Velupillai, Viveka. 2015. *Pidgins, Creoles & Mixed Languages. An Introduction* [Creole Language Library 48]. Amsterdam: John Benjamins.

PART II

Syntactic complexity and language acquisition

Constructional grounding in emerging complexity

Early COMP-*que* constructions in Spanish acquisition

Cecilia Rojas-Nieto
Universidad Nacional Autónoma de México / Instituto de Investigaciones Filológicas

A constructional grounding view to emerging complexity is based on the assumption that components of complex constructions may be used as free clauses before they are integrated into a complex frame. In child language development, constructional grounding has already been tested with positive results for some simple frames: passives and existential constructions. This chapter extends this view to complex sentence formation in early acquisition of Spanish. The analysis focuses on data pointing to a possible grounding relation between main clauses with lexical uses of future complement-taking verbs, and free, insubordinated clause constructions marked by COMP-*que*. Children have access to and may learn these independent frames from parental models. Evidence of constructional grounding is presented: Chronological, distributional and functional results point to a grounding relation between free uses of these components and complex sentences that integrate them. The paper adds to synchronic and diachronic analysis of emerging complexity, and brings developmental evidence on how children integrate earlier and independently learned clausal construction frames.

Keywords: complex constructions, constructional grounding, insubordination, free COMP-*que* constructions, complement-taking verbs

https://doi.org/10.1075/tsl.126.03nie

> If it could be demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down.
>
> (Charles Darwin)

1. Introduction

Language acquisition is an important source of evidence for the long-standing debate between determinist and constructivist theories. In determinist theories language is considered to result from dedicated biological endowment, while constructivist theories consider that language emergence is supported by general learning mechanisms, shared intentionality, and pro-social interactional ethology (Elman et al. 1996; Enfield & Sidnell 2014; Levinson 2006; Tomasello 2008).

A current arena in which determinist-universalist and emergentist-constructivist theories confront each other involves the human capacity to build recursive constructions. Recursivity has survived the gradual erosion of Universal Grammar and continues to be considered as a biologically determined universal by the minimalist program. Emergentist studies, instead, aim to account for how recursive constructions, in particular embedded constructions, emerge in language history from discourse to syntax, and are learned in ontogeny (Givón 1979, 2009; Heine & Kuteva 2007; Tomasello 2003).

In line with this constructivist pursuit, this paper will focus on early clause embedding in Spanish acquisition. It relies on the assumption that components of complex clauses may be used as main sentences before they are integrated in complex frames. This developmental pathway to clausal complexity, with "clause integration of previously free, main constructions," is presupposed in Givón's "from discourse to grammar" proposal (1979, 2009). Heine and Kuteva (2007: 214) have also proposed this pathway as one channel for the genesis of syntactic complexity. A similar process, referred as "constructional grounding," is considered to be involved in the acquisition of different constructions (Abbot-Smith & Behrens 2006; Johnson 1999, 2001; Israel, Johnson & Brooks 2000).

Early developmental studies of complex clause constructions were mainly interested in the gradual production of different dependent clauses. The developmental sequence was believed to reflect the relative complexity of the semantic notions coded by dependent clauses. However, only partial affinities between semantic load and developmental emergence were found. These unexpected results induced a search for language-specific effects on learning (Slobin 1985). Lois Bloom and collaborators (1991[1980]) offered an insightful and pioneer account in this line. More recently, inspired by converging research on the genesis and use of complex

constructions,[1] constructivist studies aim at verifying that learning complex constructions is possible. This field of study also aims at finding the conditions and processes that make such learning possible (Bod 2009; Diessel 2004; Diessel & Tomasello 2000a, 2001; Kidd et al. 2007).

Several factors have been shown to be involved in emergent complexity in child language: fine-grained impact with frequency effects of specific constructions in parental models upon a "conservative, attentive learner" (Brandt et al. 2011; Culicover 1999; Tomasello 2003); dialogical scaffolding and "spreading of complexity" among participants (Givón 2009; Ochs, Schieffelin & Platt 1979, inter alia); "starting small" effects related to children's limited processing resources (Elman 1990, 1993; Newport 1990), such as treating matrix verbs as clause modifiers or parenthetical items instead of embedding heads, or learning anchoring frames for dependent clause adjunction (Diessel 2004; Diessel & Tomasello 2000a, 2001; Rojas-Nieto 2009a).

In a general vein, some argue that items and constructions that have already been learned become grounding pieces for successive composite constructions. Such constructional grounding is understood as a process whereby certain uses of simple source construction provide the basis for children's initial hypotheses about a more difficult target construction (Israel, Johnson & Brooks 2000: 103).

This argument, already developed for simpler constructions, would also be expected to apply to complex clause constructions. The rationale in every case is that composite or complex constructions may be grounded on former uses of their component frames. Methodologically, a construction grounding analysis would imply tracing back the history of the components (for embedded clauses, complement-taking verbs and dependent clause frames), and finding how they may develop before becoming embedded combinations.

From this perspective, complement-taking verbs (hereafter CTV) are expected to have former lexical uses and to first adopt phrasal complements. Spanish clauses marked by the complementizer *que* (hereafter COMP-*que*) – prototypical exemplars of dependent clauses – are presumed to be used by children as main sentences before being embedded into a CTV. From this view, both constructions, main free uses of COMP-*que* clauses, and lexical uses of the future CTV, are proposed as grounding sources for corresponding complex constructions of matrix verbs with embedded COMP-*que* clauses.

---

1. Usage-based theory, Dialogic syntax, Grammar in interaction, Cognitive linguistics, Construction grammar, Computational modelling, inter alia, collaborate all together and feed this joint project (Bybee & Hopper 2001; Du Bois 2014; Elman 1990, 1993; Elman et al. 1996; Frank, Bod & Christiansen 2012; Goldberg 1995, 2006; Langacker 2000; Ochs, Schegloff & Thompson 1996).

Guided by these expectations, children's early constructions are analyzed considering the questions in (1).

(1) Research questions
   i. Do children use CTV as lexical verbs with phrasal arguments before embedded clause arguments?
   ii. Do children use COMP-*que* clauses as free, main constructions?
   iii. Is there any evidence (chronological, functional, conversational) to support a grounding relation between lexical uses of CTV and Free COMP-*que* constructions, and the emergence of complex clause constructions?

In what follows, the paper will refer to some previous studies on constructional grounding, and some Spanish-dependent clauses used as main sentences (§2). The corpus and methodological issues will be presented in (§3). Section (§4) is dedicated to the analysis, beginning with an overview of the earlier complex clauses produced by children (§4.1). This section continues with the presentation of the lexical uses of CTV with phrasal objects (§4.2), followed by Free COMP-*que* constructions as main clauses (§4.3). The chronology of the development associated with the proposed grounding constructions is presented in (§4.4). Under the "Bridging the gap" heading (§4.5), evidence is given to support the fact that children trade off between uses of CTV and Free COMP-*que* constructions and embedding frames along the conversation. The paper ends (§5) with some concluding remarks and further questions on emerging complexity from a constructivist, usage-based perspective.

2. Antecedents

2.1 Constructional grounding in acquisition

Earlier research on children's development of composite constructions has carefully explored whether new constructions can be related to former productions. Investigations tracing back children's constructions have shown that nearly all of the new constructions produced by children in a given time can be related to earlier ones by some simple operations such as *add*, *substitute*, *drop*, *insert*, *rearrange* and, later, *blend* (Lieven 2006, 2008; Lieven et al. 2003; Lieven, Salomo & Tomasello 2009). Most new constructions may be related to earlier ones, either as direct replicas, or by one or two operations: observed creativity mainly relies on previous attainments. The relation between previous and new productions has also been successfully tested as the target of computational modeling (Bannard, Lieven & Tomasello 2009).

More specific research on constructional grounding has analyzed particular composite constructions, such as the English or German passive or existential constructions. Such research has traced lexical components and sub-constructions in earlier uses of these constructions: e.g., German *sein* used as a copula, and as auxiliary in *sein*-passives; English adjectival uses of participles, and participles used latter as nuclear verbs in *be*-passive constructions; English *there* in a deictic use, and *there*-existential constructions (Abbot-Smith & Behrens 2006; Dąbrowska & Lieven 2005; Israel, Johnson & Brooks 2000; Johnson 1999, 2001). Results have likewise shown that the distribution and use of the lexical components and sub-frames of those elaborated constructions are compatible with a constructional grounding hypothesis (Ambridge & Lieven 2011; Frank, Bod & Christiansen 2012).

2.2 Spanish antecedents

To my knowledge, constructional grounding analysis has not been applied to Spanish acquisition, let alone considered as a possible path in complex clause development. However, some specific aspects of Spanish grammar and some by-products of developmental studies make it reasonable to expect constructional grounding effects in early clause embedding.

On the one hand, in adult Spanish, some dependent clauses are commonly used as main clauses (Gili Gaya 1983: 115). This phenomenon of *insubordination,* in Nick Evans' terms (2007), is widely spread among different languages (Mithun 2008), and it has received increasing attention in Spanish (Gras 2011, 2013; Montolío 1999; Pons 2003; Porroche 2000, Rodríguez 2008; Sansiñena 2015; Sansiñena, De Smet & Cornillie 2015). Former studies have reported that subjunctive verbs, nonfinite verb constructions – infinitives, gerunds, participles –, and COMP-*que* constructions among others, even though they are expected to occur in subordinated clauses, they may also be employed as free, main sentences (Gili Gaya 1983: 115; Luna-Traill 1980). They are common in adult colloquial speech, and children may frequently experience such marked dependent frames as main sentences.

As for children's productions, they use dependent nonfinite verb forms – mostly infinitives – as main verbs, before using auxiliaries + nonfinite verbs to form complex predicates (***a ver*** > *vamos a* ***ver*** '(let's go) to see', ***a dormir*** > *quiero/voy a* ***dormir*** '(I want/go) to sleep') (Ezeizabarrena 1997, 2002; Freudenthal et al. 2010; Serrat & Aparici 2001). Likewise, it has been found that in conversation, children use early COMP-*que* clause constructions in main sentence position, with no overt controller (Aguirre 2000[1994]; Barreña 1994, 1999).

In fact, prompted and scaffolded by the interlocutor's preceding conversational move, children produce constructions – dependent clauses included – that

complete, expand, or depend upon the interlocutor antecedent construction (Ochs, Schieffelin & Platt 1979; Givón 2009; Köymen & Kyratzis 2014; Rojas-Nieto 2009a, 2009b; Sansiñena et al. 2015). In dialogue, clauses overtly marked for their dependent status may form a complex construction with the clause uttered in the preceding move. They display a dyadic production mode that Givón (2009) considers as the "spreading of complexity".

Spreading of complexity among participants has been well and widely attested in cross-linguistic studies of child language. As for Spanish data, at between 24 and 30 months, children produce causal *porque* 'because' clauses only in answering *¿por qué*? 'why' questions (Aguirre 2000[1994]; Barreña 1999; Rojas-Nieto 1992); clausal conjuncts marked by *pero* 'but' occur in isolation in a conversational turn, forming a sequence with the interlocutor's antecedent move (Varela 2006, 2011); Spanish relative clauses may have a dialogical link without being properly included in a dependent frame (Rojas-Nieto 2009a); Spanish COMP-*que* constructions are conversationally prompted by *Qu*-questions, at least from 2;03 onward (Aguirre 2000[1994]; Barreña 1999). Similar conversational effects have been found in Dutch or German between *waarum* questions and *omdat* responses (Evers-Vermeul 2005; Evers-Vermeul & Sanders 2011; Van Veen et al. 2009, 2013), or in English *why-because* sequences, or *but* constructions (Braunwald 1997; Diessel 2004; Köymen & Kyratzis 2014; Kyratzis, Guo & Ervin-Tripp 1990).

In sum, previous studies show that children have access to models of free uses of dependent clause frames in colloquial Spanish. They also receive dialogical scaffolding to use dependent marked clauses as freestanding constructions in their conversational interventions. From both sides, children have sources from which to learn, and a supportive dialogic context to use dependent clause constructions as free sentences. We may thus suppose children benefit from these situations and produce those frames as free clauses before using them embedded into a CTV to form a whole, fleshed-out complex clause.

3. Method

Based on these antecedents, this paper will explore constructional grounding of complex sentences in early child language. The analysis will be focused on complement clauses marked by COMP-*que* and by CTV that may take clausal arguments as a direct object.

Data have been obtained from the spontaneous productions of six children – four girls (Flor, Natalia, Elia, and Tita) and two boys (Julio and Luis) – from the ETAL database (*Etapas Tempranas en la Adquisición del Lenguaje*-UNAM) (Rojas-Nieto 2007), whose names have been changed to protect their privacy. All of these

children are members of urban, Spanish monolingual, educated middle-class families from Mexico City. The children's spontaneous conversations were collected at home along free and variable interactions with parents or with close members of the family (grandparents, uncles, aunts, cousins). All are only children, except Julio, who has an older brother. Data were collected in a multimodal format – audio and video – at different time paces: five points per week for Natalia; one data point every 10 to 15 days for Tita and Flor, and one recording per month for Elia, Julio, and Luis.

3.1 Data selection

In every case, the corpus was defined in terms of each child's production of the target frames: overt embedded COM-*que* constructions. The upper age cut-off was set when the child produced five or more embedded COM-*que* constructions at a single session. This session was taken as an anchoring point for a trace-back search of the constituent frames: CTV, and Free COMP-*que* constructions. The trace-back search stopped for each child at the earlier time point when evidence of these components was no longer found.

Individual differences in development were found at both ends of the period under inspection, i.e., the anchoring point and the earlier data point. Flor and Natalia had an anchoring point around 2;03–2;04, when both girls were already producing several overtly embedded constructions in the same session. Tita's anchoring point was around 2;07. Elia and Julio had an anchoring point around 3;04–3;05. Likewise, the trace-back search stopped at a specific age for each child (see Table 1 for details).

Table 1. Data sources

Child	Type of corpus	Earlier data point – anchoring point	Number of points & corpus in hours
Nata (fem)	5 per week/45 min each	1;11,28–2;03,00	46 points–34 hrs
Flor (fem)	7–10 days/2 hrs each	2;00,00–2;03,28	14 points–28 hrs
Tita (fem)	10–15 days/2 hrs each	2;00,10–2;08,00	14 points–28 hrs
Julio (male)	Monthly based/2 hrs each	2;04,27–3;04,11	10 points–20 hrs
Elia (fem)	Monthly based/2 hrs each	2;06,22–3;04,27	10 points–20 hrs
Luis (male)	Monthly based/2 hrs each	2;02,05–(…)	10 points–20 hrs

In this search, by the end of his corpus (2;08), Luis had not yet produced any overtly embedded COMP-*que* construction, despite his independent production of CTV with lexical complements and Free COMP-*que* constructions in main clause uses. For this child, it has not been possible to set a critical anchoring point for the trace-back

search. He was dropped from the study, although the distribution of his data does not falsify a constructional grounding account.

3.2 Data presentation

For an easier presentation, the data points in tables and graphs are distributed into three periods identified by the initial of the child's given name plus an ordinal number: e.g., F1, F2, and F3, for Flor; T1, T2, and T3, for Tita, and so on. The length of each period depends upon the extension of each child's overall data. Each time cut-off covers 1 month for Flor and Natalia, about 2 months for Tita, and 3 months for Julio and Elia.

4. Analysis

4.1 Overview of early complex sentences: CTV + COMP-*que* constructions

Most aspects of the constructional grounding of embedded clause constructions present individual differences. However, along the period under study, not considering the child dropped from the study, all remaining five children produced COMP-*que* clausal constructions as an argument of a handful of CTV: *decir* 'tell, say', and *querer* 'wish, want', these were the most frequent ones (Table 2). These complex clauses already expose the usual markers that Spanish requires in dependent clauses: COMP-*que* is rarely missing, subjunctive verb forms are usually present when are required. Some CTV -*querer* 'want', *necesitar* 'need', and *gustar* 'like' - may take infinitive arguments under same subject conditions, and adopt COMP-*que* clauses, with different subjects.

Examples of early complex clauses with different CTV and COMP-*que* constructions are included below. Contrary to English complement constructions in which COMP omission is usual (Diessel 2004; Diessel & Tomasello 2000b), COMP-*que* is present in every case (see COMP in the corresponding glosses), and subjunctive verb forms are already present (2a), (3e), (4d), (5d), and (6c).[2]

2. Transcription conventions are: items in CAPITALS = voice elevation; (…) = missing segments; <… > = < assigned interpretation>; [,] pause inside an intonation unit; [.] = end of intonation unit: line changes at the end of an intonation unit. Conventional Spanish orthography is maintained for s, z, c, v, b, j, g, q, and h, and for word-segmentation. Lit. = literal version.

(2) Natalia

a. *Quieo que, que se vaya una má(s)cara.*
want.PRS.1SG COMP COMP 3SG.RFL go.SBJ.PRS.3SG a mask
Lit. 'I want that, that a mask will go.' (24,11)

b. *Se CAI!, dijo que se CAI.*
3SG.RFL fall.PRS.3SG say.PST.3SG COMP 3SG.RFL fall.PRS.3SG
Lit. It falls!, she said that it falls.' (25,4)

c. *L(u)ego vemo(s) que, peo juga(r). <puedo jugar>*
then see.PRS.1PL COMP can.PRS.1SG play:INF
Lit. 'Then we see that I can play.' (24,27)

d. *Co que tene homigas. <creo que tiene hormigas>*
believe.PRS.1SG COMP have.PRS.3SG ants
Lit. 'I think that it has ants.' (25,10)

e. *Se pa(r)ece qu' etá lloviendo.*
RFL.3SG seem.PRS.3SG COMP be.PRS.3SG rain.GER
'It seems that it's raining', 'It looks like raining.' (26,27)

(3) Flor

a. *Quieo que metes, Susi.*
want.PRS.1SG COMP put_into.PRS.2SG Susi
'I want that you put into, Susi.' (25,18)

b. *Dijites que no hicites popó.*
say.PST.2SG COMP NEG make.PST.2SG poop
'You said that you did not make poop.' (25,18)

c. *Me enojé que me quitalon mi coche.*
1SG.RFL get_mad.PST.1SG COMP 1SG.IO take.PST.3PL my car
'It made me mad that they took my car away.' (25,18)

d. *Creo que se paró el coche.*
think.PRS.1SG COMP 3SG.RFL stop.PST.3SG the car
'I think that the car stopped.' (27)

e. *Yo quiero que pinte aquí mamá.*
1SG.S want.PRS.1SG COMP paint.SBJ.3SG here Mum
'I want that mommy paints here.' (27,20)

(4) Tita

a. *¿Quiees que saca a llave?*
want.PRS.2SG COMP take_out.PRS.3SG a key
'Do you want that he takes out a key?' (28,17)

b. *Le voy a decí a, que vaya.*
3SG.IO go.PRS.1SG to say.INF to COMP go.SBJ.3SG
'I'm going to tell him to, that he should go.' (28,17)

c. *Te juro que yo no fui, creo*
2SG.IO swear.PRS.1SG COMP 1SG.S NEG be.PST.1SG think.PRS.1SG
que yo no fui.
COMP 1SG.S NEG be.PST.1SG
'I swear I was not, I think that I was not.' (29,03)

d. *Le estoy diciendo a Mimí que no rompa*
3SG.IO be_LOC.PRS.1SG tell.GER to Mimi COMP NEG brake.SBJ.3SG
mi cualo.
my frame
'I'm telling Mimi that she shall not break my frame.' (29,28)

(5) Julio

a. *Quero que se acabó.* (closing a book)
want.PRS.1SG COMP 3SG.RFL end.PST.3SG
'I want that it finished.' (34,03)

b. *Un palo, se ve que, un palo.*
a stick 3SG.RFL see.PRS.3SG COMP a stick
'A stick, one sees that, a stick.' (37)

c. *Dile que me abe el avión.*
say.IMP = 3SG.IO COMP 1SG.IO open.PRS.3SG the airplane
'Tell him that he opens the plane for me.' (40,11)

d. *Dile que los hague. <haga>*
say.IMP = 3SG.IO COMP 3PL.DO make.SBJ.3SG
'Tell him that he should make them.' (40,11)

(6) Elia

a. *Dice que no.*
say.PRS.3SG COMP NEG
Lit. 'She says that no.' = 'She says no.' (37)

b. *Me dijo Manuel que, que no viene mañana.*
1SG.IO say.PST.3SG Manuel COMP COMP NEG come.PRS.3SG tomorrow
'Manuel told me that, that he doesn't come tomorrow.' (40)

c. *Quiero que me busques mi, mi reloj.*
want.PRS.1SG COMP 1SG.IO look_for-SBJ.2SG my my watch
'I want that you look for my, my watch.' (40,25)

Children's early use of different CTV with embedded COMP-*que* clauses makes it possible to ask whether the integrated constructions – verbs and marked clauses – have been used and how, before they form, together, a complex sentence. A trace-back search will explore whether these constructions have former independent uses and whether they may be suitable grounding components of future embedding frames.

4.2 On verbs that will take COMP-*que* constructions: Developing a construction inventory

The joint inventory of CTV used by the children in early embedding frames is relatively large (10 verb types). The majority of these verbs is transitive and may take a phrasal or clausal direct object (DO) (*creer* 'believe, think', *decir* 'tell, say', *esperar* 'wait', *jurar* 'swear', *mirar* 'look', *necesitar* 'need', *querer* 'wish, want', *saber* 'know', *soñar* 'dream', *ver* 'see'). Some of these are intransitive verbs and may take phrasal or clausal subjects (*faltar* 'lack', *gustar* 'please', *parecer* 'seem'). Despite their constructional possibilities, most CTV do not take COMP-*que* constructions from the start.

Among the joint inventory of CTV produced by the children, only *decir* 'tell, say' and *querer* 'want, wish' have been used with COMP-*que* constructions by all of them. The verb form *creo* 'I suppose, I think' has been used in a handful of cases by a single child, and on only one occasion by two additional children. Certain other CTV are used only once with a COMP-*que* complement by a single child: *necesito* 'I need'; *parece* 'it seems; it looks like'; *soñaba* 'I dreamt', among others (Table 2).[3]

Table 2. Children's lexical inventory of complement-taking verbs with COMP-*que* arguments

CTV with COMP-*que* arguments		Natalia	Flor	Tita	Elia	Julio	Total
Decir que…	'to say that …'	9	19	9	12	3	52
Querer que…	'to want that…'	2	6	3	2	3	16
Creo que…	'I believe that…'	4	1	1			6
Veo que…	'I see that …'	1					1
Parece que…	'It seems that …'	2					2
Necesito que...	'I need that …'	1	1	1			3
Enojé que …	'I got mad that…'		1				1
Juro que …	'I swear that …'			1			1
Soñaba que that…	'I dreamt that…'			1			1
Hago que…	'I make that …'			1			1
Total verb types/tokens		6: 19	5: 28	7: 17	2: 14	2: 6	10: 84

CTV that have been used once or twice with a COMP-*que* clause have not been included in the trace-back search, even if they are frequently used as lexical verbs with phrasal complements. Therefore, from this point onward, only the more frequent CTV, *querer* 'want, wish' and *decir* 'tell, say', will be addressed: they are those

3. Verb items in this table are exposed with a specific inflection if they are exclusively used with this particular form. Only *querer* 'want' and *decir* 'say' are presented in the infinitive – Spanish citation form – because they are already used with different inflection forms.

that better show the gradual incorporation of different frames in children's construction inventory.

Querer *'want', 'wish'. Construction inventory*
All children use *querer* 'wish, want' as a lexical verb first, either in bare constructions (7a), or with a lexical or pronominal DO (7b–c). *Querer* is also used later as a quasi-modal auxiliary with infinitive verbs (7d). Some weeks (for Natalia and Flor) or months (for Elia, Tita, Julio) later, *querer* is used with an embedded COMP-*que* clause complement (7e).

(7) a. Chi: *Abiba. <arriba>* (pointing at a book shelf)
'Above.' (Elia 28)
Fat: *¿Arriba qué?*
'Above what?'
Chi: ***Kede.*** *<quiero>*
want.PRS.3SG
'I want.'
Fat: *¿Un libro?*
'A book?'

b. Chi: ***Tede*** ***ese,*** *papá. <quiero ese>*
want.PRS.3SG this Dad
'I want this one, Dad.' (Elia 28)

c. Chi: ***Tedo*** ***duce****. <quiero dulce>*
want.PRS.1SG sweet
'I want a candy.' (Elia 28)

d. Chi: ***tedo*** ***hacer*** ***pipí.*** *<quiero hacer pipí>*
want.PRS.1SG make.INF pee
'I want to pee.' (Elia 33)

e. Fat: *Déjalo allí, para que no te estorbe.*
leave.IMP = 3SG.DO there for COMP NEG 2SG.IO hinder.SBJ.3S
'Leave it there, so that it will not hinder you.' (Elia 41)
Chi: ***Yo quiero que me estorbe.***
1SG.S want.PRS.1SG COMP 1SG.IO hinder.SBJ.3SG
'I want that it hinders me.' (Elia 41)

Unexpectedly, and with no parallel models in adult speech, children produce the CTV *querer* 'want' adjoined to a main clause construction (8a–e) in a plain discourse sequence, with no dependency marking, however. These sequences cannot be considered incidental COMP-*que* omissions, given that COMP-*que* would be unacceptable in every case.

Moreover, clauses adjoined to *querer* 'want' in these sequences do not comply with the features and restrictions for the clausal complements of this verb. Instead,

they show some main-sentence properties: some are performative expressions, like greetings or warnings, common as free utterances (8b); instead of the expected infinitives under same-subject conditions, they may present un-embeddable imperative verb forms as in (8c), unexpected past-tense forms (8d), or other finite verb forms (8a), (8b), (8d). They also take various unexpected verb forms instead of the subjunctive verb form required with different subjects, as in (8c) and (8e).

(8) a. *Quiero* ***peino.*** (Flor 25)
want.PRS.1SG comb.PRS.1SG
Lit. 'I want I comb.' = 'I want to comb'

b. *Quiedo* ***olitas*** ***vengo.*** (Flor 27)
want.PRS.1SG now.DIM come_back.PRS.1SG
Lit. 'I want now I come.'

c. *Quie(r)o* ***dué(r)mete,*** *oso.* (Tita 29)
want.PRS.1SG sleep.IMP = 2SG.RFL bear
Lit. 'I want you sleep, bear.' = 'I want you to sleep, bear'

d. *Quie(r)o* ***saqué*** ***la*** ***vela.*** (Julio 34)
want.PRS.1SG take_out.PST.1SG the candle
Lit. 'I want I took out the candle.' = 'I want the candle to be for me.'

e. *No quiero,* ***no*** ***ab(r)ochas.*** (Elia 36)
NEG want.PRS.1SG NEG button.PRS.2SG
Lit. 'I don't want, you don't button.' = 'I don't want you button (my clothes)'

These discourse sequences constitute critical evidence of constructional grounding. For a small window of time, a free utterance, previously and independently used, may become a piece in a construction chain with *querer* 'want, wish'. These discourse sequences have no formal marking of embedding, showing at most a single intonation contour as a trace of integration (but see (8e) with an intermediate pause).

Decir *'say', 'tell': Construction inventory*

Children's constructions with *decir* 'tell, say' – the most frequent CTV in the data – also present gradual adoption of different complement types. As a speech act verb, *decir* 'tell, say' may take all sorts of "quote" as a complement. Such quoting complement constructions are represented by different types of phrases (9a–b) – not only noun phrases –, including interjections and onomatopoeic items. All of these are, in structural terms, sequential, juxtaposed elements (Diessel 2004).

(9) a. *El lepalo dijo fue-* ***AFUERA.***
The leopard say.PST.3SG ou(t) out of here
'The leopard said out of here.' (Nata 25)

b. *No me *digue* ***loca.*** <*no me digas loca*>
NEG 1SG.IO say.SBJ.2SG crazy
'Do not say to me crazy.' (Flor 28,6)

Later, these quoting frames referring to clause constructions are found again in plain discourse sequences: a variable inflected form of *decir* 'tell, say' precedes (10a–c) or follows (11a–b) a clause quotation with no dependency marking or adjustments at all.

(10) a. *(D)**icía** Bambi, (d)**icía,** voy a, aquí.*
say.IMPF.3SG Bambi say.IMPF.3SG go.PRS.1SG to here
Lit 'Bambi said, said, I go to, here.' (Nata 24)

b. *Sabina **dijo** no me pegas.*
Sabina say.PST.3SG NEG 1SG.DO hit.PRS.2SG
'Sabina said don't hit me.' (Elia 37)

c. ***Dice,** bájame por favor.*
say.PRS.3SG put_down.IMP = 1SG.DO by favour
'He said, put me down please.' (Flor 23)

(11) a. *Se ebaló Bami, (d)**ijo.*** <se resbaló Bambi, dijo>
3SG.RFL slide.PST.3SG Bambi say.PST.3SG
'Bambi slid, he said.' (Nat 24)

b. *Ciérrale, **dijo.***
close.IMP = 3SG.DO say.PST.3SG
'Close it, he said.' (Elia 41)

Clause quotations forming a discourse sequence with *decir* 'tell, say' are frequent in adult models and they represent the preferred frames for making speech reports. In child language, such sequences appear early and become as frequent and stable for *decir* as in adult speech.

Besides quoting frames, *decir* 'tell, say' also takes conventional DO: pronouns, and nouns referring to tellings/sayings: *lies, a story, a goodbye*. These genuine nominal DO are infrequent and seem to appear late (12).

(12) *Ya le dice **un ayós** a sus amigos.*
already 3SG.IO say.PRS.3SG a goodbye to her friends
'She already says **a goodbye** to her friends.' (Flor 28,6)

As for COMP-*que* clauses embedded to *decir* 'tell, say', they are first poorly elaborated, verbless constructions with isolated polarity items or single word utterances as in (13a), (13b). They soon become fully expanded clauses (13c), (13d) with rich and variable adjustments in indexical items or Tense-Aspect-Mood (TAM) inflection in verbs (13d), which is a clear evidence of a dependent status (Gast & Diessel 2012; Lehmann 1988).

(13) a. *Le vamos a decir a Susi* ***que no.***
3SG.IO go.PRS.1PL to say.INF to Susi COMP NEG
'We are going to tell Susi that no.' (Elia 37)

b. *Chocomí no, no,* ***dije que demís.***
chocomilk NEG NEG say.PST.1SG COMP "demis"
'Chocomilk no, no, I said that *demis.*' (Nata 24)

c. *Me dijo el señor* ***que no hay huevo.***
3SG.IO say.PST.3SG the man COMP NEG exist.PRS.3SG egg
'The man told me that there are no eggs.' (Flor 31)

d. *Dijo el conejo* ***que era tarde.***
say.PST.3SG the rabbit COMP be.IMPF.3SG late
'The rabbit said that it was late.' (Nata 25,9)

In sum, the construction inventory of CTV complementation develops gradually. Verbs have lexical uses with phrasal DO, or bare constructions. The DO position is the earliest slot available to variable fillers: simple lexical items, pronouns, nouns, and noun phrases (NP) with *querer* 'want', and quoted lexical items, phrase constructions, or onomatopoeias, with *decir* 'tell-say'. Both verbs progress into discourse sequences, adjoined to main sentence frames. Adopting a COMP-*que* clausal argument is a further and later step in the gradual development of a construction inventory for both verbs. These have prior lexical uses and develop a range of constructional possibilities for phrasal arguments; they also progress into discourse sequences with adjacent clause constructions. They develop, in various ways, the slot frames for taking embedded clausal objects.

4.3 Looking at Free COMP-*que* constructions

Besides former lexical uses of CTV and gradual adoption of complement constructions, it is crucial, in order to test constructional grounding, to verify that children learn and use COMP-*que* constructions as main sentences before inserting them into embedding frames. In that regard, the study has shown that all children – including Luis, the child dropped from the study – learn COMP-*que* clausal frames and use them as main sentences long before using them as arguments of CTV. Although conversational moves scaffold free uses of COMP-*que,* it is more revealing that COMP-*que* constructions occur as free standing sentences without any support from conversation, even before the corresponding dialogical pairs. Children learn COMP-*que* constructions as independent frames and use them freely, as main sentences.

Since constructions are, by definition, a form and meaning pair (Fillmore 1988; Goldberg 1995), the acquisition of COMP-*que* constructions presupposes both,

learning the linguistic frame and the semantic or pragmatic functions for which the frame may serve.

As any main sentence does, COMP-*que* constructions serve different illocutionary functions. The earlier, and most prominent, illocutionary functions children encode with Free COMP-*que* frames include: (i) desiderative, and (ii) logophoric-reiterative. Free COMP-*que* constructions may express a wish, or may be a (near) repeat of an utterance or intention, with the COMP-*que* frame indexing the repeat. Spanish grammarians tend to confer these functions directly to COMP-*que* (Demonte & Fernández 2009; Gras 2013). In my view, the constructional frame itself and some characteristic components (in particular, verb inflection) contribute to assign a particular I-function to a given COMP-*que* construction, always in interaction with the conversational context and the situation.

Desiderative functions are associated with a COMP-*que* frame, usually including a subjunctive verb (14) (see glosses). For reiterative COMP-*que* constructions (15), it is the relation with a previously spoken utterance that allows the logophoric interpretation. In addition to these dominant functions, children's COMP-*que* constructions are also used in justifying previous saying or doings (16), and in representing 'let's pretend' activities in symbolic play, with a past imperfect – *ludic co-preterit* – verb, usually (17) (Coseriu 1976; Gras 2011; Demonte & Fernández 2009; Moreno de Alba 1978; Sheridan 2007: 14).

(14) Free COMP-*que*-constructions with desiderative function

a. ***Que se vaya.***
COMP 3SG.RFL go.SBJ.3SG
'(Hope/wish) that she goes.' (Nata 23,29)

b. *No (l)a comas,* ***que*** *la pelita* ***poma.***
NEG 3SG.DO eat.SBJ.2SG COMP the dog.DIM eat.SBJ.3SG
'Don't eat it, (wish) that the doggy eats.' (Flor 24,12)

c. *La bebé* ***que*** *no* ***se ca(i)ga.***
the baby COMP NEG 3SG.RFL fall_down.SBJ.3SG
'(Wish) the baby that does not fall.' (Tita 28,17)

d. *E bebé,* ***que se tape.***
the baby COMP 3SG.RFL cover.SBJ.3SG
'(Wish) the baby, that gets covered.' (Elia 30,22)

e. ***Que me bajes.***
COMP 1SG.DO put_down.SBJ.2SG
'(Wish) that you put me down.' (Julio 33,0)

(15) Free COMP-*que* constructions indexing repetition

a. Chi: *Vedá* ***vine la noche.***
truth come.PST.1SG the night
'I came last night, right?' (Flor 25,12)

Mot: *¿Mande?*
'what?'
Chi: ***Que vine e la noche.***
COMP come.PST.1SG in the night
'(I said) that I came last night.'

b. Chi: ***Ya no*** *me jales pelo.*
already NEG 1SG.IO pull.SBJ.2SG hair
'Don't pull my hair anymore.' (Nata 24,16)
Fat: *no te lo jalo, te lo estoy*
NEG 2SG.IO. 3SG.DO pull.PTS.1SG 2SG.IO 3SG.DO be.PRS.1SG
alisando.
sleek.GER
'I don't pull it, I'm sleeking it back.'
Chi: ***Que ya no.***
COMP already NEG
Lit. 'that already not' '(say/wish) no more.'

c. Chi: ***Toy aquí*** .
be.PRS.1SG here
'I'm here.' (Tita 29,3)
Mot: ¿Mande?
what?
Chi: ***Que toy ugano.***
COMP be.PRS.1SG play.GER
'(I said) that I'm playing.'

(16) Free COMP-*que* frames as justification

a. *CUIDADO,* ***que*** *se cae.*
care **COMP** 3SG.RFL fall.PRS.3SG
'Careful, (cause) that it may fall.' (Nata 24,22)

b. *¡Uy!, ayúdame,* ***que*** *no puedo con el*
Oh! help.IMP = 1SG.IO **COMP** NEG can.PRS.1SG with the
rompetabezas.
puzzle
'Help me (cause) that I can not with the puzzle.' (Elia 40)

(17) 'Let's pretend' reading of Free COMP-*que* frames

Chi: ***Que*** *le* ***sirvía*** *agua Ana.*
COMP 3SG.IO pour.IMPF.3SG water Ana
'(Pretend) that Ana poured water to her.' (Nata 25,11)
Fat: *Pero Ana no está.*
but Ana NEG be.PRS.3SG
'But Ana is not here.'

This development in which children add Free COMP-*que* frames to their construction inventory and learn to perform the corresponding I-functions, may be considered as preparatory steps from a grounding perspective.

We have evidence of children's learning from both the frame itself and the function of desiderative COMP-*que* frames in children's creative productions. Having learned the frame with a subjunctive verb from the start, children later produce unexpected extensions of COMP-*que* frames with imperative (18a) and indicative verbs (18b), holophrastic constructions with no verb (18c), or use the frame in replies (18d). All these variations on the COMP-*que* theme convey a desiderative I-function and illustrate how children are skilled at recruiting innovative frames to carry out this function.

(18) Overextension of Free COMP-*que* constructions with desiderative function

a. Chi: ***Que ten*** *regalo.*
COMP **hold.IMP** present
'(Wish) that you hold this present.' (Julio 33,0)

b. Fat: *Duérmete un rato.*
sleep.IMP = 2SG.RFL a while
'Sleep for a while.' (Nata 24,20)

Chi: *No,* ***que veo*** *Pu e navidad.*
NEG **COMP see.PRS.1SG** Pooh in Christmas
'No, (wish) that I see Pooh in Christmas.'

c. Chi: *Sí, tele,* ***que sayos,*** ***que dinosadios***
AFF tele(vision) COMP (dino)saurs COMP dinosaurs
'Yes, TV, (wish) that dinosaurs, (wish) that dinosaurs.' (Nata 24,20)

d. Chi: *¿Quele, Papi?* (Mum to read a book)
want.PRS.3SG Dad
'Does she want, Dad?' (Nata 24,20)

Fat: *No, no creo que quiera.*
NEG NEG think.PRS.1SG COMP want.SBJ.3SG
'No, I don't think that she wants'

Chi: ***¡Que quiera!***
COMP want.SBJ.3SG
'(Wish) that she wants!'

None of these constructions would be acceptable in adult use. Children's production of these overgeneralized combinations shows that they have detected and creatively apply the COMP-*que* clause pattern to refer to variable event types or situations as 'object of desire'.

As for the logophoric function performed by Free COMP-*que* constructions, evidence on how the frame and the function are learned, can be obtained from conversation in which the COMP-*que* frame appears itself. Creative online production

of logophoric frames is the norm. This can be observed in cases such as (15a–c), in which children produce a COMP-*que* version of their previous utterance. The children repeat their own previous utterances, by recruiting a Free COMP-*que* frame to introduce the reiteration.

In either case, children have learned the frames and the associated functions. Free COMP-*que* frames with desiderative and logophoric functions are learned before children produce their embedded versions. The desiderative and logophoric functions are on a par with the meaning conveyed by CTV *querer* 'want' and *decir* 'tell'. Each COMP-*que* frames will later be combined with the corresponding verb. Both construction frames may ground the complex sentences that children will later produce.

4.4 Development of CTV frames and COMP-*que* constructions

Besides the independent and earlier learning of proposed source constructions, the way in which these constructions develop over time offers suggestive evidence for a grounding proposal (Table 3).

Table 3. Developmental chronology between lexical uses of *querer* and *decir,* free COMP-*que* frames, and embedded CTV + COMP-*que* constructions

	Bare CTV construction	**CTV + phrasal argument**	**CTV + infinitive**	**Discourse sequence**	**Earliest CTV COMP-*que***	**Free COMP-*que* desid/logoph**
NATA						
Querer	N1	N1	N1	N2	N3	N1
	23,25	23,25	23,25	24,18	25,07	23,11
Decir	N1	N1		N1	N2	N1
	23,25	23,25	nr	23,25	25,07	24,11
FLOR						
Querer	F1	F1	F1	F2	F2	F1
	23,21	23,21	23,21	25,18	25,18	24,0
Decir		F1		F2	F2	F1
	nr	24,01	nr	25,18	25,18	24,12
TITA						
Querer	T1	T1	T1		T2	T1
	24,10	24,10	24,10	Nr	28;17	24,0
Decir	T1	T2		T2	T2	T1
	24,0	27,7	nr	27,24	28,17	25,22
ELIA						
Querer	E0	E0	E1	E2	E3	E1
	28,4	28,4	32,8	36,4	41,0	30,22

(*continued*)

Table 3. (*continued*)

	Bare CTV construction	**CTV + phrasal argument**	**CTV + infinitive**	**Discourse sequence**	**Earliest CTV COMP-*que***	**Free COMP-*que* desid/logoph**
Decir	E1	E2	nr	E2	E2	E1
	33,0	36,4		37,5	37,5	32,25
JULIO						
Querer	J1	J1	J2	J3	J3	J2
	28,27	28,27	30,28	34,3	34,3	32,0
Decir	nr	J2	nr	J4	J4	J2
		32,0		37,2	38,11	33,0

Despite the presence of some individual differences that I will not focus on here, and the presence of an item-based learning style (Tomasello 1992, 2000, 2003), a general developmental profile can be proposed. All children use the two verbs with phrasal complements and they produce Free COMP-*que* constructions as main sentences before they use the same CTV with a COMP-*que* clause as DO. Using grounding constructions always precedes children's production of complex clauses in which a COMP-*que* clause is embedded as an argument of a CTV.

The relationship between Discourse sequences and genuine embedding of a COMP-*que* clause into a CTV is also a possible grounding relationship: whenever there is a time lapse between these constructional frames, Discourse sequences precede the embedding of a COMP-*que* argument. When a CTV verb occurs in the same data point in a Discourse sequence and with a COMP-*que* frame (e.g., in the case of Flor), this coincidence does not invalidate a grounding relationship: it simply does not confirm it.

Summing up, all of the five children display the same developmental path, although they are not synchronized at every step. Developmental time points of the emerging sequence are schematized below in (19).

(19) Construction inventory: development over time from a grounding perspective

T1	T2	T3 (=/> T2)
Lexical use of CTV	Discourse sequences:	CTV + COMP-*que* frames
Free COMP-*que* frames	CTV + main sentences	

4.5 Bridging the gap. Dialogical support and on-line integration

Information obtained from prior uses of CTV and Free COMP-*que* constructions and their relationship with embedded target frames, strongly supports a constructional grounding proposal on the basis of distribution, semantic-functional

affinities between grounding components and target construction, and developmental timing.

There is also evidence showing that children build a link between grounding constructions and target frames based on the enchronic organization of language in dialogue. On the one hand, CTV and COMP-*que* constructions are produced in child-adult conversations by a spreading of complexity (Givón 2009): each speaker is in charge of one piece of the complex clause structure, particularly when children respond to a question about an argument of a CTV with a COMP-*que* construction. In this way, COMP-*que* frames form an adjacency pair with a CTV (20a–d) without showing an explicit embedding relation (Aguirre 2000[1994]; Barreña 1999; Givón 2009; Sansiñena et al. 2015).

(20) a. Mot: ***¿Qué le decías*** *a, Nata, al tigre?*
what 3SG.IO tell.IMPF.2SG to Nata to.the tiger
What did you tell to, to the tiger? (Nata 24,01)
Chi: *Mm,* ***que ela, e, u moto.***
mm COMP be.IMPF.3SG mmm a monster
'That he was a monster.'

b. Chi: *No oga de domi(r)se a la cama.*
NEG time of sleep.INF = 3SG.RFL to the bed
'No, time to go to sleep to bed.' (Tita 38)
Mot: ***¿Y qué le dice?***
and what 3SG.IO tell.PRS.3SG
'And what does she tell her?'
Chi: ***Que ya i(r)se a domi(r)se.***
COMP already go.INF = 3SG.RFL to sleep.INF = 3SG.RFL
'That it's time to go to bed'

c. Chi: *Sus hermanos le dijeron feo.*
3PL.POS brothers 3SG.IO tell.PST.3PL ugly
'His brothers told him ugly.' (Elia 36,24)
Mot: ***¿Qué le dijeron?***
what 3SG.IO tell.PST.3PL
What did they tell him?
Chi. ***Que se ve feo.***
that 3SG.RFL look_like.PRS.3SG ugly
'That he looks like ugly.'

d. Chi: *Coche.* (Flor 27)
'Car.'
Mot: ***¿Qué quieres***?
what want.PRS.2SG
What do you want?

Chi: ***Que se paró el coche.***
COMP 3SG stop.PST.3SG the car
'That the car to stop.'

Composite utterances in conversation represent one prominent bridging context for children to experience CTV and Free COMP-*que* frames being put together on-line. Also, adults offer children models of CTV + COMP-*que* integration when they try to interpret child's Free COMP-*que* constructions. In these moves, adults may include a CTV that overtly encodes the I-function associated with the COMP-*que* construction performed as a main sentence in the children's previous move (21). In this way, an embedded version of a Free COMP-*que* construction is offered to the child's consideration as a complete version of the same communicative intention.

(21) Mot: *Lastima, por eso no se juega con el tenedor.*
hurt.PRS.3SG for that NEG 3SG.RFL play.PRS.3SG with the fork
'It hurts, that's the reason we do not play with the fork.'
Chi: *Por eso,* ***que me haces avión.***
'for this COMP 1SG.IO make-PRS.2SG airplane
Lit. for this (wish) that you make for me an airplane.'
(='Wish you play with the fork as an airplane with the food.')
Mot: ***¿quieres que te haga un avión?***
want.PRS.2SG COMP 2SG.IO make.SBJ.1SG an airplane
Do you want I make an airplane for you?

Children make the same type of assemblage. They take charge of the production of both CTV and COMP-*que*-constructions in the successive steps of a conversation, not integrated into a single intonation unit, and sometimes with a near repair-like flavor. For instance, in (22), the child firstly produces a Free COMP-*que* clause, which is reformulated immediately after including a CTV: this CTV makes explicit the I-function carried out by the prior Free COMP-*que* frame. Thus, a Free COMP-*que* clause becomes a complex clause construction on-line.

(22) Flor: ***Que me pongan crema a mi … a mi cuello.***
COMP 1SG.IO put.SBJ.3PL cream to my to my neck
'(Wish) that they put cream on my neck.'
Tiero que mi mi quemita, ah, aquí.
want-PRS.1SG COMP my my cream.DIM ah here
'I want that some cream here.'

The on-line relationship between CTV and COMP-*que* constructions becomes also explicit when children have a change of plans in conversation: when a CTV construction is interrupted and replaced with a Free COMP-que construction (23a–b);

or rather when a main clause is successively reframed as a COMP-*que* construction with a CTV (23c).

(23) a. *Quieo que,*
want.PRS.1SG COMP
'I want that,' (Nata 24,11)
¡Que se vaya una má(s)cara!
COMP 3SG.RFL go.SBJ.3SG a mask
'(Wish) that mask goes away!'

b. *Le voy a deci a,* ***que vaya!***
3SG.IO go.PRS.1SG to say.INF to COMP go.SBJ.3SG
'I'm going to tell him to, that he should go!' (Tita 28,17)

c. ***Se CAI!***
3SG.RFL fall_down.PRS.3SG
'She falls down!' (Nata 24,04)
Dijo ***que se CAI.***
Say.PST.3SG COMP 3SG.RFL fall.PRS.3SG
'She said that she falls down.'

In either case, dialogical events clearly show that CTV and COMP-*que* clauses used independently by the child may be recruited on-line to form a complex clause construction.

Some dialogical sequences still present an opposite movement: a prior complex clause with a CTV and an embedded COMP-*que* clause is successively dissociated by the child. Therefore, the former embedded clause becomes a main sentence when it is produced as a Free COMP-*que* construction (24).

(24) Mot: *¿Tú también la sentiste una gota?*
you too 3SG.DO feel.PST.2SG one drop
'Did you feel one drop?' (Flor 27,20)
Sí, están cayendo gotitas.
Yes be.PRS.3PL fall.GER drop.DIM.PL
'Yes, there are drops falling.'

Flor: *Yo no* ***quiero que*** *caigan gotitas,*
1SG.S NEG want.PRS.1SG COMP fall.SBJ.3PL drop.DIM.PL
'I don't want drops to fall,'
Que no *caigan gotitas.*
COMP NEG fall.SBJ.3PL drop.DIM.PL
'That drops do not fall.' (Flor 27,20)

In either case, dialogically interrelated sequences of different types join the grounding components of a complex construction and split them apart in the interactional

scenario (Givón 2009). Dialogue explicitly places the child in a position to observe or represent the change of a Free COMP-*que* construction into an embedded one, along with the possible dissociation of complex clauses into their constituent frames, and the interchange between the I-function of a free COMP-*que* construction and the semantics of a corresponding verb. What is observed once again is that for the child, CTV, Free COMP-*que* constructions, and complex embedding frames actually become related among each other. Children's manipulation on these constructional pieces, by assembling and splitting the target and the grounding constructions, is the strongest evidence supporting that both frames are involved in emergent complex clause production. These frames become available to the child as independent or joint pieces and are displayed as isolated units or in combination in order to produce or dissociate complex clause construction.

5. Concluding remarks

The results of this study provide evidence supporting the constructional grounding proposal under consideration. Distribution and chronological emergence of source constructions, Free COMP-*que* constructions, and lexical uses of CTV, represent positive support for a grounding proposal. Each constructional piece is learned by itself and used independently before becoming jointly assembled in complex constructions. The functional affinities between Free COMP-*que* constructions and CTV observable in conversations through on-line assemblage of source constructions into complex frames, and the splitting of complex sentences into source constructions, taken together and jointly, are more than suggestive evidence of a constructional grounding relation.

It has been shown that children first produce clear cases of grounding constructions in independent uses, with a gradual enlargement of constructional possibilities for CTV. The constructional development of verbs begins with the adoption of different DO, making a slot available to COMP-*que* construction. Discourse sequences of CTV and free sentences, in simple adjacent vicinity, show that children consider clausal frames by themselves, as indicated by the no overt syntactic dependency between them. This CTV is used to add a comment about the main clause (Diessel 2004). Taking clause constructions as a topic of a verbal comment represents a major step in the path through which clauses become conceptual objects that may be treated as entities in embedding frames.

As for COMP-*que* constructions, their developmental history has a starting point in their use as main free sentences. A crucial aspect for a grounding proposal arises from these free uses: as any main sentence, Free COMP-*que* clauses display an illocutionary function that, whether Desiderative or Logophoric, is associated

with the same semantic domain of specific and corresponding CTV. Logophoric COMP-*que* constructions and CTV *decir* 'say', on the one hand, and Desiderative COMP-*que* frames and CTV *querer* 'want' on the other hand, respectively correspond to the same semantic domain. When children combine these constructions in an embedding frame, an unexpected semantic absorption of the illocutionary function of the COMP-*que* clause occurs through lexical encoding with the CTV. This absorption is to a certain extent inverse to the inheritance principle used in Construction grammar. It is clearly attested along the splits and assemblages of the source frames in dialogue sequences, and strongly supports the grounding relation here proposed. The presence of an absorption process, which has not previously been referred as far as I know, raises a theoretical question. It is well known that I-functions are exclusively performed by main sentences. A normal criterion for identifying clause combining is that dependent clauses do not carry out any I-function of their own, which is exclusive of the whole complex frame (Lehmann 1988; Gast & Diessel 2012). This study has found in developmental data, as in a slow-motion film, the way in which, along conversational moves, a free sentence loses its I-function when it becomes embedded into a CTV, and how an embedded sentence acquires an I-function when it is insubordinated from the embedding frame.

The I-functions that Free COMP-*que* constructions display may have a bridging effect facilitating the embedding into corresponding CTV. Affinities between the lexical meaning of CTV and the respective illocutionary function of Free COMP-*que* constructions (Desiderative and Logophoric) are important clues for a constructional grounding proposal. Here it is useful to consider the pragmatic bootstrapping hypothesis proposed by Anat Ninio (2001) for understanding how items accomplishing the same function may be involved in the origin of elaborated constructions that include them both. Former studies maintain that constructional grounding is possible because components of the source construction share some important properties with the target construction that are required to ground (Israel, Johnson & Brooks 2000: 106). Therefore, the functional affinities between the illocutionary functions carried out by Free COMP-*que* constructions and by the CTV used in embedding frames help to understand how and why independent constructions may become integrated into an embedding frame, therefore confirming the plausibility of a constructional grounding proposal. A functionally-based distribution analysis, as the one proposed by Tomasello (2006: 288) and applied both to items and to constructions, could be invoked to understand the relationship between CTV and Free COMP-*que* constructions serving the same communicative intention.

A word of caution is needed, since the plausibility of the particular developmental path proposed here strictly depends on the constructional frames that are available to each language community, as suggested by Radical Construction Grammar. For children who have Spanish as a target language, the complexity of the

path is based on the availability of Free COMP-*que* constructions in parental models. Children may learn insubordinate frames from language practitioners; otherwise, they would need to step back and start by extracting them from the complex chains if they only appear in embedding frames.

We can end this analysis with a fragment from the initial epigraphic quote, applied now to this research space: CTV constructions present "numerous, successive, slight modifications" that may be considered preparatory for their composition with a COMP-*que* clause construction. Evidence offered along this paper shows that complex constructions produced by young children have clear foundations in antecedent attainments. As constructivist theories propose, a series of learning events in an epigenetic chain, supported by general learning mechanism can explain the emergence of complex constructions without the need of dedicated support from our biological endowment.

Acknowledgement

I thank two anonymous reviewers for insightful comments and suggestions.

Abbreviations

COMP	complementizer	3SG	third person singular
CTV	complement-taking verb	1PL	first person plural
PRS	present tense	3PL	third person plural
PST	past tense	S	subject
IMPF	imperfect	DO	direct object
SBJ	subjunctive mode in present	IO	indirect object
IMP	imperative	RFL	reflexive
INF	infinitive	NEG	negation
GER	gerundive	AFF	affirmation
1SG	first person singular	DIM	diminutive
2SG	second person singular	PL	plural

References

Abbot-Smith, Kirsten & Behrens, Heike. 2006. How known constructions influence the acquisition of other constructions: The German passive and future constructions. *Cognitive Science* 30(6): 995–1026. https://doi.org/10.1207/s15516709cog0000_61

Aguirre, Carmen. 2000 [1994]. *La adquisición de las categorías gramaticales del español*. PhD dissertation, Universidad Autónoma de Madrid.

Ambridge, Ben & Lieven, Elena. 2011. *Child Language Acquisition. Contrasting Theoretical Approaches*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511975073

Bannard, Colin, Lieven, Elena & Tomasello, Michael. 2009. Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences of the United States of America* 106(41): 17284–17289. https://doi.org/10.1073/pnas.0905638106

Barreña, Andoni. 1994. Sobre la adquisición de la categoría funcional COMP por niños vascos. In *La adquisición del vasco y del castellano en niños bilingües*, Jurgen Meisel (ed.), 231–284. Frankfurt: Vervuert-Iberoamericana. <http://dialnet.unirioja.es/servlet/articulo?codigo=2007290)>

Barreña, Andoni. 1999. La adquisición de estructuras subordinadas por un niño bilingüe vasco-español, *Lingüística para el siglo XXI. III Congreso organizado por el Departamento de Lengua Española*, 277–286. Salamanca: Universidad de Salamanca. <http://dialnet.unirioja.es/servlet/libro?codigo=386>

Bloom, Lois, Lahey, Margaret, Hood, Lois, Lifter, Karin & Fiess, Kathleen. 1991[1980]. Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. In *Language Development from Two to Three*, Lois Bloom, 261–289. Cambridge: CUP.

Bod, Rens. 2009. From exemplar to grammar: A probabilistic analogy-based model of language learning. *Cognitive Science* 33(5): 752–793. https://doi.org/10.1111/j.1551-6709.2009.01031.x

Brandt, Silke, Verhagen, Arie, Lieven, Elena & Tomasello, Michel. 2011. German children's productivity with simple transitive and complement-clause constructions: Testing the effects of frequency and diversity. *Cognitive Linguistics* 22(2): 325–357. https://doi.org/10.1515/cogl.2011.013

Braunwald, Susan R. 1997. The development of *because* and *so*: Connecting language, thought, and social understanding. In *Processing Interclausal Relationships: Studies in the Production and Comprehension of Text*, Jean Costermans & Michel Fayol (eds), 121–137. Mahwah NJ: Lawrence Erlbaum Associates.

Bybee, Joan L. & Hopper, Paul J. 2001. Introduction to frequency and the emergence of linguistic structure. In *Frequency and the Emergence of Linguistic Structure* [Typological Studies in Language 45], Joan Bybee & Paul M. Hopper (eds), 1–24. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.45.01byb

Coseriu, Eugenio. 1976. *Das Romanische Verbal system*. Tübingen: Gunter Narr.

Culicover, Peter W. 1999. *Syntactic Nuts. Hard Cases, Syntactic Theory, and Language Acquisition*. Oxford: OUP.

Dąbrowska, Eva & Lieven, Elena. 2005. Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* 16(3): 437–474. https://doi.org/10.1515/cogl.2005.16.3.437

Demonte, Violeta & Fernández-Soriano, Olga. 2009. Force and finiteness in the Spanish complementizer system. *Probus* 21: 23–49. https://doi.org/10.1515/prbs.2009.002

Diessel, Holger. 2004. *The Acquisition of Complex Sentences*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511486531

Diessel, Holger & Tomasello, Michael. 2000a. The development of relative constructions in early child speech. *Cognitive Linguistics* 11: 131–152.

Diessel, Holger & Tomasello, Michael. 2000b. Why complement clauses do not include a *that*-complementizer in early child language. In *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society*, Jeff Good & Alan C. L. Yu (eds), 86–97. Berkeley CA: BLS.

Diessel, Holger & Tomasello, Michael. 2001. The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics* 12(2): 97–141. https://doi.org/10.1515/cogl.12.2.97

Du Bois, John W. 2014. Towards dialogic syntax. *Cognitive Linguistics* 25(3): 359–410.
Elman, Jeffrey L. 1990. Finding structure in time. *Cognitive Science* 14(2): 179–211. https://doi.org/10.1207/s15516709cog1402_1
Elman. Jeffrey L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48(1): 71–99. https://doi.org/10.1016/0010-0277(93)90058-4
Elman, Jeffrey, Bates, Elizabeth, Johnson, Mark H., Karmiloff-Smith, Annette, Parisi, Domenico & Plunkett, Kim. 1996. *Rethinking Innateness. A Connectionist Perspective on Development*. Cambridge MA: The MIT Press.
Enfield, Nicholas J. & Sidnell, Jack. 2014. Language presupposes an enchronic infrastructure for social interaction. In *The Social Origins of Language: Studies in the Evolution of Language*, Daniel Dor, Chris Knight & Jerome Lewis (eds), 99–104. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199665327.003.0008
Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and Empirical Foundations*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.
Evers-Vermeul, Jacqueline. 2005. *The Development of Dutch Connectives: Change and Acquisition as Windows on Form-Function Relations*. Utrecht: Utrecht University.
Evers-Vermeul, Jacqueline & Sanders, Ted. 2011. Discovering domains. On the acquisition of causal connectives. *Journal of Pragmatics* 43(6): 1645–1662. https://doi.org/10.1016/j.pragma.2010.11.015
Ezeizabarrena Segurola, María José. 1997. Infinitivos verbales en el lenguaje de niños bilingües. In *Views on the Acquisition and Use of a Second Language: Eurosla'7 Proceedings*, Lourdes Díaz & Carmen Pérez (eds), 177–190. Barcelona: Universidad Pompeu Fabra.
Ezeizabarrena Segurola, María José. 2002. Root infinitives in two pro-drop languages. In *Acquisition of Spanish Morphosyntax*, Ana Teresa Pérez-Leroux & Juana M. Liceras (eds), 35–65. Amsterdam: Kluwer.
Fillmore, Charles J. 1988. The mechanisms of "construction grammar". *Annual Meeting of the Berkeley Linguistics Society* 14: 35–55. https://doi.org/10.3765/bls.v14i0.1794
Frank, Stefan L., Bod, Rens & Christiansen, Morten H. 2012. How hierarchical is language use? *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20121741. <rspb.royalsocietypublishing.org> (24 August, 2013). https://doi.org/10.1098/rspb.2012.1741
Freudenthal, Daniel, Pine, Julian & Gobet, Fernand. 2010. Explaining quantitative variation in the rate of optional infinitive errors across languages: A comparison of MOSAIC and the Variational Learning Model. *Journal of Child Language* 37(3): 643–669. https://doi.org/10.1017/S0305000909990523
Gast, Volker & Diessel, Holger. 2012. The typology of clause linkage: Status quo, challenges, prospects. In *Clause Linkage in Cross-linguistic Perspective: Data-driven Approaches to Cross-clausal Syntax*, Volker Gast & Holger Diessel (eds), 1–36. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110280692.1
Gili Gaya, Samuel. 1983. *Curso superior de sintaxis española*. Barcelona, España: Bibliograf.
Givón, T. 1979. *On Understanding Grammar*. New York NY: Academic Press.
Givón, T. 2009. The ontogeny of complex verb phrases. In *Syntactic Complexity. Diachrony, Acquisition, Neuro-cognition, Evolution* [Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 311–387. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.13the
Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago IL: Chicago University Press.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalizations in Language*. Oxford: OUP.

Gras, Pedro. 2011. Gramática de construcciones en interacción: Propuesta de un modelo y aplicación al análisis de estructuras independientes con marcas de subordinación en español. PhD dissertation, Universitat de Barcelona.

Gras, Pedro. 2013. Entre la gramática y el discurso: Valores conectivos de *que* inicial átono en español. In *Autour de que/El entorno de que*, Daniel Jacob & Katja Ploog (eds), 89–112. Frankfurt: Peter Lang.

Heine, Bernd & Kuteva, Tania. 2007. *The Genesis of Grammar. A Reconstruction*. Oxford: OUP.

Israel, Michael, Johnson, Christopher & Brooks, Patricia J. 2000. From states to events. The acquisition of English passive participles. *Cognitive Linguistics* 11(1–2): 103–129.

Johnson, Christopher. 1999. *Constructional Grounding: The Role of Interpretational Overlap in Lexical and Constructional Acquisition*. PhD dissertation, University of California, Berkeley CA.

Johnson, Christopher. 2001. Constructional grounding: On the relation between deictic and existential there-constructions in acquisition. In *Conceptual and Discourse Factors in Linguistic Structure*, Alan J. Cienki, Barbara J. Luka & Michael B. Smith (eds), 123–136. Stanford CA: CSLI.

Kidd, Evan, Brandt, Silke, Lieven, Elena & Tomasello, Michael. 2007. Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes* 22: 860–897. https://doi.org/10.1080/01690960601155284

Köymen, Bahar & Kyratzis, Amy. 2014. Dialogic syntax and complement constructions in toddlers' peer interactions. *Cognitive Linguistics* 25: 497–521. https://doi.org/10.1515/cog-2014-0028

Kyratzis, Amy, Guo, Jiansheng & Ervin-Tripp, Susan M. 1990. Pragmatic conventions influencing children's use of causal constructions in natural discourse. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 16: 205–214. https://doi.org/10.3765/bls.v16i0.1711

Langacker, Ronald W. 2000. A dynamic usage-based model. In *Usage-based Models of Language*, Michael Barlow & Suzanne Kemmer (eds), 1–63. Stanford CA: CSLI.

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

Levinson, Stephen C. 2006. On the human 'interactional engine'. In *Roots of Human Sociality. Culture, Cognition and Human Interaction*, Nicholas J. Enfield & Stephen C. Levinson (eds), 29–69. Oxford: Berg.

Lieven, Elena. 2006. Producing multiword utterances. In *Constructions in Acquisition*, Eve V. Clark & Barbara Kelly (eds), 83–110. Stanford CA: CSLI.

Lieven, Elena. 2008. Learning the English auxiliary: A usage-based approach. In *Trends in Corpus Research: Finding Structure in Data* [Trends in Language Acquisition Research 6], Heike Behrens (ed.), 61–98. Amsterdam: John Benjamins.

Lieven, Elena, Behrens, Heike, Speares, Jennifer & Tomasello, Michael. 2003. Early syntactic creativity: A usage-based approach. *Journal of Child Language* 30: 333–370. https://doi.org/10.1017/S0305000903005592

Lieven, Elena, Salomo, Dorothé & Tomasello, Michael. 2009. Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics* 20(3): 481–507. https://doi.org/10.1515/COGL.2009.022

Luna-Traill, Elizabeth. 1980. *Sintaxis de los verboides en el habla culta de la ciudad de México*. México: Universidad Nacional Autónoma de México.

Mithun, Marianne. 2008. The extensión of dependency beyond the sentence. *Language* 84(1): 69–119. https://doi.org/10.1353/lan.2008.0054

Montolío, Estrella. 1999. ¡Si nunca he dicho que estuviera enamorada de él! Sobre construcciones independientes introducidas por *si* con valor replicativo. *Oralia* 2: 37–70.

Moreno de Alba, José G. 1978. *Valores de las formas verbales en el español de México*. México: Universidad Nacional Autónoma de México.

Newport, Elissa L. 1990. Maturation constraints on language learning. *Cognitive Science* 14: 11–28. https://doi.org/10.1207/s15516709cog1401_2

Ninio, Anat. 2001. Pragmatic keywords and first combining verbs. *First Language* 21(3): 433–460. https://doi.org/10.1177/014272370102106309

Ochs, Elinor, Schieffelin, Bambi B. & Platt, Martha L. 1979. Propositions across utterances and speakers. In *Developmental Pragmatics*, Elinor Ochs & Bambi B. Schieffelin (eds), 251–268. New York NY: Academic Press.

Ochs, Elinor, Schegloff, Emanuel A. & Thompson, Sandra A. (eds). 1996. *Interaction and grammar*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511620874

Pons Bordería, Salvador. 2003. *Que* inicial átono como marca de modalidad. *Estudios de Lingüística* 17: 531–545.

Porroche Ballesteros, Margarita. 2000. Algunos aspectos del uso de *que* en el español conversacional (*que* como introductor de oraciones 'independientes'). *Círculo de Lingüística Aplicada a la Comunicación* 3. <https://dialnet.unirioja.es/servlet/articulo?codigo=276219&orden=0&info=link> (15 August 2013).

Rodríguez Ramalle, Teresa María. 2008. Estudio sintáctico y discursivo de algunas estructuras enunciativas y citativas del español. *Revista Española de Lingüística Aplicada* 21: 269–288.

Rojas-Nieto, Cecilia. 1992. La pregunta en la construcción del diálogo: Funciones constructivas de las preguntas a los niños pequeños. *Revista de Lingüística Aplicada* 15–16: 182–198.

Rojas-Nieto, Cecilia. 2003. Early acquisition of verb inflexion in Spanish. A usage-based account. *Psychology of Language and Communication* 7(2): 33–56.

Rojas-Nieto, Cecilia. 2007. La base de datos *ETAL*: Etapas tempranas en la adquisición del lenguaje. Origen, descripción y metas de un proyecto. *Jornadas Filológicas 2005*: 575–599. México: Universidad Nacional Autónoma de México.

Rojas-Nieto, Cecilia. 2009a. Starting small effects in the acquisition of early relative constructions in Spanish. In *The Genesis of Syntactic Complexity. Diachrony, Acquisition, Neuro-cognition, Evolution* [Typological Studies in Language 85], T. Givón & Masayoshi Shibatani (eds), 277–310. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.12sta

Rojas-Nieto, Cecilia. 2009b. Before grammar. Cut and paste in early complex sentences. In *Hispanic Child Languages. Typical and Impaired Development* [Language Acquisition and Language Disorders 50], John Grinstead (ed.), 143–174. Amsterdam: John Benjamins. https://doi.org/10.1075/lald.50.07roj

Sansiñena, María Sol. 2015. The Multiple Functional Load of *que.* An Interactional Approach to Insubordinate Complement Clauses in Spanish. PhD dissertation, Catholic University Leuven.

Sansiñena, María Sol, De Smet, Hendrick & Cornillie, Bert. 2015. Between subordinate and insubordinate. Paths toward complementizer-initial main clauses. *Journal of Pragmatics* 7: 3–19. https://doi.org/10.1016/j.pragma.2014.12.004

Serrat, Elisabet & Aparici, Melina. 2001. Morphological errors in early language acquisition: Evidence from Catalan and Spanish. In *Research on Child Language Acquisition*, Margareta Almgrem, Andoni Barreña, Ma. José Ezeizabarrena, Itziar Idiazábal & Brian MacWhinney (eds), 1260–1277. Somerville MA: Cascadilla Press.

Sheridan, Guillermo. 2007. Quijote furioso. *Paralelos y Meridianos*, 11–33. México: Pértiga.

Slobin, Dan I. 1985. Why study language crosslinguistically? In *The Crosslinguistic Study of Language Acquisition*, Vol. 1: *The Data*, Dan I. Slobin (ed.), 3–24. Hillsdale NJ: Lawrence Erlbaum Associates.

Tomasello, Michael. 1992. *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511527678

Tomasello, Michael. 2000. The item based nature of children's early syntactic development. *Trends in Cognitive Sciences* 4: 156–163. https://doi.org/10.1016/S1364-6613(00)01462-5

Tomasello, Michael. 2003. *Constructing a Language. A Usage-based Child Language Acquisition*. Cambridge MA: Harvard University Press.

Tomasello, Michael. 2006. Acquiring linguistic constructions. In *Handbook of Child Psychology*, Vol. 2: *Cognition, Perception and Language*, William Damon, Richard M. Learner, Deanna Kuhn, Robert Siegler (eds), 255–298. Hoboken, NJ: Wiley. https://doi.org/10.1002/9780470147658.chpsy0206

Tomasello, Michael. 2008. *Origins of Human Communication*. Cambridge MA: The MIT Press.

Van Veen, Rosie, Evers-Vermeul, Jacqueline, Sanders, Ted & Van Den Bergh, Huub. 2009. Parental input and connective acquisition: Growth curve analysis. *First Language* 29: 266–288. https://doi.org/10.1177/0142723708101679

Van Veen, Rosie, Evers-Vermeul, Jacqueline, Sanders, Ted & Van Den Bergh, Huub. 2013. The influence of input on connective acquisition: A growth curve analysis of English *because* and German *weil*. *Journal of Child Language* 40: 1003–1031. https://doi.org/10.1017/S0305000912000451

Varela Cortés, Vianey. 2006. *Interacción y diálogo en la adquisición de las construcciones introducidas por pero*. MA dissertation, Universidad Nacional Autónoma de México.

Varela Cortés, Vianey. 2011. El nicho discursivo de las construcciones con *pero* en la adquisición temprana del español. In *Interacción y uso lingüístico en el desarrollo de la lengua materna*, Cecilia Rojas-Nieto & Donna Jackson-Maldonado (eds), 173–208. México: Universidad Nacional Autónoma de México.

PART III

The syntactic complexity of adverbial clauses

The predicates of Luiseño clausal adjuncts

Susan Steele
University of California, Berkeley

The predicates of Luiseño clausal adjuncts are informationally unique among the predicates of its subordinate clause types in ensuring access to their temporal properties; they are morphologically unique in that this characteristic coexists with properties of form otherwise associated with predicates whose temporal properties are not accessible. The character of this uniqueness is consistent with two facts: (1) A clausal adjunct's temporality must be determined relative to that of the clause to which it is adjoined; and (2) if the subject of a clausal adjunct is not given internal to the clause, it is determined relative to that of the clause to which it is adjoined.

Keywords: clausal adjunct, complement clauses, relative clauses, informational structure, temporal and subject values

1. Introduction

Of the three types of subordinate clauses – complements, relative clauses, and clausal adjuncts (often called 'adverbial clauses') – the last is the poor stepsister in terms of the amount of attention it has received, whether in theoretical or descriptive circles.[1] By offering a formal treatment of the essential properties of adjunct clauses in Luiseño, a Uto-Aztecan language of Southern California, specifically in contrast to the essential properties of complements and relative clauses, this paper takes Cinderella to the ball.[2]

---

1. To my mind, 'clausal adjunct' better captures the character of this subordinate clause type than does 'adverbial clause'. An adjunct is a non-essential addition whose presence does not give lexical form to properties of the structure to which it is adjoined. In this sense, both a relative clause and an 'adverbial clause' are adjuncts – the second to the clause (i.e. a combination of a predicate and its arguments) and the first to a part of the clause.

2. The data in this paper are drawn largely from my work with the late Villiana Hyde, a speaker of the Rincón dialect of Luiseño. Examples are presented in the orthography of Hyde (1971),

https://doi.org/10.1075/tsl.126.04ste

The Luiseño sentences in (1) illustrate each of the three subordinate clause types, in each case with a single word, the predicate, indicated in bold. For reasons that will shortly become clear, this word is presented without an internal analysis, but because all three involve the same base, *tooya* 'laugh', the locus of our attention is reasonably obvious.[3] Since all properties identifying the clause type are part of the morphology of its predicate, this need be the only word in a subordinate clause.

(1) a. Clausal adjunct
tooyaqanik *nawitmal wi'iw.*
after.laughing girl made.wiwish
'After laughing, the girl made wiwish.'

b. Complement clause
tiiw'yax nil nawitmali ***tooyaqal.***
saw AUX girl.OBJECT laughing.OBJECT
'I saw the girl laughing.'

c. Relative clause
chaqalaqiqu$ nil nawitmali ***tooyaqalmokwichi.***
was.tickling AUX girl.OBJECT laughed.NUMBER&OBJECT
'I was tickling the girl who laughed.'

No subordinate clause can include the aux and, thus, all subordinate clauses can be distinguished from (non-imperative) main clauses on this score.[4]

(2) Main clause
tooyaqu$ upil.
was.laughing AUX
'S/he was laughing.'

---

although I use the symbol $ to replace the symbol Hyde employs for a retroflex fricative. This orthography is fundamentally IPA-based, but employs the symbols *ch* and *sh,* for an alveopalatal voiceless affricate and an alveopalatal voiceless fricative, respectively; ', for glottal stop; *ng,* for a velar nasal; and *VV*, for a long vowel.

3. The base at issue is actually *tooyax*, with /x/ deleted before the homorganic stop /q/. As simple as this case seems, the conditions under which this segment is present has excited considerable debate. See Jacobs (1975) and Davis (1973) for two somewhat related takes on the issue. Steele (1985) offers another, one I now think to be totally misguided.

4. 'Aux' is an old-fashioned term, but still the best, for Luiseño's second position clitic complex. (See Akmajian, Steele & Wasow (1979) and Steele et al. (1981) for discussion). With the exception of sentences where the predicate is an imperative, every Luiseño main clause can include an aux. In casual speech, the aux is largely absent and in stories or other extended text types, it is often found only in the first sentence or in a sentence introducing some kind of shift in perspective.

Our focus is on the morphological properties of the words functioning as predicates in each of the three subordinate clause types. The fundamental idea is that words that function as the predicate in a clausal adjunct are distinctive among the three in giving prominence to both their temporal properties and their properties of person – the first in the word's informational structure and the second in its morphological form.[5]

The discussion is based on two assumptions:

First, a word is associated with an articulated informational structure, of the sort employed in Generalized Phrase Structure Grammar and its off-shoots, composed of the set of attributes and their values. For the purposes of this paper, we attend only to the four indicated in (3), where the first three have to do with the notional properties their label indicates and the fourth has to do with a word's syntactic potential.[6]

(3) PER(SON): …
N(UMBER): …
TEMP(ORALITY): …
FIN(AL): …

The value for the attribute FIN can be entirely independent of the notional properties represented in the values for the other two attributes, but it need not be. That is, the value for FIN can include temporality, person, and number – although not all simultaneously – but not all values for FIN include the notional properties represented in the other three attributes. Where the value for FIN is independent of the informational structure's notional properties, the word's syntactic potential does not give access to its notional properties; where it is not independent, the word's syntactic potential involves – and gives access to – the notional property it represents.

Second, consistent with the evisceration of the traditional concept of a morpheme in Anderson (1992), this treatment is not based on the assumption of a one-to-one match between the values in an informational structure and the formatives that comprise the word. There is reason to think at the outset that such an approach might be a productive strategy. If the value for PER includes person and number and the word includes a base, the associated morph, one of those in (4),

5. 'Informational structure' identifies the set of attributes and their values that contain the notional and combinatory information associated with a word form. It is not to be confused with 'information structure', as employed to capture 'topic', 'focus', and the like.

6. The full set of attributes includes, in addition, SUBCAT(EGORIZATION), whose value has to do with a word's combinatory potential relative to arguments – like the value for FIN, a syntactic property.

must appear to the immediate left of the base and at the word's left edge, as shown in (5).

(4) NO CHAM
'O 'OM
PO POM

(5) $_{word}$[NO[base… = PER: 1SG
$_{word}$['O[base… = PER: 2SG
$_{word}$[PO[base… = PER: 3SG
$_{word}$[CHAM[base… = PER: 1PL
$_{word}$['OM[base… = PER: 2PL
$_{word}$[POM[base… = PER: 3PL

Further, if an informational structure includes [FIN: postposition] and the word includes a base, the associated morph, most of which are in (6), must appear to the immediate right of the base and at the word's right edge, as shown in (7).

(6) IK 'to'
NGA 'at, on, in'
NGAY 'from'

(7) …base]IK]$_{word}$ = FIN: postposition
…base]NGA]$_{word}$ = FIN: postposition
…base]NGAY]$_{word}$ = FIN: postposition

These generalizations hold regardless of any other properties of the informational structure. It is noteworthy, then, that one of the other values for PER ([PER: null]) cannot be associated with a morphological form in the absence of reference to the values for TEMP and FIN nor can other values for FIN be associated with a morphological form in the absence of reference to the values for PER, N and, arguably, TEMP. In fact, with the exception of the morphs in (4) and (6), the informational context writ large is critical to a word's morphological form. Unlike the situation in (5) and (7), there isn't a simple association between a value for an attribute and a formative.[7]

This fact underlies the non-morphemic approach taken in this paper. The conditions governing the form associated with the informational structure of a clausal adjunct's predicate are not unique among subordinate clause predicates. However, interestingly, they yield a unique form, one consistent with their syntactic property of giving prominence to both temporal properties and properties of person.

7. Insofar as both 5 and 7 require mention of a base, the association is not as straightforward as it might appear.

2. Temporal properties and a word's informational structure

The informational structure of the predicate in a clausal adjunct is distinctive on two counts, both having to do with temporality. First, its value for FIN(AL), the attribute that identifies a word's syntactic potential, must be temporal. Second, its value for TEMP, the attribute that indicates the presence or absence of notional temporality, must be accessible because it is contained by the value for FIN. Neither of these properties holds for words that function as the predicate in a relative clause; some words that function as the predicate in a complement may have both properties, but they are not obligatory to all such.

As noted, the values for the attribute FIN(AL) identify the word's syntactic potential, the role(s) a word can play within a syntactic context. For our purposes, the major distinction in terms of the value for this attribute has to do with the presence or absence of temporality ([FIN: temporal value]) or not ([FIN: no temporal value]), where both 'temporal value' and 'no temporal value' are abbreviations for the actual values. The examples in 1 are illustrative. The predicate of the clausal adjunct in 1a is specified [FIN: preceding&continous], one of the temporal values for FIN. The values for FIN in the predicates in (1b) and (1c) are non-temporal – that in the complement in (1b) is [FIN: object]; that in the relative clause in (1c) is [FIN: number&object].

The predicate in a complement is not limited to a non-temporal value; it may also be temporal. The values subsumed by the former are [FIN: object], [FIN: postposition], [FIN: person&number], and [FIN: number]; that subsumed by the latter is [FIN: infinitive& number].

(8) predicate in a complement

a. FIN: non-temporal value

tooyaqal	'laughing'
[FIN: OBJECT]	
chaqalaqiqalanga	'being tickled'
[FIN: postposition]	
notooyaxpi	'1SG will laugh'
[FIN: person&NUMBER]	
tooyaqat	'laugh'
[FIN: NUMBER]	

b. FIN: temporal value

potooyilo	'to laugh'
[FIN: infinitive&NUMBER]	

However, FIN in the informational structure of every word functioning as the predicate in a clausal adjunct has one of the temporal values. These include, in addition

to [FIN: preceding&continuous] as in (1a), [FIN: preceding&non-continuous], [FIN: continuous] and [FIN: non-continuous&number].

(9) predicate in a clausal adjunct = [FIN: temporal value]

tooyaqanik	'after laughing (for a while)'
[FIN: preceding&continuous]	
tooyanik	'after laughing (briefly)'
[FIN: preceding&non-continuous]	
notooyaqala	'1SG laughing'
[FIN: continuous]	
tooyaat	'laughing'
[FIN: non-continuous&NUMBER]	

And, FIN in the informational structure of every word that functions as the predicate in a relative clause has a non-temporal value – one of [FIN: number&object], [FIN: postposition], and [FIN: number].

(10) predicate in a relative clause = [FIN: non-temporal value]

tooyaqalmokwichi	'laughed'
[FIN: NUMBER&OBJECT]	
polovi'ivonga	'on X(s) 3SG made'
[FIN: postposition]	
tooyaxlut	'will laugh'
[FIN: NUMBER]	

The inventory of temporal values for FIN varies with the clause type; the inventory of non-temporal values includes some that are specific to a clause type and others that are shared across clause types. But the critical point for our purposes is the contrast between temporality and non-temporality in FIN's value: The informational structures of words that serve as the predicate in a clausal adjunct always include the former and, as such, they are maximally opposed to those for words that serve as the predicate in a relative clause, which always include the latter.

(11)

FIN: temporal value	FIN: non-temporal value
	predicate in a relative clause
	[FIN: NUMBER]
	[FIN: NUMBER&OBJECT]
	[FIN: postposition]
predicate in a complement	predicate in a complement
[FIN: infinitive&NUMBER]	[FIN: OBJECT]
	[FIN: NUMBER]
	[FIN: person&NUMBER]
	[FIN: postposition]

predicate in a clausal adjunct
[FIN: preceding&continuous]
[FIN: preceding&non-continuous]
[FIN: continuous]
[FIN: non-continuous&NUMBER]

This isn't to say that the informational structures of subordinate predicates where the value for FIN is non-temporal cannot indicate temporality. In fact, all of the examples in (8a) and (10) indicate temporality, but only as the value for the attribute TEMP(ORAL), not as the value for the attribute FIN. (12a) displays both values for the examples in (8a); (12b) displays both values for the examples in (10).[8]

(12) [TEMP: temp]/[FIN: no temporal value]

a. *tooyaqal* 'laughing'
[TEMP: continuous]/[FIN: OBJECT]
chaqalaqiqalanga 'being tickled'
[TEMP: continuous]/[FIN: postposition]
notooyaxpi '1SG will laugh'
[TEMP: non-continuous future]/[FIN: person&NUMBER]
tooyaqat 'laugh'
[TEMP: continuous]/[FIN: NUMBER]

b. *tooyaqalmokwichi* 'laughed'
[TEMP: continuous past]/[FIN: NUMBER&OBJECT]
polovi'ivonga 'on X(s) 3SG made'
[TEMP: non-continuous past]/[FIN: postposition]
tooyaxlut 'will laugh'
[TEMP: non-continuous future]/[FIN: NUMBER]

Recall the distinction between the value for the attribute TEMP and the value for the attribute FIN. The first is purely notional; the latter identifies a word's syntactic potential. Consistent with this distinction between the attributes TEMP and FIN is the fact that words that can function as the predicate in a complement or as the predicate in a relative clause need not include a temporal value in their informational structure. Like the last of the temporal examples in (8a), *yawaywish* in (13a) is the predicate in a complement specified [FIN: number] and like the first of the temporal examples in (10), *yawaywichi* in (13b) is the predicate in a relative

8. The temporal values for TEMP in the predicate of both relative clauses and complements include two that are simple aspectual ([TEMP: continuous] and [TEMP: non-continuous]) as well as four that are combinations of aspect and a temporal location ([TEMP: continuous past], [TEMP: continuous future], [TEMP: non-continuous past], and [TEMP: non-continuous future]), although in each kind of predicate the selection from this inventory varies with the value for FIN. The temporal values for TEMP in complements also include [TEMP: generic].

clause specified [FIN: number&object]. In neither of these cases is this attribute/value pair accompanied by an instance of [TEMP: temp]; rather, the informational structure of these words is specified [TEMP: null], to indicate the absence of notional temporality.

(13) a. [TEMP: null]/[FIN: NUMBER]
nawitmal upil ***yawaywish*** *miyqu$.*
girl AUX nice was
'The girl was nice.'

b. [TEMP: null]/[FIN: NUMBER&OBJECT]
chaqalaqiqu$ nil nawitmali ***yawaywichi.***
was.tickling AUX girl.OBJECT nice.NUMBER&OBJECT
'I was tickling the nice girl.'

The informational structures of the words in (8b) and (9) also include the attribute TEMP. But in these, the value for this attribute is contained by the value for FIN; that is, the value for TEMP must appear in the value for FIN. So, the word in (8b) is [TEMP: infinitive]/[FIN: infinitive&number] and the words in (9) are [TEMP: continuous]/[FIN: preceding&continuous], [TEMP: non-continuous]/[FIN: preceding&non-continuous], [TEMP: continuous]/[FIN: continuous], and [TEMP: non-continuous]/[FIN: non-continuous&number], respectively.

Example (14) adds the information about the values for the attribute TEMP to the distinction based on the value for the attribute FIN in (11).

(14) a. [TEMP: temp]/[FIN: temporal value]
predicate in a clausal adjunct
predicate in a complement

b. [TEMP: temp]/[FIN: no temporal value]
predicate in a complement
predicate in a relative clause

c. [TEMP: null]/[FIN: no temporal value]
predicate in a complement
predicate in a relative clause

All three of the combinations in (14) are options for the informational structure of the predicate in a complement, while the informational structure of the predicate of a clausal adjunct and that of the predicate of a relative clause remain maximally opposed. In the first, the value for the attribute indicating syntactic potential (i.e. FIN) must include temporality as must the value for the attribute indicating notional temporality (i.e. TEMP) and the latter must be contained by the former. In the second, the value for the attribute indicating syntactic potential must not include temporality and the value for TEMP may or may not, but it cannot be contained by the value for FIN.

The idea that the presence of a temporal value in FIN gives access to the notional temporality indicated in the value for TEMP – and the presence of a temporal value in FIN is possible only if the value for TEMP is also temporal – captures a fact about all clausal adjuncts, as well as about the single complement type with a predicate where the value for FIN is temporal: Their time is determined relative to another temporal reference point.[9] In the case of a clausal adjunct, this is the clause to which it is adjoined. So, whether the predicate in a clausal adjunct is [TEMP: continuous]/[FIN: preceding&continuous], as shown in (15), or [TEMP: non-continuous] /[FIN: non-continuous&number], as shown in (16), the time of the clause varies with the tense of the main clause.

(15) a. ***tooyaqanik*** *nawitmal wi'iw.*
after.laughing girl made.wiwish
'After laughing, the girl made wiwish.'
b. ***tooyaqanik*** *nawitmal wiiwmaan.*
after.laughing girl will.make.wiwish
'After laughing, the girl will make wiwish.'

(16) a. ***tooyaat*** *nawitmal wi'iw.*
laughing girl made.wiwish
'The girl made wiwish while laughing/while she laughed.'
b. ***tooyaat*** *nawitmal wiiwmaan.*
laughing girl will.make.wiwish
'The girl will make wiwish while laughing/while she laughs.'

In the case of the complement, the temporal reference point is the clause in which it is an argument.

(17) a. *looviq up* ***potooyilo.***
is.good AUX to.laugh
'It is good to laugh.'
b. *yamayk upil looviqu$* ***potooyilo.***
long.ago AUX was.good to.laugh
'It used to be good to laugh.'

Where FIN in the informational structure of a subordinate clause predicate has a non-temporal value, the temporal reference point of the clause varies, first, with whether it is accessible to the predicate in the embedding clause. That of a relative clause never is and its temporal reference point is independent of the embedding clause. The predicate in both relative clauses in (18) is [TEMP: continuous past]/

9. Steele (1990) makes a distinction between 'relative to' and 'identical with'. This distinction isn't critical to the point at issue here.

[FIN: number&object]; that in both relative clauses in (19) is [TEMP: continuous]/[FIN: number&object]. In neither case is the time of the relative clause determined relative to the main clause.

(18) a. *chaqalaqiqu$ nil nawitmali* ***tooyaqalmokwichi.***
was.tickling AUX girl.OBJECT laughed.NUMBER&OBJECT
'I was tickling the girl who laughed.'
b. *chaqalaqin nupo nawitmali* ***tooyaqalmokwichi.***
will.tickle aux girl.OBJECT laughed.NUMBER&OBJECT
'I will tickle the girl who laughed.'

(19) a. *chaqalaqiqu$ nil nawitmali* ***tooyaqati.***
was.tickling AUX girl.OBJECT laugh.NUMBER&OBJECT
'I was tickling the girl who is laughing.'
b. *chaqalaqin nupo nawitmali* ***tooyaqati.***
will.tickle AUX girl.OBJECT laugh.NUMBER&OBJECT
'I will tickle the girl who is laughing.'

A non-temporal value for FIN in the predicate of a complement, as the locus of the fact that the clause is an argument to the embedding predicate, is always accessible to this predicate. However, embedding predicates differ as to whether they 'call' its temporal property, i.e. as to whether they require that their argument is a complement or not. If they do, the time of the complement is determined relative to that of the embedding predicate; if they do not, the time of the complement is independent of that of the embedding predicate. For example, *tiiwi* 'see' can take two object-marked arguments, but, if it does, one must be a complement, while *ʼayali* 'know' requires an object-marked argument, but doesn't distinguish between a simple argument and a complement. The sentences in (20) illustrate that tense of the embedding predicate *tiiwi* 'see' provides the reference point of the time of the complement. The sentences in (21) illustrate the independence of the time of the complement relative to the tense of *ʼayali* 'know', its embedding predicate.

(20) a. *tiiwiqu$ mil hengeemali* ***tooyaqal.***
saw AUX boy.OBJECT laughing.OBJECT
'They saw the boy laughing.'
b. *tiiwin mo hengeemali* ***tooyaqal.***
will.see AUX boy.OBJECT laughing.OBJECT
'They will see the boy laughing.'

(21) a. *ʼayaliqu$ mil* ***potooyaxvoy.***
knew AUX 3SG.laughed.OBJECT
'They knew that he laughed.'

b. ˈ*ayalin* *mo* ***potooyaxvoy.***
will.know AUX 3SG.laughed.OBJECT
'They will know that he laughed.'

In sum, the informational structure of the predicate in a clausal adjunct is distinctive on two counts, both having to do with temporality. First, its value for FIN, the attribute that identifies a word's syntactic potential, must be temporal. The value for FIN in the informational structure of the predicate in a complement can be temporal, but need not be; the value for FIN in the informational structure of the predicate in a relative clause cannot be temporal. Second, its value for TEMP, the attribute that indicates the presence or absence of notional temporality, must be accessible because it is contained by the value for FIN. The value for TEMP in the informational structure of the predicate of a relative clause is inaccessible. The value for TEMP in the informational structure of the predicate of a complement can be accessible, but it need not be. And, should it be, the source of the accessibility can be either because it is contained by the value for FIN or because the embedding predicate 'calls' its value.

These two properties distinguishing the predicate in a clausal adjunct from other words that can function as a subordinate predicate are consistent with the fact that the time of a clausal adjunct must be relative to the clause to which it is adjoined.

3. Properties of form

The predicate of a clausal adjunct is distinctive among words that can function as the predicate in a subordinate clause in that its informational structure ensures access to its temporal properties. In contrast, the conditions governing the form associated with the informational structure of a clausal adjunct's predicate are part and parcel of those governing the form of subordinate clause predicates more generally. This section is concerned with this fact and its implications.

3.1 Values for PER and N

As indicated in (2), the informational structure of a Luiseño word includes, in addition to the attributes FIN and TEMP, the attributes PER(SON) and N(UMBER). Like the value for TEMP, the values for these attributes are notional. Every informational structure indicates the property of number in the value for the attribute N, so the only difference between informational structures in this regard can be the

particular value – i.e. [N: sg], [N: pl], or [N: numb].[10] The value for PER indicates a more dramatic difference – whether a word expresses person ([PER: 1, 2, or 3]) or not ([PER: null]). The values '1' and '2' are always accompanied by a number value; the value '3' needn't be. (That is, in addition to the values involving person and number in (5), there is also [PER: 3].)

The informational structure of a predicate in all three of the subordinate clause types can be specified for one of the six values for the attribute PER involving person and number. For example, the informational structures of the words in (22) are such that they can serve as the predicate in a clausal adjunct, a complement, and a relative clause, respectively.

(22) a. [PER: 1SG]/[TEMP: continuous]/[FIN: continuous]
notooyaqala '1SG laughing'
cf. ***notooyaqala*** *nawitmal* *wi'iw.*
1SG.laughing girl made.wiwish
'While I was laughing, the girl made wiwish.'

b. [PER: 3SG]/[TEMP: non-continuous past]/[FIN: OBJECT]
potooyaxvoy '3SG laughed' (cf. 21)

c. [PER: 1SG]/[TEMP: non-continuous future]/[FIN: NUMBER&OBJECT]
nochaqalaqipiy '1SG will tickle'
cf. *'ayaliq* *up* *nawitmali* ***nochaqalaqipiy.***
knows AUX girl.OBJECT 1SG.will.tickle.NUMBER&OBJECT
'He knows the girl that I will tickle.'

The informational structure of a predicate in all three subordinate clause types can also be specified [PER: null]. The examples in (23) are illustrative.

(23) a. [PER: null]/[TEMP: continuous]/FIN: continuous&preceding]
tooyaqanik '(after) laughing' (cf. 15)

b. [PER: null]/[TEMP: continuous]/[FIN: postposition]
chaqalaqiqalanga 'being tickled'
cf. *tiiwiqu$* *nil* *nawitmali* ***chaqalaqiqalanga.***
saw AUX girl being.tickled
'I saw the girl being tickled.'

10. [N: pl] indicates the possibility of individuating 'at least two'. [N: numb] indicates that individuation may be possible, but is not essential, a concept that will be characterized, somewhat inadequately, as 'at least one'. [N: sg] represents 'exactly one', but this includes two possibilities – either an individual or a collection.

c. [PER: null]/[TEMP: null]/[FIN: postposition]
yawaywinga 'in nice'
cf. *wiiwq* *up* *pokiinga* ***yawaywinga.***
is.making.wiwish AUX 3SG.house.in nice.in
'S/he is making wiwish in her/his nice house.'

3.2 Temporality and form

The form of the predicate in a clausal adjunct must include temporal morphology. The form of the predicate in a relative clause or a complement may but need not. It does not, of course, when the informational structure is non-temporal (i.e. [TEMP: null]); it also does not when the informational structure includes [TEMP: non-continuous]. However, in all three cases, should the informational structure of the predicate in a subordinate clause allow for temporal morphology, it is always immediately right adjacent to the base. The temporal words in (22) are illustrative, as shown in (24), where the temporal morphology is represented in capital letters.

(24) a. *notooyaQALA* '1SG laughing'
b. *potooyaxVOy* '3SG laughed'
c. *nochaqalaqiPIy* '1SG will tickle'

The informational structure of these words is specified [PER: person&number]. The form associated with such an informational structure includes one of the morphs in (4), located at the word's left edge and immediately left adjacent to its base, as shown in (5). The subordinate predicates in (23a) and (23b), specified [PER: null] and, thus, lacking one of these morphs, indicate that neither the value for PER nor the presence of one of these morphs has any bearing on the location of temporal morphology in a subordinate clause predicate.

(25) a. *tooyaQANIK* '(after) laughing'
b. *chaqalaqiQALAnga* 'being tickled'

The location of the temporal morphology in the words in (24) and (25) does suggest, however, a difference between the predicate in a clausal adjunct and the predicate in either a complement or relative clause. The temporal morphology in (22a) and (23a), both clausal adjunct predicates, is not only immediately right adjacent to the base, it is at the right edge of the word. The temporal morphology in (24b) (complement predicate), (24c) (relative clause predicate) and (25b) (complement predicate) is not. Upon closer inspection, this difference dissolves. Whether the informational structure includes [PER: person&number] or [PER: null], the fundamental principle governing whether temporal morphology can be at the right

edge of the word or not has to do with whether the value for FIN includes a notional property and, if so, which.

If FIN in informational structures with either PER value is specified for a non-notional property, the temporal morphology cannot be at the word's right edge. Since the value for FIN in the informational structure for a clausal adjunct predicate always involves temporality, this condition does not apply. It does apply to the complement predicate in (24b) ([FIN: object]), the relative clause predicate in (24c) ([FIN: number&object]), and the complement predicate in (25b) ([FIN: postposition]). In the first two forms, the morph I (/y/ after a vowel) is at the word's right edge; in the last, the postposition NGA (cf. 7) is at the word's right edge. As indicated at the beginning of this paper, a word associated with an informational structure specified [FIN: postposition] must include one of the postpositions at the word's right edge. A word associated with either of the two FIN values specified 'object' ([FIN: object] or [FIN: number&object]) may or may not include the morph I, but it always includes morphology to the right of the base – and to the right of temporal morphology where such is present. We return to this point below.

Our primary concern in regard to temporal morphology and the right edge of the word, then, has to do with informational structures where the value for FIN is notional only. The notional properties at issue vary with the value for PER.

If the value for FIN in an informational structure specified [PER: person& number] contains the value for TEMP, the temporal morphology must be both immediately right adjacent to the base and at the word's right edge, but this location also holds when the value for FIN contains the value for PER. No informational structure for the predicate in a relative clause has either FIN value option (cf. 12b), but the informational structure of the complement predicate shown in (17) has the first and that of the third complement predicate in (12a) has the second – and as illustrated in (26), the temporal morphology in both cases is at the right edge of the word.

(26) a. [PER: 3sg]/[TEMP: infinitive]/[FIN: infinitive&NUMBER]
potooyiLO 'to laugh'
b. [PER: 1SG]/[TEMP: non-continuous future]/[FIN: person&NUMBER]
notooyaxPI '1SG will laugh'

A slightly more complicated situation holds for informational structures of the form [PER: person&number]/[FIN: number]. Unless the informational structure is specified [FIN: pl], the location of the temporal morphology is as in (26) or (24a), i.e. both immediately right adjacent to the base and at the right edge of the word. However, if the informational structure is specified [FIN: pl] – and [FIN: pl] finds morphological form, as the morph UM – the temporal morphology need not be at the right edge of the word. (The conditions under which [FIN: pl] in

an informational structure specified [PER: person&number]/[FIN: pl] may find morphological form are beyond the scope of this paper). (27) illustrates both possibilities with forms associated with informational structures for a relative clause predicate; (28), both possibilities with forms associated with informational structures for a complement predicate. In the first, [PER: person&number]/[FIN: numb] alternates with [PER: person&number]/[FIN: pl]; in the latter, the morph UM, not the temporal morphology, is at the word's right edge.[11] In the second, both informational structures are specified [PER: person&number]/[FIN: pl] and either the temporal morphology or the morph UM can be at the word's right edge.

(27) a. [PER: 1SG]/[N: numb][TEMP: non-continuous future]/[FIN: numb]
nochaqalaqiPI '1SG will tickle'
b. [PER: 1SG]/[N: numb]/[TEMP: non-continuous future]/[FIN: pl]
*nochaqalaqiPI**m*** '1SG will tickle'

(28) [PER: 1PL]/[N:pl]/[TEMP: generic]/[FIN: pl]
a. *chamchaqalaqAX* '1PL good at tickling'
b. *chamchaqalaqAX**um*** '1PL good at tickling'

In sum, there is a single situation where temporal morphology need not be at the right edge of the word in a subordinate predicate specified [PER: person&number] and for a notional value in FIN – if the informational structure includes [FIN: pl].

In subordinate predicates whose informational structures are specified [PER: null] and for a notional value in FIN, the only situation where temporal morphology must be at the right edge of the word is when FIN is specified for a temporal value unaccompanied by any other notional property. These informational structures – both of which identify the predicate in a clausal adjunct – are illustrated in (29).

(29) a. [PER: null]/[TEMP: continuous]/[FIN: continuous&preceding]
tooyaQANIK 'after laughing (for a while)'
b. [PER: null]/
[TEMP: non-continuous]/[FIN: non-continuous&preceding]
tooyaNIK 'after laughing'

This claim depends on the analysis of the form for all other [PER: null] subordinate clause predicates, except for those specified [FIN: postposition] (as in the complement predicate in 25b) as including one of the morphs generally referred to in Uto-Aztecan circles as an 'absolutive' (L/T/SH), a form appearing to the right of the

11. The informational structure in (27a) is much more highly valued than that in (27b).

base, albeit not necessarily immediately right adjacent.[12] (Which of the absolutive forms appears in the word is a function of properties of what appears to its left). (30) provides examples of words with this form and the potential to function as the predicate in a clause adjunct, a complement, and a relative clause respectively.[13] In each of these, the temporal morphology is right adjacent to the base but not at the word's right edge. The absolutive is capitalized and the temporal morphology is italicized.

(30) a. [PER: null]/[TEMP: non-continuous]/[FIN: non-continuous&NUMBER]
tooyaaT [*tooya**an**T*] 'laughing' (cf. 16)
b. [PER: null]/[TEMP: continuous]/[FIN: OBJECT]
tooyaqal [*tooya**qala**L*] 'laughing' (cf. 20)
c. [PER: null]/[TEMP: continuous past]/[FIN: NUMBER&OBJECT]
*tooya**qalmokwichi*** [*tooya**qalmokwi**SHi*] 'laughed' (cf. 18)

Descriptive treatments of Luiseño such as Kroeber and Grace (1960) and Elliott (1999) do not identify the words in (30) as containing an absolutive. The fundamental basis of their position is the idea that an absolutive is limited to a certain grammatical class of words – as Kroeber and Grace (p. 68) put it 'the absolutive suffixes…distinguish nonverbs from verbs….' – and their concomitant assumption that words like those in (29) are verbs.[14]

In the case of words that can function as the predicate in a subordinate clause, the evidence in favor of the analysis indicated in (30) follows from a simple fact: the character of the morphology following the forms in question is as would be expected if the final segment is an absolutive.

First, we have already encountered the morph UM associated with informational structures specified [PER: person&number]/[FIN: pl]. This morph is also found in words associated with an informational structure of the form [PER: null]/[FIN: pl…]. In the first case, the plural morph appears at the right edge of the base; in the second, it appears immediately following the absolutive. No disagreement

12. There are six Luiseño absolutives, organized into long/short pairs – LA/A; TA/T; and CHA/SH. Only the short forms are possible in [TEMP: temp] words; the long forms are limited to a subset of [TEMP: null] words.

13. The analysis of *tooyaat* indicated in the brackets of (30a) identifies a morph AN. The /n/ is deleted before a word final absolutive and, thus, is absent in the word form. The analysis of *tooyaqal* indicated in the brackets of (30b) presumes that /a/ intervening between two laterals is deleted. The absolutive in *tooyaqalmokwichi* in (30c) is SH, although in anything but final position, the segment is an affricate.

14. Actually, Kroeber and Grace (1960: 150) do identify a form *a-t*, written as such presumably to indicate the presence of an absolutive. This is the surface form appearing in (30a). But they don't associate it with the predicate in a clausal adjunct.

exists as to whether the word in (31a) contains the absolutive T or that in (31b), the absolutive SH (here as CH, as noted in fn 13).

(31) [PER: null]/[FIN: PL]
  a. *hunwuTUM*
     bear.absolutive.PLURAL
     'bears' [cf. *hunwuT* 'bear']
  b. *qeengiCHUM*
     squirrel.absolutive.PLURAL
     'squirrels' [cf. *qeengiSH* 'squirrel']

The 'plural' counterparts to (30a) and (30c) behave precisely as expected, if their final segment is an absolutive.

(32) a. [PER: null]/[TEMP non-continuous]/[FIN: non-continuous&PL]
     *tooya**an**TUM* 'laughing'
  b. [PER: null]/[[TEMP: continuous past]/[FIN: PL&OBJECT]
     *tooya**qalmokwi**CHUMi* [*tooya**qalmokwi**SHUMi*] 'laughed'

The value for FIN in the informational structure of the word in (30b) doesn't include number and, hence, can't include 'plural'. But (8) contains a word of the appropriate type and its 'plural' form patterns with the words in (32).

(33) [PER: null]/[TEMP: continuous]/[FIN: PL]
     *tooya**qa**TUM* 'laugh' [cf. *tooya**qa**T* 'laugh']

Second, as noted above in regard to the words in (24b) and (24c), a word associated with an informational structure where the value for FIN includes 'object' can involve the morph I (appearing as /y/ following a vowel). This morph is present, at the right edge of the stem, only if the value for the attribute PER includes person, the situation in (24b) or (24c), or the form of the word includes an absolutive and refers to an animate object. The informational structures of the words in (34) are specified [PER: null]/[FIN: object]. Both contain an absolutive. The first is the singular form of the word in (31b), a word referring to an animate object. The second is a word referring to an inanimate object.

(34) [PER: null]/[FIN: OBJECT]
  a. *hunwutI*
     bear.OBJECT
     'bear'
  b. kut
     fire.OBJECT
     'fire'

The form of *tooyaqal,* the word in (30b) which functions as the predicate in a complement clause and is specified [PER: null]/[TEMP: continuous]/[FIN: object], is exactly as expected then, if the word is analyzed as containing an absolutive (i.e. with the analysis [tooyaqalaL] shown above). Like the word in (34b), it doesn't refer to an animate object; like the word in (34b), the morph I is not present – but the absolutive is.[15]

In sum, the temporal morphology in subordinate predicates specified [PER: null] appears at the right edge of the word only when the informational structure also is specified for a temporal value, unaccompanied by number. In all other cases, it is separated from the word's right edge, either by a postposition or by, at least, an absolutive. This last is noteworthy. There are other words with an informational structure including [PER: null]/[TEMP: temporal value]/[FIN: temporal value and number]. These function only as the predicate in a main clause. If there is temporal morphology in such a word, it is both right adjacent to the stem and, unlike (30a) or (33), at the word's right edge.[16]

(35) a. [PER: null]/[TEMP: present]/[FIN: present&NUMBER]
tooyaQ 'is laughing'
cf. *tooyaq* *up.*
is.laughing AUX
'S/he is laughing.'

b. [PER: null]/[TEMP: imperative]/[FIN: imperative&NUMBER]
tooyaxAM 'laugh! (PL)'
cf. *tooyaxam!*
laugh!
'Laugh (you all)!'

---

15. The morph I also appears in words with a long form of the absolutive, if they refer to an animate object. If they do not, the final segment of the absolutive is elided.

(a) [PER: null]/[FIN: obj]
muutaY
owl.object
'owl' (cf. *muuta* 'owl')

(b) [PER: null]/[FIN: obj]
huul
arrow.object
'arrow' (cf. *huula* 'arrow).

16. The temporal values that accompany [PER: null] in the predicate of a main clause include, in addition to those in (35), [FIN: past continuous], [FIN: simple past], [FIN: simple future], [FIN: future habitual], [FIN: past habitual], [FIN: open time], [FIN: open time habitual].

The bottom line, then, is that in terms of its properties of form, the predicate of a clausal adjunct is like other subordinate predicates. But the consequence of this fact is that a predicate in a clausal adjunct where the informational structure is [PER: null]/[TEMP: non-continuous]/[FIN: non-continuous&number] is absolutely unique: It is the only case of a form associated with an informational structure with a temporal value for FIN whose temporal morphology is separated from the word's edge by other morphology.[17]

3.3 Implications

This fact is consistent, in an interesting fashion, with another property of subordinate clauses.

If there is a subject in a subordinate clause where the informational structure of the predicate includes [PER: null], it must be determined within its embedding environment. So, the subject of the clausal adjuncts in (15) and (16) is shared with the subject of the clause to which each is adjoined; the subject of the complements in (20) is given by another argument to the embedding predicate; and the subject of the relative clauses in (18), (19) and (23c) is given by their heads. (The complement in (23b) lacks a subject.) In contrast, if there is a subject in a subordinate clause where the informational structure of the predicate includes [PER: person&number], it is, by and large, independent of its embedding environment. So, the subject of the clausal adjunct in (22a) is '1SG', as distinct from *nawitmal* 'girl', the form realizing the subject of the clause to which it is adjoined; the subject of the complements in (21) is '3SG', as distinct from the '3PL' subject of the embedding clause; and the subject of the relative clause in (22c) is '1SG', as distinct from its head *nawitmali* 'girl (object)'.[18]

Note that the claim is not that the absolutives or the morphs in (4) are themselves subjects nor is it that the value for PER directly represents a subject. The idea, rather, is that the informational structure of the clause's predicate codes how the subject of the clause is to be interpreted within a syntactic context. However, the location of the absolutive in the predicate of a clausal adjunct is consistent with what is coded in the informational structure in regard to the subject – i.e. that the

17. With some bases, the temporal morphology associated with [TEMP: simple past]/[FIN: simple past] is a left-edge phenomenon, involving reduplication or other alteration of the first mora. Jacobs (1975) offers a particularly nice exposition of this phenomenon.

18. The informational structure of the predicate in one kind of complement clause and one kind of relative clause is limited to [PER: 3number]. Where this is the case, if the clause has a subject, the situation is somewhat more complicated. The property of an independent subject holds for predicates open to any of the six values including person and number.

subject is determined relative to the clause to which it is adjoined. [FIN: temp] still gives access to the value for TEMP. The location of the absolutive after the temporal morphology reflects the accessibility of the subject directly in the word's form.

4. Conclusion

Implicit in the analysis offered in this paper is the claim that the critical properties of Luiseño clausal adjuncts are, first, that their time is dependent on that of the clause to which they are adjoined and, second, that their subject can be independent of or dependent on the subject of this clause. One explicit claim is that the first follows from the character of the informational structure of the clausal adjunct's predicate. Because the attribute that indicates their syntactic potential (i.e. FIN) must have a temporal value, it must give access to the value for the attribute TEMP. The predicate of a clausal adjunct is unique among the predicates of subordinate clauses on this score. The proposals in regard to the subject are more complex. The independence or dependence of the clausal adjunct's subject relative to its embedding domain is represented in the informational structure of the clausal adjunct's predicate – in whether the value for PER is one of those involving, respectively, person and number or indicating the absence of both. In this regard, as in the conditions describing its properties of form, the predicate of a clausal adjunct is no different from the predicates of other subordinate clauses. Thus, there are predicates whose informational structure is such that both the subject and the time of the clausal adjunct must be relative to the clause to which it is adjoined. In these, the presence of the absolutive in the word's form – interceding between the temporal morphology and the word's right edge – allows the subject to share equal billing.

References

Akmajian, Adrian, Steele, Susan & Wasow, Thomas. 1979. The category AUX in universal grammar. *Linguistic Inquiry* 10(1): 1–64.

Anderson, Stephen. 1992. *A-Morphous Morphology* [Cambridge Studies in Linguistics 62]. Cambridge: CUP. https://doi.org/10.1017/CBO9780511586262

Davis, John. 1973. A Partial Grammar of Simplex and Complex Sentences in Luiseño. PhD dissertation, Los Angeles, University of California.

Elliott, Eric. 1999. Dictionary of Rincón Luiseño. PhD dissertation, San Diego, University of California.

Hyde, Villiana. 1971. *An Introduction to the Luiseño Language*. Banning: Malki Museum Press.

Jacobs, Roderick. 1975. *Syntactic Change: A Cupan (Uto-Aztecan) Case Study* [University of California Publications in Linguistics 79]. Berkeley CA: University of California Press.

Kroeber, Alfred & Grace, George. 1960. *The Sparkman Grammar of Luiseño* [University of California Publications in Linguistics XVI]. Berkeley CA: University of California Press.

Steele, Susan, Akmajian, Adrian, Demers, Richard, Jelinek, Eloise, Kitagawa, Chisato, Oehrle, Richard & Wasow, Thomas. 1981. *An Encyclopedia of AUX: A Study in Cross-Linguistic Equivalence* [Linguistic Inquiry Monograph 5]. Cambridge MA: The MIT Press.

Steele, Susan. 1985. Changing the theme. *International Journal of American Linguistics* 51: 588–591. https://doi.org/10.1086/465984

Steele, Susan. 1990. *Agreement and Anti-Agreement: A Syntax of Luiseño* [Studies in Natural Language and Linguistic Theory 17]. Dordrecht: Kluwer.

Adverbial subordinators in Yaqui

Lilián Guerrero
IIFL-Universidad Nacional Autónoma de México

In this paper, I examine the form and function of Yaqui adverbial subordinators. In contrast to relative and complement subordinators, which are always final and bound in Yaqui, adverbial subordinators can occur in both positions, clause-final and clause-initial; the former are bound morphemes, and the latter are free particles. The adverbial markers in initial position differ from any other subordinators in two ways: they have lexical content and do not demand structural dependent features inside the linked unit. Among the final subordinators, *-kai* and *-o* are not only the most productive markers in the sample, but they also introduce several adverbial relations including manner, purpose, temporal, concessive and conditional meanings. In fact, these two subordinating morphemes show a consistent, though not obligatory distribution: while *-kai* is used when the main and dependent subjects are the same, *-o* often occurs when the subjects are different. The contrasts observed among the adverbial clauses taking final subordinators and those introduced by initial subordinators suggest that the latter may be somewhat new in the Yaqui grammar. As new adverbial clause linkage markers, they adopt the structural features of coordination instead of subordination.

Keywords: subordinators, adverbial relations, complex constructions, switch-reference, Yaqui

1. Introduction

The selection and form of the subordinating morpheme has been a crucial topic in the study of adverbial subordination, because they can specify the semantic relation between the main clause and the adverbial clause. There are two types of adverbial subordinators (Thompson, Longacre & Hwang, 2007): (i) grammatical morphemes with no lexical meaning (e.g., English *to*, as in *to buy beer*), and (ii) grammatical morphemes with lexical content (e.g., English *before*, *when*, *if*). In languages of the world, subordinating morphemes, like conjunctive elements and adpositions, may be prepositional or postpositional, free or bound (Dryer 2013).

https://doi.org/10.1075/tsl.126.05gue

Earlier works on Yaqui, a Southern Uto-Aztecan language (ISO cod: yaq), have described basic features of adverbial clauses based on subordinator selection (Lindenfeld 1973; Dedrick & Casad 1999). Lindenfeld (1973: 81) claimed that in Yaqui "there are two basic types of subordinated clauses: those marked by *-kai*, a general subordinator, and those marked by specific subordinators such as *ella'apo* 'although', *kielekun* 'so that' and the like". Therefore, the clause in (1a) taking the general subordinator *-kai* allows for multiple adverbial readings, while that one taking the specific subordinator *kielekun* in (1b) expresses a causal relation. The examples come from Lindenfeld (1973: 82–84).

(1) a. *Aapo lipti-tu-ne [tuisi tekipanoa-kai].*
3SG.NOM blind-become-POT much work-CLM
'Working hard will make him blind.'
'He will become blind if he works hard.'
'When he works hard, he becomes blind.'

b. *Empo si kusisi nooka-Ø [kielekun ne kaa enchi hikkaha-Ø].*
2SG.NOM INT fast speak-PRS CLM 1SG.NOM NEG 2SG.ACC understand-PRS
'You speak so fast that I do not understand you.'

In this paper, I offer a corpus-based study of Yaqui adverbial subordinators. The analysis of the data suggests a complex distribution and use of these clause linkage markers. First, in contrast to relative and complement subordinators, which are always final and bound, adverbial subordinators can occur in both positions: clause-final and clause-initial. The former are bound morphemes, and the latter are free particles. Second, two types of clause-final subordinators exist: markers without lexical meaning, such as *-kai* and *-o*, and markers with lexical content, like the postposition *-betchi'ibo* '(in order) to, for' as well as initial subordinators such as *kielekun* (*kielikun* ~ *kiali'ikun*) 'so that, since'. Third, the two most productive final subordinators, *-kai* and *-o*, introduce several adverbial clauses including manner, purpose, temporal, concessive and conditional meanings. These two markers show a consistent, though not obligatory distribution. While *-kai* is used when the main and dependent subjects are the same, *-o* tends to occur when the subjects are different, i.e., switch-reference marking. Fourth, the initial adverbial subordinators differ from any other linkage markers, as they do not demand features of structural dependency in the linked clause. In Yaqui, relative, complement and adverbial clauses taking final subordinators demand a genitive or accusative subject when present inside the linked clause. In contrast, initial subordinators such as *kielekun* require a nominative subject in the associated clause.

The contrasts observed among the adverbial clauses taking final subordinators and those introduced by initial markers suggest that the latter may be new in the Yaqui grammar. In fact, adverbial particles such as *kielekun* 'so that' were absent in Buelna's work (1890). As new adverbial particles with lexical content, they adopt the structural features of coordination instead of subordination.

2. Yaqui sentences

This section introduces aspects of Yaqui grammar relevant to the discussion of adverbial constructions. These include the morpho-syntax of simple clauses (§2.1), and relative (§2.2) and complement clauses (§2.3).

2.1 Simple clauses

Yaqui is a synthetic/agglutinating, head-final, primary object language (Lindenfeld 1973; Dedrick & Casad 1999; Guerrero & Van Valin 2004). It is the only Southern Uto-Aztecan language still spoken where case marking on nominals is preserved. The nominative case is unmarked (2a), and the accusative is marked by the suffix *-ta* (2b). There is no dative case but postpositions mark oblique core arguments, as seen in (2b). Determiners are morphologically marked too; when the NP is accusative, the determiners are marked by *-ka*; oblique NPs are marked by *-e*. Case markers and the plural suffix *-(i)m* are mutually exclusive in nominals and determiners.

(2) a. *U goi-Ø u-ka chu'u-ta bicha-k.*
DET coyote-NOM DET-ACC dog-ACC see-PFV
'The coyote saw the dog.'

b. *Peo-Ø u-e jamut-ta-u etejo-k.*
Pedro-NOM DET-OBL woman-ACC-DIR chat-PFV
'Pedro chatted with the woman.'

c. *Inepo u-me jamuch-im nooka-ne.*
1SG.NOM DET-PL woman-PL talk-POT
'I will talk to the women.'

Yaqui shows a rigid SOV word order, though other arrangements are possible, e.g., the logical subjects and objects can follow the verb for a specific pragmatic functions (Belloro & Guerrero 2010). There are constraints regarding the ordering of full and reduced pronouns (Table 1): while full pronouns behave as lexical elements in terms of their distribution, see the example in (2c) above, reduced nominative pronouns behave as "second position" clitics (3a), and reduced accusatives (available only for 3rd person) cliticize to the verb (3b). There is a third paradigm for pronouns

associated to postpositions (i.e., oblique arguments and adjuncts); 3rd person pronouns serving as subject tend to be covert (i.e., missing argument or zero).

Table 1. Yaqui pronominal system

	Nominative	Accusative	Oblique	Genitive
1 Sg	*inepo = ne*	*nee, ne*	*ne-*	*in, nim*
2 Sg	*empo = 'e*	*enchi*	*e-*	*em*
3 Sg	*aapo = Ø*	*apo'ik, a'a, a=*	*a-*	*apo'ik, a=*
1 Pl	*itepo = te*	*itom*	*ito-*	*itom*
2 Pl	*eme'e = 'em*	*enchim*	*emo-*	*em, enchim*
3 Pl	*bempo = Ø*	*apo'im / am*	*ame-*	*bem, bempo'im*

(3) a. *Yookoria-po = ne sewa-m jinu-k.*
morning-LOC = 1SG.NOM flower-PL buy-PFV
'In the morning, I brought flowers.'

b. *U jamut-Ø a = bicha-k jita-nenenki-wa-po.*
DET woman-NOM 3SG.ACC = see-PFV thing-RDP.sell-PASS-LOC
'The woman saw him in the market.'

The language presents a rich postpositional system (Table 2) that semantically encodes a wide range of spatial, temporal and associative meanings. Some of these postpositions are always bound while others may appear as free form in certain contexts. Most have a single form, e.g. the locative *-po* 'in, on' in (3a–b) above, but some may vary in form depending on phonological, morpho-syntactic or semantic factors (Guerrero 2019). For instance, the directional marker is *-u* when the phrase is non-final (4a), otherwise *-wi* (4b); when the noun is plural, then the forms are *-meu/-mewi* (4c). Most postpositions take a nominative complement, but some take an accusative noun, e.g. the beneficiary *-betchi'ibo* 'for'. The occurrence of *-ta* may be motivated by animacy, as seen in (4d–e). A few postpositions can serve as clausal subordinators.

Table 2. Postpositional system

Directional (sg):	*-u/-wi*	Proximal contact (sg) 'at, on'	*-t, -chi*
Directional (pl):	*-meu/- mewi*	Proximal contact (pl) 'at, on'	*-met*
Directional (toward):	*-bicha*	Proximal 'near, close'	*naapo*
Comitative:	*-mak/-make*	Positional 'beside, from'	*(be)tana*
Instrumental (sg):	*-e, -ae*	Positional '(on) the top of'	*jika*
Instrumental (pl):	*-mea*	Positional 'together'	*nau*
Benefactive, purpose:	*-betchi'ibo*	Positional 'under, beneath'	*betuk*
Locative (space & time):	*-po*	Positional 'in front of, by'	*beas*
Limiter 'until'	*tajtia*	Positional 'over'	*bepa*

(4) a. *Peo-Ø teopo-u siika.*
Peo-NOM church-DIR go.SG.PFV
'Pedro went to the church.'

b. *U-me ili usi-m saja-k eskuela-wi.*
the-PL little child-PL go.PL-PFV school-DIR
'The children went to the school.'

c. *U-me ili usi-m saja-k bem eskuela-im-mewi.*
the-PL little child-PL go.PL-PFV 3PL.GEN school-PL-DIR.PL
'The children went to their schools.'

d. *Kajlos-Ø jamut-ta-mak e'etejo-k.*
Carlos-NOM woman-ACC-COM chat-PFV
'Carlos talked with the woman.'

e. *Jeemam seboraka into kokoi siari-mak bwasai-tu-k si kia.*
liver.PL onion and chili green-COM cook-VBLZ-PFV very good
'The liver cooked with onion and green chili is delicious.'

Except for a few suppletive forms, verbs do not express person or number. The usual situation for the tense-aspect-mood (TAM) verbal system (Table 3) is to display a range of meanings, including tense-aspectual values and distinct mood/epistemic states.

Table 3. Tense-aspect-mood (TAM) verbal system

V-*Ø*	Present	RDP-V	Habitual
V-*k(a)*	Past perfective	RDP-RDP-V	Iterative, repetitive
V-*n*	Past continuative	V-*su*	Completive
V-*ne*	Potential, future	V-*'ean*	Hypothetical
V-*na*	Potential, future passive	V-*maachi*	Obligation

Broadly, a subordinate clause is understood as a clause that is (a part of) a constituent in another clause. A formal definition of a subordinate clause is a complex task since subordination can be expressed in a variety of different structures in the languages of the world (Lehmann 1988; Cristofaro 2003). As for Yaqui, there are three ways to identify a subordinate clause. These methods include: the selection of the subordinator, the coding of core arguments (especially, co-referential participants) and the coding of TAM information of the linked verb (for details, see Guerrero 2006 and further works). The description that follows focuses on the subordinator and the lexical coding of the dependent subject only as these dependency features are relevant in the discussion of adverbial subordinators.

2.2 Relative clauses

Relative clauses are expressed via nominalization in Yaqui (Álvarez 2012; Guerrero 2012). Two clause types are identified depending on the role of the head noun and the subordinator. If the head noun serves as the dependent subject, as in (5a), the modifying unit takes the linkage marker *-m(e)*. If the head noun serves as the direct or oblique argument of the verb in the linked unit, as in (5b), then the subordinating morpheme is *-'u*. Although less common, the locative postposition *-po* may introduce a locative (spatial and time) relative clause, as in (5c).

(5) a. *U o'ou-Ø$_i$ [enchi _$_i$ bicha-ka-me] siika.*
DET man-NOM 2SG.ACC see-PFV-CLM go.SG.PFV
'The man who saw you, left.'

b. *U-me o'ou-im$_i$ [em _$_i$ bicha-ka-'u] saja-k.*
DET-PL man-PL 2SG.GEN see-PFV-CLM go.PL-PFV
'The men who you saw, left.'

c. *Poso-po$_i$ [kuchu'm ane'e-Ø-po] a wo'ota-ne.*
pond-LOC fish.PL exist-PRS-LOC 3SG.ACC throw-POT
'Throw it (the fish's skin) in the pond where the fishes are.'

(Johnson 1962 1:2)

In subject relative clauses (5a), there is a missing (covert) argument which is in co-reference with the head noun.[1] In non-subject relative clauses, the dependent subject is overt and marked as genitive, when pronominal (5b), or accusative.

2.3 Complement clauses

There are four complement strategies in Yaqui. The first two strategies are very productive (the co-lexicalized and the syntactic structures), while the remaining two are limited to a few complement-taking predicates (the nominalized and the participial-like structures). The first complement consists of a V1-V2 structure, where the two events are usually joined without a subordinating morpheme, as seen in (6a–b); cognitive and speech act verbs allow the marker *-t(i)*, which immediately follows the linked verb, as in (6c).

1. In the examples, co-referential arguments are co-indexed; the '_' is for illustrative purposes only and indicates a missing argument in co-reference with a main argument. The position of '_' seeks to reflect the expected position for reduced pronouns (e.g., second position clitics for subjects), rather than full pronouns. Like nominal phrases, full pronouns rarely serve as co-referential arguments inside subordinate clauses in Yaqui.

(6) a. *Peo-Ø$_i$ Joan-ta$_j$ –$_i$ bwite-roka-n.*
Peo-NOM John-ACC run.SG-promise-PASC
'Pedro promised John to run.'
b. *Peo-Ø$_i$ Joan-ta$_j$ –$_j$ bwit-tua-k.*
Peo-NOM John-ACC run.SG-cause-PFV
'Pedro made John run.'
c. *Goyo-Ø Tibu-ta wakas-ta etbwa-ne-t-'ea-n.*
Goyo-NOM Tibu-ACC cow-ACC steal-POT-CLM-think-PASC
'Goyo thought Tibu would steal the cow.'

In a co-lexicalized complement, the main and dependent subjects may, but need not be the same. Consequently, verbs demanding identical subject (actor control verbs) code the subject once as a nominative argument; this is the case of phrasal, modals and psych-action matrix verbs (6a). Verbs demanding the main object to be the same that the dependent subject (object control verbs), code the co-referential argument once as an accusative argument; this is the case of causative and jussive verbs (6b). Speech-act, propositional attitude, and mental matrix verbs (6c) take an accusative dependent subject.

The syntactic complement takes a clause marked by the subordinating morpheme *-'u*, as in (7a). Regardless of the identity of the subjects, the linked subject needs to be overt and marked accusative. There are few examples of the locative postposition *-po* introducing the complement of some mental predicates. In some complements taking *-po*, the subject comprises a genitive pronoun when co-referential with the main subject, as shown in (7b).

(7) a. *Peo-Ø [kaba'i-m enchi jinu-ka-'u] suale-n.*
Peo-NOM horse-PL 2SG.ACC buy-PFV-CLM believe-PASC
'Pedro believed that you bought the horses.'
b. *Inepo si majae-n [ka nim loteria-ta yo'o-ne-po].*
1SG.NOM INT be afraid-PASC NEG 1SG.GEN lottery-ACC win-POT-CLM
'I was afraid I would not win the lottery.'

The last two complement strategies are restrictive and limited to a few predicates, e.g., *bo'obicha* 'expect', *teenku* 'dream/imagine', some uses of *'ean* 'believe/have the feeling of', as well as direct perception. The nominalized strategy depicted in (8a) is marked by *-m(e)* plus the accusative suffix *-ta*, and it requires different subjects.[2] The participial strategy in (8b) takes *-kai* and requires identical subjects. In the former, the dependent subject must be accusative; in the latter, the subject must be omitted.

2. See Guerrero (2012) for a detailed discussion on the distinction between relative clauses and complement clauses taking *-'u* and *-me* in Yaqui.

(8) a. *Tuuka beako nepo [Peo-ta enchi kuna-ka-m-ta]*
yesterday night 1SG.NOM Peo-ACC 2SG.ACC marry-PFV-CLM-ACC
teenku-k.
dream-PFV
'Last night, I dreamed of Pedro marrying you.'

b. *Tuuka beako Lupe-Ø$_i$ [Peo-ta $_i$ kuna-kai] teenku-k.*
yesterday night Lupe-NOM Peo-ACC marry-CLM dream-PFV
'Last night, Lupe dreamt of (herself) marrying Pedro.'

Relative and complement clauses make use of final and bound subordinators. Despite Lindenfeld's (1973:102) finding for the Arizona variant, the loanword *ke* 'that' does not serve as a major complement subordinator. Using postpositions as complement subordinators is exceptional and limits to the locative *-po*.[3] Since final and bound morphemes introducing relatives and complements lack lexical meaning, there are functional overlaps: *-me* marks subject relative clauses and the uncommon nominalized complement; *-'u* introduces non-subject relatives and the syntactic complement. As we will see next, *-kai* is, in principle, an adverbial marker but it can mark a participial complement, as shown in (8b).

3. Adverbial subordinators, some generalities

Thompson et al. (2007:238) propose two types of subordinating morphemes: (i) grammatical morphemes with no lexical content (e.g., English *to* as in *to buy beer*), and (ii) grammatical morphemes with lexical meaning (e.g., English *before*, *when*, *if*). Adverbial markers with lexical content denote semantically richer notions between the two linked units, such as comparison, temporality, causality, purpose, and circumstance (Nordstrom 2010). Yet, the categorical status of these subordinators is an open question. They can be considered as adpositions (along with adverbial particles), or as complementizers (along with elements such as *that, whether, if*). According to Kortmann (1997:241), adverbial subordinators differ from complementizers because they assign a theta-role to their complements, hence they determine the interclausal semantic relation. Because of this difference, adverbial subordinators can be treated as a predicative adpositions (Jolly 1992).

Based on a sample of 659 languages of the world, Dryer (2013) explore the form and distribution of adverbial subordinators. The author found that using free particles at the beginning of the adverbial clause is the preferred strategy (398 languages

3. The subordinator *-'u* may be historically related to the directional *-u* shown in (4a–b). Synchronically, *-'u* does not behave as an adposition anymore, since it demands a predicate or clause as a complement. In contrast, *-po* may combine with a nominal, a predicate or a clause.

of his sample), followed by separated words at the end of the dependent clause (96 languages), or over one adverbial subordinator (93 languages). In comparison, using suffixal adverbial subordinators is slightly less common (64 languages).

There are two types of adverbial subordinators in Yaqui: free particles at the beginning of the adverbial clause, and suffixal subordinators. Simply put, Yaqui makes use of the most common strategy cross-linguistically to introduce adverbial clauses (initial particles) and the less common strategy (final and bound morphemes). The first column in Table 4 lists the Yaqui adverbial subordinators documented by Lindenfeld (1973) for the Arizona variant; the second column includes those found in the Sonoran variant by Dedrick & Casad (1999).

Table 4. Adverbial subordinators in Yaqui

Lindenfeld	Dedrick & Casad	Adverbial clause type
-ka(i)	*-ka(i)*	general subordinator
-o	*-o*	general subordinator '*when/if*'
	-patchi	'upon' clause
	-tahtia	'until' clause
	-po	locative clause
	-amcha	simulfactive 'as if' clause
-bae-kai	*-pea-kai*	'in order to' purpose clause
-betchi'ibo		copulative 'is for X' purpose clause
-tek		hypothetical conditional clause
	'o'oben	concessive 'even though' clauses
(nasuk) huni'i	*huni'i*	'*in spite of/even/although*' clause
ella'apo		'although' clause
kielekun		causal clause
bwe'ituk, porke	*bwe'ituk*	reason clause
parake, pake		goal, purpose clause
kwando		'when' clause
si		'if' conditional clause
sino		'if not, else' conditional clause

Diessel (2001, 2005) has established interesting facts about the ordering of the main and the adverbial clauses. He found that languages of the world use either, adverbial clauses both before and after the main clause, or adverbial clauses preceding the main clause; languages preferring final adverbial clauses are less common in his sample. When both orders are possible in the same language, the position of the adverbial unit can be motivated by several factors. One factor is word order. While adverbial clauses that precede the main clause only occur in OV languages, adverbial clauses that are often pre- and post-posed occur in both VO languages and a

significant minority of OV languages (Diessel 2001: 442). Yaqui shows a rigid SOV order but 64% of the adverbial clauses in my sample are final (Table 5).[4]

Table 5. Distribution between main and adverbial units in Yaqui in corpus

	Adverbial-main	Main-adverbial	Total
Final subordinator	233	236	469
Initial subordinator	4	187	191
	237 (36%)	423 (64%)	660 (100%)

Diessel suggests a strong correlation between the ordering of the main and adverbial units and the position of the subordinator. In Diessel's sample, adverbial clauses including a final subordinator precede the main clause, while clauses marked by an initial subordinator are oftentimes found in both initial and final position regardless of the order of verb and object (Diessel 2001: 443). As for Yaqui, this correlation is shown in Table 5. Where clauses with initial subordinators prefer to be final, adverbial clauses taking final markers show no strong preference. In fact, Yaqui follows a mixed order pattern (Diessel 2001: 444–445) based on the semantic relation of the two units: temporal and manner clauses are initial, while cause/reason and purpose are final (Table 6). For some adverbial relations, there are too few examples to make any generalization.

Table 6. Order of main and adverbial units based on their semantics

	Adverbial-main	Main-adverbial	Total
Reason	1	130	131
Causal	3	55	58
Restriction		2	2
Conditional	5	3	8
Temporal	148	24	172
Purpose	35	185	220
Manner	33	21	54
Other relations	12	3	15
	237	423	660

In what follows, I examine the most common adverbial subordinators in Yaqui based on its distribution and use.

4. The analysis is based on a corpus comprised of several narratives from the Sonoran variant; the sample includes traditional and folk stories, life stories and short conversations (approximately 4,900 clauses). Though sentences from direct elicitation are also analyzed to illustrate specific purpose, they are not included in Tables.

4. Yaqui adverbial subordinators

4.1 The sample

The adverbial subordinators found in my corpus are listed in Table 7. Some markers are the same than those mentioned in Lindenfeld and Dedrick & Casad (Table 4 above); others mentioned before seldom occur in the sample (e.g., *-patchi* 'upon', *-amcha* 'as if', *o'oben* 'even though'). Using Spanish loanwords such as *porke*

Table 7. Adverbial constructions in Yaqui taking a subordinator

Adverbial relation	Initial sub-ordinator	No. clauses	Final sub-ordinator	No. clauses	Double sub-ordinator	No. clauses	Total clauses in the sample
Reason	*bweituk*	97					131
	porke	34					
Causal	*kiali'ikun*	33	*-betchi'ibo*	25			58
	por eso	1					
Concessive	*ella'apo*			2			2
Purpose							220
final			*-betchi'ibo*	92			
purpose of intention			*-bae-kai*	103			
purpose of motion			*-se-kai*	25			
Temporal							172
sequential			*-o*	60	*ketun ke -o/-kai*	10	
			-kai	27	*-su-o/-kai*	8	
simultaneous			*-o*	30			
			-kai	33			
delimitative			*-tajtia*	4			
Conditional			*-o*	4	*si -o*	1	8
			-tek-o	1	*si -tek-o*	2	
Manner/ circumstance							54
means			*-kai*	22			
being/having circumstance			*-kai*	24			
negative circumstance			*-kai*	8			
Other relations			*-kai*	10			15
			-o	5			
		165		475		21	660

'because', *pake/parake* '(in order) to', and *kwando* 'when', is much less common in the Sonorant variant; the examples of *porke* in my sample come from young bilingual speakers (Guerrero 2017). Notice that some adverbial relations can be expressed by alternative structures, e.g. cause, purpose and temporal clauses, whereas others make use of a single structure, e.g. manner clauses. Because of the nature of the sample, some adverbial relations are uncommon, e.g., conditional clauses.

Both, Lindenfeld and Dedrick and Casad include the particle *juni'i*, commonly glossed as 'also, yet, even, although, in spite of', as a subordinating morpheme. However, while clauses taking *juni'i* are very common in the sample, this adverbial particle is hard to analyze at this point of the study. On one hand, this particle may appear at different positions in a sentence; in the fragment in (9a), there are three examples of *juni'i* in different positions. On the other hand, *juni'i* can occur by itself, without any nominal, predicate or clausal as a complement (9a). In (9b), this particle follows a *-kai* clause but, rather than an 'even though' adverbial reading, the sentence is interpreted as a manner clause. In (9c), *juni'i* is the only particle that allows a concessive reading. Because of its own complexity, adverbial clauses taking *juni'i* are not included in this analysis.

(9) a. *Bweta jachim-po juni'i ju'u in koba-po jippue-'u, jiba*
CLM what-LOC yet DET 1SG.GEN head-LOC have-CLM always
junaman aane-Ø, in jiapsa-po naate-kai. Jaibu bu'u-m
over.there exist-PRS 1SG.GEN heart-LOC begin-CLM already big-PL
wasuktia-m sim-su-k juni'i, ju-me waati-m into ju-me
year-PL go.SG-COMP-PFV even DET-NOM memory-PL and DET-PL
tenkui-m jiba in takaa-po jiapsa. Ketun = ne au
dream-PL always 1SG.GEN fruit-LOC heart yet = 1SG.NOM self
waate-Ø juna-ka yeu = sika-'u kia iliki-k juni'i
remember-PRS DEM-ACC out=go.SG.PFV-CLM INT little-ACC even
kaa kopta-Ø.
NEG forget-PRS
'[When I become 14 years old, my life changed]. But my thought (lit. what I have in my head) is over there, since the beginning of my life. Even though several years had passed, my memories and my dreams still live inside of me (lit. always the fruit of my heart). I still remember what happened, I don't forget even the littlest things.' (Buitimea; Pesio betanaa: 1–7)

b. *[Junama'a si tebae koko-ka] juni'i te ama jooka.*
over.there INT hunger die.PL-CLM yet 1PL.NOM there live.IMPFV
'Dying of hunger, we still lived over there.' (Silva; HVH: 204)

c. *Santa iklesia-po ne-t na-t yaja-k, ju-me'e*
Holly church-LOC 1SG-LOC together-LOC arrive.PL-PFV DET-PL
tetekia-ka-me, ju-me'e Cantoora-m, Kiiyojteim, num bea kaa
RDP.work-PFV-CLM DET-PL singer-PL Kiiyojteim, there DM NEG
ne jeewite-k juni'i.
1SG.NOM accept-PFV even
'As for the persons in charge, the singers or Kiiyojteim, they all reunited in the church with me although I did not accept.' (Silva; HVH: 47–8)

The description of Yaqui adverbial subordinators begins with specific subordinators with lexical content, and then moves to the distribution of *-kai* and *-o*, the two most productive but most complex adverbial subordinators.

4.2 Specific subordinators

There are two types of subordinators that mark adverbial clauses for their semantic relationship to the main clause. The first, mentioned by Lindenfeld, involves adverb-like initial particles. In my corpus, the two most common initial subordinators introduce a specific semantic relation (Guerrero & Belloro 2014; Guerrero 2017): *bweituk* 'because' (occasionally *porke*) expresses the motivation or explanation for the event described in the main clause, as depicted in (10a); *kiali'ikun* denotes the result or consequence of the event described in the main unit, as in (10b).

(10) a. *Ju-ka tampareo-ta yaati-tua-'e [bwe'ituk kompae-Ø*
DET-ACC drummer-ACC play-cause-IMP CLM compadre-NOM
a'abo weye-k].
here walk.SG-PFV
'Stop the drummer because my *compadre* is coming.'
(Buitimea; dream: 155)

b. *U yoeme-Ø$_i$ kaa a = eya-k [kiali'ikun ___$_i$ kaa*
DET man-NOM NEG 3SG.ACC = trust-PFV CLM NEG
a = mabeta-k].
3SG.ACC = accept-PFV
'The man didn't trust him, so he didn't accept [the food].'
(Silva et al; fox: 17)

The adverbial particle *ella'apo* 'although, it does not matter' belongs to this group too, and it introduces some sort of restriction subordination (Lindenfeld 1973: 85). In the sample, there are only two examples involving *ella'apo*. The clause in (11a) comes from Lindenfeld's work; the one in (11b) comes from the corpus. Notice that *ella'apo* in (11b) is followed by *juni'i*.

(11) a. *Itepo aman kat-ne [ella'apo ili usi-Ø ko'okwe-Ø].*
1PL.NOM there go-POT CLM little child-NOM sick-PRS
'We will go there although the child is sick.' (Lindenfeld 1973: 85)

b. *Ne am = jikkaja, kaa tua in ju'une, [ella'apo juni'i*
1SG.NOM 3PL.ACC = listen NEG INT 1SG.GEN know CLM even
intok = bea ne am-eu chai-pea juni'i].
DM = DM 1SG.NOM 3PL-DIR yell-desire even
'I listen to them, in my limited conscience; even so, I want to yell them.'
(Buitimea; malagüero: 158)

Furthermore, *bweituk, kiali'ikun* (and *ella'apo*) stand out among Yaqui subordinators, not only because of their position and lexical meaning, but because they demand a nominative subject in the adverbial clause. In (10a) above, the main and the linked subjects are different, then the linked subject is overt and takes nominative case, e.g. *compadre*; in (10b), the two subjects are the same, and there is a co-referential missing argument in the linked unit. In the adverbial clauses in (11), the subject is nominative. Nominative subjects are disallowed altogether in relative, complement and several other adverbial clauses in Yaqui.

Although they do not occur in this corpus, Lindenfeld shows that temporal clauses introduced by *kwando* 'when', and purpose clauses taking *parake* 'in order for' demand nominative subjects. Two of her examples are illustrated in (12).

(12) a. *[Kwando em papa-Ø yepsa-k] ne lihta-tu-ne.*
CLM 2SG.GEN father-NOM arrive.SG-PFV 1SG.NOM ready-VBLZ-POT
'When your father arrived, I will be ready.' (Lindenfeld 1973: 83)

b. *Am-eu = te chachae-ne [parake bempo itom bit-ne].*
3PL-DIR = 1PL.NOM RDP.call-POT CLM 3PL.NOM 1PL.ACC see-POT
'We will keep calling them in order for them to see us.'
(Lindenfeld 1973: 84)

The second type of specific subordinators involves postpositions: the locative *-po*, and the directional *-u/-wi* can introduce locative clauses. It is still unclear if a formal distinction between locative relatives, as depicted in (5b) above, and locative adverbials, as in (13), exist, since both may restrict the identity of a head noun and delimit the event within a time/location frame in discourse.

(13) a. *Naman ba'a-po bauba'a-po [kaa wawake-n wasuktia-m-po].*
DEM water-LOC lake-LOC NEG RDP.dry-PASC year-PL-LOC
'For years, the water in the lake never dried out.'
(Silva et al; turtle&coyote: 2)

b. *'aman jo'a-k [bwia-ta bweji-wa'a-wi].*
there home-PFV dirt-ACC dig-PASS-DIR
'He lives there where they are digging dirt.' (Dedrick & Casad 1999: 387)

Yaqui has several syntactic strategies to express purpose clauses (Guerrero 2012, 2017) and one of these structures takes the postposition *-betchi'ibo* 'for'. Purpose clauses taking *-betchi'ibo* as a subordinator are the only purpose linkage allowing both, same- and different-subjects. When the subjects are the same, as in (14a), there is a missing argument in co-reference with the main subject; when the subjects are different, as depicted in (14b), the dependent subject must be overt and take the accusative case.

(14) a. *Peo-Ø$_i$ a'abo siika [$\__i$ ji'i-bwa-betchi'ibo].*
Peo-NOM here go.SG.PFV something-eat-CLM
'Pedro came here in order to eat.'
b. *Peo-Ø$_i$ enchi$_i$ a'abo bittua-k [enchi$_i$ ji'i-bwa-ne-betchi'ibo].*
Peo-NOM 2SG.ACC here send-PFV 2SG.ACC something-eat-POT-CLM
'Peter sent you here in order for you to eat.'

Guerrero (2007) was the first one to notice the use of *-betchi'ibo* to indicate cause, beneficiary and purpose. Álvarez (2015) recently examines the syncretism of this postposition synchronically and historically. A detailed analysis contrasting purpose and causal sentences (Guerrero 2014, 2017) shows that *-betchi'ibo* may express a direct cause–effect relation. Of note, cause clauses taking *-betchi'ibo* structurally differ from those clauses introduced by *bweituk* and *kiali'ikun*. Whereas the initial causal subordinators demand a nominative subject, as seen in (10) above, a *-betchi'ibo* clause takes a genitive subject, when pronominal (15a), or an accusative subject, when nominal (15b).[5]

(15) a. W*ikosa-ta = ne$_i$ baajta-k [in$_i$ ousi ji-bwa-ka-betchi'ibo].*
bell-ACC = 1SG.NOM loosen-PFV 1SG.GEN a.lot thing-eat-PFV-CLM
'I loosen my bell because I ate a lot.'
b. *In maala-Ø omte-Ø [puato-ta jamte-ka-betchi'ibo].*
1SG.GEN mother-NOM angry-PRS plate-ACC break-PFV-CLM
'My mother is angry because the plate is broken.'

That is, the Yaqui language has two types of subordinators denoting the semantic relation between the main and the adverbial clauses. First, *bweituk* 'because' and *kiali'ikun* 'so that', introduce reason and causal clauses, respectively; *ella'apo* 'although' introduces restriction clause. These initial subordinators demand a nominative subject in the adverbial unit. Second, some postpositions can serve as adverbial subordinators. This is the case of the locative *-po*, the directional *-u*, and

5. In §3.2 we saw that genitive subjects in the dependent unit is a feature of relatives rather than complement or adverbial clauses. Thus, the *-betchi'ibo* cause clauses in (15) may be a case of nominalized-like adverbial causal clause.

the beneficiary *-betchi'ibo*. The postpositions *-po* and *-u* introduce locative clauses, while *-betchi'ibo* mark purpose and direct cause structures. As happens with relative and complement subordinators, postpositions serving as adverbial subordinators require non-nominative subjects, i.e. genitive or accusative subjects.

4.3 General subordinators

Following Nordstrom (2010: 95), the neutral term general subordinator means subordinators without rich lexical content, as opposed to adverbial (specific) subordinators. The general subordinator only has a meaning about the clause it is heading, whereas the adverbial subordinator denotes under what circumstances the matrix event takes place, e.g., temporality, causal or conditional circumstances.

In Lindenfeld's work, *-kai* is described as carrying an undifferentiated meaning of subordination, and so it may express several types of semantic associations at once between two statements. In opposition, she refers to *-o* as a specific subordinator and so it is glossed as 'when/if' (Lindenfeld 1973: 82–83). In my sample, both *-kai* and *-o* introduce several adverbial clauses, suggesting both can be treated as general subordinators.

According to Lindenfeld, *-kai* (or *-ka* when the clause is non-final) can introduce a manner clause, as depicted in (16a), a 'goal' purpose clause, as in (16b), and a clause expressing a more general adverbial meaning, as in (16c).

(16) a. *Wa-ka kuta-ta aapo kaa chukta-ne [junen aa joa-kai].*
DEM-ACC stick-ACC 3SG.NOM NEG cut-POT like.this able do-CLM
'He cannot cut this stick by doing it this way.' (Lindenfeld 1973: 89)

b. *Inepo in koarto-wi kibake-k [supe-te-bae-kai].*
1SG.NOM 1SG.GEN room-DIR enter.SG-PFV shirt-put_on-want-CLM
'I went to my room in order to get dressed.' (Lindenfeld 1973: 82)

c. *Aapo lipti-tu-ne [tuisi tekipanoa-kai].*
3SG.NOM blind-become-POT much work-CLM
'Working hard will make him blind.'
'He will become blind if he works hard.'
'When he works hard, he becomes blind.' (Lindenfeld 1973: 82)

Additionally, the author includes *-kai* clauses functioning as a gerund clause.[6] In (17), the dependent unit is additionally marked by the accusative *-ta*. The predicative

6. In previous works, I have argued that it is not *-kai* which allows the purposive meaning, but the occurrence of a co-lexicalized V1-V2 structure inside the dependent unit (Guerrero 2012, 2013, 2017). That is, in the example in (a), *-bae* 'want' follows the intended event, e.g. they wanted to rest; without *-bae*, the clause is interpreted as a simultaneous clause (a′). In (b), there

element in Lindenfeld's examples consists of nouns and (derived) adjectives, but not full verbs, e.g., some sort of secondary or depictive predication.

(17) a. *Empo lottila-ta-kai kaa yi'i-ne.*
2SG.NOM tired-ACC-CLM NEG dance-POT
'Being tired you will not dance.'
b. *Junu kuta-Ø teta-ta-kai kaa beeti-ne.*
DEM stick-NOM stone-ACC-CLM NEG burn-POT
'This stick, being (made of) stone, will not burn.' (Lindenfeld 1973: 82)

Lindenfeld points out that *-o* introduces temporal and conditional relations. In temporal clauses (18a), the loanword *kwando* 'when' may also introduce the adverbial unit. In conditionals (18b), the particle *si* 'if' can be present. The author claims that a time/conditional clause as in (18c), can be marked by *-tek*, some sort of hypothetical verbal suffix.

(18) a. *[(Kwando) ju-ka o'o-ta yepsa-k-o] itepo saja-k.*
CLM DET-ACC man-ACC arrive.SG-PFV-CLM 1PL.NOM go.PL-PFV
'When the man arrived we left.' (Lindenfeld 1973: 81, 83)
b. *Aapo aman wee-ne [(si) enchi aman siik-o].*
3SG.NOM there go.SG-POT CLM 2SG.ACC there go.SG.PFV-CLM
'She will go there if you go.' (Lindenfeld 1973: 83)
c. *Aapo lipti-tu-ne [chikti ta'apo tekipanoa-tek-o].*
3SG.NOM blind-VBLZ-POT every day work-hypothetical-CLM
'He will become blind if ever he works every day.' (Lindenfeld 1973: 83)

---

is a motion-*cum*-purpose linkage; without *-se* 'move in order to', the clause is interpreted as a manner clause in (b′).

(a) *Wate$_i$ into [juya-ta-betukun $\__i$ kopan-bae-ka] aman rukte-k.*
some DM tree-ACC-under shelter-want-CLM there approach-PFV
'Then some people approached there to rest under a tree (lit. Wanting to rest).'

(a′) *Wate$_i$ into [juya-ta-betukun $\__i$ kopan-ka] aman rukte-k.*
some DM tree-ACC-under shelter-CLM there approach-PFV
'Then some people approached there and rest under a tree (lit. Moving & resting).'

(b) *Bwite-k [$\__i$ au esso-se-ka] Waimam-mewi.*
run.SG-PFV 3SG.REF hide-move.PURP.SG-CLM Guaymas-DIR.PL
'He ran to Guaymas to hide himself (lit. Moving to hide).' (Johnson 1962; Cajeme: 34)

(b′) *Bwite-k [$\__i$ au esso-ka].*
run.SG-PFV 3SG.REF hide-CLM
'He ran, all the while hiding himself.'

In my sample, *-kai* introduces a simultaneous temporal clause (19a), a sequential clause (19b), a clause expressing the mean or circumstance of the main event, as in (19c), and negative circumstance, e.g. the 'without' clause in (19d).

(19) a. *Peo-Ø$_i$ muuku-k [Unison-po $_{-i}$ estudiaroa-kai].*
Peo-NOM die.SG-PFV Unison-LOC study-CLM
'Pedro died while studying at the Unison.'

b. *[Bea sechupti $_{-i}$ pensasaroa-ka] ne$_i$ aman siika.*
MD suddenly think-CLM 1SG.NOM there go.SG.PFV
'Then, after I suddenly thought [about it], I went there.'
(Guerrero; Lalo: 182)

c. *Into = ne$_i$ ousi tomi-yo'o-k [$_{-i}$ yeewe-kai].*
DM = 1SG.NOM a.lot money-win-PFV play-CLM
'And then, I won a lot of money by playing.' (Buitimea; chapayeca: 83)

d. *[Kat] = e'e$_i$ [$_{-i}$ ji'i-bwa-ka] to'o-ne.*
NEG.IMP = 2SG.NOM something-eat-CLM lie-POT
'You will go to bed without eating.' (Dedrick & Casad 1999: 394)

I found examples which can be understood as a clause expressing how the event described in the main clause is realized, as seen in (20a), or a clause expressing being/having a quality, as depicted in (20b–c). Notice that the linked verb in (20b–c) comprises a stative verb without *-ta*. Yet, more data is needed to show generalizations in form and use of structures like those in (20).

(20) a. *¡Kubaji-make am-eu chai-tebo! = ti ne-u te'e-ka [omte-kai].*
drum-INST 3PL-DIR call-order = CLM 1SG-DIR insist-PFV be.mad-CLM
"Order to call them with the drum", he insisted me with anger / being mad.'
(Buitimea; malagüero: 42)

b. *Naa = bea emfermera-Ø$_i$ [kea $_{-i}$ majae-ka] bea am = bwise-Ø,*
DM = DM nurse-NOM just be.scare-CLM DM 3PL.ACC = take-PRS
u-me ili mampusiam.
DET-PL little finger.PL
'And then, the nurse, with fear/being scare, takes the little fingers.'
(Guerrero; Lalo: 321)

c. *Bea ba'a-po yeu = saja-k [kaa tajo'ore-kai].*
DM water-LOC out = go.PL-PFV NEG be.dress-CLM
'They left the water without being dressed.' (Buitimea; lagartas: 38)

Likewise, the final subordinator *-o* may introduce simultaneous (21a) and sequential (21b) temporal clauses, conditional (21c), and counter-factive (21d) clause types.

(21) a. *[Unison-po nee estudiaroa-k-o] Peo-Ø muuku-k.*
Unison-LOC 1SG.ACC study-PFV-CLM Peo-NOM die.SG-PFV
'When I was studying at the Unison, Pedro died.'

b. *Naa = bea = te pakun-bichaa yeu = saja-k*
then = MD = 1SG.NOM outside-toward out = go.PL-PFV
[a = puntaroa-su-k-o].
3SG.ACC = suture-COMP-PFV-CLM
'Then, we left [the hospital] after they sutured him.' (Guerrero; Lalo: 358)

c. *[Jiba enchi nee suaati-bicha-o] ne ya'ura-ta*
always 2SG.ACC 1SG.ACC bother-see-CLM 1SG.NOM authority-ACC
tejwaa-ne.
tell-POT
'If you keep bothering me, I will tell the authorities.'
(Silva et al.; turtle & coyote: 59)

d. *[Kaa nee$_i$ 'u'utte bwite-o] nee$_i$ bwij-'ea-n.*
NEG 1SG.ACC strong run.SG-CLM 1SG.ACC grab-should-PASC
'If I hadn't run fast, he would have grabbed me.' (Dedrick & Casad 1999: 395)

This subordinator may introduce a similarity manner clause, as shown in (22a). The adverbial clause in (22b) may be interpreted as either, a temporal, causal or reason clause.

(22) a. *[Junak = bea kaa tua jiak-nok-wa-o junama jo'ara-po],*
then = DM NEG INT Yaqui-speak-PASS-CLM there house-LOC
ne yo-nok-ta ta'a-taite-k.
1SG.NOM yori-word-ACC learn-begin-PFV
'As Yaqui was not spoken in that house, then I started to learn Spanish (lit. The word of the yoris, meaning the foreigners).'
(Buitimea; mundo de sueño: 77)

b. *[a'a kuna-wa siik-o] puh-bahiya-taka juni'i bwaana.*
3SG.ACC husband-GEN go.SG.PFV-CLM face-swell-being even cry
'When/since/because her husband left, her face was swollen from crying.'
(Dedrick & Casad 1999: 395)

In oral narratives, these two general markers prefer to encode a temporal relation. In my sample, 152 out of 180 adverbial clauses marked by *-kai* and *-o* introduce a temporal clause (Table 8). They can express simultaneous relations (complete and partial overlap), and sequential relations (posterior and anterior); there are narrative 'when' clauses in the corpus too (see Declerck 1997).

Table 8. Temporal clauses marked by general subordinators (Guerrero 2014)

	Simultaneous reading			Sequential readings			Narrative *when*-clauses
	complete overlap	point of coincidence		posterior (*before*)	anterior (*after*)		
-kai clauses (94)	17	16	33	–	60	60	2
-o clauses (65)	15	15	30	5	22	27	9
	32	31	62	5	82	87	12

Sequential clauses can take additional markers to specify the semantic association they establish with the main unit. Thus, *before*-clauses are highlighted by the initial adverbial particles *ketun* 'yet' and *ketun ke* 'not yet', besides *-o*/*-kai*, as illustrated in (23a). *After*-clauses can take the completive suffix *-su* (historically related to *ansu* 'finish') directly attached to the dependent verb, as in (23b). However, in my corpus, there are only 18 specific temporal clauses (10% of the sample of temporal clauses), suggesting that speakers prefer to use the general clauses over the more specific expressions.

(23) a. [*Kee Sulumai-ta*$_i$ *bwij-wa-o*] *aapo*$_i$ *enchi juya-m*
ADV.NEG Sulumai-ACC capture-PASS-CLM 3SG.NOM 2SG.ACC brush-PL
nasuk e'e-ria-k.
middle hide-APPL-PFV
'Before Sulamai was captured, she was able to hide you inside the brush.'
(Buitimea; torokoroyi: 78)

b. *U'u chu'u-Ø*$_i$ *[*$_{-i}$ *ji-bwa-su-ka] muku-k.*
DET dog-NOM thing-eat-COMP-CLM die.SG-PFV
'When the dog stopped eating, he died.'

In sum, both *-kai* and *-o* can be general subordinators. In the adverbial domain, the two can introduce temporal, manner, concessive, and counter-factive clauses. In this sample, only *-o* introduces conditional clauses, while only *-kai* can express means, positive and negative circumstances related to the main event. Occasionally, multiple subordinating morphemes may occur to emphasize the semantic relation between the main clause and the adverbial clause. Like relatives and complements, adverbial clauses marked by final subordinators demand accusative subjects in the linked unit. Even though nominative subjects are allowed in adverbial clauses taking initial subordinators, they are ruled out here.

4.4 General subordinators serving as switch-reference marking?

Neither Lindenfeld nor Dedrick and Casad draw attention to a consistent, though not obligatory distribution among the two general subordinators and the identity of the subjects. *-Kai* is used when the main and dependent subjects are the same, as in (24a), while *-o* occurs when the subjects are different, as seen in (24b). As for the Sonoran Yaqui, I have not seen cases of *-kai* with different-subjects, but one may find examples (three in my narrative sample) of same-subject *-o* clauses, as depicted in (24c).

(24) a. *Maria-Ø$_i$ Fermin-ta$_j$ bicha-k [kafe-ta $\__i$ saake-kai].*
María-NOM Fermín-ACC see-PFV coffee-ACC toast-CLM
'María saw Fermín when (she) was toasting coffee.'

b. *Maria-Ø$_i$ Fermin-ta$_j$ bicha-k [kafe-ta $\__j$ saake-o].*
María-NOM Fermín-ACC see-PFV coffee-ACC toast-CLM
'María saw Fermín when (he) was toasting coffee.'

c. [*Ju-ka Sulumai-ta*$_i$ *omotria-u yepsa-k-o*] $\__i$ *jaibu kaa*
DET-ACC Sulumai-ACC brush-DIR arrive.SG-PFV-CLM already NEG
enchi tea-k.
2SG.ACC find-PFV
'When Sulumai got back to the brushes, she couldn't find you.'
(Buitimea; toorokoyori: 117)

The distribution of *-kai* and *-o* in the clauses in (24a–b) resemble a canonical switch-reference system. Formally, switch-reference is almost always a verbal category showing whether two arguments have identical reference (Haiman & Munro 1983); functionally, it is a device for referential tracking (i.e., limiting possible range of nominal reference). The role of switch-reference systems is to avoid ambiguity of reference, in particular, across-clauses. Using these two Yaqui general subordinators as switch-reference is promising although it is limited to the domain of adverbial subordination.[7] There is another pair of examples below in (25). Besides *-kai/-o*, the particle *juni'i* emphasizes the concessive meaning of these clauses.

(25) a. *Ta [ama ejkuela-ka juni'i] kaa aa nooka.*
but there study-CLM even NEG able speak
'But, even though they study there, they were not able to speak [our language].' (Félix; HVC: 332)

7. The distribution of *-kai/-o* satisfies another common property of switch-reference systems, the fact that same-subject clauses are usually the reduced version of the corresponding different-subject counterpart. In addition to the missing argument, the verb inside a *-kai* clause must be unmarked for tense-aspect-mood.

b. *Empo* *a = uuse-k* *[pochi-lai-k* *a*
2SG.NOM 3SG.ACC = son-have short-PTCP-have 3SG.ACC
bwasia-k-o *juni'i].*
tail-have-CLM even
'You have a son, even though he has a short tail.' (Silva et al.; wildcat: 20)

Using *-kai/-o* establishes particular co-reference patterns and lexical coding of linked subject.[8] It is the case that the two subjects are different in *-o* adverbial clauses. Therefore, the linked subject can be overt or covert (when 3rd person); when it is overt, it must be marked as accusative case. In *-kai* clauses, there is a missing argument in co-reference with the main subject. When the adverbial clause is final, as in (19a, c) above, the co-referential subject is overtly expressed in the main clause; when the adverbial clause is at the beginning of the sentence, the typical situation is that the lexical subject is coded in the following clause, as in (26a). There are a few cases where the subject is extra-posed to the left of the sentence for pragmatic purpose; see the nominative subject in (26b).

(26) a. *[bea sechupti* $_{-i}$ *pensasaroa-ka] ne*$_i$ *aman siika.*
MD suddenly think-CLM 1SG.NOM there go.SG.PFV
'And, when I suddenly thought [about it], I went there.'
(Guerrero; Lalo: 255)

b. *Dios-Ø*$_i$ *into [*$_{-i}$ *a = bicha-ka]* *a-u* *chai-tebo-k.*
God-NOM DM 3SG.ACC = see-CLM 3SG-DIR call-order-PFV
'And God, after seeing this, he ordered to call him.'
(Silva et al; Wildcat: 36)

b′. *Dios-Ø*$_i$ *into a-u* *chai-tebo-k* *[*$_{-i}$ *a = bicha-ka].*
God-NOM DM 3SG-DIR call-order-PFV 3SG.ACC = see-CLM
'And God ordered to call him when he saw this.'

That the subject *Dios* 'God' is nominative rather than accusative, indicates that it serves as a constituent of the main unit. I asked my consultant what would happen if the linked unit in (26b) occurs at the end of the sentence, and he provided me with the alternative structure in (26b′), with the lexical subject in the main unit and the co-referential missing argument in the adverbial clause.

For the Arizonan variant, Lindenfeld claims that in *-kai* clauses "its subject is often not identical to that of the matrix sentence" (1973: 87). The author uses the example in (27a) to explain this generalization. She argues that *inepo* 'I' is the subject of *tekipanoa* 'work', while *in mala* 'my mother' is the subject of the linked verb 'die'; thus, there are two different subjects and still there is a *-kai* clause. That

8. As a reviewer suggested, the use of *-kai* and *-o* as switch-reference markers may follow a pragmatic, rather than syntactic control relation.

would contradict the findings in my corpus. However, my consultant provides a different interpretation of the same sentence. Accordingly, in (27b) 'my mother' is the subject of the two verbs inside the linked verb, *muku* 'die' and the phrasal verb *naate* 'begin', thus, the subjects of the two verbs are the same.

(27) a. *Inepo [in mala muku-k-naate-kai] tekipanoa.*
1SG.NOM 1SG.GEN mother-Ø die.SG-PFV-begin-CLM work
'I have been working ever since my mother died.'

b. *Inepo [in mala$_i$ muku-k [$_{-i}$ naate-kai]] tekipanoa.*
1SG.NOM 1SG.GEN mamá-Ø morir-PFV empezar-CLM work
'I have been working since my mother become to be death (lit. begin to die)'

Still, this sentence (presumably from direct elicitation) is strange. On one hand, there is a full clause in the middle of the main clause; the subject of *tekipanoa* 'work' is a nominative subject extraposed to the left (the same that (26b) above). On the other, the dependent subject does not take the accusative case. In Lindenfeld's examples, adverbial clauses marked by native final subordinators (no loanwords) take accusative case. Yet, at least for the Sonoran variant, *-kai* clauses in the corpus involve same subjects. The possibility of *-kai* taking different subjects (or not) deserves a study on its own (Guerrero, in preparation).

5. Adverbial subordinators in Cahita languages: Initial or final subordinators?

Bweituk and *kiali'ikun* stand out among the Yaqui subordinators not only because of their position and lexical meaning, but because they demand a nominative subject. Interestingly, these adverbial particles were not mentioned in the earlier works on Yaqui. In Buelna's work (1890), the adverbial clauses were marked by final subordinators, including those denoting cause and reason relations. Buelna argued that "en oraciones de dos supuestos [different subjects], la persona que hace, sobre que cae o apela cualquiera de las dichas partículas se pone en acusativo" (1890: 68).[9]

First, Buelna noticed that causal and reason clauses were marked by final subordinators. The clause linkage markers *teca* ~ *tuca* 'because' introduce same-subject reason clauses; *teca* is used with verbs in present and future, as in (28a), while *tuca* is used elsewhere. The subordinator *ituca* ~ *tuco* 'because' marks different-subject clauses, the former for present and imperfective, as illustrated in (28b–c), and the latter elsewhere.

9. The examples in this section come from Buelna (1890: 68–76); the orthography is respected from the original examples. The morphological gloss of the examples is mine.

(28) a. *Emchi ne$_i$ noctehoa [emchi $_{-i}$ eria teca].*
2SG.ACC 1SG.NOM teach 2SG.ACC love CLM
'I teach you because I love you.' (Sp. Porque te amo, te enseño)

b. *Emchi ne vuie [emchi ka teopa-u quivaque ituca].*
2SG.ACC 1SG.NOM yell 2SG.ACC NEG church-DIR enter CLM
'I yelled at you because you are not inside the church.' (Sp. Te riño porque no entras a la iglesia)

c. *Buiteca ne [emchi netz eriac tuco].*
run.IMPFV 1SG.NOM 2SG.ACC 1SG.ACC love-PFV CLM
'I ran because you loved me.' (Sp. Porque me amaste, huí)

Buelna's examples show how these constructions reflect the lexical coding of the linked subject observed nowadays: in same-subject constructions, there is a missing argument in the linked unit in co-reference with the main subject; in different-subject clauses, the dependent subject is overt and marked as accusative. Historically, these final subordinators do not distinguish between reason and causal relations, i.e., they denote a general consequence relation. In fact, Buelna recognized that *teca* may denote other semantic relations too, such as conditionals, as in (29a), and temporal clauses, in (29b).

(29) a. *Emchi = ne$_i$ hiocori ciai, [emchi $_{-i}$ eria teca].*
2SG.ACC = 1SG.NOM help SUBJ? 2SG.ACC love CLM
'I would help you if I would love you.' (Sp. Te socorriera, si te hubiera amado)

b. *Emchi = ne$_i$ hioco-ree, [quehe $_{-i}$ emchi eria teca].*
2SG.ACC = 1SG.NOM help-PFV not.yet 2SG.ACC love CLM
'I helped you before I love you.' (Sp. Te socorrí antes de que te amara)

For general subordinators, the linkage marker *-cari*, according to Buelna (1890: 69), marks same-subject adverbial clauses and is translated as *como* 'like', *cuando* 'when', *aunque* 'although', *después* 'after'. Some examples are shown in (30).

(30) a. *Eria-naua = 'e aeri-cari.*
love-POT.PASS = 2SG.NOM love-CLM
'As you love, you would be loved.' (Sp. Como ames, tú serás amado)

b. *Mue-naque ne aeria-cari.*
die-POT 1SG.NOM love-CLM
'I will die when I would be loved.' (Sp. Me moriré, cuando yo ame)

c. *Ca = te$_i$ eria-naua [mautz $_{-i}$ aeria-cari].*
NEG = 2SG.NOM love-POT.PASS although love-CLM
'You won't be loved although you love.' (Sp. Aunque tú ames, no serás amado)

The linkage marker *yo ~ co ~ caco* introduces different-subject clauses such as *como* 'like', *cuando* 'when', *aunque* 'although', *después* 'after', *si* 'if', and *antes* 'before'. The distribution of these morphemes depends on the TAM of the linked verb (Buelna 1890: 70): *yo* for the present and future, as in (31a, d), *co* for the past, as in (31b), and *caco* for the pluperfect in (31c).

(31) a. *Emchi = ne eria-naque, [emchi netz eria-yo].*
2SG.ACC = 1SG.NOM love-POT 2SG.ACC 1SG.ACC love-CLM
'I will love you, if you love me.' (Sp. Si me amares, yo te amaré)

b. *Ca ne emchi lulutiria-naque, [mautz emchi eria-ua-co].*
NEG 1SG.NOM 2SG.ACC forgive-POT although 2SG.ACC love-PASS-CLM
'I will not forgive you, even though you would be loved.' (Sp. Aunque hayas sido amado, yo no te perdonaré)

c. *Emchi = ne eriai [quehe emchi a eria-caco].*
2SG.ACC = 1SG.NOM love not.yet 2SG.ACC 3SG.ACC love-CLM
'I loved you before you could love him/her.' (Sp. Antes que hubieses amado, ya yo te amaba)

d. *Suroc = te taite-c, [emchi sim-su-co].*
get.sad = 1PL.NOM begin-PFV 2SG.ACC go-COMP-CLM
'We became sad after you left.' (Sp. nos entristecimos después que te fuiste)

e. *Anevebac [vitzi ka unaua].*
whip.PFV CLM NEG strong
'Whip [him] but not too strong.' (Sp. Azótele pero no mucho)

Whereas the linked subject in the *-cari* clauses in (30) is covert, the subject inside the *yo/-co/-caco* sentences in (31) consists of an accusative pronoun. Notice that *-cari* and *-yo/-co/-caco* can co-occur with initial adverbs to specify the semantic relations among the two clauses; this is the case of *mautz(i)* 'although, it does not matter' in (30c) and (31b), the adverbial *quehe(ri)* 'not yet, before' (31c) marking a before-clause, *vitzi* 'although, but' in (31e). The initial particle *mautz(i)* was not preserved in Yaqui grammar; the actual *ketunke ~ ke* is historically related to *quehe(ri)* 'not yet' (see the example in (29b)); the same may be true for *bweta* and *vitzi* 'but'.

According to Buelna's work, adverbial subordinate clauses were marked by final subordinators; some of these subordinators were bound (e.g., *-cari, -yo/-co/-caco*), and others were free but final (e.g. *teca, ituca, tuco*). These subordinating morphemes were general rather than specific linkers. The use of *-betchi'ibo* 'in order to' and *-po* 'in, on' as subordinators was not mentioned in Buelna's work either.

6. Final comments

The aim of this paper was to explore the complexity of the adverbial subordinators in Yaqui. In contrast to relative and complement clauses that take final and bound subordinators without lexical meaning, adverbial clauses can take both final and initial subordinating morphemes. While subordinate sentences – including relatives, complements and adverbials marked by final and bound subordinators – demand non-nominative subjects, adverbial clauses taking the initial linkage markers such as *bweituk* 'because', *kiali'ikun* 'so that', and *alle'apo* 'although' require nominative subjects. Historically, these adverbial clauses took final subordinators and accusative subjects, meaning the syntactic structures denoting reason, causal and concessive relations are relatively new in the Yaqui grammar. Minimally, there are two aspects that need to be addressed in future studies: the possibility of *-kai* taking different subjects and the use of *-kai* and *-o* as switch-reference markers. This phenomenon is observed in languages from the Northern branch (Hill 2012), but rarely mentioned in Southern languages.

Acknowledgements

This study was partially supported by the UNAM-DGAPA-PAPIIT grants (IN401816, IN400919). I thank two anonymous reviewers for their comments and suggestions on the first version of this document. I thank two anonymous reviewers for their comments and suggestions on the first version of this document.

Abbreviations

ACC	accusative	GEN	genitive	PL	plural
ADV	adverbial	IMP	imperative	POT	potential
APPL	applicative	IMPFV	imperfective	PTCP	participial
CLM	clause linkage marker	INST	instrument	PRS	present
COM	comitative	INT	intensifier	PURP	purpose
COMP	completive	LOC	locative	RDP	reduplication
DAT	dative	NEG	negation	REF	reflexive
DEM	demonstrative	NOM	nominative	SG	singular
DET	determiner	PASC	past continuous	SUBJ	subjunctive
DIR	directional	PASS	passive	VBLZ	verbalizer
DM	discourse marker	PFV	perfective		

Narrative corpus

Buitimea Valenzuela, C. 2007. *Peesio betana nottiwame, Regreso de Hermosillo*. Hermosillo: Universidad de Sonora.
Félix Armendáriz, Rolando. Ms. Crescencio's Life story (HVC).
Guerrero, L. Ms. *El tremendo Lalo* (Lalo).
Johnson, J. 1962. *El idioma yaqui y textos*. México: INAH.
Silva Encinas, M. C. 2004. Anexo IV (HVH). In *La secuencia temporal en el discurso narrativo yaqui*. MA thesis. Hermosillo: Universidad de Sonora.
Silva Encinas, M. C. et al. 1998. *Jiak nokpo etejoim, Pláticas en Lengua Yaqui*. Hermosillo: Universidad de Sonora.

References

Álvarez González, Albert. 2012. Relative clauses and nominalization in Yaqui. In *Relative Clauses in Languages of the Americas*: *A Typological Overview* [Typological Studies in Language 102], Bernard Comrie & Zarina Estrada-Fernández (eds), 65–96. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.102.04gon
Álvarez González, Albert. 2015. Sincretismo entre beneficiario, causa y propósito: Evolución de las funciones de *betchi'ibo* en yaqui. *Amerindia* 37(2): 69–96.
Belloro, Valeria & Guerrero, Lilián. 2010. Sobre la función discursiva de la posposición de constituyentes en yaqui. In *Análisis lingüístico: enfoques sincrónico, diacrónico e interdisciplinario*, Rosa María Ortiz Ciscomani (ed.), 71–95. Hermosillo: Universidad de Sonora.
Buelna, Eustaquio. 1890[1989]. *Arte de la lengua cahita por un padre de la Compañía de Jesús*. México: Editorial Siglo XXI.
Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.
Declerck, Renaat. 1997. *When-clauses and Temporal Structure*. London: Routledge.
Dedrick, John M. & Casad, Eugene H. 1999. *Sonora Yaqui Language Structures*. Tucson AZ: University of Arizona Press.
Diessel, Holger. 2001. The ordering distribution of main and adverbial clauses: A typological study. *Language* 77: 343–365. https://doi.org/10.1353/lan.2001.0152
Diessel, Holger. 2005. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43(3): 449–470. https://doi.org/10.1515/ling.2005.43.3.449
Dryer, Matthew S. 2013. Order of adverbial subordinator and clause. In *The World Atlas of Language Structures Online*, Matthew D. Dryer & Martin Haspelmath (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/94> (27 June 2014).
Guerrero, Lilián. 2006. *The Structure and Function on Yaqui Complementation*. Munich: Lincom.
Guerrero, Lilián. 2007. Estructuras argumentales alternativas: Las cláusulas aplicativas en yaqui. In *Mecanismos de voz y formación de palabra*, Zarina Estrada-Fernández, Albert Álvarez González, Lilián Guerrero & María Belén Carpio (eds), 177–204. México: Plaza y Valdez/ Universidad de Sonora.

Guerrero, Lilián. 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

Guerrero, Lilián. 2013. Controller-controlee relations in purposive constructions: A construction-based account. In *Linking Constructions into Functional Linguistics. The Role of Constructions in Grammar* [Studies in Language Companion Series 145], Brian Nolan & Elke Diedrichsen (eds), 1–22. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.145.01gue

Guerrero, Lilián. 2014. Temporal adverbial sentences in Yaqui. Paper presented at the *Conference on Coordination and Subordination in Lisbon*. Universidad de Lisboa, Portugal.

Guerrero, Lilián. 2017. On purpose and causal adverbial clauses in Yaqui. *International Journal of American Linguistics* 83(4): 679–718.

Guerrero, Lilián. 2019. Propiedades formales y funcionales de las posposiciones en yaqui. In *Adposiciones y elementos de su tipo en lenguas de América*, L. Guerrero (ed.), 515–562. México: UNAM.

Guerrero, Lilián & Belloro, Valeria. 2014. Clause-linking strategies expressing reason and adversative relations in Yaqui. *Paper presented at the Syntax of the World Language 6*. Pavia, Italy.

Guerrero, Lilián & Van Valin Jr., Robert. D. 2004. Yaqui and the analysis of primary object languages. *International Journal of American Linguistics* 70(3): 290–319.

Haiman, John & Munro, Pamela. 1983. *Switch-Reference and Universal Grammar* [Typological Studies in Language 2]. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.2

Hill, Jane. 2012. Cupan temporal adverbial clauses and switch reference in Uto-Aztecan perspective. *Paper presented at Taller de los Amigos de las Lenguas Yutoaztecas* 2012. Cholula, Puebla.

Jolly, Julia A. 1992. Preposition assignment in English. In *Advances in Role and Reference Grammar* [Current Issues in Linguistic Theory 82], Robert D. Van Valin Jr. (ed.), 275–310. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.82.08jol

Kortmann, Bernd. 1997. *Adverbial Subordination: A Typology and History of Adverbial Subordinators Based on European Languages*. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110812428

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

Lindenfeld, Jacqueline. 1973. *Yaqui Syntax*. Berkeley CA: University of California Press.

Nordström, Jackie. 2010. *Modality and Subordinators* [Studies in Language Companion Series 116. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.116

Thompson, Sandra A., Longacre, Robert & Shin Ja Hwang. 2007. Adverbial clauses. In *Language Typology and Syntactic Description*, Vol. 2. Timothy Shopen (ed.), 237–300. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.005

Van Valin Jr., Robert. D. 2005. *Exploring the Syntax-Semantics Interface*. Cambridge: CUP.

PART IV

The diachrony of syntactic complexity

Grammaticalization of the linking devices with *ka* in Purepecha

Claudine Chamoreau
CNRS (SEDYL/CELIA – CEMCA)

This paper describes, in synchrony, the use, form, and position of the linking devices with *ka* in coordinating phrases and clauses and in subordinate clauses. This study also explores the relation between the different occurrences of *ka* and investigates a diachronic common source. Taking into account such parameters as weight, cohesion, and variability (Lehmann 2002), I advance the hypothesis that there have been two routes of grammaticalization. These two routes include the free coordinator *ka* which has given rise to a dependent marker *-ka* that forms subordinators, and the subjunctive mood *-ka* which seems to be the grammaticalization of the assertive mood marker *-ka* used for speech-act participants.

Keywords: grammaticalization, Purepecha, coordination, subordination, linker, coordinator

1. Introduction

Coordination and subordination represent two types of relation that are encoded by complex constructions (Bril & Rebuschi 2006). For Haspelmath (2007: 1), "The term *coordination* refers to syntactic constructions in which two or more units of the same type are combined into a larger unit and still have the same semantic relations with other surrounding elements." Coordination applies to a combine of words, as in (1a) to the coordination of nouns and verbs phrases, in (1b) for the noun phrase, and in (1c) to the coordination of clauses. The members may be connected by means of a linking device[1] called a coordinator, by a coordinating conjunction, or by coordinating connectives, such as 'and', 'but', 'or'.

1. I am grateful to two anonymous reviewers who helped me to improve this paper, in particular for making the terminology more precise. The term "linking device" is used with a general meaning when I refer to the general function of *ka,* whatever the context it appears in. When I

https://doi.org/10.1075/tsl.126.06cha

(1) a. Peter **and** Maria sleep **and** snore.
b. My wife **and** my three young girls went to the restaurant.
c. He woke up early **but** he did not get up.

In contrast, subordination is a type of asymmetrical relation, restricted to clauses, in which the subordinate clause is dependent and subordinate to the main clause. The subordinate clause may be linked to the main one by means of a subordinator, also known as a subordinating conjunction, that may be called a relative subordinator or relative marker for the relative clause, as in (2a), a complementizer for the complement clause, as in (2b), or a subordinator for the adverbial clause, as in (2c). The form of the subordinator depends on the type of subordinate clause (Cristofaro 2003; Thompson et al. 2007), in particular when they are grammatical morphemes with lexical content (e.g., English *before*, *when*, *if*).

(2) a. The girl [**who** was crying] was hungry.
b. She suggested [**that** he leave].
c. He gets up [**when** the sun rises].

As illustrated above, coordination and subordination are generally expressed by means of two distinct types of linking device. (These relations may also be showed by juxtaposition with no overt markers: see Cristofaro 2003; Haspelmath 2007; Longacre 2007; and Thompson et al. 2007.) However, in certain languages, the links have a similar appearance, in the sense they have a common formal component even if they are not identical.

Purepecha, an isolate language spoken in Mexico, is an example of this: linking devices with *ka* implies a relation between at least two units. Coordination is expressed by *ka*, a coordinator that on its own establishes a conjunctive relation between two units, as in (3a) for the noun phrase and in (3b) for clauses. In coordination, *ka* constitutes a conjunction in the traditional sense, as it is a free and invariable morpheme (Giacalone Ramat & Mauri 2011: 654). Its position is always before the second unit. Moreover, *ka* may receive adverbial enclitics (Chamoreau 2014) to create other coordinators, such as *ka* = *ru* in (3c) for adversative coordination.

---

refer to the marker in coordinating phrases or clauses, I call it the "coordinator" (Haspelmath 2007: 50). In subordinating clauses, the situation is more complex. The marker *-ka* suffixed on the verb is always called the "subjunctive" in accordance with the tradition in Purepecha descriptions (see below in Section 4 the explanation for this choice). When the marker *-ka* forms a complex marker that appears at the beginning of the clause, it is generally called the "subordinator" because regardless of the type of subordinate clause the function of this marker is the same: it introduces the subordinate clause and indicates the left boundary of the clause. However, in contexts in which it is relevant to distinguish the different types of subordinator, it may be called the "relative subordinator" for relative clauses, the "complementizer" for complement clauses, and the "subordinator" for adverbial clauses.

(3) a. *pawani t'iri-a-ka = ni juchi jinkonikwa-o*
tomorrow eat-IRR-ASS1/2 S = S1SG POSS1SG sister-RES
***ka** mimi-o.*
and brother-RES
'Tomorrow, I will eat at my sister's **and** brother's house.'

b. *jwata tsakapu kw'ani-ku-xa-p-ti **ka** tsakapu ma*
hill stone throw-APPL3O-PROG-PST-ASS3S and stone one
kwhiripu anta-s-p-ti.
people reach-AOR-PST-ASS3S
'The hill was throwing stones **and** one stone reached people.'

c. *ni-ra-s-ti **ka = ru** kinse diya = ksï, pastori-icha*
go-FT-AOR-ASS3S and = other fifteen day = 3PL shepherd-PL
arhi-ra-sïn-ti yamintu ireta.
say-CAUS-HAB-ASS3S all village
'He went out **but** for fifteen days the shepherds make noise in the entire village.'

In contrast, clausal subordination is expressed by two markers *-ka*: one beginning the subordinate clause and another at the end of the verb that is generally positioned at the end of the clause. In this type of clause, *-ka* is no longer a conjunction, as in the case of coordination, because it always depends on another element. Thus, the marker *-ka* presents a dependent form, a suffix, to indicate embedding and subordinate clauses. At the start of the subordinate clause, *-ka* forms different subordinators in combination with another element: after the demonstrative pronoun *inte* 'this' (elided as *in*) to form the relative subordinator *inka* with the relative clause, as illustrated in (4a); after the adverb of manner *isï* 'thus' to build the complementizer *iska* for the complement clause, as in (4b); and after the instrumental postposition *jimpo* to form the subordinator *jimpoka* for a reason clause, as in (4c) (see below Section 4, where Table 3 presents a list of the various subordinators). At the end of the clause, the marker *-ka* suffixed on the verb is referred to as the "subjunctive" mood in the Purepechan tradition because it appears on the verb in the position occupied by the morphemes of this category (Capistrán 2002; Friedrich 1984; Monzón 2004; Wares 1984). The presence of *-ka* on the verb is a morphological constraint for all subordinate finite clauses; it is impossible to use another mood in this clause (Chamoreau 2009: 103–105). In this paper, I continue to use the label "subjunctive" even if it does not correspond to the traditional notion of the subjunctive as opposed to the assertive mood. The constraint of *-ka* on the verb and that the tense and the aspect display distinctive and reduced forms (compare the forms for the aorist aspect and past tense in (3b) and (4c)) indicate that subordinate finite clauses in Purepecha are less finite than main finite clauses (Chamoreau 2016).

(4) a. *tsimanhi-e-x-ti = ksï = nha inte werantupinti-cha*
two-PRED-AOR-ASS3S = 3PL = EV DEM orphan-PL
[***inka*** *= ksï = nha táte-empa-ni ka ama-mpa-ni no*
SUB = 3PL = EV father-KPOSS3-OBJ and mother-KPOSS3-OBJ NEG
*ka-nko-rhe-nka-ø-**ka***].
have-INTS-body-IT-AOR-SBJV
'They said that there were two orphans **who** did not have a father and mother.'

b. *arhi-x-ka* [***iska*** *= ri yóntani jo-nkwa-pirin-**ka***].
say-AOR-ASS1/2 S SUB = S2SG late come-CENTRIP-COND-SBJV
'I said **that** you should come back late.'

c. *jwata tsakapu kw'ani-ku-xa-p-ti* [***jimpoka***
hill stone throw-APPL3O-PROG-PST-ASS3S SUB
*jinche-p-**ka***].
earthquake-AOR.PST-SBJV
'The hill was throwing stones **because** the earth trembled.'

The first aim of this paper is to describe, in synchrony, the use, form, and position of the linking devices with *ka* in coordinating phrases and clauses and in subordinate clauses. The second aim is to explore the relation between the different occurrences of *ka* and to investigate a diachronic common source. Taking into account such parameters as weight, cohesion, and variability (Lehmann 2002), I advance the hypothesis that there have been two routes of grammaticalization: (1) the free coordinator *ka* has given rise to a dependent marker *-ka* that forms subordinators; (2) the subjunctive mood *-ka* seems to be the grammaticalization of the assertive mood marker *-ka* used for speech-act participants (henceforth SAP).

The structure of this paper is as follows. Section 2 gives basic grammatical information about Purepecha. The coordinator *ka* is described in Section 3. In Section 4, I explore the different context in which *ka* occurs in a subordinate clause. The paper closes by proposing a hypothesis about two possible scenarios of grammaticalization.

2. Basic typological characteristics of Purepecha

Purepecha (formerly known as Tarascan) is classified as a language isolate and is spoken in the state of Michoacan in Western Mexico by approximately 110,000 people (Chamoreau 2009, 2012). Classifying Purepecha within the Mesoamerican linguistic area is still debated, but generally it is not classified as a Mesoamerican language, as it possesses few of the characteristics shared by these languages (Smith-Stark 1994).

Purepecha is an agglutinative and synthetic language, and is almost exclusively suffixing. Although bare stems exist, there is a very productive derivational system in which a basic stem can take voice, causative, locative, positional, directional, and adverbial derivative suffixes. Inflectional suffixes follow the stem to mark aspect, tense, irrealis, mood, and person (Chamoreau 2009, 2017).

Purepecha has nominative-accusative alignment and is a case-marking language in which the nominal subject has no overt marker. The object is generally marked by the objective case marker *-ni*. This morpheme encodes the object of a transitive verb, such as *misitu-ni* 'the cat' in (5), and both objects of a ditransitive verb, such as *inte-ni wantantskwa-ni* and *Puki-ni*, in (6). The presence or absence of the objective case marker depends on different hierarchies: (i) the inherent semantic properties of the referent (human, animate); (ii) properties related to grammatical features (definite, count noun vs. mass noun, generic vs. specific, etc.); and (iii) pragmatic strategies (topic, focus) (Chamoreau 2009). Typically, the objective case marker *-ni* indicates that the noun phrase is characterized as individuated.

(5) *jo celia ata-x-ti ima-eri* ***misitu-ni.***
yes Celia beat-AOR-ASS 3SDEM-GEN cat-OBJ
'Yes, Celia beat her cat.'

(6) *celia arhi-x-ti* ***inte-ni*** ***wantantskwa-ni*** ***Puki-ni.***
Celia tell-AOR-ASS3S DEM-OBJ story-OBJ Puki-OBJ
'Celia told Puki this story.'

Purepecha is an SV and SVO constituent order language as illustrated by examples (5) and (6). This is the basic order in the region of Lake Patzcuaro (Capistrán 2002; Chamoreau 2009: 55–58). Other orders indicate specific pragmatic properties. Studies of constituent order in the other regions do not as yet exist. However, Purepecha exhibits traits of a SOV language: (i) tense, aspect, irrealis, and modal markers following the verb; (ii) postpositions; (iii) only suffixes; (iv) only enclitics; (v) case markers; (vi) main verbs preceding inflected auxiliaries; and (vii) positional variation of the head noun in the noun phrase that reveals that final head nouns precede non-final head nouns (compare examples (6) and (7)). SVO and SOV constituent orders were attested in the sixteenth century, and the former has increased since then. The change is most likely due to areal contact. Prior to the Conquest there were speakers of other languages in this territory, Nahuatl (Uto-Aztecan family) and Otomi (Otopamean family), two languages with verb-initial structure. The change probably began under the influence of these languages; Spanish, an SVO language, continued the process, for example by introducing prepositions (Chamoreau 2007).

Purepecha displays the predominance of dependent-marking, for example with the pronominal enclitic, as in (7), and the genitive case, as in (8). Subject and object pronouns are expressed by pronominal enclitics. Table 1 presents the two paradigms of pronominal enclitics, the subject and object enclitics.

(7) *no = **ri** xïpa-ku-a-xa-p-ka.*
NEG = S2SG steal-APPL3O-O3PL-PROG-PST-ASS1/2S
'You were not stealing them.'

(8) *nanaka-echa-**eri** jawiri sési ja-rha-a-ti.*
girl-PL-GEN hair well be.there-FT-IRR-ASS3S
'The girls' hair is beautiful.'

Table 1. Pronominal enclitics in Purepecha

	Subject	**Object**
1	*ø / =ni*	*=reni (=rini) / =ts'ïni*
2	*=re (=ri)*	*=kini / =kxïni*
3	*ø*	*ø*
1PL	*=ch'e (=ch'i) / =kxï**	*=ts'ïni***
2PL	*=ts'ï*	*=kxïni*
3PL	*=kxï*	*=kxïni*

* For the first person plural, the difference between = *ch'e* and = *kxï* (or = *ksï*) today exhibits a dialect variation (Chamoreau 2009: 64) that reveals a diachronic change: in the sixteenth century only = *kuch'e* (the marker that has been grammaticalized in *ch'e*) was used.
** The first person object enclitic = *ts'ïni* and the second person object = *kxïni* are always used when the subject is plural.

Independent and main clauses are the most finite types of clause: participants are expressed by a noun, as in (9), by an independent pronoun, as in (10), or by a pronominal enclitic, as in (7) for subject and (9) for object. The predicate, a verb as in (9) or a noun as in (10), must be marked by a mood. It may also be modified by an aspect and a tense, such as the progressive aspect and past tense in (7), or by an irrealis marker, as in (8).

(9) *juchi tata = **rini** kwane-xïn-ti jiwatsï k'éri-ni.*
POSS1SG father = O1SG lend-HAB-ASS3S coyote old-OBJ
'My father lends me to the old coyote …'

(10) *jucha isï = sï mi-te-s-p-ka ima ts'irakwa jimpo.*
S1PL.IND thus = FOC open-SUP-AOR-PST-ASS1/2S ART.DEF cold INS
'We, thus, knew it for the cold.'

Purepecha distinguishes four different paradigms: aspect, tense, irrealis, and mood (henceforth ATIM). These suffixes have various dialectal allomorphs; I present in

Table 2 the forms attested in the village of Jaracuaro (for the different allomorphs see Chamoreau 2009 and Friedrich 1984). There are four aspect markers (aorist – the narrative non-marked aspect – habitual, progressive, and continuative), one tense marker (past), two irrealis markers (irrealis-future, conditional), and five mood markers (assertive, interrogative, imperative, exclamative, and subjunctive). Table 2 presents the morphemes as they occur in independent or main clauses.

Table 2. ATIM markers in independent or main clauses

Aspect	Tense	Irrealis	Mood
AOR *-x**	PAST *-p/-an*	IRR *-a*	ASS1/2 S *-ka* / ASS3S *-ti*
HAB *-xïn*	PRS *unmarked*	COND *-pirin*	INT *-ki/-i/-ø*
PROG *-xa*			IMP.SG *-ø* / IMP.PL *-e*
CONT *-xam*			EXCL *-k'a*
			SBJV *-ka*

* In aorist and habitual aspect markers, the palatal *x* is pronounced as an alveolar *s* (*-s* for aorist aspect, *-sïn* for habitual aspect). This is a dialectal variation.

These morphemes have a fixed order of occurrence at the end of the predicate: (aspect) + (tense) + (irrealis) + mood (Chamoreau 2009, 2017). Mood is obligatory in a finite verb. Aspect has to occur with tense, and with mood. The irrealis marker occurs directly with mood (aspect and tense are impossible). The assertive mood is *-ka* for SAP first and second persons, as in (11a), (11b), (11d), (11e), and *-ti* for third persons, as in (11c), (11f). Interestingly, in this mood two of them – the first person singular and the third person singular – may have a zero marker when they function as the subject. The third person always has a zero marker. For the first person there is no difference in using the covert marker or the overt = *ni*. In the sixteenth century the use of *ø* was the most frequent (Chamoreau 2014). We may observe in (11) the paradigm with the verb *kara-* 'write' and the habitual aspect *-xïn*:

(11) a. *kara-xïn-**ka** / kara-xïn-**ka** = **ni*** 'I write'
b. *kara-xïn-**ka** = **ri*** 'You (sg.) write'
c. *kara-xïn-**ti*** 'He writes'
d. *kara-xïn-**ka** = **ksï*** 'We write'
e. *kara-xïn-**ka** = **ts'ï*** 'You (pl.) write'
f. *kara-xïn-**ti** = **ksï*** 'They write'

Purepecha distinguishes between an assertive mood and an interrogative mood in independent and main clauses. Compare (12) with (13), which has the interrogative mood *-ø* (this is the allomorph used after the irrealis). This language contrasts these two moods, which occur in independent clauses, with the so-called subjunctive mood *-ka*, which codifies the verb in a dependent clause. Compare (14a) with (15a) and (14b) with (15b).

(12) *ni-**a**-**ka** = **ri**.*
go-IRR-ASS1/2 S = S2SG
'You will go.'

(13) *ni-**a**-**ø** = **ri**.*
go-IRR-INT = S2SG
'Will you go?'

(14) a. *anchi-kuri-x-ka.*
work-REFL-AOR-ASS1/2 S
'I worked.'
b. *anchi-kuri-x-ti.*
work-REFL-AOR-ASS3S
'He worked.'

(15) a. *arhi-x-ka = ri* *[exka = ni* *anchi-kuri-ø-**ka**].*
say-AOR-ASS1/2 S = S2SG SUB = S1SG work-REFL-AOR-SBJV
'You said that I worked.'
b. *arhi-x-ka = ri* *[exka* *anchi-kuri-ø-**ka**].*
say-AOR-ASS1/2 S = S2SG SUB work-REFL-AOR-SBJV
'You said that he worked.'

3. Coordinator *ka*

This section is divided into two sub-sections. In the first (3.1), I introduce the general properties of the coordinator, showing its use in interphrastic contexts. In the second (3.2), I describe the use of the coordinator in interclausal contexts.

3.1 General properties and interphrastic uses

The coordinator *ka* is a free and independent element that is used to link two noun phrases, as in (16), or two verb phrases, as in (17). As illustrated in (16) and (17), *ka* on its own establishes the coordination of two units that are functionally equivalent. The coordinator *ka* expresses the conjunctive coordination that may be translated by 'and' in English.

(16) a. *t'u* ***ka*** *watsï-ti* *mak'u = ts'ï* *ja-rha-x-ka.*
S2SG.IND and son-KPOSS2 similar = S2PL be.there-FT-AOR-ASS1/2 S
'You and your son are similar.'

b. *pawani t'iri-a-ka = ni juchi jinkonikwa-o*
tomorrow eat-IRR-ASS1/2 S = S1SG POSS1SG sister-RES
ka *mimi-o.*
and brother-RES
'Tomorrow, I will eat at my sister's and brother's house.'

(17) *kwara-tsi-x-ti* ***ka*** *kaka-rhu-x-ti.*
fall-ground-AOR-ASS3S and break-nose-AOR-ASS3S
'He fell to the ground and broke his nose.'

When over two units are coordinated, *ka* always appears before the last element, as with *ka piri-mpa* in (18). The other units are juxtaposed, as with *Alicia, Celia, Emilio* in (18):

(18) *Alicia, Celia, Emilio* ***ka*** *piri-mpa ampuxï* ***juka-tsï-a-x-ti.***
Alicia Celia Emilio and sister-KPOSS3 louse have-TOP-O3PL-AOR-ASS3S
'Alicia, Celia, Emilio and his sister have lice.'

The coordinator *ka* is a free and independent element because it may function as host for an adverbial enclitic (Chamoreau 2014) to build other coordinators. In (19), the complex coordinator *ka = ru* 'and = then' is an adversative that may be translated by 'but' and in (20) *ka = teru* 'and = other' expresses disjunction that may be translated by 'or'.

(19) *k'e-xa-ti* ***ka = ru*** *marhua-ta-xa-ti.*
grow-PROG-ASS3S and = then employ-CAUS-PROG-ASS3S
'He grows but he continues to use them (diapers).'

(20) *chi kawayu urapiti-x-ki* ***ka = teru*** *turhipiti-x-ki.*
POSS2SG horse be.white-AOR-INT and = other be.black-AOR-INT
'Is your horse white or is it black?'

Nowadays these two complex coordinators are forsaken, and are replaced by the Spanish coordinators 'pero' (also pronounced *peru*) for the adversative, as in (21), and 'o' for disjunction, as in (22). These borrowings confirm the implicational hierarchy (but > or > and) proposed by Matras (1998: 301–305, 2007: 54–56; see also Chamoreau 2007: 470–471). Purepecha, like many languages, has borrowed the elements 'but' and 'or', but not 'and'.

(21) *mis-kurhi-s-ti* ***peru*** *piri-xa-ti.*
be.sad-REFL-AOR-ASS3S but sing-PROG-ASS3S
'He is sad but he is singing.'

(22) *ni-ra-a-ø ama-mpa-nkuni* ***o*** *imeri tata-nkuni.*
go-FT-IRR-INT mother-KPOSS3-COM or POSS3SG father-COM
'Will she go with her mother or with her father?

Whatever the coordinator (*ka, ka = ru, ka = teru, peru, o*), its position is always the same: the coordinator is grouped with the following phrase (or the last one, when several elements are coordinated, as in (18)), not with the preceding one. In Purepecha, it is much more natural to make a pause before rather than after a coordinator.

3.2 Interclausal uses

The various coordinators introduced above, and in particular the conjunctive coordinator *ka*, are attested in interclausal coordination. Two different contexts exist in Purepecha: either the coordinated clauses are finite and functionally equivalent (explored in Section 3.2.1) or the coordinated clauses are not functionally equivalent. In this context, the second clause is non-finite and depends on the first one, the main clause. This interesting process is found in chain-medial clauses (described in Section 3.2.2).

3.2.1 *Coordinated finite clauses*

When two clauses are coordinated, they may have the same reference, as in (25a), or switch reference, as in (23) and (24). The two clauses may have the same tense, aspect, irrealis, or mood markers, as in (24) and (25b), but the opposite is also possible, as in (23) and (25a). The coordinator links two independent and finite clauses. Each clause can be autonomous and independent. The two coordinated clauses are functionally equivalent, having the same semantic function, with possible syntactic autonomy since they contain at least one argument and each clause has a finite verb (Haspelmath 2007). These behaviors are illustrated in examples (23), (24), and (25) with the different coordinators described above (in Section 3.1).

(23) *jwata tsakapu kw'ani-ku-xa-p-ti* ***ka*** *tsakapu ma kw'iripu*
hill stone throw-APPL3O-PROG-PST-ASS3S and stone one people
anta-s-p-ti.
reach-AOR-PST-ASS3S
'The hill was throwing stones **and** one stone reached people.'

(24) *chi kawayu urapiti-x-ki* ***ka = teru*** *ima animalu*
POSS2SG horse be.white-AOR-INT and = other DEM animal
turhipiti-x-ki.
be.black-AOR-INT
'Is your horse white **or** is this animal black?'

(25) a. *ni-ra-s-ti* ***ka*** *= ru* *kinse* *diya = ksï,* *pastori-icha*
go-FT-AOR-ASS3S and = then fifteen day = S3PL shepherd-PL
arhi-ra-sïn-ti *ireta.*
say-CAUS-HAB-ASS3S village
'He went out **but** for fifteen days the shepherds make noise in the village.'

b. *mi-ti-xïn-ka* *kara-ni* ***peru*** *= ni* *no* *u-xïn-ka* […].
open-SUP-HAB-ASS1/2 S write-NF but = S1SG NEG may-HAB-ASS1/2 S
'I know how to write **but** I cannot […].'

Using coordinators between two clauses reinforces that these elements are grouped with the following unit and not the preceding one. The coordinator *ka* (and the other coordinators) may function as the host for pronominal enclitics, as in (26), although today, as in (27), the enclitic often appears after a constituent located after the coordinator.

(26) *ampuxï* *juka-tsï-a-x-ti = kxï* ***ka*** *= kxï* *menkhu*
louse have-TOP-O3PL-AOR-ASS3S = S3PL and = S3PL always
katsï-tsï-ni *ja-rha-sïren-ti.*
scratch-TOP-NF be.there-FT-HAB.PST-ASS3S
'They had lice and they always scratched their heads.'

(27) *ka* *jini = kxï* *ni-ra-x-ti* *jurimpitkwa.*
and there = S3PL go-FT-AOR-ASS3S straight.ahead
'and there they have gone straight ahead.'

3.2.2 *Coordinated non-finite chain-medial clauses*

Within discourse coherence, a clause-chain is characterized as the "smallest unit of coherent multi-propositional discourse," one with "the tightest, most continuous cross-clausal coherence links" (Givón 2001: 355). Chain-medial clauses "carry the bulk of sequential new information in the chain and display the highest cross-clausal coherence. Their grammatical marking is the most minimal, since most threads of thematic coherence (topical referents, temporality, aspectuality, modality, perspective) remain the same" (Givón 2001: 356). Chain-initial and chain-final clauses are the most finite type, while chain-medial clauses are the least finite (the degree of finiteness of chain-grounding clauses is often unpredictable, although usually they are nominalized phrases). The correlation between degree of finiteness and clause-types within the chain in discourse shows that the more referential ones and those with thematic predictability – corresponding to the highest degree of cross-clausal coherence and continuity – display less finiteness.

Non-finite chain-medial clauses in Purepecha are constructions used to facilitate thematic, referential, and aspectual continuities in discourse (Chamoreau 2016). Such strategies correlate with reduced finiteness. When the subject is

the same in the discourse, reference tracking is always possible and easy; in a chain-medial clause the referent functions as the subject. In the chain-initial clause, as in (28a), the referent is introduced by the definite noun phrase *acheti-echa* 'the men' and the verb is marked by aspect. In chain-medial clauses, as in (28b) and (28c), the pronominal enclitic = *ksï* is attested and attached to the verb. Referential and tense-aspect-mood continuities are crucial to understanding the use of these constructions. In non-finite chain-medial clauses in Purepecha, ATIM are not expressed but are recoverable. These clauses use a non-finite *-ni* marker. The aspect of the narratives is usually the aorist, as in (28a). Non-finite medial clauses depend on the chain-initial clause, the independent clause, for its ATIM and subject reference. Dependent non-finite clauses are connected to independent finite clauses by means of the coordinator *ka*, as in (28b) and (28c).

(28) a. *xasï = ksï = nha kustakwa jinkoni acheti-echa pa-s-ti,*
next = S3PL = EV music COM man-PL take-AOR-ASS3S
'They said that then the men took her with music,
b. *ka jikwa-ra-ni = ksï = nha ya,*
and bathe-CAUS-NF = S3PL = EV NOW
and they said that they bathed her,
c. *ka ampa-tsi-ku-ni = ksï = nha ya.*
and be.clean-LOW-NCS-NF = S3PL = EV NOW
and they said that they combed her.'

When *ka* is used in chain-medial clauses, the construction is ambiguous as between coordination and subordination. The morpheme *ka* appears with the form for coordinated clauses, that is, as a free and independent element, but the clauses that make up the clause chaining are not functionally equivalent with the first clause, and chain-medial clauses do not possess syntactic autonomy. The chain-medial clause is coordinated with the chain-initial clause and depends on it: the initial clause contains ATIM and argument reference, while coordinated chain-medial clauses have reduced finiteness (Longacre 2007: 375). In Purepecha, these clauses use the non-finite *-ni* marker. The coordinator *ka* is repeated at the beginning of each chain-medial clause, as in (28). But in certain chain-medial clauses the clauses are juxtaposed, as in (29b) and (29c). The coordinator *ka* shows the end of a chain-medial clause, as in (29d). Using *ka* in a chain-medial clause is relevant because it exhibits a high level of thematic and tense, aspect, irrealis, and mood continuities (Chamoreau 2016).

(29) a. *Teremendo anapu-echa kutsu-sïraam-ti,*
Teremendo origin-PL tan-HAB.PST-ASS3S
'Those from Teremendo tanned (leather),

b. *ima-echa noampe u-ni,*
DEM-PL NEG do-NF
they did not do anything,
c. *no = ksï sesi xama-ra-nte-ni,*
NEG = S3PL well smell-CAUS-nose-NF
they did not smell good,
d. *ka no = ksï tsipi-ni.*
and NEG = S3PL be.happy-NF
and they are not happy.'

To sum up: Purepecha has a free and independent coordinator *ka* that may occur alone between words, phrases, and clauses. Only one occurrence of *ka* is necessary to coordinate two units. In the case of various words, phrases or clauses are coordinated, *ka* only occurs once, before the last coordinated unit (except in certain occurrences of chain-medial clauses). The other units are juxtaposed. With words, phrases, and finite clauses, the coordinated units are functionally equivalent and possess possible syntactic autonomy. This is not the case in the contexts of coordination of chain-medial clauses that are coordinated but dependent on the chain-initial clause. In Purepecha, there is a clear distinction between coordination (linkage of two independent clauses, where both are coordinated), chain-medial clause (linkage of an independent clause with a dependent clause, where the latter depends on the former and both are coordinated) and subordination (linkage of an independent clause with a dependent clause, where the latter is embedded in the former; see Section 4). Literature has also described the second construction as a medial verb or as co-subordination (Foley & Van Valin 1984; Haspelmath 1995).

4. *ka* used in subordinate finite clauses

Purepecha is one language in which subordinate clauses are usually finite. The verb has markers of tense, aspect, irrealis, and mood. Pronominal enclitic and switch reference are possible as illustrated in (30). Non-finite subordinate clauses also exist for same reference complement clauses, as in (31), and purpose clauses, as in (32). In non-finite subordinate clauses, the verb is marked by the overt non-finite marker *-ni*, no tense, aspect, irrealis, or mood, no subordinator, no pronominal enclitic, usually same-subject and same intonation contour (for more details, see Chamoreau 2016).

(30) *arhi-nha-sïren-ti* *tanako,* [***jimpoka*** = *kxï kw'iripu-echa*
say-PAS-HAB.PST-ASS3S Tanaco SUB = S3PL person-PL
*tanha-li-sïren-**ka***].
join-body-HAB.PST-SBJV
'That place was called Tanaco **because** people were gathering together.'

(31) *ero-ta-xa-ka* [***k'ama-ta-ni*** *prontu*].
hope-CAUS-PROG-ASS1/2 S finish-CAUS-NF soon
'I hope to finish soon.'

(32) *Kumicho* *incha-parha-ku-x-p-ka* [***para***
Ocumicho enter-LONG.EXT-NCS-AOR-PST-ASS1/2 S for
*eskwela **arhi-t'a-a-ni***].
school say-IT-O3PL-NF
'I had entered into Ocumicho to teach them.'

In this section, our aim is to describe the different elements that occur with the marker *-ka* in finite subordinate clauses. As introduced in Section 1, subordination in finite clauses is expressed by the presence of *-ka* in two positions: one that begins the subordinate clause and another at its end, suffixed on the verb that is usually positioned at the end of the clause. In this clause, *-ka* is no longer a conjunction as with coordination, because it always depends on another element. The dependent markers *-ka* shows clausal embedding. At the beginning of the subordinate clause, the marker *-ka* forms a complex subordinator with another different element, delimiting the clause and indicating the type of clause: it may introduce a relative clause, a complement clause, or an adverbial clause. These markers are listed in Table 3 and described in Section 4.1.

At the end of the clause, the marker *-ka* on the verb has been referred to as a "subjunctive" mood in the Purepechan tradition because it appears on the verb in the position occupied by the morphemes of this category (Capistrán 2002; Friedrich 1984; Monzón 2004; Wares 1984). However *-ka* on the verb is a morphological constraint for the subordinate clause regardless of the type of clause; therefore it is impossible to use another mood in this type of clause (Chamoreau 2009: 103–105). In this paper, I keep the label "subjunctive" even if it does not correspond to the traditional notion of the subjunctive as contrasted with the assertive mood. This marker *-ka* is described in Section 4.2.

Table 3. Subordinators*

Relative clause		*inka / inki / enka / enki / =nka / =nki***
Complement clause		*iska / iski / eska / eski*
Adverbial clause	Temporal / Condition	*eka / eki*
	Locative / Temporal	*inka / inki / enka / enki*
	Manner	*iska na / iski na / eska na / eski na*
	Reason	*jimpoka / jimpoki*
	Concessive	*nak'iruka / nak'iruki*
	Hypothetical	*peeka / peeki*
	Purpose	*parake / paraki*

* As explained in footnote 1, I use the label "subordinator" to refer to all the markers that occur in the different types of subordinate clause. In Table 3, I present various dialectal forms because they are relevant for diachronic explanations (see Section 4.1).

** The variation between *-ka* and *-ki* is currently present in Purepecha as a dialectal variation (Chamoreau 2007, 2009: 268–273). The variation between *-ki* and *-ka* has not been explored. In the sixteenth century, *-ki* seems to have been the most frequent form attested for the subordinators (Gilberti 1987[1558]: 35), while *ka* was used for the coordinator. The problem is that for the sixteenth century, data from only the Eastern area exist: this is the area in which nowadays *-ki* is used with more frequency to build the subordinator. By contrast, in the Western area *-ka* is generally used to build the subordinator and no data from the sixteenth century are known. In this paper, I only use examples with *-ka*. Otherwise, the vowel variation and change between *e* and *i* is well attested in Purepecha (Chamoreau 2009).

4.1 *-ka* at the beginning of the subordinate clause: Different kinds of subordinator

The subordinator constitutes the first element of the subordinate finite clause. As listed in Table 3, the subordinators possess various forms depending on the kind of subordinate clause they introduce. The marker *-ka* always combines with another element to form the subordinator. These two elements form a complex element analyzed in synchrony as a single morpheme. Despite this relational outcome, the source of the element with which *-ka* forms a subordinator is often transparent.

In relative clauses, the marker may have various forms: *inka* as in (33), *enki, enka, inki, =nka* as in (34), and = *nki*. The latter two forms function as enclitics and may appear after a noun, independent pronoun, deictic pronoun, or demonstrative, as in (34). They elide the first vowel in comparison with the full forms *inka* or *inki*. The source of the form that may be analyzed as *in-ka* is the form *inte-ki* (sometimes transcribed as *jinte-ki*, with initial velar that has been lost) attested in the sixteenth century (see note ** after Table 3 for the variation between *-ki* and *-ka*). The element *inte* is a distal demonstrative pronoun (Gilberti 1987 [1558]: 25), as in *inte werantupinti-cha* in (33). In the subordinator *inka*, the demonstrative pronoun has lost the unaccented segment *te* that appears just before *-ki*.

The use of the demonstrative pronoun as the source for the relative subordinator is cross-linguistically widespread (Heine & Kuteva 2002: 113–115).

(33) *tsimanhi-e-x-ti = ksï = nha inte werantupinti-cha* [***inka*** *= ksï = nha*
two-PRED-AOR-ASS3S = S3PL = EV DEM orphan-PL SUB = S3PL = EV
táte-empa-ni ka ama-mpa-ni no
father-KPOSS3-OBJ and mother-KPOSS3-OBJ NEG
*ka-nko-rhe-nka-ø-**ka***].
have-INTS-body-IT-AOR-SBJV
'They said that there were two orphans **who** did not have a father and mother.'

(34) [*ima = **nka** jini ja-ø-**ka***] *juchiti mimi-i-x-ø-ti.*
DEM = SUB there be.there-AOR-SBJV POSS1SG brother-PRED-AOR-ASS3S
'That one **who** is there is my brother.'

In complement clauses, the marker is *iska*, as in (35), or *iski, eska, eski* (see Table 3 above). The source for *is-* is the adverb of manner, *isï* 'thus'. The current form is a grammaticalization of the marker *isï-ki* attested at the beginning of complement clauses in the sixteenth century (Gilberti 1987 [1558]: 135–136). The two variations between *e/i* and *-ki/-ka* are also present in the various forms (see note ** after Table 3). The evolution from an adverb of manner to a complementizer is another well-known process of grammaticalization (Heine & Kuteva 2002: 258).

(35) *wanta-x-ti = kxï* [***iska** no = kxï ukhuri-echa biajari-ø-ka*].
tell-AOR-ASS3S = S3PL SUB NEG = S3PL opossum-PL travel-AOR-SBJV
'They told **that** the opossums have not traveled.'

In adverbial clauses, different subordinators exist depending on the contexts of use. The temporal subordinator is *eka*, as in (36). This subordinator is also used for condition, as in (37). The use of the same marker for these two contexts has been cross-linguistically described and is due to an absence of distinction of degrees of expectability (Thompson et al. 2007: 257–258). The source for the first part of the marker *e* seems to be the proximal demonstrative pronoun *i* 'this', displaying the same variation between *e/i* as described above in note ** after Table 3. In the sixteenth century, the form was *iki* (Gilberti 1987 [1558]: 35, Lagunas 1983 [1574]: 56).

(36) [***eka** Rosita-ri ama-mpa chem-empa-o nia-nts'a-ni*
SUB Rosita-GEN mother-KPOSS3 house-KPOSS3-RES return-IT-NF
*ja-p-**ka***] *imeri tapichu no ixe-pa-nts'a-s-p-ti.*
be.there-AOR.PST-SBJV POSS3SG uncle NEG see-CENTRIF-IT-AOR-PST-ASS3S
'**When** Rosita's mother had returned to her house, her uncle did not see.'

(37) *wanti-ku-a-ka = kxï tsikata-ni* [***eka** = ri ju-pirin-ka*].
kill-FT-IRR-ASS1/2 S = S1PL chicken-OBJ SUB = S2SG come-COND-SBJV
'We will kill the chicken **if** you would come.'

For the locative adverbial clause, the same form as for the relative marker is used, *inka:* compare the marker *inka* in the relative clause in (33) and the locative adverbial clause in (38). Nevertheless, the source for the locative subordinator is different. For the marker in the relative clause, the source is the demonstrative *inte*. For the locative subordinator, the source is the deictic pronoun *jini*, and the grammaticalization has moved from *jini-ki* attested in the sixteenth century (Medina Plaza 1998 [1575]: 49 [81]), to the marker *inka* (or *inki*), as in (38). In Purepecha, demonstrative pronouns and deictic pronouns are related, as they are built on the basic deictic form *ji* (Chamoreau 2004).

(38) *ima = nha incha-tse-nt'a-x-p-ti* [***inka*** *itsï ja-p-ka*].
DEM = EV enter-SUP-IT-AOR-PST-ASS3S SUB water be.there-AOR.PST-SBJV
'They said that he submerged himself **where** there was water.'

The subordinator for the manner-adverbial clause is built with the juxtaposition of the complementizer *eska* (see above, example 35) and the interrogative pronoun *na* 'how', as in (39):

(39) *Rosa-ita, pawani pawani, jarhintku jawa-ra-sïn-an-ti urhu-ni,*
Rosa-DIM day day early get.up-MID-HAB-PST-ASS3S grind-NF
[***iska na*** *imari tarha-mpa-iri arhi-p-ka*].
SUB how POSS3SG mother.in.law-KPOSS3-GEN tell-AOR.PST-SBJV
'Rosita, all the days, got up early to grind **how** her mother-in-law told her.'

In the adverbial reason clause, the subordinator is *jimpoka*, as in (40): the instrumental postposition *jimpo* and *-ka*.

(40) *no = teru = chk'a anta-nku-x-ti* [***jimpoka*** *= ni yontki = t'u*
NEG = other = certainly gain-INTS-AOR-ASS3S SUB = 1 before = too
a-rha-ni unta-ø-ka].
divide-MID-NF begin-AOR-SBJV
'This is no longer sufficient **because** before I also began to drink.'

Another example is the concessive subordinator, *nak'iruka,* as in (41). This is a complex form built with the morpheme *nak'i* 'which' and the adverbial enclitic for politeness = *aru,* that is *nak'i* = *aru* 'which = politeness'. The element *-ka* is attached to this complex form.

(41) *kw'iripu xarha-narhi-sïn-ti eka ikia-ni ja-ø-k'a,*
People show-PRINC-HAB-ASS3S SUB be.angry-NF be.there-AOR-SBJV
[***nak'iruka*** *no wanta-ni ja-a-ka*].
SUB NEG tell-NF be.there-IRR-SBJV
'People show when they are angry **although** they will not tell.'

The subordinator for the hypothetical adverbial clause *peeka,* as illustrated in (42), is borrowed by Purepecha from the Spanish *puede que* 'may be that'. The behavior of this element correlates with the other types of subordinator in Purepecha, as it is built as a complex form with *-ka*.

(42) *arhi-sïren-ti = ksï* [***peeka*** *Maria arhi-nha-am-ka*].
tell-HAB.PST-ASS3S = S3PL SUB María tell-PAS-HAB.PST-SBJV
'They told **that maybe** she is named María.'

The last subordinator I have found in the data appears in the purpose finite clause with switch reference *paraki*. Compare the example in (43) with the example in (32). In the purpose finite clause, as in (43), the marker is borrowed from Spanish, *para que*, and appears as *paraki* or *parake* in Purepecha (this is a dialectal variation). I have never found the form *paraka*. A possible hypothesis is that this is a recent borrowing and that *ke* (or *ki*) is borrowed from Spanish because the form and the function are similar to the Spanish particle *que* (which may be why *paraka* has not been found). Nevertheless, another possibility is convergence or syncretism between the Spanish *que* and the native Purepecha element *ki* (see note ** after Table 3 above). Convergence or syncretism between the two elements might have been favored because they presented a similar form and functioned in similar contexts. This topic has not yet been studied. The subordinator *paraki* is well integrated in Purepecha; it may thus be considered a complex form (and not two morphemes, as in Spanish) because it is impossible to introduce another morpheme (suffix or enclitic) between *para* and *ki*. As illustrated in (43), the enclitic pronoun always attaches to the end of the subordinator.

(43) *ju-ø* *kokwani* ***paraki*** *= ri xe-a-ka.*
come-IMP.SG quickly SUB = S2SG see-IRR-SBJV
'Come quickly **so that** you see him.'

In summary, at the beginning of a subordinate finite clause the subordinator is always a complex form built with the form *-ka* (or *-ki*) in combination with different kinds of element that enable recognition of the subordinate clause. In this complex form, *-ka* has a dependent form. In synchrony, the combination of the two elements is analyzed as one morpheme. However, on the diachronic level we may propose the hypothesis that the two elements were distinguishable and functioned as two separated morphemes, but that then the two morphemes were reanalyzed as one element in which each lost its autonomy. This process is transparent in the elements *peeka* and *paraki* borrowed from Spanish.

4.2 *-ka* at the end of the subordinate clause: A compulsory suffix on the verb

In subordinate finite clauses, *-ka* modifies the predicate. All the predicates in subordinate finite clauses must have this suffix. Although the suffix is described as a "subjunctive" mood (Capistrán 2002; Friedrich 1984; Monzón 2004; Wares 1984), it does not have the characteristics of the morphemes of this category as it is compulsory and cannot be opposed to some other mood. Moreover, cross-linguistically the subjunctive mood can be opposed to an assertive (or indicative) mood at the semantic level, as in French in (44), in which the opposition between (44a) with the indicative and (44b) with the subjunctive expresses the degree of reality of the event. In (44a) the speaker indicates that he knows that this house (with red walls) exists, but in (44b) the use of the subjunctive mood expresses a doubt about the existence of such a house. In Purepecha, this opposition is impossible because the assertive mood always appears in main and independent clauses, and the subjunctive mood is the only one attested in subordinate finite clauses.

(44) a. *Je cherche une maison qui*
s1SG look.for.PRS.INDC.1SG ART.INDF.FEM.SG house REL
a *des mur-s rouge-s.*
have.PRS.INDC.3SG ART.INDF.PL wall-PL red-PL
'I am looking for a house that **has** red walls.'

b. *Je cherche une maison qui*
s1SG look.for.PRS.INDC.1SG ART.INDF.FEM.SG house REL
ait *des mur-s rouge-s.*
have.PRS.SBJV.3SG ART.INDF.PL wall-PL red-PL
'I am looking for a house that **would have** red walls.'

The label "subjunctive" may be explained by three factors: (1) in subordinate finite clause, the predicate must be marked by this morpheme. So this marker is associated with subordination, as is cross-linguistically true of the subjunctive; (2) this morpheme only appears in subordinate clauses and is the only mood that can be used; and (3) this morpheme is positioned at the end of the verb, in the slot for the mood.

As Purepecha is traditionally a SOV language (see above, Section 2), the verb usually appears at the end of the clause. Nowadays this is not always the case because Purepecha is changing to SVO. Nevertheless, subordinate finite clauses seem to be more conservative and present more SOV order than independent and main clauses (see examples in Section 4.1). The final position of the verb in the clause is significant, because *-ka* appears at the end, to close and delimit this subordinate clause. This type of clause always exhibits this schema: [SUB with *-ka* … V-*ka*]. The

double *-ka* marking delimits the clause, underlining its subordinate and embedding status. The role of the morpheme *-ka* on the verb is to signal the end of the subordinate clause; it has no semantic meaning as with the other moods.

In subordinate finite clauses, tense and aspect markers are present but in a distinctive and reduced form (the irrealis and conditional retain the same form), as presented in Table 4.

Table 4. ATIM markers in independent and subordinate clauses

	Main and independent clauses	**Subordinate finite clauses**
aorist	*-x*	*-ø*
aorist – past	*-x-p*	*-p*
habitual	*-xïn*	
habitual - past	*-xïn-an* / *-xïren*	*-am*
progressive	*-xa*	*V-NF AUX*
progressive – past	*-xa-p*	
continuative	*-xam*	
continuative – past	*-xam-an*	
irrealis	*-a*	
conditional	*-pirin*	

The aorist aspect is codified *-x* in an independent or main clause, and *-ø* in a subordinate clause, as illustrated in (40), while the aorist aspect and past tense are *-x-p* in an independent or main clause, and *-p* in a subordinate clause, as shown in (36).

The habitual aspect marker is *-xïn* in an independent or main clause, as in (41), and retains the same form in a subordinate clause, as in (45). In independent clauses, habitual aspect and past tense markers appear as *-sïnan*, as in (39), or as *-sïren*, as in (42) (a lot of formal variations exist for these morphemes when they appear together: see Friedrich 1975: 184–185). In the subordinate finite clause two forms have been found: *-sïren*, as illustrated in (30), and a reduced form, *-am*, as in (42).

(45) *ari-xïn-ti* [***iska*** *cho-narhi-**xïn**-ka*].
say-HAB-ASS3S SUB be.afraid-PRINC-HAB-ASS1/2 S
'He (always) says **that** I am afraid.'

The progressive aspect marker is *-xa* in an independent and main clause, as in (19). In a subordinate clause, this aspect has an analytic form: it is built with the non-finite verb and the *ja-* 'be there' auxiliary, as in (46). Note that the auxiliary

appears with the aorist aspect. The progressive aspect and past tense present a similar process: the markers for independent and main clauses are *xa-p-*, as in (23), whereas in dependent clauses an analytic construction is attested, as in (47). In this context the *ja-* 'be there' auxiliary is marked by the aorist aspect and the past tense, encoded as *-p-*. In these two contexts, the subjunctive mood presents the form *k'a*.[2]

(46) *kwhiripu xarha-narhi-sin-ti* [*eka* ***ikia-ni*** ***ja-ø-k'a***].
people show-PRINC-HAB-ASS3S SUB be.angry-NF be.there-AOR-SBJV
'People show when they are angry.'

(47) *no = kxï nia-ntsha-x-p-ti* [*jimpoka = kxï* ***t'iré-ni*** ***ja-p-k'a***].
NEG = S3PL come.back-IT-PST-ASS3S SUB = S3PL eat-NF be.there-AOR.PST-SBJV
'They didn't come back because they were eating.'

The irrealis and conditional present a similar form in independent and subordinate clauses: the irrealis is *-a*, as in (37) in an independent clause and (43) in a subordinate clause, and the conditional is *pirin-*, as illustrated in (48) in an independent clause and (37) in a subordinate clause.

(48) *Pacanda anapu-echa pa-**pirin**-ti.*
Pacanda ORIG-PL take-COND-ASS3S
'Those from Pacanda should carry it.'

To sum up: At the end of a subordinate clause *-ka* always appears suffixed to the verb, as a morphological constraint. It appears in the same position as the mood, after the aspect, tense, and irrealis markers; this is why this morpheme is traditionally treated as a subjunctive mood. However, in fact its role is to show the end of the subordinate clause and it has no semantic meaning, unlike the other moods. In subordinate clauses, the aspect and tense markers appear in reduced form. These two characteristics, the marker *-ka* requirement and the reduced aspect and tense markers, show that subordinate clauses have less finiteness than main clauses (Chamoreau 2016).

2. For the continuative aspect and the continuative aspect with the past tense, the process is the same as for the progressive aspect and the progressive aspect with the past tense (see Wares 1974: 96–97 and Friedrich 1984: 73).

5. Conclusion: Two hypotheses about the source and grammaticalization of *ka*

Purepecha distinguishes between coordination (linkage with a free and independent *ka* of two independent clauses), chain-medial clauses (linkage with a free *ka*, which may be repeated, of an independent clause with a dependent clause, the latter depending on the former), and subordination (linkage of an independent clause with a dependent clause, the latter being embedded in the former; the dependent markers *-ka* are attested). As described in this paper, in synchrony three markers with *ka* are attested in Purepecha: the coordinator, the marker that builds the subordinator, and the subjunctive mood. One question remains: Does a relation among these three morphemes exist?

The main, shared function of the first two morphemes is to link units, although each *ka* performs a different type of linkage. Thus to postulate a common source for these linking devices might make sense. Nevertheless, although the diachronic relation between the coordinator *ka* and the subordinator with *-ka* is likely, as the main function of both is linkage, the diachronic relation between these two and the subjunctive *-ka* is not so evident. This morpheme has two faces: it appears in the slot of the mood at the end of the verb, but its presence is compulsory in subordinate clauses and it has no semantic meaning (unlike the other moods). Its function is to signal the end of the embedded and subordinate clause. For this morpheme, I suggest a different route of evolution: the source seems to be the assertive mood marker for the SAP *-ka*. In the route of grammaticalization from assertive to subjunctive, the use of *-ka* extends to a new compulsory context of use, and has been desemanticized (no choice is possible, no modal meaning is possible). These processes have led it to look more like a linking device than a mood, and thus to adopt the main function of delimiting a clause and to play a role (in combination with the subordinator) in the process of linking a subordinate clause with a main clause.

Taking into account such parameters as weight, cohesion, and variability (Lehmann 2002) and analyzing the correlation between them, I advance the hypothesis that there have been two routes of grammaticalization. First, the free coordinator *ka* has given rise to the dependent and compulsory marker *-ka* that appears in the subordinator. The diachronic change illustrates a shift from marking coordination and independence to showing coordination and dependence and ultimately subordination and embedding. This process is depicted by the form of *ka*, which appears twice. I suggest five stages for this route of grammaticalization. Second, another route of grammaticalization leads from the assertive mood marker for SAP *-ka* to the subjunctive mood marker *-ka* (see Table 5).

Stage 1. Coordinator of two words or noun phrases
On the syntagmatic level, the syntactic scope of the coordinator *ka* is to link two functionally equivalent words or phrases. It is a free and independent element. It groups with the second unit it links to, or to the last one when several units are coordinated. Just one occurrence of *ka* suffices to link two or more units. On the paradigmatic level, *ka* presents the semantic feature of conjunctive coordinator, which may be distinguished from other types of coordinator (disjunctive or adversative). It thus belongs to a paradigm of coordinators. The coordinator *ka* may receive enclitics that change its meaning (to build other coordinators). Thus the possibility of a choice of coordinator exists depending on communicative intention.

Stage 2. Coordinator of verb phrases
Using the coordinator to link verb phrases presents the same features as the coordinator of noun phrases (see stage 1, above). However, it makes up the first stage of grammaticalization as verbs constitute another context of use for the coordinator *ka*. In a coordinated verb phrase, phrases have high thematic reference as well as tense, aspect, irrealis, and mood continuities.

Stage 3. Coordinator of finite clauses
The following stage involves the use of *ka* as a clausal coordinator, which often incorporates thematic reference and tense, aspect, irrealis, and mood continuities. In this stage, the coordinator *ka* maintains the features presented above for stage 1. However a change occurs at the syntactic level, as it links clauses and not phrases. Following Heine & Kuteva (2002: 83), coordinators used with NP appear to provide one source for clause-connecting markers ('and').

Stage 4. Coordinator of chain-medial clauses
This is an intermediate stage: On the paradigmatic level, its features are similar to those described for the coordinator in stage 1, but on the syntagmatic level the changes are more significant. Here *ka* remains a free and independent coordinator, but the two coordinated units are not functionally equivalent. Chain-medial clauses depend on the independent clauses, that is, chain-initial clauses, for their ATIM reference. A second syntactic feature is that *ka* is more bound: it often occurs at the head of each coordinated clause (various *ka* are used, not just one with the last coordinated clause). Therefore, the repetition of *ka* thus indicates the dependency of the clauses.

Stage 5. Marker at the beginning of a subordinate clause
The marker *-ka* is a bound element that loses its autonomy. It occupies a fixed slot and is always attached to another lexical or grammatical element to build a

subordinator. It always appears in the same position, at the beginning of a subordinate clause. It thus links two functionally non-equivalent clauses. The presence of this morpheme is compulsory at the beginning of each subordinate clause. On the paradigmatic level, *-ka* displays desemanticization, as it constitutes part of a complex morpheme and does not have a semantic feature on its own (it is combined with different types of element). The consequence is a change of paradigm (or decategorization for Heine & Kuteva 2007: 32–53), in which *-ka* no longer belongs to the paradigm of coordinators but to the paradigm of subordinators. There is no semantic choice: *-ka* is always suffixed to another element, whatever the communicative intention.

The specific route of the subjunctive mood

In a subordinate clause, *-ka* is always a suffix and appears twice (at the beginning, forming the subordinator, and at the end, on the verb), delimiting the embedded subordinate clause. In the second occurrence, *-ka* is suffixed to the verb and modifies it. The presence of the suffix *-ka* is compulsory because of the type of clause; no syntactic or semantic choice is possible, unlike with the other moods. For this morpheme, I suggest a different route of evolution: the source seems to be the assertive mood marker for the SAP *-ka*, because both appear on the predicate in the slot of the mood after aspect and tense markers. It is seen as a member of the paradigm of mood because of its position at the end of the verb, but it has a specific position in this paradigm as it is not possible to contrast it with another mood. In Purepecha, the assertive mood may be contrasted with the interrogative mood, and the imperative mood may be contrasted with the exclamative mood (see Chamoreau 2009: 100–108). In the route of grammaticalization from assertive to subjunctive, the use of *-ka* extends to a new compulsory context of use, and it has been desemanticized (no choice is possible, no modal meaning is possible). These processes have led it to look more like a linking device than a mood and thus to adopt those devices' main function of delimiting a clause and to play a role in the process of linkage (in combination with the subordinator) of a subordinate clause with a main clause. The result is that it is more closely associated with the syntactic configuration of embedding – marking the end of this type of clause and indicating that the clause delimited by the two occurrences of *-ka* is subordinated to a main clause – than with a semantic modal meaning. This is why its position in this paradigm is marginal, as it cannot be contrasted with another mood. Cross-linguistically, according to Bybee et al. (1994: 236), the subjunctive mood marker may have the indicative mood marker as its source. They explain that "subjunctive uses occur near the end

of grammaticalization paths; whether they are from indicatives or from modal elements, their restriction to subordinate clauses comes late in their development. A related point is that they are more semantically reduced".

In Table 5, I present the five stages of grammaticalization from the coordinator to the subordinator (built with *-ka* and another element) in the left-hand column and the stage from the assertive mood marker for the SAP to the subjunctive mood on the right.

Table 5. Stages of grammaticalization of *ka*

Noun phrase coordinator	Assertive mood marker for SAP
> Verb phrase coordinator	
> Finite clause coordinator	
> Chain-medial clause coordinator	
> Subordinator (*-ka* with another element)	> Subjunctive mood marker

The route of grammaticalization from coordination to subordinator has been cross-linguistically demonstrated (see for example Harris & Campbell 1995: 290). This process is not uncommon in several languages. This grammaticalization appears to be part of a more general process whereby markers of phrase coordination change into markers of clause coordination that then give rise to subordination markers (Heine & Kuteva 2002: 43). In Purepecha, the route includes an intermediate stage, coordination of a dependent clause. Interestingly, the last stage of the first route of grammaticalization and the grammaticalization of the subjunctive mood create two dependent markers *-ka* that indicate subordination and embedding. The doubling of *-ka* is a way to encode the high degree of dependence of this type of clause.

Acknowledgments

This research was made possible through financial support from the French Center for American Indigenous Language Studies, CELIA (CNRS-INALCO-IRD), the French Center for the Structure and Dynamics of Languages, SEDYL (CNRS-INALCO-IRD), and the French Center for Mexican and Central American Studies, CEMCA (CNRS-MEAE). This research would not have been possible without the support of Teresa Ascencio Domínguez, Celia Tapia, and all our Purepecha hosts.

Abbreviations

AOR	aorist
APPL	applicative
ART	article
ASS	assertive
CAUS	causative
CENTRIF	centrifugal
CENTRIP	centripetal
COM	comitative
COND	conditional
DEF	definite
DEM	demonstrative
DIM	diminutive
EV	evidential
FOC	focus
FEM	feminine
FT	formative
GEN	genitive
HAB	habitual
IMP	imperative
IND	independent
INDC	indicative
INDF	indefinite
INS	instrumental
INT	interrogative
INTS	intensive
IRR	irrealis
IT	iterative
KPOSS	kinship possessive
LONG.EXT	long exterior area
LOW	lower area
MID	middle
NCS	no coreferential subject
NEG	negation
NF	non-finite
O	object
OBJ	objective case
PAS	passive
PRINC	principal area
PST	past
PL	plural
POSS	possessive
PRED	predicativizer
PROG	progressive
PRS	present
REFL	reflexive
S	subject
SBJV	subjunctive
SG	singular
SUB	subordinator
SUP	superior area
TOP	top area

References

Bril, Isabelle & Rebuschi, Georges. 2006. Coordination, subordination et co-jonction: Faits linguistiques et concepts. *Faits de Langue*s 28: 7–20.

Bybee, Joan, Perkins, Revere & Pagliuca, William. 1994. *The Evolution of Grammar*. Chicago IL: University of Chicago Press.

Capistrán, Alejandra. 2002. Variaciones de orden de constituyentes en p'orhépecha. Topicalización y focalización. In *Del Cora al Maya Yucateco. Estudios lingüísticos sobre algunas lenguas indígenas mexicanas*, Paulette Levy (ed.), 349–402. México: UNAM.

Chamoreau, Claudine. 2004. Localización de espacios utilizando deícticos en purepecha. In *VII Encuentro Internacional de Lingüística en el Noreste*, Isabel Barreras Aguilar & Mirna Castro Llamas (eds), 531–554. Hermosillo: UniSon.

Chamoreau, Claudine. 2007. Grammatical borrowing in Purepecha. In *Grammatical Borrowing in Cross-Linguistic Perspective*, Yaron Matras & Jeanette Sakel (eds), 465–480. Berlin: Mouton de Gruyter.

Chamoreau, Claudine. 2009. *Hablemos purepecha*, *Wanté juchari anapu*. Morelia: Universidad Intercultural Indígenas de Michoacán / IIH-UMSNH / IRD / CCC-IFAL / Grupo Kw'anískuyarhani de Estudiosos del Pueblo Purépecha.

Chamoreau, Claudine. 2012. Dialectology, typology, diachrony and contact linguistics. A multi-layered perspective in Purepecha. *Sprachtypologie und Universalienforschung (STUF – Language Typology and Universals)* 65(1): 6–25.

Chamoreau, Claudine. 2014. Types of enclitic in Purepecha. In *Patterns in Meso-American Morphology*, Jean-Léo Léonard & Alain Kihm (eds), 119–143. Paris: Michel Houdiard Editeur.

Chamoreau, Claudine. 2016. Non-finite chain-medial clauses in the continuum of finiteness in Purepecha. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada Fernández (eds), 83–104. Amsterdam: John Benjamins.

Chamoreau, Claudine. 2017. Purepecha. A polysynthetic but predominantly dependent-marking language. In *Handbook of Polysynthesis*, Michael Fortescue, Marianne Mithun & Nicholas Evans (eds), 667–695. Oxford: OUP.

Cristofaro, Sonia. 2003. *Subordination*. Oxford: OUP.

Foley, William & Van Valin Jr., Robert D. 1984. *Functional Syntax and Universal Grammar*. Cambridge: CUP.

Friedrich, Paul. 1975. *A Phonology of Tarascan*. Chicago IL: University of Chicago Press.

Friedrich, Paul. 1984. From meaning to sound. In *Handbook of Middle American Indians*, Vol. 2, Munro S. Edmonson (ed.), 56–82. Austin TX: University of Texas Press.

Giacalone Ramat, Anna & Mauri, Caterina. 2011. The grammaticalization of coordinating interclausal connective. In *Handbook of Grammaticalization*, Heiko Narrog & Bernd Heine (eds), 656–667. Oxford: OUP.

Gilberti, Maturino. 1987[1558]. *Arte de la lengua de Michuacan*. Morelia: Fimax Publicistas Editores.

Givón, T. 2001. *Syntax*. Amsterdam: John Benjamins.

Harris, Alice & Campbell, Lyle. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511620553

Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In *Converbs in Cross-Linguistic Perspective*, Martin Haspelmath & Ekkehard König (eds), 1–55. Berlin: Mouton de Gruyter.

Haspelmath, Martin. 2007. Coordination. In *Language Typology and Syntactic Description*, Vol.2, Timothy Shopen (ed.), 1–51. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.001

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.

de Lagunas, Juan Baptista. 1983[1574]. *Arte y Diccionario con otras obras en lengua Michuacana*. Morelia: Fimax.

Lehmann, Christian. 2002[1982]. *Thoughts on Grammaticalization*, 2nd rev. edn. Erfurt: Universität Erfurt.

Longacre, Robert. 2007. Sentences as combinations of clauses. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, Timothy Shopen (ed.), 52–150. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.007

Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36(2): 281–331. https://doi.org/10.1515/ling.1998.36.2.281

Matras, Yaron. 2007. The borrowability of structural categories. In *Grammatical Borrowing in Cross-Linguistic Perspective*, Yaron Matras & Jeanette Sakel (eds), 31–73. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110199192.31

de Medina Plaza Juan. 1998[1575]. *Dialogo sobre la naturaleza*. Zamora: El Colegio de Michoacán y Fideicomiso Teixidor.

Monzón, Cristina. 2004. *Los morfemas espaciales del p'urhépecha*. Zamora: El Colegio de Michoacán.

Smith-Stark, Thomas. 1994. Mesoamerican calques. In *Investigaciones lingüísticas en Mesoamérica*, Carolyn Mackay & Verónica Vázquez (eds), 15–50. Mexico: IIF-UNAM.

Thompson, Sandra, Longacre, Robert & Hwang, Shin Ja. 2007. Adverbial clauses. In *Language Typology and Syntactic Description*, Timothy Shopen (ed.), 237–300. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619434.005

Wares, Alan. 1984. *Tarascan Verb Inflection*, 93–100. Mexico: SIL.

Syntactic nominalizations in Pima Bajo

Diachronic diversity

Zarina Estrada-Fernández
University of Sonora

Pima Bajo is a Uto-Aztecan language from northwestern Mexico, traditionally spoken in the central part of the states of Chihuahua and Sonora. It is the most endangered language in the Uto-Aztecan family, a situation partially responsible for the loss of some remarkable features observed in complex clauses, in particular, syntactic nominalizations. This paper focuses on syntactic nominalization constructions in Pima Bajo involving three nominalizing suffixes, *-dam*, *-kɨg*, and *-ka*. These nominalization strategies are consonant with Comrie (2011) and Malchukov (2004), especially in that they have a mixed status, combining nominal as well as verbal properties. Additionally, these constructions should not be considered discrete, but gradual along a continuum. Moreover, the relevant properties that are observed within the different types of nominalizations mixed or not, follow a hierarchical organization, since not all of them are relevant within the different instantiations of this kind of constructions.

Keywords: clausal nominalizations, mixed status, subordination, Pima Bajo

1. Introduction

The notion of nominalization refers to two types of processes, the first one, lexical in nature, and the second one, syntactic. Lexical nominalization implies a derivative process by which elements of a given category, verbal or adjectival, and even nominal, change to function as nominal elements, that is, nouns (Comrie & Thompson 2007). In the examples in (1) the basic (first column) and derivative forms (third column) have been provided; within the derived forms different nominalizing suffixes appear: *-dam* 'agentivizer', *-k(a)* 'stative', *-i* 'nominalizer', *-mag*/*-dag* 'adjectivizer', *-di* ~ *-id* 'possessive', *-ab* 'directional', and -(*t*)*am* 'locative':

https://doi.org/10.1075/tsl.126.07est

(1)	*naat*	'to finish'	*naata-**dam***	'maker, builder'
	mɨlia, *mɨr*(PFV)	'to run'	*mɨr-**dam***	'runner'
	vo'o	'to lean'	*voo-**k***, *vo'o-**ka***	'abdomen, stomach'
			*vo-**i***, *vo'-**i***, *voh-**i***	'road'
	mua'a	'to kill'	*mua'a-**k***	'sharp'
	tu'a	'to throw'	*tu'a-**k***, *tua'-**k-am***[1]	'outside'
			tu-***k-mag***	'darkness'
	ko'a	'to eat'	ko'a-***dag***	'food'
	duda, duud(PFV)	'to rain'	*duu-**k***	'rain'
			*dud-**(d)ag***	'autumn'
	vui	'towards'	*vui-d*[2]*-**(d)am***	'torn'

Syntactic nominalization, in turn, involves action nominal constructions, i.e., a derived noun or verb being headed by a noun (Comrie 2011; Koptjevskaja-Tamm 2013), and clausal nominalizations, where the whole clause shows some properties of nominal elements.[3] In this paper, I will concentrate in clausal nominalizations only, and will restrict the term of syntactic nominalizations to only this kind of constructions.

The aim of this contribution is two-fold. First, this work shows that Pima Bajo, a Uto-Aztecan language of the Tepiman branch, has different ways to encode syntactic nominalization, in particular, those involving the suffixes *-dam*, *-kɨg*, and *-ka*. Second, it demonstrates that in this language syntactic nominalizations correspond to mixed categories (Malchukov 2004), since they show nominal as well as verbal properties. I depart from an assumption, supported in various functional studies, that nominalization should be explained as a continuum, or scale, resulting from the different diachronic stages that contributed to the encoding of such constructions along a period of change. All the nominalizations can be arranged along this continuum depending on their different degrees of nominalization. In this way, constructions with the highest degree of nominalization will show more morpho-syntactic nominal properties and therefore will be located on the most nominalized extreme of the continuum. In turn, those that are found on the opposite end of the scale will display no nominal properties and thus be considered not nominalized at all, and for this reason more clausal-like. Moreover, the morpho-syntactic properties that are relevant to support this continuum may be organized

1. Where *-k* is a stative suffix, and *-am* a locative.

2. Applicative *-id*.

3. See Genetti et al. (2008: 4) for a definition of these two types of nominalizations.

in a hierarchy (Noonan 1985;[4] Moyse-Faurie 2016), which seems to be the case in Pima Bajo as I will discuss at the end of Section 4.

The structure of this contribution is as follows: after Section 1, where a brief introduction to the topic of nominalization is given, Section 2 provides an overview of some basic information concerning the language and its speakers. Section 3 introduces some of the notions that will be useful to set the stage for a discussion of nominalization in Pima Bajo. Section 4 deals with the nominalizing suffixes of Pima Bajo, the contexts where they occur and their functions. Finally, Section 5 offers the conclusions and final remarks.

2. Sociolinguistic information and grammatical properties of Pima Bajo

Pima Bajo (ISO 639–3: pia), Tohono 'O'odham,[5] Northern Tepehuan and Southern Tepehuan, have been classified by Miller (1983: 121) and Dakin (2004) as languages of the Tepiman branch of the Uto-Aztecan or Uto-Nahua family. Within Tepiman, Pima Bajo is the language exhibiting the highest degree of obsolescence. This is caused above all by the scattered arrangements of their settlements, the influence of radio and television, and the old discriminatory behavior towards indigenous people that creates in them a sense of linguistic denial or rejection of their language.

Currently, the Pima Bajo population is composed of approximately 4,000 people, of whom less than 741 are speakers of the language (Hope 2006). Most of the Pima Bajo, particularly, those who still speak the language, live in small villages and settlements dispersed in the lands and canyons of the Western Sierra Madre, in the central area bordering the states of Sonora and Chihuahua. However, it is now possible to find members of this group living in some of the major cities of the Mexican states mentioned above, such as Ciudad Obregón and Hermosillo, in Sonora, or cities of Chihuahua, like Cuauhtémoc, Madera, Casas Grandes, Chihuahua or Ciudad Juárez, places where they go in search of work and a better life. Traditional communities where the Pima Bajo still congregate during the Easter holiday and Saint Francis Day on October 4th are Maycoba, in Sonora, and Yepachi, in Chihuahua.

Pima Bajo has five vowels: /a/, /i/, /ɨ/, /o/, /u/, and their corresponding long vowels; although in the everyday use of the language the long vowel phonemes have lost this distinctive feature. The consonant system is composed of fourteen segments:

---

4. Noonan's (1985: 57) implicational hierarchy is represented as follows, where the left end corresponds to the least nominalized construction and the right end to the most nominalized one:

Subject agreement and Mood > Tense > Aspect > Voice, Valency, Object agreement.

5. This language and Akimel 'O'odham, or Pima, were previously known as Papago.

/p/, /t/, /k/, /ʔ/, /b/, /d/, /g/, /v/, /s/, /h/, /m/, /n/, /l/, /r/, and two glides: /y/ and /w/. Morphologically, it is an agglutinative language with words that tend to have one, two or three morphemes; there are no case markers on nouns; oblique participants are marked by a series of particles that are half grammaticalized as postpositions or oblique case suffixes. The language is head marking, which is observed in noun phrases where the possessed entity is marked by one or two suffixes, the alienable suffix *-ga* or the third person possessive *-di* ~ *-r*. The possessor is usually encoded as a non-subject pronoun prefixed onto the possessed noun. Like Tohono O'odham, it shows a relatively free order of major clause constituents, although the basic unmarked order is APV (SOV). As for the order of other constituents, Pima Bajo can be characterized as consistent with the predictions for SOV languages observed in Greenberg (1963), except for the order of relative clauses, which is not prenominal but follows the head noun.

Although Pima Bajo lacks case marking morphology on nouns, a pair of determiners, unlike nouns, alternate their forms according to the syntactic function, subject or object that they play within the clause. The elements illustrated in (2a) function as determiners, or articles, when they occur to the left of a noun or an adjective plus noun, or as demonstrative pronouns if they appear alone replacing the full nominal phrase. There is also a pair of distal demonstratives 'that', in (2b). The subject and object forms are provided in (2) and an example where the determinant functions as subject and object is shown in (3):

(2) a. ***ɨg*** 'this.SBJ' ***ɨk*** 'this.OBJ'
b. ***id*** 'that.SBJ' ***ik*** 'that.OBJ'

(3) ***ɨg*** *kɨl* ***ɨk*** *gogos* *mua'a*.
DET.SBJ man DET.OBJ dog kill.PROS
'The man will kill the dog.'

Regarding personal pronouns, the language has four sets of forms, three to encode the subject, and only one to encode the other remaining grammatical functions. The independent pronouns usually encode emphatic subjects; the preverbal reduced pronouns, which do not necessarily occur adjacent to the verb, also encode the subject. The dependent subject enclitic pronouns are restricted to the right edge of the connective *ko* within several dependent clause types. Contrastingly, the set of prefixed pronouns encode all grammatical functions except the subject. The full set of pronouns of Pima Bajo is provided in Table 1.

Table 1. Pronouns in Pima Bajo

	Subject pronouns			Non-subject pronouns
	Independent	Reduced	Dependent clitic pronouns	
1SG	aani	aan	=an	in-
2SG	aapi	aap	=ap	am-
3SG	hɨg	hɨg	Ø	a- / Ø-
1PL	aatim	aat	-at	tit-
2PL	aapim	aap	-am	mɨm-
3PL	hɨgam	hɨgam	Ø	a- / Ø-

The alignment system of this language is nominative-accusative, which is illustrated in (4a–b) in the contrast between the nominative, or subject determiner, ***ɨg*** 'DET.SBJ' vs. the accusative, or object determiner, ***ɨk*** 'DET.NSBJ'. Meanwhile (4c–d) show the nominative-accusative contrast of personal pronouns:

(4) a. ***ɨg*** *in-gaag-li-ar.*
DET.SBJ 1SG.NSBJ-search-APPL-TERM
'This (person) searched it for me.'

b. ***aan*** ***ɨk*** *gaag-li-ar.*
1SG.SBJ DET.OBJ search-APPL-TERM
'I searched this for somebody.'

c. ***aap*** *tɨmitim* ***in****-niar.*
2SG.SBJ tortillas 1SG.NSBJ-buy.PFV
'You bought me tortillas.'

d. *takav* ***aan*** ***am****-tɨh.*
yesterday 1SG.SBJ 2SG.NSBJ-throw.PFV
'Yesterday I threw you.'

The most basic or prototypical finite clause in Pima Bajo (Estrada-Fernández 2016) has a verb showing an aspect marker, either 'perfective' which is encoded by truncating the verbal root, as in (5a–b), or imperfective, zero marked, as well as other aspectual suffixes, e.g. *-im* 'continuous', *-va* 'completive', *-tad* 'remote', *-ia* 'probability' or *-hag* 'prospective', among others, as in (5c–d).[6]

(5) a. *ɨg* *gogis* ***muuk****.*
DET.SBJ dog die.PFV
'The dog died.'

6. Estrada-Fernández (2014: 67–75).

b. *kafee mees-tam* ***dah***.
coffee table-LOC be_placed.IMPF
'the (cup of) coffee is on the table.'
c. *Hoan a = ko'i-****im***.
John UNSP.OBJ = eat-CONT
'John is eating something.'
d. *h~~i~~gi ilvah da'i-****va*** *~~i~~k iskon.*
DEM girl pick_up-COMPL DET.OBJ ribbon
'That girl picked up a ribbon.'

A prototypical finite clause has one or two participants encoded as full noun phrases or personal pronouns. A basic clause can also show a locative, comitative, instrument, directional or a beneficiary oblique argument with its postpositional head as either a free or a suffixed element attached to the end of a noun, as in (6a–d).

(6) a. *am-t~~i~~k g~~i~~i bisikle-****tam*** *g~~i~~s = it.*
LOC-DIR fall.PFV bycicle-LOC fall = SS
'Up there (he) fell from the bycicle.'
b. *lii oob ventaan hain hod-****kad***.
DIM person window break.PFV stone-INST
'The boy broke the window with a stone.'
c. *~~i~~g k~~i~~l oidig-****tav*** *hi.*
DET.SG.SBJ man town-DIR go.PFV
'The man went to town.'
d. *o'okosi o'ob-****v~~i~~in*** *duv.*
IT~old-woman people-COM come.PFV
'The old woman came with the people. '

3. Clausal nominalization

Clausal nominalization has been analyzed as the process through which a prototypical verbal clause, either a full sentence – with subject – or verb phrase – without subject – **becomes a noun phrase** (Givón 1990: 498). Other authors, such as Hopper and Thompson (1984: 747) for example, focused on the **ambivalent** nature of nominalizations to affirm that, as verbs, nominalizations can have aspect or mood suffixes and occur with some of their core arguments. Furthermore, nominalizations that are closest to nominal elements may show **genitive or possessive agreement** and **determiners**. Moreover, for the majority of languages, nominalizations are

commonly marked with some special morphology that is not found in prototypical nouns or verbs, morphology that is recognized to be "nominalizing."[7]

Malchukov (2004) characterizes clausal nominalizations as being **mixed categories** since they show nominal and verbal properties in which a particular construction may show certain markers that make it more or less similar to a noun. The author considers that the option [–D/+R], that is, less D(ecategorization) and more R(ecategorization), can be useful to explain the status of these constructions.[8] This representation implies, that depending on the type of the clausal nominalization construction, the verb or head of the construction does not decategorize, that is, it does not change from verb to noun even though it functions in a syntactic construction that shows some morphology usually associated with nominal elements. In other words, nominalizations are less decategorized because the verb does not change to a noun, and more recategorized because the construction of which it forms part has more or less noun-related properties.

Other authors, like Lehmann (1984) and Givón (1990), also consider that clausal nominalizations are gradual or scalar. Both these authors, as well as Bisang (2001, 2016), relate clausal nominalization to the finite or non-finite character of the clause. Lehmann (1984), for example, finds empirical validation for the gradual or scalar status of nominalizations within the rich variety of clausal constructions that function as nominals in different languages. Givón (2001: 24) defines clausal nominalization as the process by which a finite verbal clause – full or without subject – behaves as a noun phrase.[9] The author, returning to the proposal of Hopper and Thompson (1984), suggests that nominalizations should be organized in terms of a scale, according to their morphosyntactic properties, as mentioned in (7) or (8), and in Section 4, above. This continuum places different types of constructions between two extremes, based on their morphosyntactic properties. One extreme, which is the least similar to a clause, corresponds to the highest degree of nominalization, while the opposite extreme, which is the least similar to a noun, corresponds to the lowest degree of nominalization.

Furthermore, Givón (2009) considers that during the process of becoming a nominalized construction, a prototypical finite verbal clause will gradually lose more verbal morphosyntactic properties. The list of properties provided in (7)

---

7. Nominalizations "like V's… may show aspect or mood and take arguments; like N's, they may take possessive markers or determiners; being neither, they also require, in most languages, special morphology which is found neither on prototypical N's or V's-what we call "nominalizing" morphology" (Hopper & Thompson 1984: 747).

8. In this contribution, I will not use Malchukov's schemas.

9. "Nominalization is the process via which a *finite verbal clause* – either a complete clause or a subject-less verb phrase – is converted into a *noun phrase*" (Givón 2001: 24).

illustrates the gradual adjustment of a finite construction into the least finite, i.e. nominalized construction.

(7) **Adjustment from the prototype finite verbal clause to the prototype noun phrase:** (Givón 2009: 67)
   a. The verb becomes a head noun.
   b. The verb acquires *nominal* morphology.
   c. The verb loses tense-aspect-modal marking.
   d. The verb loses pronominal agreement marking.
   e. The subject and/or object assume genitive case-marking.
   f. Determiners may be added.
   g. Adverbs are converted into adjectives.

Other authors, like Bisang (2001), prefer to deal with the properties and processes of clause relations considering the notion of finiteness. In this way, Bisang (2001) mentions that among the main syntactic features that determine the degree of finiteness of a clause are those listed in (8), some of them mentioned by Givón (1990: 852–891). Note that he adds the topic makers as a pragmatic feature.

(8) **Main syntactic features coding finiteness:**
   a. Verbal inflections:
      i. Tense-aspect-mood (TAM)
      ii. Pronominal concordance
      iii. Nominalizing affixes
   b. Nominal inflections:
      i. Subject and object case markers
      ii. Articles, determiners
   c. Pragmatic features: (Bisang 2001: 1401)
      i. Topic markers

Despite the expected cross-linguistic differences, nominalized constructions tend to exhibit the properties listed in (9), although the last one is restricted to object complement clauses (cf. Givón 2009: 68):

(9) a. the subject of the verb is encoded with genitive case,
   b. the verb acquires nominal morphology, and
   c. the clause occurs marked with object case.

However, since nominalizations are not discrete but gradable, they do not necessarily comply with all the features in (9), and for this reason, the richness of nominalized constructions will depend on the different properties that occur within them. The latter point will be considered in our analysis of the different types of clausal nominalizations observed in Pima Bajo.

4. Syntactic nominalizations in Pima Bajo

Syntactic nominalizations in Pima Bajo are encoded by less than a handful of derivational suffixes: (a) the nominalizing suffix *-dam*, (b) the relativizer suffix *-kɨg*, and (c) the nominalizing stative suffix *-ka*. In addition, there are two nominalizing relics which may also emerge in certain constructions. The resulting nominalizations occur in different syntactic contexts. This typological internal diversity is now in decay since only one of them, the suffix *-ka*, is highly productive and frequent in spontaneous speech. In this section, I will discuss syntactic nominalizations obtained by adding each of the aforementioned suffixes, as well as partially nominalized or not nominalized at all, alternative constructions.

4.1 The suffix *-dam*

In this language, one of the most common mechanisms to derive agent lexical nominalizations is the suffix *-dam*, although its use in syntactic nominalized constructions is scarce. The suffix *-dam*, as shown in (10), derives agentive nouns from verbs; in all those elements the basic root is a verb that changes into a nominal element when the suffix *-dam* attaches to it. This data was previously given in (1).

(10) *gaga-**dam*** hunt-NMLZ ‘hunter’
*naata-**dam*** make-NMLZ ‘maker’, ‘builder’
*mɨr-**dam*** run-NMLZ ‘runner’

The suffix *-dam* also attaches to the verb of a complement construction. In (11), the derivative suffix occurs attached to the complement of the verb *tukgi-dam* ‘faint-NMLZ’. In the example, the nominalized verbal complement corresponds to a non-finite construction, since the nominalized verb is unable to bear any TAM (tense-aspect-mood) morphology. The nominalized verb denotes an event that is in the process of being performed (i.e., imperfective or continuous event). Moreover, since the subject of the main verb, *tadag* ‘feel’, is identical to the subject of the verbal complement, there is no need to code a subject on the complement:

(11) *aan* *[tukgi-**dam**]* *tad-ag.*
1SG.SBJ faint-NMLZ feel-PROS
‘I feel like fainting.’/‘I feel fainting.’

The analysis of other types of nominalizations in Pima Bajo allows us to propose that nominalized constructions like the one in (11), *tukgi-dam* ‘fainting’, which functions as the verbal complement of a control verb, *tadag* ‘to feel’, must be considered towards the margin of the most nominalized syntactic constructions since,

within that context, the nominalization displays the following properties: (i) it semantically refers to an event, 'to faint', although in examples in (10) it can also refer to an argument, (ii) its subject keeps its semantic identity with that of the control verb, *tadag* 'feel'; in other words, there is a co-referential relationship (zero anaphora) between the subject of the main clause and that of the verbal complement, (iii) the nominalized verbal complement is embedded within the main construction and occupies the object position between the subject pronoun, *aan* '1SG.SBJ', and the predicate *tadag* 'feel', if we consider that the basic or neutral order of the language is SOV; and finally, (iv) the nominalized verb shows no TAM morphology, but a nominalizing suffix. All these features are consistent with nominalized constructions.

However, nominalized verbal complements, like the one illustrated in (11), are rare in daily speech. Instead, other non-finite, partially subordinated constructions like the one illustrated in (12) are preferred. Note that in (12), the verb has a nominalizing suffix *-a*,[10] but the subordinator *ko* introduces the verbal complement. In this respect, Pima Bajo shows the opposite situation observed by Drude (2011: 187) in Awetí, where nominalization is most frequent than subordination in spontaneous speech and texts.

(12) *aan* *a = dumat-va* ***[ko*** *kav* *tɨs-di-**a**]*.
1SG.SBJ MID = learn-COMPL SUB horse ride-APPL-NMLZ
'I learned to ride a horse.'

Example (12) illustrates a construction with a predicate, the verb *dumat* 'to learn', requiring a verbal complement. However, only one of the properties of the verbal complement in (12) corresponds to a non-finite nominalized clause: (i) the suffix *-a*. The other two morpho-syntactic properties, (ii) the clausal connective, *ko*, which occurs at the beginning of the verbal complement, and (iii) the adjunct position of the verbal complement, since it no longer occurs embedded in the object position, but placed on the rightmost position of the whole construction, are usually found in non-nominalized finite subordinate constructions like the one provided in (13):

(13) *Peier* *mat* ***[k-at*** *kav* ***mua]***.[11]
Pedro know.PFV SUB-1PL.SBJ horse kill.PFV
'Peter knew that we killed the horse.'

10. Other Uto-Aztecan languages from northwestern Mexico such as, for example, Yaqui, Guarijio, and Tarahumara, also have instances of nominalized constructions with the archaic suffix *-a* (cf. Estrada-Fernández & Villalpando, forthcoming).

11. The subordinator *ko* loses its final vowel when the clitic pronoun *at* '1PL.SBJ' attaches to it.

Finite subordinate clausal complements, like the one illustrated in (13), are currently competing with, and replacing fully nominalized constructions such as the one in (11), as well as other partially nominalized constructions, like the one illustrated in (12). The three different constructions provided in (11), (12) and (13), demonstrate that nominalization in Pima Bajo is a gradual phenomenon where different degrees of nominalization, from fully nominalized, to partial nominalized or non-nominalized at all, are observed. An important factor that explains the highest degree of nominalization in constructions like (11) is the semantic nature of the main verb 'feel' in the complement clause, a perception verb. For verbs that are semantically different, as for example, verbs of knowledge, like *maat* 'to know' in (14), the language prefers a non-embedded finite construction in an adjunct position. The set of properties that each construction, from (11) to (14) shows, is also relevant to claim that these properties are crucial for the organization of the constructions within the continuum of nominalization.

(14) *aan* *maat* *[tud-**ag**].*
1SG.SBJ know.IMPF dance-PROS
'I know how to dance.'

However, other alternative constructions involving the same verb are also possible. In Example (15), the complement verb in the nominal complement clause occurs coded as a noun; that is, with a basic or neutral form of a verb (i.e. infinitival), which is recognized by the final vowel *-i*, in (15a), or the non-tensed unmarked form, *o'os* 'to write' (also meaning 'something written'), in (15b). Observe that the main verb has the clitic marker for same-subject, *=ti ~ =it*. This marker can attach to any element of a main or complement clause:

(15) a. *Huaan si'* *maat = it* *[**nɨ'i**].*
Juan INT know = SS sing.NMLZ
'John knows how to sing.' (lit. John knows singing)
b. *aap* *maa = it* *[**o'os**].*
1SG.SBJ know = SS write.NMLZ
'You know how to write.' (lit. You know writ*i*ng)

More nominalized verbal complements like those illustrated in (11), are being replaced by more subordinate-like constructions as those illustrated in (12), (13) and (14), or by half nominalized ones in adjunct postposed position as those in (15). The most relevant features that confirm this change are: (i) the presence of a subordinator *k(o)*, (ii) the use of a 1st or 2nd person clitic dependent pronoun that attaches to the subordinator to mark the dependent subject, particularly when the subjects of the two clauses are non-coreferential, and (iii) the occurrence of the same-subject marker = *ti*, which is not obligatory. The constructions illustrated in

(11) to (15), demonstrate that nominalization in Pima Bajo is a gradual phenomenon. Example (12) must be considered to be in the middle of the continuum, since the verb at the final position shows a nominalizing suffix. Based on these examples it can be argued that a diachronic shift from nominalizations towards finite subordinate constructions is an ongoing process in Pima Bajo. Drude (2011) also argues that in Awetí, a change from nominalization to subordination is observed and that the subordination markers have their source in nominalizing suffixes, which is not the case in Pima Bajo.

Moreover, nominalized constructions by means of the suffix *-dam* also occur in temporal adverbial clauses, and in adversative and concessive constructions, as is shown in (16) and (17). In (16a–b) the temporal clauses are introduced by a conjunction *kuanda* 'when', which is borrowed from Spanish. The verb in the temporal adverbial clauses is nominalized since it has the following properties: (i) the verb is nominalized by means of the suffix *-dam*, and (ii) the agent of the eating is coded by means of the non-subject pronoun *in-* '1SG.NSBJ', which in Pima Bajo is the equivalent form of an accusative or genitive marker. Furthermore, the aspect marker, *-(i)m* 'continuous aspect', in (16b), evidence that the construction is mixed, since fully nominalized verbal elements, like the one in (11), do not occur with aspect markers. The occurrence of the Spanish conjunction *kuanda* 'when' in both adverbial clauses in (16) confirms the mixed nature of such constructions since both examples show a nominalized verb plus a temporal conjunction.

(16) a. *aap am sudag-tam gahi-vuus-an, [kuanda* ***lii****-him-****dam****]. Kova!*
2SG.SBJ LOC water-LOC side-cross-IRR when DIM-go-NMLZ EVI
'Certainly! You can cross the river when it runs low.'

b. *kova-in vuihim-(i)d-a [kuanda* ***in*** *= ko'i-****m****-****dam****]!*
EVI-IMP bother-APPL-PROS when 1SG.NSBJ = eat-CONT-NMLZ
'I hope you do not bother me when I'm eating!'

Furthermore, as I previously mentioned, some adversative constructions, like (17a), or concessive constructions, like (17b), can also be nominalized by the suffix *-dam*. The nominalized constructions correspond to a dependent or non-finite construction (marked inside the brackets in the examples):

(17) a. *[uus am ṫisa-di-a taa-****dam****], aan im apod.*
tree LOC climb-APPL-PROB like-NMLZ 1SG.SBJ NEG can
'I want to climb the tree, but I can't.'

b. *aan am himi-a [ṫimosa duuk-****im****-****dam****].*
1SG.SBJ LOC go-PROB even.though rain-CONT-NMLZ
'I'll go even though it is raining.'

Examples in (17) show the following properties: (i) the presence of the derivative suffix, *-dam*, and (ii) the occurrence of an aspectual suffix, *-im* 'continuous', in (17b), giving evidence of the mixed nature of nominalizations. These nominalizations again show that they must be considered as occupying an intermediate position between partially nominalized and non-nominalized but subordinate clauses. Most frequently, adverbial clauses, like the examples in (18), never have verbs bearing a nominalizing suffix on the verb, such as *duvi-kat* 'come-REM', in (18a), *vɨta-tɨv-an* 'down-hit-IRR', in (18b) or *buahk-an* 'carry.IRR', in (18c):

(18) a. *huaan kos-kat kia'a [aap = koi ab duvi-kat].*
Juan sleep-REM still 2SG.SBJ = LIM DIR come-REM
'Juan will still be sleeping when you arrive.'

b. *koi vɨta-tɨv-an dɨvɨra-tama…*
LIM down-hit-IRR land-LOC
'(Immediately) as soon as he hit the land…'

c. *hɨ'ɨkɨd aan nuukad a'an, [pake in = buahk-an] …*
INTER 1SG.SBJ have.IMPF wings for 1SG.NSBJ = carry-IRR
'That's why I have wings, so they can take me…'

4.2 The nominalizing suffix *-kɨg*

The second nominalizing suffix that I discuss is restricted to relative constructions, as in (19)–(20). In these examples, the verb takes the suffix *-kɨg*, which has its diachronic origin in the combination of a stative suffix *-ka* and an emphatic demonstrative *hɨgai* (also an independent third person pronoun or determiner).[12] Three properties distinguish subject relatives in (19) from object relatives in (20): (a) the ellipsis of the subject argument (zero anaphora) when it keeps a co-referential relation with the head noun of the relative construction, in all the constructions inside the brackets in (19); (b) TAM morphology, as for example, the stem truncation for perfective in (19a), the continuous suffix *-im*, in (19b), or a zero morpheme for imperfective in (19c); and finally (c) the possibility of encoding a core argument required by the semantics of the nominalized verb, the object noun *sudag* 'water' for the verb *nukad* 'to have' in (19c). The intermediate or mixed status of nominalizations in (19) is demonstrated by the occurrence of the properties just mentioned.

12. This is a structural pattern that Pima Bajo has grammaticalized as a result of language contact with Seri and probably with other Hokan (non-Uto-Aztecan) languages (Estrada-Fernández 2012).

(19) Subject relative clauses:
a. *ɨg a'an [gɨi-**kɨg**] vɨg.*
DET.NOM feather fall.PFV-REL red
'The feather that fell is red.'
b. *gogis kɨi kɨl [hink-im-**kɨg**].*
dog bite.PFV man shout-CONT-REL
'The dog bit the man who was shouting.'
c. *in-mak-in ɨk boteii [sudag nukad-**kɨg**]!*
1SG.NSBJ-give-IMP DET.OBJ bottle water have.IMPF-REL
'Give me the bottle that has water!'

However, object relative constructions such as those provided in (20) should be considered further nominalized in comparison with those in (19). The property that motivates this analysis is the obligatory accusative/possessive notional subject, i.e., genitive/possessive agreement, in all the examples in (20), where the subject of the relative construction is marked in Pima Bajo with a non-subject pronoun. This property shows that the construction, ***in**-nɨid-kɨg* 'my seeing' equivalent to 'that I saw', in (20a), or ***am**-niar-kɨg* 'your buying' or 'what you bought', in (20b), have lost all the characteristic properties of a clause, being themselves similar to possessive phrases. Thus, nominalized constructions with a non-subject pronoun encoding the notional subject are nearest to the extreme of the scale corresponding to the highest degree of nominalization. However, (20c) differs from the two previous constructions, (20a–b), in that the verb of the relative construction still retains its object argument, *gogos* 'dog'.

(20) Object relative clauses:
a. *okosi [**in**-nɨid-kɨg] nɨ'i-im.*
woman 1SG.NSBJ-see.IMPF-REL sing-CONT
'The woman who I saw was singing.'
b. *gogos [**am**-niar-kɨg] si' lii.*
dog 2SG.NSBJ-buy.PFV-REL INT small
'The dog that you bought is small.'
c. *ɨg kɨl [gogos in-mua-**kɨg**] vuus.*
DET.SG.SBJ man dog 1SG.NSBJ-kill.PFV-REL go_out.PFV
'The man whose dog I killed went out.'

4.3 The nominalizing suffix *-ka*

The third nominalizing suffix, *-ka*, is highly frequent and multifunctional since it can have adjectival or adverbial functions, among others. As an adjectival nominalizing device, *-ka* occurs in verbs functioning as a predicate of a relative construction,

as in (21), where the adjectival or nominalized relative construction – between brackets – follows the head noun *okosi* 'woman'. This nominalization is not frequent in spontaneous speech in Pima Bajo, and it must be considered historically related to the ancestor of the related relative constructions discussed in Section 4.2.

(21) *okosi [ikis vakin-**ka**] hɨgai si' gɨg.*
woman clothes wash-NMLZ DEM INT big
'The woman (who is) washing clothes is very big.'

The construction in (21) shows most of the prototypical properties of syntactic nominalizations: (a) the verb is modified by a nominalizing suffix whose diachronic origin is related to a posture verb *kaat* 'to be lying', (b) the subject of the nominalized construction, while maintaining its identity with the head noun of the main clause, is elided (encoded as zero anaphora), (c) the deverbal nominalized verb *vakin* 'wash' lacks TAM morphology, and (d) the nominalized verb can retain its capability of having argument requirements. In (21), the nominalized verb retains its object argument, *ikis* 'clothes', and therefore the construction must be considered a mixed category.

Furthermore, in Pima Bajo, the suffix *-ka* may encode stative events in temporal adverbial constructions conveying anteriority, as in (22a), or simultaneity, as in (22b). In these instances, TAM morphology is observed in the verb occurring just before the stative suffix.[13]

(22) a. *aan [duv-ia-**ka**] duuls maa lii oob.*
1SG.SBJ arrive-PROB-STAT candy give.PFV DIM person
'When I arrived, I gave the boy (some) candy.'
b. *lii oob [mɨr-**ka**][14] nɨ'i.*
DIM person run.PFV-STAT sing.IMPF
'The boy ran while (he) was singing.'

Alternatively, a simultaneous temporal clause may lack the stative suffix, as illustrated in (23). In these instances, the simultaneity of the events in both clauses is inferred from the continuous aspect marker *-(')im*, and the occurrence of the identical or same subject clitic marker *=ti ~ =it*:

(23) a. *okosi a-ko'i-'im = **it** hi-'im.*
woman MID-eat-CONT = SS GO-CONT
'The woman is eating while walking.'

---

13. For a better understanding of the adverbial function of the suffix *-ka*, I have decided to adopt the gloss STAT 'stative' rather than labeling it a gerund nominalizer.

14. The imperfective or non-perfective form of the verb 'to run' is *mɨlia*.

b. *kɨl himi-'im = **it** nɨ'i-'im.*
man go-CONT = SS sing-CONT
'The man was walking while singing.'

Moreover, the examples in (24) show the use of the same-subject clitic *=ti/=it* and the suffix *-ka*: The same-subject marker appears either clitiziced into the subject pronoun in (24a), at the finite verb in (24b), or at the nominalized verb in (24c). This demonstrates the clitic nature of this marker. Meanwhile, the stative suffix *-ka*, in (24a) and (24b–c), respectively marks anterior or simultaneous events.

(24) a. *aat = **it** [hugi-**ka**] bo'o.*
1PL.SBJ = SS eat-STAT sleep.PFV
'Once we ate, we went to sleep.'
b. *kɨl [nɨ'ia-**ka**] uus gɨv = **it**.*
man sing-STAT firewood cut.PFV = SS
'Singing, the man cut the firewood.'
c. *okis [ha'a uu-**ka = it**] tutk-im.*
woman pot carry-STAT = SS dance-CONT
'The woman is dancing carrying a pot (on her head).'

Contrastingly, if a construction has different subjects, a subordinator *ko* will head the adverbial temporal clause, as in (25), so this element also implies a different subject:

(25) *aat vaaki-'im [**ko** kɨl am duv].*
1PL.SBJ taking_a_bath-CONT SUB.DS man LOC come.PFV
'We were taking a bath when the man arrived.'

Further evidence showing that adverbial constructions with the stative suffix *-ka* are categorically mixed as those previously shown for complement and relative constructions, are locative adverbial clauses as in (26), where an oblique complement is part of the adverbial construction:

(26) *[mul-tam huaha-**ka**] bɨhk.*
mule-LOC carry-STAT take.PFV
'(They) took (him/her) carrying him on the mule.'

In (26), the locative complement, *mul-tam* 'on the mule', leads itself to ambiguity concerning which of the predicates it corresponds to. However, if it is accepted that nominalizations are mixed constructions, and being *huaha-ka* a dependent and a nominalized verb, we should expect that the peripheral argument would be part of the adverbial construction, which is represented inside the brackets.

Regarding the diachronic origin of the suffix *-ka*, authors like Langacker (1977: 83–84) propose a different historical origin. This author supports his claim based on the comparative analysis of several Uto-Aztecan languages and declares that the suffix *-ka* is a reflex of an old (archaic) accusative case marker *-kV, which used to appear with nominal or adjectival modifiers. According to this author, the suffix currently still survives in several Uto-Aztecan languages, as in Northern Paiute, in (27a), and Tarahumara, in (27b), where apparently it appears as an independent prenominal case particle, respectively, *ka* or *ke*, as in (27b).[15]

(27) a. Northern Paiute (Langacker 1977: 84, ex. (61))
ka *taba*
ACC sun
'the sun.ACC'

b. Tarahumara (Langacker 1977: 84, ex. (62))
ke *pegro ne a-re ripura.*
ACC Peter I give-PST axe
'I gave the axe to Peter.'

The same diachronic origin may be proposed for Yaqui, where it may be argued that a reflex of the old Proto-Uto-Aztecan accusative marker shows up either as a full suffix or as a shortened form that has lost its final vowel. The full accusative form may occur at the determiner *uka* 'DET.SG-ACC' in (28a) and (28b), and a phonologically reduced suffix *-k*, in the accusative adjective *siali-k* 'green-ACC' in (28b).[16]

(28) Yaqui
a. *Ta si u-**ka** aukam-ta bette-si a*
but INT DET.SG-ACC problem-ACC be.difficult-INT 3PL.ACC
machia-k-o…
seem-PFV-COND
'But if (they think) that the problem seems to be very difficult…'

b. *inepo u-**ka** siali-**k** bicha-k.*
1SG.NOM DET.SG-ACC green-ACC see-PFV
'I saw the green one.'

However, Langacker's (1977) analysis about the presence of a case marker suffix in Northern Paiute and Tarahumara, doesn't seem to be the initial source of the

15. Recent research concerning the information structure of Tarahumara makes us think that this particle has grammaticalized to currently mark only the topic (Estrada-Fernández 2018).

16. Data from Yaqui were obtained during my fieldwork on this language.

nominalizing suffix *-ka* that appears in adverbial constructions in Pima Bajo.[17] As I have previously mentioned, I consider that the stative suffix *-ka* has its origin in a posture verb *kaat* 'to be lying'. Evidence from Guarijío (Miller 1996; and Félix 2007), are useful to elucidate between both the stative source for the suffix *-ka* observed in Pima Bajo and the homophonous form that according to Langacker is related to the accusative marked.

In Warihío, also a Uto-Aztecan language but from the Taracahitan branch as well as Yaqui, and Tarahumara, two distinct forms may be related to the suffix *-ka* observed in Pima Bajo. One of them has its origin in a positional verb, *kahtí* 'to be sitting' or 'be_placed' (Miller 1996: 88), which is illustrated in (29a) and its reduced voiced form provided in (29b), where the initial consonant of the verb is voiced into a /g/ when it dissimilates from the last syllable of *mehká* 'far':

(29) a. *Sausé = ga witú tabalási = ci* ***kahtí*** *= ra.*
Saint_José = EMPH down card_board = LOC be_sitting = REP
'Saint José was there sitting down on the table.' (Miller 1996: 98, ex. 1)

b. *Mehká = **ga** = ra waʔá rerú.*
far = be_sitting = REP there down
'(He(she/it) was far away sitting down there.' (Miller 1996: 87, ex. 57)

Observe that (29a) also shows an instance of an emphatic clitic *=ga* at the end of the nominal *Sausé* 'Saint José'. This second element is the one that is related to the accusative marker described by Langacker. Meanwhile, the clitic *=ga* in (29b) is related to the stative verb *kahti* 'to be sitting'.

To sum up, in this section I have shown that Pima Bajo has three nominalizing suffixes to encode syntactic nominalizations: *-dam*, *-kɨg* and *-ka*. The suffixes, with certain overlapping situations, seem to be restricted to different functional domains: verbal complement clauses and temporal adverbial constructions in the case of the *-dam*, relative constructions for *-kɨg* and an old kind of relative constructions and adverbial constructions for *-ka*. According to the predictions that nominalizations are mixed constructions, we have seen that Pima Bajo indeed follows such behavior. For nominalizations encoded by means of the suffix *-dam*, I hypothesized that its limited productivity at the discourse level anticipates a diachronic shift whereby this type of clauses will be replaced by finite constructions.

Recall that in Section 1, I mentioned that clausal or syntactic nominalizations in Pima Bajo seem to follow a hierarchical organization of features. A similar situation

17. I acknowledge this comment to Walter Bisang and Andrej Malchukov (see also Estrada-Fernandez, in press).

is observed by Moyse-Faurie (2016)[18] in some Oceanic languages. The hierarchy that I propose below explains which features will occur in more and less nominalized constructions in Pima Bajo. In this view, not only nominalized constructions should be ordered along a continuum that goes from more to fewer nominalized constructions, but I also claim that the properties that occur in mixed constructions are those that motivate their more or less nominalized character since not all features occur in any type of construction. The ordering of the properties observed in this hierarchy is highly dependent on the properties available for each different nominalized construction. Table 2 shows the organization of features relevant for the nominalized constructions I discussed in this paper. The less nominalized construction will show an overt subordinator and no nominalizing suffix on the verb. The more nominalized construction will have a notional subject encoded as

Table 2. Hierarchical organization of features in Pima Bajo

Properties	*-dam*		*-kig*	*-ka*	
More nominalized	**Verbal complements**	**Adverbial**	**Relative**	**Relative**	**Adverbial**
i. genitive/possessive subject*	–	+	+	–	NR
ii. core or peripheral argument	–	+	+	+	+
iii. embedded position	+	–	+	+	+
iv. nominalizing suffix	+	+	+	+	+
v. subordinator (complementizer), temporal conjunction	–	–	–	–	–
vi. TAM morphology**	–	+	+	–	+
Less nominalized					

* Only when there is a co-referential relation (NR = not relevant in same subject constructions).

** TAM morphology is a prototypical property of finite verbs. Thus, when a clause starts to shift into a nominalized clause, it can immediately substitute TAM morphology for a nominalizing suffix and this change may be accompanied by the presence of a subordinator, as it was shown in (12). TAM morphology may be preserved even in the presence of a subordinator.

18. Moyse-Faurie's (2016: 198) typological implicational hierarchy is the following:

Table 1. Nominal versus verbal features depending on the nominalization strategies

		More verbal features	**More nominal features**
Tense-aspect		±	±
Article	Specific	±	±
	Non-specific	–	+
Affix	Prefix	±	±
	Suffix	–	+
Agent as possessor		–	+

accusative or genitive (non-subject pronoun) and possibly other properties that make those nominalizations be considered mixed categories. It is important to observe that in Table 2 there is no mention of a determiner, a feature that has been considered an important property of nominalizations in other languages (cf. Mithun 2016). The reason for this absence is that in several Uto-Aztecan languages – especially Tarahumara and Pima Bajo – the determiner does not occur at all in nominalized constructions, perhaps because in the common use of the language such an element is highly optional.

5. Final remarks

The analysis presented in this work allows us to conclude that Pima Bajo is a language that has at least three types of nominalized syntactic constructions. The different nominalization strategies observed in this language are consonant with what Comrie (2011), Givón (2001), Lehmann (1984), and especially with what Malchukov (2004) have pointed out. That is, that nominalized constructions combine nominal and verbal properties. Thus, I have shown that in Pima Bajo, nominalizations differ among them depending on the presence or absence of the properties listed in Table 2. The most nominalized constructions are those that have a genitive/possessive/non-subject pronoun encoding the logical subject, but this does not apply to same-subject constructions. In the same vein, the occurrence of an argument required by the nominalized verb seems to be associated with a certain degree of nominalization.[19] Further research needs to be done to clarify this property of nominalizations since the argument depends on the semantics of the verb. Table 2 also shows that in any combination of clauses, if a subordinator appears within the dependent clause, the use of a nominalizing suffix will be avoided, except in (12), where the subordinator *ko* and the nominalizing suffix *-a* illustrate perhaps the most mixed construction. This is what is happening in most combinations of clauses at the discourse level. In Pima Bajo, the scarce use of nominalizing clauses in most instances of spontaneous speech seems to indicate that such constructions are being replaced by alternative, more finite clauses.

19. See Nikitina (2008), which addresses the topic of hybrid syntactic constructions (mixed nominalizations in our own terms) and mentions that "language specific-properties" vary regarding which nominalization may take an object or oblique-peripheral arguments.

Acknowledgements

I would like to thank Bernard Comrie for having supported me to visit over several years the Max Planck Institute for Evolutionary Anthropology, and for his valuable insights concerning this contribution. I also want to acknowledge Pilar Valenzuela, Swintha Danielsen, Ana Fernández Garay, and Claudine Chamoreau for the comments they made to improve this work. Any error is my own responsibility.

Abbreviations

1, 2, 3	first, second, third person
ACC	accusative
APPL	applicative
COM	comitative
COMPL	completive
COND	conditional
CONT	continuous
DEM	demonstrative
DET	determiner
DIM	diminutive
DIR	directional
DS	different subject
EMPH	emphatic
EVI	evidential
IMPF	imperfective
IMP	imperative
INCH	inchoative
INST	instrument
INT	intensive
INTER	interrogative
IRR	irrealis
IT	iterative
LIM	limitative
LOC	locative
MID	middle
NEG	negative
NMLZ	nominalizer
NOM	nominative
NSBJ	non-subject
OBJ	object
PFV	perfective
PL	plural
POSS	possessive
PROB	probability
PROS	prospective
PST	past
RDP	reduplication
REL	relativizer
REM	remote
SBJ	subject
SG	singular
SS	same subject
STAT	stative
SUB	subordinator
TAM	tense-aspect-mood
TERM	terminative
UNSP	unspecified

References

Bisang, Walter. 2001. 100. Finite vs non-finite languages. In *Language Typology and Language Universals. An International Handbook*, Vol. 2, Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds), 1400–1430. Berlin: Walter de Gruyter.

Bisang, Walter. 2016. Finiteness, nominalization, and information structure: Convergence and divergence. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada-Fernández (eds), 13–41. Amsterdam: John Benjamins.

Comrie, Bernard. 2011. Action nominals between verbs and nouns. *Rivista di Linguistica* 23(1): 7–20.

Comrie, Bernard & Thompson, Sandra A. 2007[1985]. Lexical nominalization. In *Language Typology and Syntactic Description*, Vol. 3: *Grammatical Categories and Lexicon*, Timothy Shopen (ed.), 334–381. Cambridge: CUP. https://doi.org/10.1017/CBO9780511618437.006

Dakin, Karen 2004. Prólogo. In *Diccionario yaqui-español y textos. Obra de preservación lingüística*, Zarina Estrada Fernández, Crescencio Buitimea Valenzuela, Adriana E. Gurrola Camacho, María E. Castillo Celaya & Anabela Carlón Flores (eds), 13–25. México: Plaza y Valdez.

Drude, Sebastian. 2011. Nominalization and subordination in Aweti. *Amerindia* 35: 189–218.

Estrada-Fernández, Zarina. 2012. From demonstrative to relative marker to clause linker. Relative clause formation in Pima Bajo. In *Relative Clauses in Languages of the Americas. A Typological Overview* [Typological Studies in Language 102], Bernard Comrie & Zarina Estrada-Fernández, 127–146. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.102.06est

Estrada-Fernández, Zarina. 2014. *Gramática de referencia de pima bajo*, Vol. 1. Hermosillo: Universidad de Sonora.

Estrada-Fernández, Zarina. 2016. Exploring finiteness and non-finiteness in Pima Bajo. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada-Fernández (eds), 43–68. Amsterdam: John Benjamins.

Estrada-Fernández, Zarina. 2018. Estructura de la información y expresión de tópico y foco en lenguas yuto-aztecas de la Sierra Tarahumara. *Revista Línguas Indígenas Americanas (LIAMES)* 18(1): 7–24.

Estrada-Fernández, Zarina. In press. Grammaticalization in Uto-Aztecan languages from northwestern Mexico. In *Areal Patterns of Grammaticalization*, Walter Bisang & Andrej Malchukov (eds). Berlin: De Gruyter.

Estrada-Fernández, Zarina & Villalpando-Quiñónez, Jesús. Forthcoming. Clause combining and nominalization in Tarahumara: Towards a diachronic source of clausal nominalizing suffixes. Submitted to *Amerindia*.

Félix Armendáriz, Rolando. 2007. *A Grammar of River Warihío*. Munich: Lincom.

Genetti, Carol, Bartee, Ellen, Coupe, Alexander R., Hildebrandt, Kristine & Lin, You-Jin. 2008. Syntactic aspects of nominalization in five Tibeto-Burman languages of the Himalayan area. *Linguistics of the Tibeto-Burman Area* 3(1–2): 97–143.

Givón, T. 1990. *Syntax. A Functional-Typological Introduction*, Vol II. Amsterdam: John Benjamins.

Givón, T. 2001. *Syntax. An Introduction*, Vol. 1. Amsterdam: John Benjamins.

Givón, T. 2009. *The Genesis of Syntactic Complexity*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.146

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, Joseph H. Greenberg (ed.), 73–113. Cambridge MA: The MIT Press.

Hope, Margarita. 2006. *Pimas*. México: Comisión Nacional para el Desarrollo de los Pueblos Indígenas. Col. Pueblos Indígenas de México Contemporáneo.

Hopper, Paul J. & Thompson, Sandra A. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60(4): 703–752. https://doi.org/10.1353/lan.1984.0020

Koptjevskaja-Tamm, Maria. 2013. Action nominal constructions. In *The World Atlas of Language Structures Online*, Matthew S. Dryer & Martin Haspelmath (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/62> (10 September 2014).

Langacker, Ronald W. 1977. *Studies in Uto-Aztecan Grammar*, Vol. I: *An Overview of Uto-Aztecan Grammar*. Dallas TX: Summer Institute of Linguistics/University of Texas at Arlington.

Lehmann, Christian. 1984. *Der relativisatz*. Tübingen: Gunter Narr.

Malchukov, Andrej L. 2004. *Nominalization/verbalization: Constraining Typology of Transcategorial Operations*. Munich: Lincom.

Miller, Wick R. 1983. Uto-Aztecan. In *Southwest. Handbook of North American Indians*, Alfonso Ortiz (ed.), 113–124. Washington DC: Smithsonian Institution.

Miller, Wick R. 1996. *Guarijío: Gramática, Textos y Vocabulario*. México: Universidad Nacional Autónoma de México.

Mithun, Marianne. 2016. Shifting finiteness in nominalization: From definitization to refinitization. In *Finiteness and Nominalization* [Typological Studies in Language 113], Claudine Chamoreau & Zarina Estrada-Fernández (eds), 297–322. Amsterdam: John Benjamins.

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.

Nikitina, Tatiana V. 2008. *The Mixing of Syntactic Properties and Language Change*. Stanford CA: Stanford University.

Noonan, Michael. 1985. Complementation. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, Timothy Shopen (ed.), 42–141. Cambridge: CUP.

Syntactic complexity and grammaticalization in Toba language (Guaycuruan)

Cristina Messineo
Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET) / Universidad de Buenos Aires

This chapter analyzes phasal, modal and verbs of motion which, in Toba language (Guaycuruan family), have grammaticalized starting from two types of constructions: completive clauses with phasal and modal verbs, and serial verb constructions (SVCs) with motion verbs. In the domain of phasal and modal completive clauses, the process of grammaticalization involves a change of the grammatical status of the phasal or modal verb. The verb loses the agreement person marker, a common process in other domains of the language. This change constitutes the sole evidence of the co-evolution of form and meaning. In this domain, phasal verbs develop into either aspectual auxiliaries (inceptive or resultative) or as words with a prepositional function. The latter is the only conventionalized strategy that Toba has to express what is encoded through adpositions in other languages. In serial verb constructions (SVCs) with motion verbs, these cover a wide range of meanings of meanings associated with movement (direction, path, location, etc.). In such contexts, one of the verbs, the directional verb is grammaticalized as progressive aspect auxiliary, while the second verb, the locative *weta* does so to a durative aspect auxiliary. Thus, serial verb constructions with motion verbs also explain the origin of directional and locative suffixes from verbs. The synchronic coexistence of serial and complex verb constructions with an identical function favors the argument of different diachronic paths of grammaticalization within languages for which no historical documents are available as Toba.

Keywords: grammaticalization, complex predicates, completive clauses, serial constructions, verbs of motion, Toba, Guaycuruan family

https://doi.org/10.1075/tsl.126.08mes

1. Introduction

The aim of this paper is to analyze the grammaticalization paths that are observed in completive, i.e., verbal complement clauses, with phasal and modal verbs, as well as serial verb constructions (SVCs) with verbs of motion. The analysis of the grammaticalization changes observed in these constructions show that they are among those with a higher degree of semantic integration, since in these constructions one of the verbs develops into an auxiliary, a preposition or an affix.

Diachronic typology, according to Givón's view, is fundamental to show two main grammaticalization routes that have led to clause union in the languages of the world (2009: 93). The first route involves the embedding of a clause into a verb phrase as a verbal complement, while the second corresponds to the condensation of a clause chain into a single serial-verb construction. Both routes of grammaticalization show how this type of constructions, verbal complement clauses with complex predicates and serial-verb clauses, develop starting from two paratactic clauses. Later, these constructions go through an intermediate stage and finally end in a grammaticalized complex predicates.[1]

The aim of this paper is to contribute to recent studies about grammaticalization of complex clauses (Heine & Kuteva 2002; Heine & König 2005; Bisang 2009), by analyzing the phenomena in Toba, as well as to contribute to the understanding of the indigenous languages of the Gran Chaco region.

Messineo (2010) argues that the formal criteria characterizing multi-verb constructions in most of the European languages are not entirely applicable to Toba, for several reasons. First, the syntactic complexity phenomena in this language are not associated with explicit markers of linking or syntactic dependency. Second, given the synchronic coexistence of the grammaticalized forms and the structures that originated them, it is difficult to explain notions as coordination or subordination in absolute or discrete terms. Third, in regard to grammaticalization processes, it is also difficult to find evidence for the co-evolution of form and meaning in the domain of syntactic complexity (Heine & König 2005; Bisang 2009). As a final point, due to the polysemy and polyfunctionality of certain verbs, some verbs appearing within complex predicates are difficult to be interpreted out of context without the encyclopedic knowledge of the world that native speakers have (Heine & König 2005). This situation can be illustrated through the following example, where polyfunctionality of the verb *wotaike* 'to want' varies according to the different contexts. The verb can be interpreted as a modal verb, as in (1a) or as an aspectual marker, in (1b):

1. Estrada-Fernández and Ramos Bierge (2009) offer a detailed analysis of the two routes of grammaticalization proposed by Givón (2009) for the occurrence of auxiliary verbs in Pima Bajo and Northern Tepehuan (Uto-Aztecan languages from northwestern Mexico).

(1) so *lashinek* ***wotaike*** (*da) chek so ladetak so kiʒok.*
DD chuña[2] 3.want DD 3.eat DD food DD tiger[3]
a. 'The chuña wanted to eat the tiger's food.'
b. 'The chuña was about to eat the tiger's food.'

The paper is organized as follows: Section 2 provides general information about the Toba language: genetic affiliation, current habitat of its speakers, dialectal and sociolinguistic situation and main typological features. Section 3 shows a global scenario of completive clauses with phasal and modal verbs and serial verb construction with motion verbs in Toba. Section 4 analyzes the processes of grammaticalization documented in the domain of completive clauses and its development into aspectual auxiliaries or as words with a prepositional function. Section 5 deals with serial verb constructions (SVCs) and their development into aspectual auxiliaries or affixes. Finally, Section 6 includes the conclusions.

2. An overview of Toba language

Toba (*Qom lʔaqtaqa*) belongs to the Guaycuruan linguistic family, together with Pilagá, Mocoví, Kadiwéu, and other extinct languages, such as Abipón and Mbayá. Its dialectal situation is diverse and complex. It could be described in terms of dialectal chains (Kaufman 1990: 70) in which each link of the chain, geographically connected to the next, shows a different degree of intelligibility that gradually diminishes at the edges of the area. In the province of Chaco, Argentina, four major ethno-dialectal areas are identified: *Dapigemlʔek* (Northwest), *NoʔolgaGanaq* (Mid-North), *LʔañaGashek* (Mid-South), and *Takshek* (Southeast).

The Toba ethnic group comprises approximately 70,000 persons. Approximately, 60% of the Toba group can speak and/or understand the language.[4] Toba-speaking groups were traditionally nomadic hunter-gatherers, but at present they are all settled in rural and semi-urban communities occupying a reduced territory compared to pre-columbian times. Currently, cities from Argentina where the Toba are settled are Resistencia, Rosario, Buenos Aires and La Plata.

---

2. A terrestrial bird of the *Cariamidae* family and a character in Chaco mythic narratives.

3. Another mythical character of Toba narratives, also called *jaguar* or *yaguareté* (*FELIDAE, Pantheraonca*).

4. Source: ECPI 2004–2005 (*Encuesta Complementaria de Pueblos Indígenas* [Complementary Indigenous Peoples Survey]), conducted by the INDEC (*Instituto Nacional de Estadísticas y Censos* [National Statistics and Census Institute]).

The main typological characteristics of Toba language are the following: the language is characterized by its morphological complexity of both nouns and verbs, since Toba is an agglutinative, mildly polysynthetic language. It makes a distinction between alienable vs. inalienable possessed nouns, and attributive possession is encoded following a head-marking pattern. Additionally, nouns are inflected for number and gender.

Morphologically, the language shows an active/inactive system, which is supported by three different sets of personal prefixes which alternate their forms according to the agentivity/affectation of the subject participant: (a) active or agentive (controller or initiator of the event expressed by the verb), (b) middle or semi-reflexive (agent affected by the action of the verb) and (c) inactive (patient affected by the action of the verb). Some of this personal prefixes are discontinuous.

Table 1. Sets of personal prefixes

	Active	**Active/affected**	**Inactive**
1 SG	*s-*	*ñ- (<i-n-)*	*ʒ- (<i-d)*
1 PL	*s-…q*	*ñ-…q*	*qad-*
2 SG	*ʔaw-*	*ʔan-*	*ʔad-*
2 PL	*qaw-…i*	*qan-…i*	*qad-…i*
3 SG	*i- / d- / Ø-*	*n-*	*i- / n-*
3 PL	*i- / d- / Ø-…ʔ*	*n-…ʔ*	*i- / n-…ʔ*

Toba verbs do not inflect for tense or mood. On the contrary, the language shows aspectual distinctions: continuous, durative and progressive,[5] and non-continuous. It also has suffixes expressing direction, position, reflexivity and reciprocity. Some of the directional suffixes also function as applicatives. Other features of the language are the absence of adpositions, and the lack of a separate class of adjectives (Messineo 2008).

Typologically unusual but characteristic of the languages of Chaco is the system of six demonstrative determiners which, besides localizing a referent in time and space, encodes deixis (near, distant or absent), and position (standing, sitting, lying down). Some of such deictic classifiers, in particular *da:* 'standing', precede relative and completive clauses and also occur between multi-verbal constructions.

5. Aspect is indicated in the verb by means of the suffixes *-ta* 'durative' and *-tak* 'progressive'. While the former occurs in verbs expressing temporary states (e.g. location), the latter applies to verbs denoting dynamic processes or situations implying the achievement of a final objective (Messineo 2003: 76–83).

Finally, the basic word order of constituents is VO for transitive sentences and VS for intransitives. When O is an independent first or second-person pronoun, it precedes the verb (OV).

The data presented in this contribution are mainly from the Toba dialect *Da pigemlʔek* spoken in the northwestern area of the province of Chaco, Argentina. The material comes from elicitation and from texts representing different genres: advice, prayers, and narratives, among others.

3. Syntactic complexity in Toba

Toba lacks a clear formal strategy to distinguish between different types of complex clauses and multi-verbal constructions. For instance, the phenomenon of syntactic complexity in this language is not necessarily related to the occurrence of explicit linkers or markers of syntactic dependency, such as connectives and subordinators, nor to non-finite verbal morphology, among others. Previous studies have argued for a remarkable tendency toward clausal adjacency of two or more clauses without any conjunction or subordinator (cf. Vidal 2002; Messineo & Porta 2009; Messineo 2010; Censabella & Carpio 2012).

Furthermore, it is important to highlight that verbs in complex clauses are full forms (except for specific cases of nominalization), that is, they have all the argument requirements and markers of a finite verb. Thus, main, embedded and dependent clauses are morphologically similar in terms of verbal inflection, word order of constituents, type of negative marker, etc., which are good evidences to characterize them as paratactic constructions.

Finally, deictic classifiers, especially *da* 'standing', can function as either relativizers or complementizers. Although the presence or absence of the classifier can be optional, the absence of this element reinforces the clausal integration and contributes to the grammaticalization of complex predicates.

3.1 Complement clauses with phasal and modal verbs

This section provides a general characterization of completive clauses in Toba, focusing mainly on phasal and modal verbs that introduce complements. The Toba language presents two types of completive clauses: nominalized clauses and finite verb clauses. In both cases the demonstrative determiner *da* 'standing'[6] introduces

6. For a discussion about the development of the complementizer *da* from the demonstrative, see Vidal (2002).

the complement, even though some utterance predicates can also occur with the classifier *i* 'lying down' (cf. Messineo 2010). With the exception of a few specific cases,[7] the presence of the complementizer is not mandatory. As we will see in the following sections, its presence or absence depends on the degree of syntactic integration in the construction.

Generally, nominalized completive clauses occur with predicates that refer to the phase of an event ('to start', 'to finish', etc.) and when the subjects in the main and completive clause are correferential (SS). The verb in the completive clause has a non-finite form (i.e. nominalized) and carries the possessor marker in agreement with the subject marker of the verb in the main clause:

(2) *Juan i-me* ***da*** *na-paGagen-aGak.*
Juan 3A-finish DD 3POSS-learn-NMLZ
'Juan has finished studying.' (lit. 'He has finished his learning.')

Completive clauses with finite verbs can have a different subject (3) or an identical subject (4). In both cases, they act as an independent sentence, both morphologically and syntactically speaking.

The verb in the main clause and the verb in the completive clause carry the person, number, and directionality marker, among other verbal affixes, and the complement clause presents the same order of constituents of an independent clause. According to the classification of predicates that take complements (Noonan 1985; Cristofaro 2003), all types of predicates can be found in Toba.[8]

(3) *so Juan enapek da hek a-so María.*
DD Juan 3.say DD 3.leave F-DD María
'Juan said that Maria left.'

(4) *ñi maestro y-amaqten da n-vi' nte'eta.*
DD teacher 3A-think DD 3MID-come early_in_the_morning
'The teacher thinks he will come early in the morning.'

Phasal and modal verbs that introduce complements have a morphosyntactic behavior similar to the one found in the rest of the completive clauses in the language:

---

7. These exceptions include descriptive verbs like the verb *ishit* 'power', and adjectives conveying judgement or evaluation (*no'on* 'it is good', *tagiziko* 'it is weird', *yoqta'a* 'it is important', etc.), which obligatory need the deictic classifier *da* to introduce a complement clause (cf. Messineo 2009).

8. These complement-taking predicates are: (i) modals ('can' 'to be able to', 'must'); (ii) phasals ('to start', 'to finish', 'not to stop'), etc.; (iii) manipulatives ('to command', 'to force', 'to persuade'); (iv) desideratives ('to wish', 'to want' etc.); (v) perception ('to see', 'to hear', etc.); (vi) knowledge ('to learn', 'to know', 'to understand', etc.); (vii) propositional attitude ('to think', 'to believe', 'to imagine', etc.); and (viii) utterance ('to say', 'to tell', 'to ask', etc.).

the demonstrative determiner *da* works as a complementizer and both verbs, the verb in the main clause and the verb in the completive clause, have finite forms:

(5) a. *nache cheta'age da n-lo-shigem qaltaq n-qowagaGagen.*
COOR 3.start DD 3MID-look-DIR COOR 3-pray
'Then, he started to look up and pray.'

b. *qomiʔ se-mata-q da sa-chige-naq na lapat.*
1PL.PRO 1A.finish-PLS DD 1A-eat-PLS DD meat
'We finished eating meat.'

c. *sa-wotaike da ana-ʔaqtaGan-eʔ a-ñi ad-teʔe.*
1A-want DD 2MID-talk-COM F-DD 2POSS-mother
'I want you to talk to your mother.'

d. *se-shit-aq da qami se-tawan-a-ʔ-q.*
1A-be_able-PLS DD 2PL.PRO 1A-help-PLS-PLO-PLS
'We can help you.'

However, complement clauses with phasal and modal verbs present two different structures in Toba: (a) less integrated clauses and/or (b) more integrated constructions. This phenomenon can be analyzed taking Givón's typological perspective into account (2009: 83–84) which states that complementation is one of the diachronic precursors of clause union and, thus, it is the most favorable scenario for the grammaticalization of one of the two predicates in the anteceding clauses. The author proposes a correlation between the semantic integration of events and the functional and structural features of clause union, illustrated by the following features in English: expression of the co-referent argument (zero vs. present); grammatical relations (single sets vs. two distinct sets); adjacency of the two verbs (co-lexicalization vs. separation); finite verb morphology (presence or absence on the complement verb); adjacency of the two clauses (presence or absence of complementizer); intonation contours (joint vs. separate).

On the other hand, Cristofaro (2003: 118–120) states that complement clauses with phasal and modal predicates show a higher level of semantic integration in comparison with constructions with other complement-taking predicates, in particular those introduced by utterance, manipulative, perception and knowledge predicates. She argues that, from a conceptual point of view, phasal and modal predicates do not encode separate events, but they are not a portion or a part of the event either. The phasal verbs 'to end' and 'to start' function as aspectual operators on the verb in the completive clause while the modal predicates 'to be able' and 'to want' make up modal filters of the state of affairs to which they apply.

In Toba, completive phasal and modal clauses constitute semantic-syntactic configurations favorable to clause union and at the same time they contribute to

the grammaticalization of the main verb. The phenomenon of clause union in Toba involves the interaction of the following properties:

1. Expression of correferentiality of the Subject (SS) in both verbs
2. Adjacency of the two verbs through the elision of the complementizer
3. Negation operator covering the entire construction and not just a single verb
4. The multiple lexical predicates fall under joint intonation contour.

To illustrate this phenomenon, we are showing different examples of the modal verb *wotaike* 'to want, to desire' when it takes a complement. The examples are shown starting with clauses with a lesser degree of semantic integration moving toward clauses with a higher degree of semantic integration. In Section 3.2 we analyze the grammaticalization process of this verb and of other phasal and modal verbs in contexts of maximal clause union.

In (6), the verbs in the construction have different subjects (DS), each of them expressed by a specific person marker (1st and 2nd). At the same time, each of the verbs can be negated independently (7a and 7b):

(6) *sa-wotaike da ana-ʔaqtaGan-eʔ a-ñi ya-teʔe.*
1A-want DD 2MID-talk-COM F-DD 1POSS-mother
'I want that you talk to my mother.'

(7) a. *sa sa-wotaike da ana-ʔaqtaGan-eʔ a-ñi ya-teʔe.*
NEG 1A-want DD 2MID-talk-COM F-DD 1POSS-mother
'I do not want that you talk to my mother.'

b. *sa-wotaike da sa ana-ʔaqtaGan-eʔ a-ñi ya-teʔe.*
1A-want DD NEG 2MID-talk-COM F-DD 1POSS-mother
'I wish you did not talk to my mother.'

Constructions illustrated in (8) show identical (same) subject (Property 1). Different from what we observe in (7) they are associated with only one assertion, as shown by the fact that when the negation operator comes before the modal verb, it affects the whole of the construction and not just one verb (8b) (Property 3). On the contrary, (8c) is ungrammatical:

(8) a. *sa-wotaike da ñ-aʔaqtaGaneʔ a-ñi y-ateʔe.*
1A-want DD 1MID-talk F-DD 1POSS-mother
'I want to talk to my mother.'

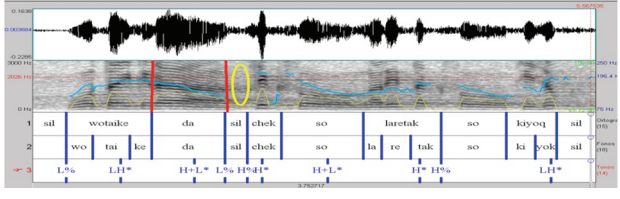
b. *sa sa-wotaike da ñ-aʔaqtaGaneʔ a-ñi y-ateʔe.*
NEG 1A-want DD 1MID-talk F-DD 1POSS-mother
'I do not want to talk to my mother.'

c. **sa-wotaike da sa ñ-aʔaqtaGaneʔ a-ñi y-ateʔe*

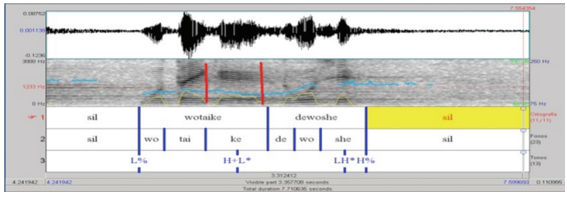
Lastly, the presence or absence of the demonstrative (Property 2) is directly associated with the prosody (Property 4). In terms of the intonational contour, when the demonstrative occurs before a complement clause, the intonational analysis shows two separate tonal phrases: the one that contains the matrix clause and the demonstrative and the one that contains the complement. On the contrary, if the demonstrative is absent, the whole construction is integrated into a single intonational contour (Property 4):

The following graphics show the acoustic behavior of two types of completive clauses with and without a demonstrative (Messineo & Renato 2013):

(9) so *lashinek* ***wotaike da chek*** *so ladetak so kiyoq.*
DD chuña 3.want DD 3.eat DD food DD tiger
'The chuña **wanted** to eat the tiger's food.'



(10) *a-da alo' n-pelek-te-ge-ñi da lapat* ***wotaike de-woshe.***
F-DD woman 3MID-cut-DUR-DIR-DIR DD meat want 3-cook
'The woman is cutting the meat that she is about to cook.'



In (9), we can see a descending tone (HL*) in the demonstrative *da* followed by a prominent lengthening of the syllables (in the figure, we can see it between two red bars, followed by a silence, which is marked by the yellow circle). At the same time, the fundamental frequency (FO) marks two separate prosodic phrases: *wotaike da* and *chek so ladetak so kiyoq*, in which the demonstrative becomes part of the modal verb by way of the intonation.[9]

However, in (10), we can see a larger prosodic integration between the modal verb and its complement (*wotaike dewoshe*), represented by one single tonal phrase marked by two silences.

9. For a more in-depth analysis of the prosodic dimension of complements in the Toba language, see Messineo & Renato (2013).

Examples such as (10), where we can see the highest syntactic integration, constitute scenarios favorable to the grammaticalization of phasal and modal verbs. The characteristics of this phenomenon will be discussed in the following section.

3.2 Grammaticalization of phasal and modal verbs

The grammaticalization of the phasal and modal verbs that take complements presents in Toba the following features. First, they involve a change in the grammatical status of the phasal and modal verb, which is desemanticized, i.e. it loses its lexical meaning and evolves into a grammatical element, among them: aspectual auxiliary, prepositional word and negation. Second, in the formal level, the more grammaticalized phasal and modal verbs appear in a fixed form, identical to the third person singular form, regardless of the person and number of the subject.[10] Therefore, this form has completely lost its deictic and referential value, which indicates that it is grammatically distinct from finite verbs in the language.[11]

The following section presents phasal (3.2.1) and modal (3.2.2) verbs sensitive to grammaticalization processes in Toba. The synchronic co-existence of the lexical source and the more grammaticalized forms reveal the different phases or moments of the processes taking place.

3.2.1 *Phasal verbs*

Toba has two intransitive verbs of movement and path (*chigoqchigiña* 'to come from, appear' and *ivitta* 'to arrive at') that can occur either as finite verbs or as grammaticalized verbs serving as phasal auxiliaries or prepositional words. The following examples show the lexical and grammaticalized forms of both verbs coexisting synchronically in the language. In (11) and (12), the lexical verbs *chigoqchigiña* 'to

---

10. The use of fixed forms of the third person also occurs in other domains such as the attributive and possessive compounds. The first two take the third person verbal marker (*i-*/*n-* o Ø-), while in the compounds, the head presents the morphological marker of the third person possessive (*l*(V)-) (Cúneo 2013: 237).

11. The same phenomenon has been documented for other languages, such as Maybrat (Papuan) (see footnote 13) and Pilagá (Guaycuruan). An interesting feature of Pilagá is that a similar phenomenon occurs with the verb 'to want', 'to desire', but the fixed form is the one of the first person (Vidal 2002: 169–170):

(i) *hadam'e se-take da' ek.*
DEM.F 1A-want COMP 3-go
'She wants to leave.'

(ii) *da'm'e se-take y-e'et di' l-ona-naGak.*
DEM 1A-want 3A-do DD 3POSS-work-NMLZ
'He wants to do his job.'

come from, appear' and *ivitta* 'to arrive at' occur independently with the subject marker expressed by the personal prefix:

(11) ***sa-chigoqchigiña*** *neʔena Chaco.*
1A-come_from DD Chaco
'I come from there, from Chaco.'

(12) ***s-ivitta-Ga*** *na qad-maʔ so Chaco.*
1A-arrive_at-PLS DD 1PL.POSS-home DD Chaco
'We arrive at our home in Chaco.'

On the contrary, in (13) and (14) the verbs do not receive the person marker that appears in the second predicate, which suggests that they have been grammaticalized as aspectual auxiliary forms (inceptive and resultative, respectively) without inflection of the subject argument:[12]

(13) ***chigogoqchigiña*** *da* ***ñi-yen*** *da se-la-ʔa da ya-qaya.*
AUX:$_{\text{INCEP}}$ DD 1MID-cry DD 1A-see-DIR DD 1POSS-brother
'I started to cry when I saw my brother.'

(14) *ivitta da* ***qa-y-ʔot*** *da l-eʔenaGat ʒiʔiʒi alwa TaGaki …*
AUX:$_{\text{RES}}$ DD IMP-3A-give DD 3POSS-call DEM land Olla…
'They finally came to call that land Olla …'

Finally, as shown in (15) and (16), the verbs *chigoqchigiña* and *ivitta* may also precede a nominal phrase, in which case they behave as prepositions:[13]

(15) ***chigogoqchigiña*** *so viʔi 2007 da* ***ñi-yom.***
PREP:since DD year 2007 DD 1MID-drink
'I have been drinking since 2007.'

(16) *neʔena qa-i-sok na shiyaGawa-pi* ***chigoqchiguiña*** *ze ʔoonolek*
DEM IMP-3A-vaccinate DD man-PL PREP:from DD one
ivitta *ʒe 60 viʔi-yi deʔeda l-llaGa.*
PREP:to DD 60 year-PL DEM 3POSS-age
'Here individuals from the age of 1 to 60 are vaccinated.'

---

12. According to the typology of auxiliary verb constructions among the languages of the world (Anderson 2006), the resulting construction corresponds to the LEX-headed type, which consists of an uninflected or fixed form of an auxiliary verb and a lexical verb with all the inflectional morphology.

13. Dol (1999: 87–88) has made a similar observation. In Maybrat (Papuan), there is a word *ae* 'at' which is morphologically and syntactically a verb but can serve as either a main verb or a preposition. When *ae* is used prepositionally, it always occurs with a third person singular feminine subject prefix, regardless of the person, number, and gender of the subject, indicating that it is grammatically distinct from normal verbs.

3.2.2 *Modal verbs*

Modal verbs like *ishit* 'to be able' and *wotaike* 'to want' also undergo different processes of grammaticalization. When used as a lexical verb, *ishit* 'to be able' implies 'ability.' In this case, it always occurs with subject inflection, and the presence of the demonstrative *da* with a complementizer function is obligatory:

(17) *seʔeso shiyaGawa* ***i-shit*** *a* ***n-ʔemge.***
DEM man 3A-be_able DD 3MID-walk.fast
'That man can walk fast.'

(18) ***se-shit-aq*** *da qamiʔ se-tawan-a-ʔ-q.*
1A-be.able-PLS DD 1PL.PRO 1A-help-PLS-PLO-PLS
'We can help you.'

The lexical source, however, coexists synchronically in the language with epistemic modal uses, as can be seen in (19), where the verb *ishit* can appear either inflected or taking the fixed form of the third person:

(19) *chaʔaʒe da ñ-omgi na lataGa ayem* ***se-shit/ishit***
CON DD 1MID-drink DD alcoholic beverage 1PRO 1A-be_able/be_able
da se-wet na i-qaik.
DD 1A-ache DD 1POSS-head
'Because if I drink alcoholic beverage, it is possible that I get a headache.'

Finally, there is a deontic modal use of the verb *ishit* when it is combined with the negative operator *sa*. The modal verb loses the person marker by being fixed in the third person and is grammaticalized in the prohibitive:

(20) ***sa-ishit*** *da ʔa-dowa a-ka shipkai ʔalo.*
NEG-be_able DD 2A-marry F-DD indecent woman
'You must not marry an indecent woman.'

The verb *wotaike* behaves in a similar way. In canonical completive clauses with the lexical meaning of 'to want' or 'to desire', it has person inflection:

(21) *da woʔoka* ***ʔa-wotaike*** *da* ***ʔan-aʔaqtaGan-eʔ*** *nache nache-ka ʔonolek.*
DD INDF.PRO 2A-want DD 2MID-talk-COM CON CON-DD one
'If you want to talk with a man, then only do it (talk) with one.'

(22) *wataGanaq mashi qa-n-qoʔona ʒi Espinillo* ***qa-y-wotayke*** *da*
soldier immediately IMP-3MID-arrive DD Espinillo IMP-3A-want DD
qa-y-oda-wek *ltaGa ʒi qom.*
IMP-3A-drive-DIR:out again DD people.Toba
'The soldiers who arrived in Espinillo immediately wanted to drive out the Toba people again.' (Lit. 'They had the intention of driving out the people.')

On the other hand, the use of the auxiliary as a marker of ingressive aspect or imminent future is observed in complex constructions in which the modal verb is not inflected to indicate person. Like in the rest of the grammaticalized forms, the verb *wotaike* is fixed in the third person. As a formal correlate of the clausal integration, the two (modal and main) verbs appear next to one another without the demonstrative *da* intervening between them:

(23) *ayem* ***wotaike sa-chek*** *ʒi* *nawok lapat.*
1PRO want 1A-eat DD grilled beef
'I am about to eat grilled beef.' (Lit. 'want, I eat')

(24) so *tegesan enapek iʒik* *ʔa-lew-tapigiʔ* *aahhh …* *mashi*
DD crow 3.say my_friend 2I-die-PROG aahh immediately
wotaike ʒi-lew.
want 1I-die
'The crow said: My friend! Are you dying? Aahhh … I'm about to die…'
(Lit. 'want, I die')

5. Serial verb constructions (SVCs) with motion verbs in Toba

As a general feature in Toba, an SVC is a sequence of two or three verbs that behave as one. That is to say SVCs describe a single event, share aspect and polarity and exhibit grammatical subject agreement. None of the verbs in the series is subordinated to or works as a complement of another one. On the contrary, they are all finite verbs that are juxtaposed with no explicit subordination or coordination marker. In all cases, the verbs in the series are lexically autonomous, which means they can make sentences on their own in non-serial constructions (Messineo & Cúneo 2010). Serial constructions with motion verbs mainly involve directional (also known as 'deictic'; cf. Givón 1991: 139), path and locative verbs, and the order in the sequence is iconic with the temporal or sequential order of the movement. Directional or deictic verbs are intransitive, and they specify the direction of the movement related to a starting point. As such, they occupy the first slot in the sequence:

(25) *ʔam* ***ʔad-yiʔigelaq***[14] *ʔa-(a)y-ge* *ʒi* *Pampa Aguará.*
2PRO 2I-come_back 2A-go-DIR:towards DD Pampa Aguará
'You came back to Pampa Aguará.' (Lit. 'You came back, you went to Pampa Aguará.')

14. Other directional verbs documented in SVCs of movement are: *hek* '3.leave', *anak* '3.come', *ñigelaq* '3.go back there', *ʒigelaq* '3.come back here', among others.

Path verbs indicate the direction of the movement in relation to a reference point.[15] When they do not combine with locative verbs, they occupy the last position in the serial sequence, following directional verbs, and they introduce an obligatory locative nominal phrase, which semantically works as the reference point of the movement (see (26)). Last, locative verbs encode the final point of a change or of a movement process undergone by the figure, and, therefore, when they do not occur in grammaticalized constructions they occupy the last position in the serial construction:[16]

(26) *koʔollaGa neʔena layipi koleetegaʔ* ***weta-lek*** *a-na alwa.*
in_the_past DEM people 3.wander 3.be-DIR:on F-DD earth
'In the past, those people wandered around the Earth.' (Lit. 'they wandered, they were on the Earth'.)

Verbs of transport and manner can also participate in movement SVCs. The former lexicalize sub-events which refer to the way in which the movement is realized:

(27) *mashe* ***hek petañi t-ay-aGasom*** *ʒe lacheoge.*
already 3.go 3.walk 3A-go-DIR:towards_the_water DD river
'She/he's already left walking towards the river.' (Lit. 'He/she goes, walks, goes towards the water.')

Verbs of transport are transitive verbs whose object is the moving entity. They can combine with another verb of transport, in which case the last verb of the sequence must carry a directional suffix showing the final point of the movement:

(28) ***n-ache*** *a-so l-llikta* ***no-da-Gasom***
3MID-bring_carrying F-DD 3POSS-canoe 3 M-bring-DIR:towards_the_water
a-ʒi tala.
F-DD river
'He/She brings his/her boat to the river.' (Lit. 'She brings it by carrying it, she brings it to the river.')

Other characteristic features of SVCs in Toba are:

---

15. One of the most utilized path verbs in SVCs is *ta(y)/ta(h)/-y-* 'go' and its derivated forms *tayge*'3.go.towards', *taya* '3.go.there.close', *tashigem*'3.go.up', *tayaGasom* '3.go.towards.the.river', etc.

16. This type includes the following verb forms: *weta-* 'be', *paʔa-* 'be' y *neta-* 'inhabit' and their derivates by means of directional suffixes: *wetalek* 'be on', *wetangi* 'be inside', *paʔashigem* 'be above', *paʔagiñi* 'be below', *netalek* 'inhabit', etc.

(i) They present multiple lexical predicates that are integrated in a single intonation contour. This prosodic feature allows an SVC (29) to be distinguished from an asyndetic coordination (30):

(29) *ashek* ***sh-iʔako*** ***s-ay-a*** *ʒi tala.*
1A.leave 1A-go_fishing 1A-go-DIR:towards DD river
'I am going to the river to fish.' (Lit. 'I'm going, I'm fishing, I'm going towards.')

(30) *so nataGala* ***t-aya-oga*** *ʒiʔʒi noʔonaGa* ***t-ay-lek*** *ʒi-m*
DD chief 3A-go-DIR:outside DEM field 3A-go-LOC:over DD-FOC
le-ʔenaGat da nede.
3POSS-name DD Nede
'The chief went to that field; he passed through (that place) called Nede.'

The same prosodic guideline applies to discourse, where the verbs of an SVC are grouped in a single discursive unit – the line[17] – while the coordinate clauses are grouped in various lines. The following example is an excerpt from a story that contains a relative clause (lines a and b), two SVCs (one on line b and the other on line e) and two nominal coordinate phrases (lines c and d):

(31) a. *woʔo so-m shiyaGawa ipiʔaGayk*
EX DD-FOC man hunter
'Once upon a time there was a hunter'
b. *da* ***hek*** ***neʔ-epe*** ***weta-gi*** *a-ñi aviaq ltadaik*
DD 3MID-leave 3MID.go_hunting 3.be-LOC F-DD big forest
'who would go hunting in the big forest'
c. *nache somaʒi ya-komna-lo so-wa l-ʔogoGonaqte* ***cheqnek***
CON 3PRO 3A-take-PLO DD-PAU 3POSS-weapon bow
'so he took his weapons, his bow'
d. *qataq* ***so npon tadegek***
CON DD club jacarandá
'and his jacaranda club'
e. *nachi* ***hek*** ***t-ay-aGaama*** *a-so tala*
CON 3.leave 3MID-go-DIR:towards_the_water F-DD river
'and went out towards the river'
f. *qaq koʔollaGa i-vida-Gaama a-so tala*
CON in_the_past 3-reach-DIR:towards_the_water F-DD river
'and when he reached the river'

17. The rhetorical structure of Toba discourse is based on the line (Messineo 2004). Each line is an intonation unit, which comprises of one or more phonological phrases but is limited by pauses longer than those used for such phrases. It corresponds to a full unit of content. According to Chafe (1994), speakers regulate the flow of information so as to essentially introduce only one new idea (an event, a state or a referent) per prosodic unit.

g. *na-GayaGañi i-wataʔa da woʔo ka lʔaGayaGak …*
3MID-stop 3A-wait DD EX DD sign
'he stopped and waited for a sign…'

(ii) They lack a sole locus for the finite morphology. The co-referentiality of the argument S is expressed through explicit concordance markers on each of the verbs of the series (see examples (32) and (33)):

(32) *ayem* ***ashek*** ***so-ʔota-shigem*** ***s-ay-ge*** *deʔeda*
1PRO 1A.leave 1A-be_in_position-DIR:up 1A-go-DIR:towards DD
Las Palmas.
Las Palmas
'I am riding to Las Palmas.' (Lit. 'I'm leaving, I'm on a horse, I'm going to Las Palmas.')

(iii) The verbs in the series share subject and object:

(33) ***ñ-ache*** *a-na i-llikta* ***ñ-oda-Gasom*** *a-ʒi tala.*
1MID-bring_carrying F-DD 1POSS-chalana 1MID-bring-DIR:water F-DD river
'I am carrying my *chalana* [boat] to the river.' (Lit. 'I'm carrying my *chalana*, I'm taking it to the river.')

(iv) Unlike other typical or compact constructions (for example, those described by Pawley (2009: 136) in Kalam), the verbs of the series can be non-continuous. The following elements can interrupt the sequence of the serialized verbs:

If V_1 is intransitive, the subject can be placed between V_1 and V_2 following the VS order of the language:

(34) *nache ltaq* ***ni-ʔigelaq*** *so qom* ***t-ay-ge*** *da Espinillo.*
CON again 3MID-return DD Toba_people 3A-go-DIR:towards DD Espinillo
'And once again the Toba people returned to Espinillo.' (Lit. 'they returned, they went to'.)

In spontaneous discourse, if the subject is a nominal phrase, it can appear in a discontinuous form between two verbs in an SVC, as observed in (35). This phenomenon is unquestionably related to the flow of information and to the tension between relevant information that is to be placed at the beginning of the line and the order of constituents for intransitive clauses, in which the verb is placed at the beginning (VS):

(35) *qalota so* ***hek*** *qom t-ay-ge da shew.*
QUA DD 3A.leave Toba_people 3A-go-DIR:towards DD north
'Many Qom people went to the north.' (Lit. Many of them, they left, Qom, they went north.')

Thus, a nominal object phrase can follow the first verb of a directional SVC:

(36) *qa-y-ache* *seʔeso* *lo-ʔokyaGak* ***qa-y-do-oga***
IMP-3A-take_carrying DEM 3POSS-body IMP-3A-carry-DIR:out
ʒi *napaʔalpi.*
DD graveyard
'They took his/her body to the graveyard.' (Lit. 'They carried his/her body, they took it out to the graveyard.')

An adverb of time can follow the first verb in the series:

(37) *nache* ***hek*** *koʔollaGa* ***t-ay-ge*** *da* *Pampa Aguará.*
CON 3.leave in_the_past 3A-go-DIR:towards DD Pampa Aguará
'And he went to Pampa Aguará.' (Lit. 'He left, he went to Pampa Aguará.')

A locative adjunct can also intervene between two verbs when the first of the two is positional:

(38) *ashek* ***so-ʔota-shigem*** *na* *kaayo* ***s-ay-ge***
1A.leave 1A-be_in_position-DIR:on.top DD horse 1A-go-DIR:towards
deʔeda *Las Palmas.*
DEM Las Palmas
'I am riding to Las Palmas.' (Lit. 'I'm leaving, I'm on a horse, I'm going to LP.')

(v) Aspect and negation are indicated on the first verb of the construction, as observed in (39) and (39a). In contrast, they cannot be marked separately on each of the verbs (39b) is ungrammatical:

(39) *ayem* ***ñi-ʔep-tak*** ***so-ʔotai-gi*** *a-na* *aviaq.*
1PRO 1MID-go_hunting-PROG 1A-be-DIR:inside F-DD forest
'I am hunting in the forest.'
a. *saishit* *da* ***ñe-ʔep-tak*** ***s-ache*** *na* *yo-ʔoGonaGat.*
NEG DD 1MID-go_hunting-PROG 1A-carry DD 1POSS-weapon
'I am not hunting with my weapon.' (Lit. 'I'm not hunting, I'm not carrying my weapon.')
b. **saishit* *da* ***ñeʔep-tak*** *saishit* *da* ***sache*** *na* *yoʔoGonaGat*

(vi) The scope of negation and temporal operators applies to the entire construction. In an SVC such as (40), the negative operator (*saishit da*) affects all three verbs in the series, which proves that this is a clause with multiple lexical predicates that encode a single event, a single affirmation. In contrast, the constituents of an asyndetic coordination can be negated alone, as illustrated in (41):

(40) *so ipiaGyak saishit da* ***hek*** ***petañi t-ay-aGaama*** *so tala.*
DD *hunter* NEG DD *3.leave 3.walk 3-go-*DIR*:towards_the_water* DD river
'The hunter did not go walking to the river.'

(41) *so ipiaGaik* ***i-vida-Gaama*** *so tala* ***na-GaỹaGañi***
DD hunter 3A-reach-DIR:towards_the_water DD river 3MID-stop
saishit da ***i-wataʔa*** *da la-ʔaGayaGak.*
NEG DD 3A-wait DD 3POSS-sign
'The hunter got to the river, stopped (but) did not wait for the sign.'

In semantic terms, an SVC constitutes a macro-event.[18] Therefore, the influence of the temporal operators on the entire construction and not on each of the sub-events proves that this is a single event. In (42a), the scope of the adverb *koʔollaGa* applies to all the verbs in the series, while (42b), though grammatically correct, is not an acceptable sentence in Toba. The sub-events *hek* 'leave', *tayge* 'go towards' and *ikaalek* 'follow' are not semantically "individuated", but presented as parts of an event unanalyzed in terms of criteria of duration and location in time, precisely in the sense that the "leaving" sub-event is not accessible to operators of temporal position or duration at the exclusion of the 'going towards' or 'following', and vice-versa:

(42) a. *nache* ***hek***$_{1}$ *koʔollaGa*$_{1, 2, 3}$ ***t-ay-ge***$_{2}$ *da*
CON 3A.leave in_the_past 3A-go-DIR:towards DD
Pampa Aguará ***i-kaa-lek***$_{3}$ *so misionero.*
Pampa Aguará 3A-follow-DIR DD missionary
'Afterwards$_{1, 2, 3}$ he went to Pampa Aguará following the missionary.'
(Lit. 'He left$_{1}$, he went to$_{2}$Pampa Aguará, he followed$_{3}$ the missionary.')

b. **nache qoɔoʔoGoñi*$_{1}$ ***hek***$_{1}$ ***t-ay-ge***$_{2}$ *ka avit*$_{2}$ *da*
CON morning 3.leave 3A-go-DIR:towards DD afternoon DD
Pampa Aguará ***i-kaa-lek***$_{3}$ *so misionero ka pe*$_{3}$
Pampa Aguará 3-follow-DIR DD missionary DD night
*'Then he left$_{1}$ in the morning$_{1}$ and went to$_{2}$ Pampa Aguará in the afternoon$_{2}$ and followed$_{3}$ the missionary at night$_{3}$.'

18. Authors such as Bohnemeyer et al. (2007: 497) consider that the formal features (i.e syntactic or prosodic) of linguistic complexity are specific to each language and are thus not comparable as measures of segmentation. Therefore, their basic premise is the notion of the macro-event, which consists of dealing with a part of the continuum of experience and perception as if it were an entity. The macro-event has a temporal limitation; it presents a simple set of semantic arguments and consists of parts that can sustain different relationships with the whole. These parts, referred to as sub-events, are the minimal entities that can appear lexicalized in a language but which do not constitute events on their own.

(vii) In some cases, the last verb of the series acquires a grammatical function, since it allows the argument structure to be expanded. Note that (43b), (44b), (45b) and (46b) are ungrammatical, since the incorporation of an argument (patient, locative, instrumental or benefactive) requires a serial verb construction when the first verb is intransitive:[19]

(43) a. *ayem koʔlloGa* ***sh-aʔako*** ***se-laike*** *qalota niyaq.*
1PRO in_the_past 1A-go_fishing 1A-search QUA shad
'I fished a lot of shad.' (Lit. 'I fished, I searched for a lot of shad.')
b. **ayem koʔlloGa sh-aʔako qalota niyaq*

(44) a. *ʔam* ***ʔa-lache*** ***ʔa-(a)y-ge*** *da Espinillo.*
2PRO 2A.move 2A-go_to-DIR:to DD Espinillo
'You moved to Espinillo.' (Lit. 'You moved, you went to E.')
b. **ʔam* ***ʔalache*** *da Espinillo*

(45) a. ***ñe-ʔepe*** ***sa-che*** *na yo-ʔoGonaGat.*
1MID-go_hunting 1A-take_carrying DD 1POSS-weapon
'I hunt with my weapon.' (Lit. 'I hunt, I carry my weapon.')
b. **ñiʔepe na yoʔoGonaGat*

(46) a. ***ne-ʔepe*** ***ne-ʔep-lek*** *na l-laqpi.*
3MID-go_hunting 3MID-go_hunting-APPL DD 3POSS-family
'He hunts for his family.' (Lit. 'He hunts, he hunts for his family.')
b. **neʔepe na l-laqpi*

5.1 Grammaticalization of motion verbs

The grammaticalization of motion verbs in contexts of verb serialization involves a change in the semantic and grammatical status of the first verb in the series (DIR or LOC), which acquires functions of the continuous aspect (progressive and durative). Unlike canonical SVCs, here the aspectual verb occupies the V_1 slot in the sequence following the AUX-V order. In this context, the verb that expresses lexical meaning has to take a progressive or durative aspect marker. The verbs of the series reveal subject co-referentiality, retaining the person concordance marker of the canonical SVCs.

Some of the directional and locative verbs that are grammaticalized in contexts of verbal serialization are *hek* 'to go, to leave', *anak* 'to come', *tayge* 'to go towards'

19. The verbs *yaʔako* 'to fish' and *niʔepe* 'to hunt' are intransitive. As such, they require a serial verb construction (cf. (43) and (45)) or an applicative suffix (cf. (46)) to incorporate an object argument. The closest English translation would be 'I go fishing' or 'I go hunting.'

and *weta* 'to be in'. Each of these verbs may function as a verb with lexical autonomy, as one occupying the first slot in a serial verb construction, or as an auxiliary.

The following examples show the verb *hek* (and its irregular forms for second person singular and plural) functioning as an autonomous verb (47), taking in a movement serial verb construction (48) and acting as an auxiliary verb which provides the grammatical content (progressive aspect) of the construction (49):

Lexical verb: 'TO LEAVE'

(47) *mashi* ***hek*** *so Juan.*
recently 3A.leave DD Juan
'Juan left recently.'

Serial verb: 'TO GO/TO LEAVE'

(48) *ayem* ***ashek*** ***s-ay-ge*** *deʔeda Las Palmas.*
1PRO 1A.leave 1A-go-DIR:towards DEM Las Palmas
'I leave to Las Palmas.' (Lit. 'I leave, I go to Las Palmas.')

Aspectual verb: PROGRESSIVE

(49) *saishit da* ***qo-ke-tak*** ***an-aʔaqtaGa-teg-eʔ.***
NEG DD 2A-leave-PROG 2MID-talk-PROG-COM
'Don't keep talking.' (Lit. (Don't) you are leaving, you are talking.)

It should be observed that, in the latter case (49), apart from keeping co-referential person inflection with the main verb, the verbs carry a progressive aspect suffix, which is typical of lexical verbs. Under Anderson's typology (2006: 145), the example in (50) is a construction with auxiliary verb in which both the lexical and auxiliary verb present double subject and TAM inflection.[20]

A similar process affects the verb *anak* 'to come', which, coming from a serial verb construction as the one in (51), is also grammaticalized as a double subject inflection and progressive aspect (52):

Lexical verb: 'TO COME'

(50) *nache woʔo so naʔaq koʔollaGa nache* ***anak*** *a-so waqajñi.*
CON EX DD day in_the_past CON 3.come F-DD star
'One day there came a star.'

20. According to Anderson (2006: 145), serial verbs are one of the sources for this type of constructions. The difference between this and a typical serial verb construction like (41) is that the verb *hek* 'to leave' loses its lexical meaning and acquires the grammatical meaning of progressive aspect.

Serial verb: 'TO COME'

(51) ***ñ-anak*** ***sa-chigaqaʔ-a*** *ñi aviaq.*
1MID-come 1A-come_from-LOC DD forest
'I come from the forest.' (Lit. 'I come, I come from')

Aspectual verb: PROGRESSIVE (FROM THE PAST)

(52) *nache* ***ñ-anak-to-ʔot*** ***se-vigaGaa-tak*** *seʔeso-wa viʔi-ʒi*
CON 1MID-come-PROG-DIR:below 1A-play-PROG DEM-PAU time-PL
ivittaʔa da ayem yi.
until DD 1PRO adult
'I used to play (an instrument) in those years until I grew up.' (Lit. 'I am coming from below, I am playing.')

The locative verb *weta-*[21] 'to be' may occur in the language as a lexical verb (53), as a part of a SVC of motion – in which it canonically occupies the last slot of the sequence – (54) or as an auxiliary verb with the grammatical meaning of a durative aspect (55). In this grammaticalization process, apart from losing its semantic locative weight and acquiring grammatical meaning (durative aspect), the verb occurs in initial position of the sequence, in agreement with the canonical order AUX-V:

Lexical verb: 'TO BE'

(53) *so waGayaGa* ***weto-ʔot*** *a-da epaq.*
DD fox 3.be-DIR:under F-DD tree
'The fox was under the tree.'

Serial verb: 'TO BE'

(54) *neʔena qom* ***t-ay-ge*** ***weta-ʔa.***[22]
DEM Toba_people 3A-go-DIR:towards 3A.be-LOC:there
'Those (Toba) people went to that place.' (Lit. 'they go to, they are there'.)

Aspectual verb: DURATIVE

(55) *woʔoka* ***weta-gi*** ***ne-ʔep-tak.***
INDF.PRO 3.be-LOC:inside 3MID-go_hunting-PROG
'(They) were hunting.' (Lit. 'They are inside (the forest), they are hunting.')

---

21. The verb *weta* consists of a basis formed by the root *we* followed by the durative aspect suffix *-ta*, to which one or more directional suffixes are added.

22. As locative verbs, they occupy the last position in the serial construction, which is iconic with the temporal order of the movement.

Finally, directional serial verb constructions represent the ideal environment for directional and locative verbs to develop into verbal affixes. This behavior of the directional verbs that present SVCs as a source of directional suffixes has it also been documented in languages of Africa (Heine & Kuteva 2002), Australia (Dixon 2006) and Mesoamerica (Zavala 2006). In Toba, the directional verb *tay-* 'to go' and the locative verb *weta-* 'to be at/in' are potentially the origin of directional suffixes when used as part of an SVC:

Table 2. Grammaticalization of locative and directional verbs into verbal suffixes

weta-wek	'to be out'	>	*-wek*	'outwards'
weta-shigem	'to be up'	>	*-shigem*	'upwards'
weta-ñi	'to be down'	>	*-ñi*	'downwards'
tay-ge	'to go towards'	>	*- ge*	'towards'
taya-oga	'to go out'	>	*-aoga*	'out, outside'
ta-ngi	'to go in'	>	*-ngi*	'on the inside'
tay-aGama	'to go towards the water'	>	*-aGaama*	'towards the water'

On the other hand, serialization and suffixation can synchronically coexist in Toba, fulfilling the same function. Note that (56) and (57) are consecutive lines of a single story. In (56), the two directional phases are encoded through an SVC: direction (V_1) and arrival to destination (V_2). In contrast, in (57) the same movement is expressed through a complex verb (V-DIR), and the suffix *-aGaama* 'towards the water' marks the end of the journey:

(56) *nache* ***hek*** ***tay-aGaama*** *a-so tala*
CON 3.leave 3MID-go-DIR:towards_the_water F-DD river
'and (the hunter) left for the river'

(57) *qaq koʔollaGa i-vida-**Gaama*** *a-so tala …*
CON in_the_past 3A-reach-DIR:towards_the_water F-DD river
'and when he reached (the river)…'

6. Conclusions

This work has examined two favorable contexts for the formation of complex grammaticalized predicates in Toba language: completive clauses with modal and phrasal verbs and serial verbs constructions with motion verbs. Both types of construction show two different grammaticalization pathways along which certain lexical verbs (phasal, modal and motion verbs) become more grammaticalized elements. The grammaticalization of phasal verbs is observed when these verbs change into either aspectual auxiliaries – inceptive or resultative – or prepositional elements. The

latter is the only conventionalized method that Toba has to express what in other languages is encoded by means of different adpositions. As regards to the modal verb *ishit* 'to be able', it becomes an epistemic modal auxiliary with the meaning of 'ability'. Similarly, when combined with the negative operator *sa*, the modal verb changes into a specialized deontic element that currently constitutes the way of expressing the prohibitive negation. In the case of *wotaike* 'to want', 'to desire', this verb develops into an ingressive aspect or into an immediate future auxiliary. The main argument to argue for the grammaticalization of phasal and modal verbs comes from the coexistence of inflected and non-inflected forms of the verbs. A similar situation is observed with the verb *ishit* 'to be able', which currently have two coexisting forms in the language (the inflected and the non-inflected forms).

In the case of serial verb constructions (SVCs) with motion verbs, recall that these verbs cover a wide and varied set of meanings associated with movement: direction, path, location, etc, where each of these meanings are encoded as different verbs and yield a serial verb construction. (SVCs). In these constructions, the directional verbs are grammaticalized as progressive aspect auxiliaries, while the locative verb *weta* 'to be at/in' does so as a durative aspect auxiliary. The grammaticalization of a verb denoting locative meaning into a verbal auxiliary is the result of a metaphorical transfer from the spatial to the temporal domain (Heine & Kuteva 2002: 3 and 157–159). Although this verb fulfills a grammatical function, i.e. aspect, it preserves the formal features of a lexical category and retains the agreement morphology of its source within a SVC. Finally, serial verb constructions with motion verbs also constitute the origin of directional and locative suffixes from a verbal source. The synchronic co-existence of forms that have a lexical meaning and a grammatical meaning allows the different diachronic phases of grammaticalization to be identified when dealing with languages for which no historical documents are available.

This work has shown that syntactic complexity in Toba reveals different stages of grammaticalization. Although, in some cases, the lexical sources have turned into suffixes, it is generally the case that the grammaticalization of complex clauses and predicates retains, quoting Heine & König (2005), an "essentially lexical outfit". It is possible to assert – as argued within the framework of the theory of grammaticalization – that not all the functional categories are developed or grammaticalized in affixes (Bisang 2009), nor do they reach a final point of the process as derivative or inflectional affixes (Heine & König 2005). It is also evident that certain lexical categories that appear simple on the surface can cover an ample and varied range of meanings and grammatical functions (Bisang 2009).

To a certain extent, the difficulty of describing the phenomena of syntactic complexity in absolute or discreet terms depends on this panorama. Canonical notions of coordination, subordination and serialization are not entirely satisfactory for

describing processes in transit, or for explaining incomplete or early phases of grammaticalization within the domain of syntactic complexity (Heine & König 2005).

Acknowledgements

The research was carried out thanks to the financial support of CONICET (National Council of Scientific and Technical Research of Argentina) and the University of Buenos Aires (UBACyT Project 2016–2019).

Abbreviations

1	first person
2	second person
3	third person
A	active participant or agent
APPL	applicative
AUX	auxiliary
COM	comitative
COMP	complementizer
CON	discourse connector
COOR	coordinate conjunction
DD	demonstrative determiner
DEM	demonstrative
DIR	directional suffix
DUR	durative
EX	existential
F	feminine
FOC	focus marker
I	inactive participant
IMP	impersonal
INCEP	inceptive aspect
INDF	indefinite
LOC	locative
MID	middle or semi-reflexive participant
NEG	negation
NMLZ	nominalizer
PAU	paucal
PL	plural
PLO	plural object
PLS	plural subject
POSS	possessive marker
PREP	preposition
PRO	independent pronoun
PROG	progressive aspect
QUA	quantifier
RES	resultative aspect
SG	singular
V	verb

References

Anderson, Gregory D. S. 2006. *Auxiliary Verb Constructions* [Oxford Studies in Typology and Linguistic Theory]. Oxford: OUP. https://doi.org/10.1093/acprof:oso/9780199280315.001.0001

Bisang, Walter. 2009. On the evolution of complexity: sometimes less is more in East and mainland Southeast Asia. In *Language complexity as an evolving variable*, Geoffrey Sampson, David Gil, and Peter Trudgill (eds), 34–49. Oxford: Oxford University Press.

Bohnemeyer, Jürgen, Enfield, Nicholas, Essegbey, James, Ibarretxe-Antuñano, Iraide, Kita, Sotaro, Lüpke, Friederike & Ameka, Felix. 2007. Principles of event segmentation in language: the case of motion events. *Language* 83(3): 495–532. https://doi.org/10.1353/lan.2007.0116

Censabella, Marisa & Carpio, Belén. 2012. Clauses as noun modifiers in toba (Guaycurú). In *Relative Clauses in Languages of the Americas: A Typological Overview* [Typological Studies in Language 102], Bernard Comrie & Zarina Estrada-Fernández (eds), 173–190. Amsterdam: John Benjamins.

Chafe, William. 1994. *Discourse, Consciousness and Time. The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago IL: University of Chicago Press.

Cristofaro, Sonia. 2003. *Subordination* [Oxford Studies in Typology and Linguistic Theory]. Oxford: OUP.

Cúneo, Paola. 2013. *Formación de palabras y clasificación nominal en el léxico etnobiológico en toba (guaycurú)* [Lincom Studies in Native American Linguistics 68]. Munich: Lincom.

Dixon, Robert W. M. 2006. Serial verb constructions: Conspectus and coda. In *Serial Verb Constructions: Cross-linguistic Typology*, Alexandra Aikhenvald & Robert M. W. Dixon (eds), 338–350. Oxford: OUP.

Dol, Philomena. 1999. *A Grammar of Maybrat: A Language of the Bird's Head, Irian Jaya, Indonesia*. Leiden: University of Leiden.

Estrada-Fernández, Zarina & Ramos Bierge, Stefanie. 2009. Tipos de cláusulas completivas en pima bajo y tepehuano del norte: Gramaticalización de verbos auxiliares. Paper presented during the *Conference on Indigenous Languages of Latin America IV*, Austin (Texas), 29-31 October.

Givón, T. 1991. Some substantive issues concerning verb serialisation: Grammatical vs cognitive packaging'. In *Serial Verbs: Grammatical, Comparative and Cognitive Approaches* [Studies in the Sciences of Language Series 8], Claire Lefèbvre (ed.), 137–184. Amsterdam: John Benjamins. https://doi.org/10.1075/ssls.8.06giv

Givón, T. 2009. *The Genesis of Syntactic Complexity*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.146

Heine, Bernd & Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511613463

Heine, Bernd & König, Christa. 2005. Grammatical hybrids: Between serialization, compounding and derivation in !Xun (North Khoisan). In *Morphology and its Demarcations*, Wolfgang U. Dressler, Dieter Kastovsky, Oskar E. Pfeiffer & Franz Rainer (eds), 81–96. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.264.06hei

Kaufman, Terence. 1990. Language History in South America. What we know and how to know more. In *Studies in Lowland South American Languages* [Current Issues in Linguistic Theory 264], Doris Payne (ed.), 13–73. Austin TX: Universidad of Texas Press.

Messineo, Cristina. 2003. *Lengua toba (guaycurú). Aspectos gramaticales y discursivos* [Lincom Studies in Native American Linguistics 48]. Munich: Lincom.

Messineo, Cristina. 2004. Toba discourse as verbal art. *Anthropological Linguistics* 46(4): 216–238.

Messineo, Cristina. 2008. Entre nombres y verbos. Categorización de los conceptos de propiedad en toba. In *Estudios en lingüística y antropología. Homenaje a Ana Gerzenstein*, Cristina Messineo, Marisa Malvestitti & Roberto Bein (eds), 157–166. Buenos Aires: Editorial de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires.

Messineo, Cristina. 2010. Coordinación, complementación y serialización en toba (guaycurú). ¿Qué marca la diferencia? In *Actas del II Encuentro de Lenguas Indígenas Americanas* y *II Simposio Internacional de Lingüística Amerindia* (ALFAL) *17 al 19 de septiembre de 2009*, Marisa Censabella & Raúl González (comps). Resistencia: Chaco.

Messineo, Cristina & Cúneo, Paola. 2010. Construcciones seriales en toba (guaycurú). *Amerindia. Revue d'Ethnolinguistique Amérindienne* 33–34: 217–248.

Messineo, Cristina & Porta, Andrés. 2009. Cláusulas relativas en toba (guaycurú). *International Journal of American Linguistics* 75(1): 49–68. https://doi.org/10.1086/598206

Messineo, Cristina & Renato, Alejandro. 2013. La dimensión prosódica de la complejidad sintáctica. Un examen de la integración de predicados de complemento en toba. Paper presented at the *III Encuentro de Lenguas Indígenas Americanas* (ELIA III), Bariloche (Argentina), 15–17 May.

Noonan, Michael. 1985. Complementation. In *Language Typology and Syntactic Description*, Vol. 2: *Complex Constructions*, Timothy Shopen (ed.), 42–140. Cambridge: CUP.

Pawley, Andrew. 2009. On the origins of serial verb constructions in Kalam. In *Syntactic Complexity: Diachrony, acquisition, neuro-cognition, evolution* [Typological Studies in Language Companion Series 85], T. Givón & Masayoshi Shibatani (eds), 119–144. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.85.05ont

Vidal, Alejandra. 2002. Oraciones complejas en pilagá (guaycurú). In *Temas de lingüística aborigen*, II, Ana Fernández Garay & Lucía Golluscio (eds), 161–185. Buenos Aires: Editorial de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires.

Zavala, Roberto. 2006. Serial verbs in Olutec (Mixean). In *Serial Verb Construction*, Alexandra Aikhenvald & Robert M. W. Dixon (eds), 273–300. Oxford: OUP.

From discourse to syntax

The use of the discourse marker *bwe* in the creation of interclausal connectives in Yaqui

Albert Álvarez González
Universidad de Sonora

This paper shows how Yaqui has recruited the element *bwe,* a discourse connective of discontinuity (that is, a spoken discourse marker that introduces a topic-shift), in order to participate in the creation of two new interclausal connectives: the cause/reason adverbial connective *bwe'ituk* and the adversative connective *bweta.* I argue that these two interclausal connectives come from a recent formation process that combines the discourse marker *bwe* and linguistic elements associated with the strategies used in the past for marking cause/reason clauses and adversative clauses in Yaqui, and that the recruitment of *bwe* from discourse to syntax is functionally motivated, since *bwe* can be viewed as a thematic reorientation device used in discourse, and *bwe'ituk* and *bweta* can be viewed as thematic reorientation devices used in syntax. A comparison between the two evolutionary processes originating the two *bwe-* interclausal connectives is also provided, showing that contrary to the *bweta* formation, the creation of *bwe'ituk* implies a process of explicitness-driven maturation and two syntactic rearrangements referring to the encoding of the adverbial clause subject and to the position of the connective within the adverbial clause, which are clear consequences of the discourse origin of *bwe*.

Keywords: discourse marker, interclausal connectives, cause/reason adverbial clause, adversative clause, thematic reorientation, grammaticalization

1. Introduction

This paper aims to show and explain the evolutionary path through which a discourse marker of the Yaqui language (Uto-Aztecan) has been recruited for interclausal connectivity purposes. The main point of this study is to propose that two Yaqui interclausal connectives (*bwe'ituk* and *bweta*) are the result of a recent formation process that combines a discourse marker (*bwe*) and linguistic elements

https://doi.org/10.1075/tsl.126.09gon

associated with the strategies used in the past for marking cause/reason clauses and adversative clauses in Yaqui. I will argue that these formations are functionally motivated by the fact that *bwe* is a discourse connective of discontinuity, that is, a discourse marker that introduces a topic-shift. Based on this connecting function, the element *bwe* has been recruited from discourse to syntax, in order to participate in the creation of two new interclausal connectives that also correspond to thematic reorientation devices: the cause/reason adverbial connective *bwe'ituk* and the adversative connective *bweta*.

The information is organized as follows. In Section 2, I will present some general morphosyntactic features of Yaqui. Section 3 will illustrate the current uses of two *bwe*-formed syntactic connectives in Yaqui: *bwe'ituk* that serves to introduce cause/reason adverbial clauses and *bweta* that is an adversative conjunction. In Section 4, I will exemplify the cause/reason adverbial clauses and the adversative clauses documented in *Arte de la lengua cahita*, which contains the first available description of Yaqui from the first half of the seventeenth century. Based on the comparison of these past and present-day constructions, I will propose that a discourse marker *bwe* has been used for creating two interclausal connectives. In order to defend this hypothesis, I will focus in Section 5 on the different uses of the discourse marker *bwe*, trying to determine the discursive features of this particle that can explain its involvement in the creation of new syntactic connectives. In addition to their interclausal connecting uses, *bwe'ituk* and *bweta* also present nowadays some interesting discursive uses that will be explained in Section 6 and that can be viewed as a possible intermediate stage in the formation process proposed here, in particular in the case of the *bwe'ituk* formation. Lastly, the final remarks will insist on the main aspects of the evolution presented in this study.

2. Yaqui language

Yaqui is spoken in northwestern Mexico in the state of Sonora by almost 17,000 speakers and in the bordering state of Arizona by approximately 500 speakers.[1] This language belongs to the Taracahitan branch of the Sonoran group within the Southern Uto-Aztecan languages. Table 1 presents the different languages of this southern division of the Uto-Aztecan family.

1. At the beginning of the twentieth century, this community settled in the USA from its original homeland in the south of the neighboring state of Sonora in Mexico, fleeing persecution by the Mexican dictator, Porfirio Diaz.

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*
- c. Tubar: †Tubar

Corachol-Aztecan
- a. Corachol: Cora, Huichol
- b. Aztecan:
  - – †Pochutec
  - – General Aztec (or Nahuatl): Pipil, Aztec (many varieties)

As shown in Table 1, Yaqui is part of a sub-group named Cahita[2] that also includes Mayo and Tehueco (nowadays, extinct). Although it is not uncommon to use the term "languages" when referring to Yaqui, Mayo and Tehueco, they are structurally very similar, so it is possible to recognize them as three varieties of the same language: the Cahita language, as named in the first available documentation of this linguistic group (see Section 4.1).

Yaqui is an agglutinative language with a very predominant use of suffixes and postpositions. Its alignment system is nominative-accusative, as can be seen from the pronoun system illustrated in Table 2.

Table 2. Pronominal system in Yaqui

	Nominative	**Accusative**	**Possessive**
1SG	*inepo, =ne*	*nee*	*in, nim*
2SG	*empo, ='e*	*enchi*	*em*
3SG	*aapo, Ø*	*aapo'ik, a=*	*aapo'ik, a=*
1PL	*itepo, =te*	*itom*	*itom*
2PL	*eme'e, ='em*	*enchim*	*em, enchim*
3PL	*bempo, =mme, Ø*	*aapo'im, am*	*bem, bempo'im*

Examples (1) and (2) show that the basic order in the transitive construction is SOV. This unmarked order of constituents tends to change to OSV with pronominal

2. This exonymic glossonym comes from the word *kaita* meaning 'nothing'.

arguments. In noun phrases, the nominative case is unmarked, whereas the accusative case is marked by the suffix *-ta* as in (1), with the exception of plural objects (2), because there is an incompatibility in Yaqui between the accusative marker and the plural marker (differential object marking). Determiners are optional, especially in object position.

(1) *U yoeme-Ø uka kari-ta jinu-k.*
DET man-NOM DET.ACC house-ACC buy-PFV
'The man bought the house.'

(2) *U yoeme-Ø u-me kari-m jinu-k.*
DET man-NOM DET-PL house-PL buy-PFV
'The man bought the houses.'

The data for this study are from several sources. Historical data come from the *Arte de la lengua cahita,* a colonial grammar written in the first half of the seventeenth century by an anonymous Jesuit and edited by Eustaquio Buelna in 1890, and from the letters written by the Yaqui leader Juan Bandera between 1830 and 1832, and published by Dedrick (1985). Synchronic data of Yaqui come from texts included in Silva et al. (1998), Estrada et al. (2004), Buitimea (2007), and Estrada & Álvarez (2008).

3. Two *bwe*-formed syntactic connectives in Modern Yaqui

Nowadays, Yaqui has two syntactic connectives presenting the same initial syllable *bwe-*: *bwe'ituk* and *bweta*. The former is used to introduce a clause that provides the cause/reason for which the situation denoted in the associated clause is carried out, and the latter corresponds to an adversative connective that introduces a clause that is in opposition with the first clause.

3.1 *Bwe'ituk* as a cause/reason adverbial connective

Examples (4) and (5) illustrate the use of *bwe'ituk* as a cause/reason adverbial connective, introducing an adverbial clause that has respectively the same subject and a different subject comparing to the main clause.

Same subject:

(4) *Baanu'u-ta te tapunia-bae [**bwe'ituk** te*
water_bottle-ACC 1PL.NOM fill-DES **because** 1PL.NOM
ke'u-bae].
go_to_the_wood-DES
'We are going to fill the water bottle because we are going to the wood.'

Different subject:

(5) *Inepo in yo'owam baisae [**bwe'ituk** bempo kaba'i-ta*
1SG.NOM 1SG.POSS parents thank **because** 3PL.NOM horse-ACC
nee miika-k].
1SG.ACC give-PFV
'I thank my parents because they gave me a horse.'

Three important features can be mentioned from these examples:

- The adverbial connective *bwe'ituk* is located in the adverbial clause-initial position, after the main clause.
- There is no switch-reference system associated with the use of *bwe'ituk*, in the sense that we have the same construction independently if the subject of the adverbial clause is identical or distinct to the subject of the main clause. In both cases, the connective *bwe'ituk* is used in the same way.
- Following the terminology proposed by Stassen (1985), the adverbial clause introduced by *bwe'ituk* is balanced, that is, the structure is alike to the structure of the main clause: the subject is in the nominative case, the object in the accusative, and the verb is finite and does not present any Tense-Aspect-Mood restriction.

3.2 *Bweta* as an adversative connective

Examples (6) and (7) illustrate the use of *bweta* as an adversative connective, introducing a clause that presents respectively the same subject and a different subject comparing to the main clause.

Same subject:

(6) *Junue juya-t ne ja'amu-bae **bweta** ne kaa aawe.*
DEM tree-LOC 1SG.NOM climb_up-DES **but** 1SG.NOM NEG know_how
'I want to climb up this tree but I don't know how.'

Different subject:

(7) *Joan nee Peesio-u nunu-k* ***bweta*** *ne e'e*
Juan 1SG.ACC Hermosillo-DIR invite-PFV **but** 1SG.NOM NO_EMP
au jiia-k.
3SG.OBL say-PFV
'Juan invited me to Hermosillo but I said to him no.'

The same three features mentioned above for the use of *bwe'ituk* are present in the use of *bweta*: the adversative connective is in the adversative clause-initial position, after the first clause, there is no switch-reference system, and the adversative clause is balanced (Stassen 1985).

In the following sections, I will propose that both present-day interclausal connectives of Yaqui are the result of a recent connective-formation process that combines a discourse marker (the element *bwe*) and linguistic elements associated with the strategies used in the past for introducing cause/reason clauses and adversative clauses in Yaqui. In order to defend this hypothesis, I will first present the old structural means used in the past for this kind of constructions and second, I will focus on the different uses of the discourse marker *bwe*, trying to identify the reasons that have motivated the recruitment of this discursive particle for the aforementioned interclausal connectivity purposes.

4. The evolution of cause/reason and adversative connectives in Yaqui

Firstly, I describe the cause/reason adverbial clauses and the adversative clause clause existing in Colonial Cahita and secondly, I present my hypotheses about the formation of the current cause/reason and adversative connectives in Yaqui.

4.1 Cause/reason adverbial clauses and adversative clause in Colonial Cahita

The earliest known description of the Cahita language is given in the *Arte de la lengua Cahita escrita por un Padre de la Compañía de Jesús.* Although the first version of this colonial grammar was printed in Mexico City in 1737 by Francisco Xavier Sanchez, it was probably compiled prior to 1650 (Dedrick & Casad 1999: 3; Álvarez 2018). This version was later edited and published in 1890 by Eustaquio Buelna, who acknowledged in his introduction (Buelna 1890: X) that the Cahita language is represented by three dialect variants: Yaqui, Mayo and Tehueco. The same assumption is made by the own author of the *Arte* in the information provided to the reader (Buelna 1890: 5), where it is said that, in spite of their differences,

Yaqui, Mayo and Tehueco can be considered as the same language. I will refer to it here as Colonial Cahita.

The linguistic forms documented in the *Arte* come from Tehueco but the original author was very careful to point out, all along the *Arte,* the existing differences between Tehueco and the other two Cahita variants. In that respect, Buelna (1890: XI) admits that these differences are very few in number, thus, it is possible to use the linguistic information provided in the *Arte* as comparative data in order to identify the evolution undergone by cause/reason adverbial and adversative clauses in Yaqui.

4.1.1 *Cause/reason adverbial clauses in Colonial Cahita*

From example (8) to example (13), I present the constructions provided in Buelna (1890: 72) for illustrating the cause/reason adverbial clauses from Colonial Cahita.

Same subject:

(8) *Emchi ne noctehoa, [emchi eria* ***teca****].*
2SG.ACC 1SG.NOM teach 2SG.ACC love **SUB**
'Porque te amo, te enseño.'[3]
'I teach you because I love you.'

(9) *Emchi ne vuiu-c, [emchi eria-c* ***tuca****].*
2SG.ACC 1SG.NOM quarrel-PFV 2SG.ACC love-PFV **SUB**
'Porque te amé, te reñí.'
'I quarreled you because I loved you.'

(10) *Emchi ne veb-naque, [emchi eria-naque* ***teca****].*
2SG.ACC 1SG.NOM whip-FUT 2SG.ACC love-FUT **SUB**
'Porque te amaré, te azotaré.'
'I will whip you because I will love you.'

Different subject:

(11) *Tuurisi ne ane, [emchi netz eria* ***ituca****].*
well 1SG.NOM behave 2SG.ACC 1SG.ACC love **SUB**
'Porque me amas, procedo bien.'
'I behave well because you love me.'

---

3. In the examples taken from the *Arte*, I have left the original translation into Spanish proposed by the own author of the *Arte*, whereas the morphological segmentation, the gloss and the English translation are mine. The same has been made for the different discursive examples presented below in this paper, which mainly come from narrative texts included in Silva et al. (1998), Estrada et al. (2004) or Buitimea (2007).

(12) *Buite-ca ne [emchi netz eria-c **tuco / tuca**].*
run_away-PFV 1SG.NOM 2SG.ACC 1SG.ACC love-PFV **SUB SUB**
'Porque me amaste, me huí.'
'I run away because you loved me.'

(13) *Tuurisi e a-naque [netz emchi eria-naque **ituca**].*
well 2SG.NOM behave -FUT 1SG.ACC 2SG.ACC love-FUT **SUB**
'Porque te he de amar, has de vivir bien.'
'You will behave well because I will love you.'

In all these examples, the cause/reason clause is marked by a particle located in the final position within the adverbial clause and conveying distinctions associated with tempo-aspectual meanings and switch reference:

- *teca* was used when the subject of the main clause and the subject of the adverbial clause co-referred in present (Example 8) and future situations (Example 10)
- *ituca* was also used in present (Example 11) and future situations (Example 13) but in cases without co-reference between the subjects of the main clause and of the adverbial clause.
- *tuco* was used when the subject of the adverbial clause was different from the subject of the main clause but only in past situations (Example 12)
- *tuca* was used in co-referential and past situations (Example 9) but it could also be used instead of *tuco*, for non co-referential and past situations (Example 12).

If we compare the cause/reason adverbial clauses illustrated in (8)–(13) from Colonial Cahita with the cause/reason adverbial clauses of Modern Yaqui exemplified in (4) and (5), we can observe several interesting differences:

- In Colonial Cahita, there was a system of switch-reference, not anymore. The marking of the old cause/reason adverbial clauses was different depending on whether the subjects of the adjacent clauses co-referred. As seen above, this distinction is no longer relevant nowadays because *bwe'ituk* is used for both, same subjects (SS) and different subjects (DS).
- This old switch-reference system was accompanied by tempo-aspectual distinctions. Indeed, there were different markers depending on the tense-aspect situation: for present and future, *teca* (SS) and *ituca* (DS); for past, *tuca* (SS/DS) and *tuco* (DS). Nowadays, this differential marking of the cause/reason adverbial clause does no longer exist since the connective *bwe'ituk* serves for all tempo-aspectual situations. In that respect, it can be observed that

the neutralization of the switch reference system was apparently incipient in Colonial Cahita because *tuca* could be used in past situations for both SS and DS.[4]

– The old cause/reason adverbial clause exhibited a higher degree of syntactic integration in relation to the main clause. Contrary to the present-day balanced adverbial clause introduced by *bwe'ituk*, the old cause/reason adverbial clause was deranked (Stassen 1985) since it showed some differences with regard to an independent clause, particularly in the encoding of the dependent clause subject: zero anaphora for same subject, accusative marking for distinct subject. Recall that nowadays the subject of the adverbial clause introduced by *bwe'ituk* is in the nominative case (see Examples (4) and (5) for both, same subjects and different subjects, respectively).
– The old adverbial connectives (*teca, tuca, tuco, ituca*) were in the adverbial clause-final position, whereas the present-day adverbial connective *bwe'ituk* is in the adverbial clause-initial position.

4.1.2 *The multifunctionality of teca in Colonial Cahita*

Apparently, the final particle *teca* was not only used to mark cause/reason clauses in Colonial Cahita. According to the information provided in the *Arte* (Buelna 1890: 66–68), this element was also used to introduce other adverbial clauses, like conditional clauses (14) and (15), purpose clauses (16) and (17), and temporal posteriority clauses ('before' clause) in (18).

Conditional clause in Colonial Cahita (Buelna 1890: 66–68)

(14) *Emchi-ne hiocori eiai, [emchi eria* ***teca****].*
2SG.ACC-1SG.NOM help try 2SG.ACC love **SUB**
'Te socorriera, si te amara ó te hubiera amado.'
'I would help you, if I loved you.'

(15) *[Ca-ne emchi eria-****tec****], ca emchi mica-na.*
NEG-1SG.NOM 2SG.ACC love-**SUB** NEG 2SG.ACC give-OPT
'Si yo no te amara, no te lo diera.'
'If I didn't love you, I wouldn't give it to you.'

4. As shown in Álvarez (2009), the neutralization of this switch-reference system corresponds to a generalized process in Yaqui adverbial clauses. Indeed, almost all adverbial clauses in Colonial Cahita exhibited a switch-reference system, which has been lost in Modern Yaqui. Some relics of this old switch-reference system can still be found in temporal adverbial clauses with the markers *-ka(i)* for SS, *-o* for DS (Álvarez 2009; Guerrero this volume).

Purpose clause in Colonial Cahita (Buelna 1890: 82)

(16) *uaquim ne iepsa-c [misa-ta bit-naque **teca**].*
here 1SG.NOM arrive-PFV mass-ACC look-FUT **SUB**
'He llegado aquí á oir misa.'
'I have arrived here to hear the mass.'

(17) *Teopa-u ne quivaque [Dios-ta eria-naque **teca**].*
church-DIR 1SG.NOM enter God-ACC love-FUT **SUB**
'Entro en la Iglesia à amar a Dios.'
'I enter into the church to love God.'

Before clause in Colonial Cahita (Buelna 1890: 68)

(18) *emchi ne hiocore-c, [quehe emchi eria **teca**].*
2SG.ACC 1SG.NOM help-PFV not_yet 2SG.ACC love **SUB**
'Te socorrí antes de que te amara.'
'I helped you before I loved you.'

These different adverbial clauses show that *teca* was not an explicit marker of interclausal causality, but a general adverbial clause connective that required contextual information to be activated in order to infer the appropriate semantic adverbial relation between the two clauses. This marker seems to have been used, in fact, more generally for backgrounding the situation denoted in the dependent clause to the situation expressed in the main clause, and for linking both clause situations in terms of consecutiveness, sequentiality. If the adverbial clause marked by *teca* was not expressing a future situation, the dependent clause situation was then anterior to the main clause situation (before, cause/reason and conditional interpretations). If this adverbial clause was marked by the future marker *-naque*, the dependent clause situation was posterior to the main clause situation (purpose interpretation) or anterior in the case of temporal co-referentiality between both clause situations (cause/reason interpretation, as in (10) and (13)).

This sequentiality may be then interpreted literally as temporal[5] but also metonymically as causal, conditional or purposive, based on the idea that causes/reasons and conditions are usually anterior to their consequences, and that purposes are prospective consequences, posterior to the main event (Croft 1991). As pointed out by Heine & Kuteva (2002: 291), the grammaticalization from TEMPORAL to CAUSAL and from TEMPORAL to CONDITIONAL are instances of a "widespread process whereby spatial and temporal markers are grammaticalized in specific contexts

5. In the case of before adverbial clauses, the temporal interpretation of anteriority is explicitly expressed by the element *quehe* 'not yet' located in the initial position within the adverbial clause in (18).

to markers of "logical" grammatical relations such as adversative, causal, concern, concessive, and conditional relations." Kortmann (2001) has also shown that temporal relations represent the most common source domain for cause, condition and purpose adverbial connectives. Moreover, the syncretism between temporal, causal, conditional and purposive meanings is also explained because conditions may be conceptualized as hypothetical causes/reasons that would chronologically precede the main event, whereas purposes may be conceptualized as prospective causes/reasons that would chronologically follow the main event.

The multifunctionality of the connective *teca* shows a semantic underspecification and concomitant context-driven interpretations, that is, the activation of pragmatic inferences in order to correctly interpret the corresponding meaning of the adverbial clause from the context. This need for pragmatic enrichment from context illustrates a kind of hidden complexity, that is, a complexity created by economy and represented by a language structure that (i) does not force the speaker to overtly express grammatical categories that are part of its grammatical inventory (lack of obligatory categories) and (ii) has multifunctional markers whose concrete meaning must be inferred from context (Bisang 2009, 2014).

Another fact that is interesting to note is the possibility illustrated in (15) to have the adverbial clause before the main clause, something very common cross-linguistically for conditional clauses (Ford & Thompson 1986; Diessel 1996; Diessel 2001) and that is iconically motivated since conditions are anterior to their consequences. In this case, we can observe that the adverbial marker *teca* seems to be more integrated to the dependent verb (if we consider the transcription used in the *Arte*) and, more importantly, its final vowel *a* is elided, something that is very pervasive in Colonial and Modern Cahita where the final vowels of bi-syllabic morphemes are lost if not in final position of the sentence.

Additionally, all these adverbial clauses marked by *teca* confirm two features that have been mentioned above in the description of the cause/reason adverbial clauses in Colonial Cahita: (i) *teca* was apparently only used as a same-subject adverbial connective since the subject of the adverbial clause is always the same subject than the subject of the main clause, (ii) *teca* was apparently incompatible with past adverbial clauses, since all *teca* examples shows present and future adverbial clauses (recall that in past cause/reason adverbial clauses as in (9) and (12) the particle changes to *tuca*).

4.1.3 *Adversative clauses in Colonial Cahita*

As shown in (19) and (20) from the *Arte* (Buelna 1890: 125), the adversative connective was *tepa* in Colonial Cahita. Notice in these examples that the third singular person subject is not overtly expressed in Colonial Cahita, something that remains the same nowadays in Yaqui.

(19) *ca aeua-c,* ***tepa*** *hulen noca.*
NEG do-PFV **but** like this tell
'No lo hizo, pero lo dice así.'
'He didn't do it, but he tells it like this.'

(20) *ca a-veba-c* ***tepa*** *a-buiu-c.*
NEG 3SG.ACC-whip-PFV **but** 3SG.ACC-quarrel-PFV
'No lo azotó pero lo riñó.'
'He didn't whip him but he quarreled him.'

If we compare with adversative clauses of Modern Yaqui, we can observe that there is no structural difference, except for the change of the connective (*tepa* vs. *bweta*). The adversative connective *tepa* was in the adversative clause-initial position, after the first clause, there was no switch-reference system, and the adversative clause was balanced.

4.2 The *bwe'ituk* and *bweta* formations

In the light of the past strategies used for cause/reason clauses and adversative clauses in Colonial Cahita, it is possible to consider that the present-day cause/reason and adversative connectives of Yaqui seem to be the result of the combination of a monosyllabic element *bwe-* and a form that can be easily related to the old strategies. These formation hypotheses are then the followings:

(21) **Formation hypotheses:**
a. Formation of the current cause/reason adverbial connective:
bwe'ituk < *bwe* + *ituc(a)*
b. Formation of the current adversative connective:
bweta < *bwe* + *t(ep)a*

The proposal in (21a) implies two main syntactic rearrangements in the expression of the new cause/reason adverbial clause: (i) the change of the adverbial connective position from dependent clause final position to dependent clause initial position, (ii) the change in the subject marking of the dependent clause from zero (SS) or accusative (DS) marking to nominative marking, with the concomitant loss of the old switch-reference system. This formation also implies the apocope of the vowel *a* from the old final particle *ituca*. This vowel elision that has already been illustrated in Colonial Cahita with the example (15), is still a very frequent phenomenon in Yaqui when suffixed words are in positions others than the final position.

On the contrary, the proposal in (21b) is not associated with syntactic rearrangements, only with the phonetic erosion corresponding to the syncope of the old adversative connective from *tepa* to *ta*.

Two major facts support these formation hypotheses: (i) the existence of a discourse marker in Yaqui that has the form *bwe*, (ii) the co-existence of *bweta* and *ta* for adversative connectivity purposes in Modern Yaqui. This synchronic variation exemplified in (22) with the lexicographic information provided in Estrada et al. (2004) represents a strong evidence in favor of the idea that *bweta* is the result of the proposed combination and a good support for postulating the combined origin of *bwe'ituk*.

(22) a. **bweta** *nexo*. Pero. *nexus*. But. *Cf*. **ta**. (Estrada et al. 2004: 73)
b. **ta** *nexo*. Pero. *nexus*. But. *Cf*. **bweta**. (Estrada et al. 2004: 177)

About the dating of the *bwe'ituk* and *bweta* formations, it is possible to state that *bwe'ituk* has been created before *bweta*. Indeed, the first available evidence of *bwe'ituk* appears in the *Cartas de Juan Bandera* written between 1830 and 1832 and published by Dedrick in 1985, whereas *bweta* seems to be very recent since neither Johnson (1962), nor Lindenfeld (1973), nor Dedrick and Casad (1999) mention the existence of *bweta*, registering only *ta* or *táa* for 'but'.[6]

In the following section, the different functions of the discourse marker *bwe* will be presented, trying to understand why it has been recruited for the creation of these two new interclausal connectives.

5. The discourse marker *bwe*

Unfortunately, we have no information about old uses of discourse markers or interjections in Cahita. In that respect, at the end of the *Arte* (Buelna 1890: 126), it is pointed out that with reference to interjections, "*no se hallan en* ***los Artes****, el uso las dará*".[7]

The situation is different for current uses of discourse markers. For instance, in the bilingual Yaqui-Spanish dictionary published by Estrada et al. (2004), several discourse markers are registered as lexical entries. One of them is the particle *bwe* whose lexical entry is given in (23).

---

6. The study of Johnson (1962) is based on a manuscript from 1940, and the grammar of Dedrick & Casad (1999) is based on the data obtained by Dedrick who lived with the Yaquis for about 20 years between the 1940s and the 1960s. The text of Ambrosio Castro present at the end of their grammar dates back to between 1941 and 1944, and only shows the use of *ta* as an adversative connective. In both works, the connective *bweta* is never mentioned.

7. "They are not found in **the Artes**, the use will give them".

(23) **bwe** *part*. ¡ah!, ¡a poco!, ¡bueno!, ¡este!, ¡pues! *Cf*. **abwe**.
bwe *part*. ah!, oh really!, well!, er! um! *Cf*. **abwe**
(Estrada et al. 2004: 73)

The information provided in this lexical entry refers to various potential equivalents in Spanish but also to the coexistence of a longer form of this particle, the form *abwe*.[8] The example (24) corresponds to the exemplification of its use given in Estrada et al (2004: 73). It shows *bwe* functioning as a discourse marker in the utterance initial position, which conveys a discursive meaning associated with a surprise.

(24) ***¡Bwe** jiba yepsa-k!*
PART always arrive-PFV
'¡A poco siempre llegó!'
'Oh really, s/he arrived at last!'

The information provided in the lexical entry corresponding to *abwe*, which is given in (25), insists on this initial position and on the fact that it is a conversational marker. The example of the *abwe* use given by Estrada et al. (2004: 49) is proposed in (26). It is very similar to the use of *bwe* exemplified in (24).[9]

(25) **abwe** *part*. Partícula que se utiliza al inicio de una conversación. Equivale en español a: 'a poco', 'bueno', 'este', 'pues'. *Cf.* **bwe.** (Estrada et al. 2004: 49)

---

8. Regarding the origin of *bwe,* it seems plausible and tempting to consider that the discourse marker *bwe* comes from Spanish *pues* and *abwe* could be a loan from *ah pues*. Indeed, several studies suggest that, in intense language contact situations, the borrowing of discourse markers is quite common because discourse markers are a frequent locus for code-switching (Pfaff 1982; Brody 1987; Myers-Scotton 1993), which may serve as a trigger for borrowing (Myers-Scotton 1993; Torres 2002). However, in his brief study *Los elementos de la lengua cahita,* Lionnet (1977: 50) proposed that *bwe* corresponds to an old imperative form of the verb 'look' (the long form *abwe* could be then the result of the *a-* "3SG.ACC" pronominal prefixation, rendering the meaning 'look at it/him/her'). This possibility is interesting since, if true, this verbal origin would imply that the evolution presented in this paper is not only a case of grammaticalization but also of pragmaticalization in a first stage: IMPERATIVE VERB > DISCOURSE MARKER > INTERCLAUSAL CONNECTIVES. Unfortunately, there is no trace of a contemporary verb form *bwe* meaning 'look' or 'see' (the two current visual perception verbs in Yaqui are *bicha* 'see' and *bichu* 'look'), no attestation of this verb in the *Arte,* and no evidences of this old imperative use (Lionnet 1977 proposed this origin without providing any evidence). For reasons of space, I will leave the question of the *bwe* origin open for further investigation. Suffice it to say at this moment that, although the external source is tempting, it is preferable to be cautious since an internal source cannot be totally excluded for now.

9. Both forms of this discourse marker seem to be equivalent. My Yaqui informants consider them perfectly interchangeable, although the short form *bwe* is much more frequent in discourse.

(26) ***Abwe*** *jiba a'abo noite-k.*
PART always here come-PFV
'Pues siempre vino.'
'Well s/he came here at last.'

5.1 The discursive functions of the particle *bwe*

The importance of discourse contexts for explaining the evolutionary paths implied in many grammaticalization processes has been largely pointed out, not only in a general way (see, for instance, the discourse-based model of grammaticalization proposed by Waltereit and Detges 2007) but also more specifically like for example in the domain of interclausal connectivity. In that respect, Givón (2009) has proposed that the genesis of clausal conjunctions is constrained by the discourse context that frames its emergence. Therefore, it is crucial to examine here the different discourse contexts in which the element *bwe/abwe* may appear in Yaqui discourse, in order to understand the motivations and the constraints implied in the diachronic rise of the interclausal connectives *bwe'ituk* and *bweta*.

The first comment that is important to make here is that in texts the particle *bwe* only appears in speaking situations, that is, in conversations between characters. This situation clearly confirms that *bwe* is a conversational marker, as pointed out in Estrada et al. (2004) for *abwe*. But, beyond this interpersonal function, what are the different and more specific discourse functions associated with this conversational marker?

According to the following examples, the particle *bwe* seems to function as a discourse organizer that structures the information and that introduces a new comment distinct to the prior discourse: it is usually the first element of the answer to a question like in (27) and (28). It always appears in an initial position and serves then as part of a reactive intervention that introduces a new discursive unit.

(27) *¿jita ket nee betana nattemae-k, jaboi?*
something too 1SG.ACC from ask-PFV grandfather
– '¿Y qué más te preguntaron de mí, abuelo?'
– 'And what else did they ask you about me, grandfather?'
– ***Bwe*** *jachim-po juni'i kaa nee enchi empolai-k su'utoji-sae.*
PART how-LOC though NEG 1SG.ACC 2SG.ACC alone-ACC leave-order
– '**Ah pues**, que no merecías que yo te dejara solo.'
– '**Oh well**, that you didn't deserve that I leave you alone.'
Mi abuelo y yo (Buitimea 2007: 59)

(28) – *¿Jaisa into bea yee ko'oko-si jooa?*
– how and then someone pain-ADVZ do
– '¿Y cómo es que hace daño?'
– 'And how is it that it hurts?'
– ***Bwe**, yee ko'oko-si jooa yet-et tajte-ko.*
PART someone pain-ADVZ do someone-on touch-TEMP_SUB
– '**Pues**, hace daño cuando alcanza a pegarle a alguien.'
– '**Well**, it hurts when it reaches to hit someone.'
Mi abuelo y yo (Buitimea 2007:65)

Many other uses also imply exclamative utterances and the *bwe* particle functions then more like an interjection, that is, as an exclamation operator that allows carrying out an expressive speech act, as in (29).

(29) – *¿Jaisa ayu-ka kom a wike-k?*
how do-SUB down 3SG.ACC get_down-PFV
– '¿cómo le hizo para bajarlo?'
– 'how did you do to get it down?'
– ***¡Bwe** bu'u tekil-ta! Tua aapo'ik ta'a-'u jiba*
PART too_much work-ACC truly 3SG.NOM know-REL always
– '¡**Todo un ritual**! Que solo él entendía. (**Lit. Pues mucho trabajo!…**)'
– '**A whole ritual**! That only he understood. (**Lit. Well, a lot of work! …**)'
Mi abuelo y yo (Buitimea 2007:66)

In other examples, it functions as a reactive particle that serves to denote surprise like in (30).

(30) – *¿Jabetasa empo maala-k intoko? -ti neu jiia-k.*
who 2SG.NOM mother-POSS and like_this 1SG.OBL say-PFV
– '¿Y quién es tu mamá? – me preguntó.'
– 'And who is your mother? – s/he asked me.'
– ***¡Bwe** mala-ta!*
PART mother-ACC
– '¡**Pues** mamá!'
– '**Well** mom!' *La viejita viuda* (Buitimea 2007:95)

It can also be used as a phatic operator like in (31). It serves then as a support for the elocution in informal speech (as a kind of '*muletillas*', a conversational tic like 'er, um'). Notice in (31) the repetition of *bwe* that expresses the speaker's hesitation.

(31) –*¿Jaisaka empo kaa a'abo siika tuuka?*
why 2SG.NOM NEG here come.PFV yesterday
– '¿Por qué no viniste ayer?'
– 'Why didn't you come yesterday?'

–¡***Bwe…bwe*** *ousi ne tekil-ek-an.*
PART…PART a lot 1SG.NOM work-POSS-IMPFV
– '¡**Pues**… **pues** tenía mucho trabajo.'
– '**Well**… **well** I had a lot of work.' *La viejita viuda* (Buitimea 2007: 107)

All the discursive uses exemplified so far present the particle *bwe* as the initial element of a reply. This is its most frequent use. However, it can be used in a different way. For instance, another important use of this discourse marker corresponds to its use as a contrastive focus, introducing information that is contrary to the presupposition elaborated by the interlocutor. Example (32) seems to illustrate a similar use but a little different. Martin and Portalés (1999) have named this use '*enfocador de la alteridad*' ('otherness focus'), that is, a use that serves to reinforce the positive image of the speaker. This type of marker is usually employed at the beginning of a reactive intervention that implies a sort of opposition, disagreement with the interlocutor. It signals the position of the speaker with regard to the hearer, in the sense of attracting him/her to the sphere of the speaker and indicating his/her point of view.

(32) – *Bwe'ituk ne jakko juni kee enchim bicha-n.*
Because 1SG.NOM when also yet_no 3PL.ACC see-IMPFV
– 'Porque yo nunca las había visto.'
– 'Because I had never seen them.'
– *ti-ne ameu jiia-k.*
like_this-1SG.NOM 3PL.OBL say-PFV
– 'les dije.'
– 'I told them.'
– ***Bwe*** *itepo ala enchi ta'a.*
PART 1PL.NOM yes_EMP 2SG.ACC know
– '**Pues** nosotras sí te conocemos.'
– '**Well** we do know you.' *Mi abuelo y yo* (Buitimea 2007: 56)

Example (33) seems to illustrate a discursive use of *bwe* as a consecutive connective. It allows then introducing a part of discourse as a consequence of the prior discourse. This example is interesting because *bwe* is not strictly associated with a turn at talk in (33) as it is in the previous examples. It seems to function here more appropriately as an interclausal connective, linking two independent clauses. Indeed, although the Spanish and English translations of the first clause refer to a temporal adverbial clause, it is in fact a non-promotional passive construction (Lit. *it was put on me the headpiece*). Another interesting point is the fact that the particle *bwe* is not located in (33) in a turn-initial position. On the contrary, it is located in a turn-medial position (Clayman 2012) but it is still in the initial position in relation to the new reactive comment.

(33) – *Ta juka chomo-ta ne-t yecha'a-wa-k.*
but DET.ACC penacho-ACC 1SG-on put-PASS-PFV
– 'Pero cuando me pusieron el penacho.'
– 'But when they put me the headpiece.'
– ***Bwe*** *¡e'e! ti ne kaa into jiu-bae.*
PART NO_EMP like_this 1SG.NOM NEG and say-DES
– '**Pues** ya no quise decir que no.'
– '**Well** I didn't want to say no anymore.' *Los coyotitos* (Buitimea 2007: 147)

From all these previous examples, the element *bwe* appears to be a discourse/pragmatic marker mainly used in conversations, which presents the prototypical features of discourse markers that have been pointed out, among others, by Schiffrin (1987), Fraser (1990), or Onodera (2011): (i) *bwe* is used in the initial position of an utterance, which is usually associated with an intense focal character and a topic-shift. Its most frequent use is as an introducer of a reply, (ii) *bwe* signals the speaker's view/attitude/judgment with respect to the relationship between the chunks of discourse that precede and follow it, introducing in most cases a new reactive comment. It then conveys an intersubjective meaning and it has a metatextual, interactional function. In a general way, the element *bwe* can be considered as a discourse-connecting conjunction ("contextual coordinates" in Schiffrin's (1987) terms) that introduces a topic-shift.

5.2 The functional motivations of the *bwe* recruitment

Based on these discourse features associated with *bwe*, it is possible to propose that the element *bwe* has been recruited for introducing cause/reason adverbial clauses and adversative clauses in Yaqui because *bwe* is a discontinuity connective (topic-shifter), serving as a thematic reorientation device with a cataphoric orientation. The same cataphoric thematic discontinuity is strongly associated with cause/reason adverbial clauses and adversative clauses, since causal and adversative situations usually represent new topics in the sense that in both types, the interclausal connective provides pragmatically presupposed information that links the following information to information that is already in the hearer's knowledge store and that is expressed in the first main clause. Based on this connecting function, I propose that the element *bwe* could move from a metatextual function to a sentential function, from discourse to syntax, in order to participate in the creation of the causal and adversative connectives in Yaqui.

The topic shifting associated with the discourse marker *bwe* would explain why it is the old particle *ituca* that has been combined with the element *bwe* in order to create *bwe'ituk*, instead of the other particles (*teca*, *tuca*). Indeed, this particle

ituca was used in different subject situations, that is, in situations that introduce new referents, new participants in relation to the main clause. As referents ('participants') are one of the most important sub-elements of thematic coherence (Givón 2001, Ch. 18), the introduction of new referents is clearly associated with thematic discontinuity, thus with topic shifting. The combination of *bwe* and *ituca* could be based on this shared discursive function of discontinuity.[10]

As for the combination of *bwe* with the adversative connective *t(ep)a*, the same function of discontinuity also seems to be relevant. Indeed, as pointed out by Givón (2009), various studies have shown a strong statistical association between the adversative conjunctions and cataphoric referential discontinuity in different languages (English in Hayashi 1989; Sup'ire in Carlson 1987; Polish in Frajzyngier 1988 and Green Hmong in Li 1988). Like in the case of *bwe'ituk*, the presence of the discourse marker *bwe* in the new adversative connective *bweta* seems to be associated with this discontinuity function, in order to reinforce the topic-shift present *per se* in the old adversative connective.

6. Discursive uses of *bwe'ituk* and *bweta* in Yaqui

The discursive function mentioned above for the element *bwe* can also be observed nowadays in some current uses of *bwe'ituk* and *bweta* in Yaqui, showing that the discourse/pragmatic function of *bwe* has not been totally eliminated by the process of interclausal connectives formation. These uses, which can be viewed as a possible intermediate stage in the evolutionary process proposed here, represent strong evidences in favor of our formation proposals.

6.1 Discursive uses of *bwe'ituk*

Concerning the use of the cause/reason connective, Dedrick and Casad (1999: 397) have pointed out that "Because clauses are commonly marked by the introducer *bwe'ituk* and often occur as the initial clause in a complex sentence… preceding the main clause of the sentence." This comment is interesting because in our corpus, the position of the adverbial clause introduced by *bwe'ituk* is usually after the main clause, contrary to what Dedrick and Casad (1999) mentioned. This pre-posed location is cross-linguistically unusual since causal clauses typically follow the main

10. Consequently, it could be proposed that in the chronology of the appearance of the new causal connective *bwe'ituk*, it should have been firstly used with different subjects cause/reason adverbial clauses and thereafter, with same subjects cause/reason adverbial clauses.

clause (Diessel 2001: 445–446). This initial position could be associated with the discursive origin of *bwe* and could indicate an intermediate stage in the process of grammaticalization. Moreover, in some examples proposed by these scholars, the meaning of *bwe'ituk* is not clearly causal and it seems to correspond more to a discourse marker, as in (34) and (35).

(34) ***Bwe'ituk*** *ne kaa enchim saja-ko bwana-ka matchu-nee.*
because 1SG.NOM NEG 2PL.ACC go.PL-SUB_COND cry-SUB wake_up-FUT
'Because if you don't go, I'll wake up crying tomorrow.'
(Dedrick & Casad 1999: 393)

(35) ***Bwe'ituk*** *waa'a weye-me ji'i-bwa-bae-ka weye.*
because DEM go-NMLZ thing-eat-DES-SUB go
'Because that one is on the move, he is going around looking for food.'
(Dedrick & Casad 1999: 397)

First, it is important to note that in these two examples, we do not have a main clause and an adverbial clause introduced by *bwe'ituk*, but just a construction that presents a cause/reason expressed by a complex sentence that is interpreted as an adverbial cause/reason clause. In other words, in (34), the clause *I'll wake up crying tomorrow* is not the main clause but one part of the cause/reason clause, the other part being the conditional clause *If you don't go*. In (35), a more appropriate English translation would be: 'because that one who is moving, is going around wanting to eat something'. In both examples, the element *bwe'ituk* introduces the justification for something (the state of affairs denoted by the supposed main clause) that does not appear but was presumably present in the previous chunk of discourse, like if it was the response to a why-question in a conversation. This use is clearly linked to the use of the discourse marker *bwe* mentioned above, since *bwe'ituk* is used here in conversation (note the two pronominal forms referring to the speech act participants in (34)), and it is used in the initial position of an utterance, with an intense focal character and a topic-shift, introducing an event that is interpreted as a cause/reason of a previous unmentioned situation.

Another interesting example comes from the letters of Juan Bandera (Dedrick 1985: 146), which represents the first evidence of the *bwe'ituk* formation from the beginning of the eighteenth century. Again, we can observe the same initial position of the element *bwe'ituk* and the discursive function rather than the function of interclausal connection. Although we do not have strictly speaking a conversation in (36), this example also illustrates a communicative interaction since it comes from a written communication between several Yaqui participants (Captain Ignacio Buitime'a and Juan Domingo Husakame'a, who sign the letter and Juan Bandera who is the addressee).

(36) *'emchi-te tebotua lioh-ta bahi team-po.*
2SG.ACC-1PL.NOM greet God-ACC three name-LOC
'We greet you in the three names of God.'
'itepo, 'inim pueplo-po hoome-m ko'okorim-po,
1PL.NOM here town-LOC native-PL Cocorim-LOC
'we, the natives of this town of Cocorim,'
bat-naata-ka 'itepo toopa-m,
first-begin-SUB 1PL.NOM troop-PL
'beginning with the troops,'
'ae-t cha'a-ka 'itepo komunim chiktia.
3SG.INST-LOC lay-SUB 1PL.NOM community all
'including all the community with this.'
***Bwe'ituk**-te 'emchi temae 'inim-wakini*
because-1PL.NOM 2SG.ACC ask here-neighborhood
'(We write to you) because we ask you here'
huka'a lutu'uria-ta hu'unak'eeria-bae-kai.
DET.ACC truth-ACC know-DES-SUB
'wanting to know the truth.'

In this case, the element *bwe'ituk* is used again in the initial position of a new utterance without the presence of the supposed main clause written in parentheses in the Dedrick's English translation. It introduces the cause/reason for writing the letter, and comes after the initial greetings. The element *bwe'ituk* serves then to introduce a change of topic, which is associated with a strong focus function, something that fully corresponds to the use of the discursive particle *bwe*. This discursive function here seems to be better rendered by a translation like: 'Well, we ask you here…'.

Additionally, we can observe that the cause/reason introduced by *bwe'ituk* in (36) is not really factual, that is, the clause does not really express the cause/reason that explains an effect like for example in *the floor is wet because it rained*, but rather it expresses the justification of the speech act corresponding to the event of writing. Thus, the cause is here metadiscursive.[11]

In evolutionary terms, the use of *bwe'ituk* illustrated in (36) might correspond to an intermediate stage in the development of the cause/reason adverbial connective from the discursive particle *bwe*. This same discursive use of *bwe'ituk* is present in (34) and (35), with the same initial/focal position and the introduction of a clause interpreted as the cause/reason of the previous discourse. It is also interesting to note that this stage seems to involve a case of 'insubordination' (Evans 2007), since the three examples (34), (35) and (36) appear to be formally subordinate clauses,

11. Thompson and Longacre (1985: 203) named this type of adverbial clauses that provide the motivation for uttering the main clause, as *speech act adverbial clauses* or *speech act qualifiers*.

but they are used alone, without a main clause.[12] The *bwe'ituk* evolution could thus imply at this stage the conventionalization of the main clause elision (as in (36)) and the grammaticalization of interactive presuppositions (Evans 2009). As Evans (2007) has pointed out, insubordination occurs in situations where a high degree of intersubjectivity between speaker and hearer can be presupposed and one of the functions of insubordination is precisely to signal presupposed material, such as contrastive focus or discourse contrast (Evans 2007, 2009).

Another interesting point to comment here is that this intermediate stage in the evolution of *bwe'ituk* seems to involve not just discourse > syntax (where *bwe* is concerned), but also syntax > discourse (where the *ituca* particle is concerned). As Traugott (1988, 1995) has proposed, pragmatic strengthening usually occurs in the early stages of grammaticalization, in which meanings tend to shift toward greater subjectivity, that is, they become increasingly associated with speaker attitude, especially metatextual attitude toward discourse flow. The early stages of the *bwe'ituk* evolution seem to show this increase of pragmatic significance and subjective expressiveness.[13]

---

12. Interestingly, Cecilia Rojas (this volume) shows that in child language development, components of complex constructions may be used as free, insubordinated clauses before they are integrated into a complex frame. Based on the interlocutor's preceding conversational move, children produce constructions – dependent clauses included – that complete, expand, or depend upon the interlocutor antecedent construction. So, in dialogue, clauses overtly marked for their dependent status may form a complex construction with the clause uttered in the preceding move, displaying a dyadic production mode that Givón (2009) considers as the "spreading of complexity". Here, we can see a clear parallelism between ontogeny (language acquisition) and diachrony (historical change) in the genesis of syntactic complexity.

13. In this evolution from the old cause adverbial clause marked by the clause-final particle *ituca* to the new one marked by the clause-initial *bwe'ituk*, the question remains of how *ituca* could dissociate from final position of the subordinate clause. Analogy with *bweta* has to be excluded, since *bwe'ituk* most probably came first. An anonymous reviewer suggests that this change had to be originated via a reanalysis of the clause division as follows: Subordinate clause *ituca* + main clause > subordinate clause + *ituca* (discourse function) main clause > subordinate clause + *bwe* + *ituca* main clause, with somewhere in there the identity of main and subordinate clause getting switched (e.g. 'in order to get full + he ate a lot' > 'he got full + because he ate a lot'). Another interesting possibility would be an *ituca* clause used by itself via 'insubordination' (Evans 2007), with the *ituca* particle becoming associated with the following main clause. This issue clearly deserves more attention and requires addressing more precisely not only the role of 'insubordination' but also the syntax and origin of *ituca.* These aspects will be the subjects of a forthcoming study.

6.2 Discursive uses of *bweta*

In the same way, it is interesting to notice that *bweta* can also be used as a discourse marker when it occupies the initial position of the utterance as shown in (37).[14]

(37) – *Wate bwia-me-u bichaa ne tekil-ta jariu-bae.*
others land-PL-DIR towards 1SG.NOM job-ACC look_for-DES
– 'Voy a buscar trabajo para otras tierras.'
– 'I'm going to look for a job in other lands.'
– *ti ne au jiia-n.*
like_this 1SG.NOM 3SG.OBL tell-IMPFV
– 'le decía.'
– 'I told him.'
–¡***Bweta*** *empo into jakko juni'i kee inim yeu siime!*
but 2SG.NOM and when yet NEG here outside go
– '¡**Pero** tú nunca has salido de aquí!'
– '**But** you have never left from here!'
El mundo de los sueños (Buitimea 2007: 100)

In (37), *bweta* has a clear expressive function since it serves to introduce an utterance that signals the surprise and/or the doubt the speaker feels about the plausibility of what have been said in the previous discourse.

The example (38) is interesting because it presents three uses of the connective *bweta*. The discourse here is not a conversation but a narrative about the actions of a bird. This extract shows a sequence of actions where a bird hesitates and finally decides to do something unexpected. This sequentiality is reinforced by the use of temporal adverbs (*seechiti* 'suddenly', *sep* 'soon', *intuchi* 'again', *ian* 'now').

(38) *Seechiti jikau ne'e-ka,* ***bweta*** *intuchi sep notte-k.*
suddenly to_up fly-PFV but again soon return-PFV
'De pronto voló, pero luego regresó.'
'Suddenly he flew, but later returned.'
Intuchi am bitchu-taite-k ***bweta*** *ian ala tua ameu*
again 3PL.ACC see-INCH-PFV but now yes_EMP truly 3PL.OBL
batuba-bae-k.
decide-DES-PFV
'Los empezó a ver de nuevo, pero ahora sí se decidió.'
'He started to look at them again, but now he did make a decision.'

14. The use of *bweta* as a discourse marker has already been pointed out by Hernández Doode (2002: 106–108). In her brief analysis, she proposes that, besides its syntactic use as an adversative conjunction, *bweta* presents other uses as a discourse marker introducing a pragmatic contrast, an "antiorientation", i.e., something that is contrary to expectations.

¡bweta *am* *nuksim-bai-kai ju-me* *bwaji-m!*
But 3PL.ACC take-DES-SUB DET-PL underpants-PL
¡A llevarse los calzones! (Lit. **Pero** llevándose los calzones!)
To take away the underpants! (Lit. **But** taking away the underpants!)
Las lagartas (Buitimea 2007: 75)

In the first two uses, *bweta* acts as an adversative connective between two clauses. The first *bweta* introduces an action that is the opposite action with regard to the one denoted in the first clause. The second *bweta* is also used for introducing a clause in opposition with the previous clause since it notifies that the bird has made a decision, at last. This change of mind is reinforced by the temporal adverb *ian* 'now' and the affirmative particle *ala*.

Contrary to the first two uses, the third *bweta* functions more as a discourse marker. It appears at the utterance-initial position and it serves to indicate the surprise caused by the decision the bird made. The exclamation mark is the reflection of this expressive function in the punctuation. Additionally, it is interesting to observe that, contrary to what happens in the uses as an adversative connective, the clause introduced by the third *bweta* is a dependent clause (marked by the subordinator *-kai*). Therefore, as seen above for *bwe'ituk* (see Examples (34) to (36)), the discursive function of *bweta* seems to be associated again with insubordination (Evans 2007), where the *bweta* clause becomes used on its own in relation to the discourse/pragmatic context.[15]

This discursive function associated with the utterance-initial position of the adversative connective is not exclusive of *bweta*, it can also be observed for its short counterpart *ta*. An example can be found in (33). This situation is not a surprise since adversative connectives have a strong intersubjective function *per se* (Giacalone Ramat & Mauri 2011) and it is very common that they are used as discourse markers (Schiffrin 1987).[16] It seems then that the use of the adversative connective as a discourse marker is not really due to the recruitment of the

15. In Example (37), the *bweta* clause is also used on its own in relation to the discourse/pragmatic context but, strictly speaking, this use does not correspond to a case of insubordination since in this case, this is not a subordinate clause that is used as a stand-alone, independent main clause, but a coordinated clause. Kuteva (2017) has proposed the term 'incoordination' for this independent use of formally coordinated clauses. The third use of *bweta* in (38) could thus be considered as an instance of 'incosubordination', a type of elliptical contructions combining incoordination and insubordination.

16. The coercion of clause-linking particles (especially adversatives) to become discourse markers has also been described by Traugott. See e.g. Traugott (2004), citing Onandera (1995), where Japanese clause-final adversative construction *-te* 'GERUND' + *mo* 'but' becomes the clause-initial discourse marker *demo* 'but' used in present-day Japanese to claim floor and change sub-topic.

discursive particle *bwe*, which appears to potentiate something that is in a certain way already present.

7. Final remarks

This paper has shown how Yaqui has created two interclausal connectives out of a discourse marker, illustrating an evolution from discourse to syntax in which the notion of discontinuity connective is central. Indeed, all uses of (free and bound) *bwe* correspond to a connecting device associated with a thematic re-orientation: *bwe* is used as a discursive connective of discontinuity whereas *bwe'ituk* and *bweta* can be viewed as syntactic connectives of discontinuity. In this connecting function, the position of these connectives is always fixed between the two discourse parts that are connected, and the orientation is always cataphorical, that is, the connectives are always at the left margin of the new topic discourse, then functioning as a clause-initial reorientation device.

The recruitment of the discourse marker *bwe* for the formation of interclausal connectives in Yaqui has thus been motivated by this thematic reorientation function. But each case is particular, since the reasons, the syntactic implications and the timing of both formations are different.

In the case of *bweta*, this formation seems to be very recent (probably, during the last fifty years) and it coexists with a reduced form of the old adversative connective (*ta* particle from Colonial Cahita *tepa*). The uses of these two different adversative connectives are similar. Comparing to the Colonial Cahita structure, neither syntactic rearrangement, nor functional innovation seem to be associated with the new adversative connective *bweta*, since there is no structural differences between both constructions. Additionally, the present-day version of the old adversative marker in Yaqui (the adversative connective *ta*) also presents some uses as a discourse marker. The recruitment of *bwe* appears, in fact, to be only motivated by an increase in subjectivity, something that is typical in the grammaticalization process of adversative connectives, as Hopper and Traugott (1993) have pointed out. The discourse marker *bwe* has then been recruited to expressively reinforce the original adversative meaning conveyed by the conjunction *t(ep)a*, something that is clearly motivated by the "deeply intersubjective function of adversative connectives" (Giacalone Ramat & Mauri 2011:659).[17]

In this case, the evolution is from syntax to discourse; as seems to be the case of *bweta* in (37) and (38) as well as of *ta* in (33).

17. In this regard, Giacalone Ramat and Mauri (2011:659) have pointed out that "Adversative connectives are crucial to the expressive potential of speakers, and therefore speakers are

The situation is different in the case of *bwe'ituk*. The genesis of this new cause/reason connective is less recent (the first appearance is from 1830). This formation implies that the meaning of the clause-combining device has become totally explicit. If we consider the multifunctionality of the final particle *teca* in Colonial Cahita, the evolution associated with the *bwe'ituk* formation was from a multifunctional structure (temporal, conditional, purpose and causal interpretations) to a monofunctional structure (only causal interpretation),[18] from economy (hidden complexity) to explicitness (overt complexity). This change thus illustrates an evolutionary process of explicitness-driven maturation (Dahl 2004; Bisang 2013), resulting in a language structure that forces the speaker to overtly express certain grammatical categories (obligatoriness) and that provides a rich inventory of fine-grained grammatical categories. This maturation process might be associated with the written language since, as Mithun (1988: 357) and Kortmann (1997: 46, 2001: 850) have pointed out, explicit linking devices are especially frequent in written language (recall that the first appearance of *bwe'ituk* comes from a letter). The contrary is usually true in spoken discourse. As Giacalone Ramat and Mauri (2011:657) have correctly stated, "in spoken discourse, the situational context (intonation, extra-linguistic cues, etc.) helps in defining the nuances that language may miss, but in written texts language is the only tool available to establish and infer interclausal relations" (cf. also Meillet 1958 [1921]). The planned language associated with the written use has probably influenced the creation and the consolidation of this new explicit marker of interclausal causality in Yaqui.[19] Interestingly, in spoken discourse, the Spanish-borrowed connective *po(r)ke* is sometimes used, whereas it never appears in the Yaqui written texts that have been consulted for this study.

Additionally, we can observe that the recruitment of a spoken discourse marker (the conversational marker *bwe*) in search of explicitness was not just accompanied by more overt complexity, but also by the loss of the intersubjective/pragmatic meaning associated with the discursive uses of *bwe*, which has reduced its scope from discourse to sentence, although some residual uses of *bwe'ituk* as a discourse marker are still found.

---

constantly in search of new and expressive ways of conveying contrast, determining a high synchronic intra-linguistic variation and a quicker renewal."

18. This evolution is in accordance with the general tendency that has been observed in the semantic development of adverbial conjunctions, that is, the tendency away from polysemy to monosemy (Kortmann 2001: 849).

19. In the consolidation process, the influence of literacy is fundamental. As Miller (2006) has argued, complex constructions are acquired after the age of seven and much later, through the strong influence of literacy.

The creation of this new explicit cause/reason connective has also caused two syntactic rearrangements in the expression of cause/reason adverbial clauses, which affect: (i) the encoding of the adverbial clause subject (from zero (SS) or accusative (DS) to nominative marking), and, therefore, the old switch-reference system that has been lost, (ii) the position of the connective within the adverbial clause (from clause-final to clause-initial position).

The first syntactic change corresponding to the loss of the switch-reference system is probably associated with the recruitment of *bwe* since this discourse marker is usually combined with nominative subjects, introducing independent clauses. This change in the adverbial clause subject marking clearly indicates a lower degree of syntactic integration, since the evolution was from a deranked adverbial clause to a balanced adverbial clause, implying then less syntactic dependency, less embedding, and, therefore, a change from more to less clause integration. If we consider now that explicitness implies a higher degree of semantic dependency (only one possible interpretation of the interclausal meaning relationships, that is, semantic monofunctionality, monosemy), the decrease of syntactic dependency is then associated with the increase of semantic dependency.

As for the second syntactic modification, the change from clause-final to clause-initial position is obviously caused by the recruitment of *bwe* that, in accordance with its function as a discourse topic-shifter, always occupies the chain-initial position. The first uses of *bwe'ituk* were probably in this chain-initial position,[20] where topic shifting usually occurs (Givón 2009). Nowadays, the most frequent uses of *bwe'ituk* are in chain-medial position, that is, after the main clause, in the adverbial clause-initial position, before the new cause/reason topic. This adverbial clause-initial position goes, however, against the general tendency that states that OV languages tend to employ adverbial conjunctions in clause-final position (Kortmann 1997: 71, 2001: 852). As seen above, this apparently inconsistent initial position is due to the grammaticalization path that has created *bwe'ituk* out of the discourse marker *bwe*.

Lastly, it has been shown that the development of the *bweta* and *bwe'ituk* particles involves not only an evolution from discourse to syntax (where *bwe* is concerned), but also, in the early stages, an apparent evolution from syntax to discourse (where the *ta* and *ituca* particles are concerned). However, this syntax > discourse part of the transition, in which the process of insubordination (Evans 2007) seems to play an important role,[21] does not really represent a counterexample

20. Recall the comment made by Dedrick and Casad (§6.1) about the initial position of the *bwe'ituk* adverbial clause.

21. As Rojas (this volume) points out, cross-linguistic studies of child language have shown that syntactic complexity is spread out among participants involved in conversations, and that

to the 'unidirectionality' of grammaticalization, since as Traugott (2004) has convincingly demonstrated, early grammaticalization implies pragmatic strengthening and subjectification. This paper has revealed that the early combination of the *ta* and *ituca* particles with the discourse marker *bwe* has precisely this function of reinforcing the speaker attitude and the pragmatic purposes of contrast.

Acknowledgements

Two anonymous reviewers are acknowledged for their constructive comments and suggestions. Any remaining errors are, of course, my sole responsibility.

Abbreviations

ACC	accusative	EMP	emphatic marker	PASS	passive
ADVZ	adverbializer	FUT	future	PFV	perfective
CAUS	causative	IMPFV	imperfective	PL	plural
COM	comitative	INCH	inchoative	POSS	possessive
COND	conditional	INST	instrumental	RED	reduplication
COMP	completive	LOC	locative	REL	relativizer
DAT	dative	NEG	negation	RES	resultative
DEM	demonstrative	NOM	nominative	SG	singular
DES	desiderative	NMLZ	nominalizer	SS	same subject
DET	determiner	OBL	oblique	SUB	subordinator
DIR	directional	OPT	optative	TEMP	temporal
DS	different subject	PART	particle	VBZ	verbalizer.

insubordinated clauses are used before complex clauses in child language development. For instance, concerning the cause/reason and adversative connectives, Rojas mentions the case of Spanish in which, at between 24 and 30 months, children produce causal *porque* 'because' clauses only in answering *¿por qué*? 'why' questions, and clausal conjuncts marked by *pero* 'but' occur in isolation in a conversational turn, forming a sequence with the interlocutor's antecedent move. Similar conversational effects have been found in Dutch or German between *waarum* questions and *omdat* responses, or in English *why-because* sequences, or *but* constructions. Here again, ontogeny (language acquisition) and diachrony (historical change) seem to represent two parallel developments in the genesis of syntactic complexity.

References

Álvarez González, Albert. 2009. Origen y desarrollo de marcadores de cláusulas adverbiales en yaqui. *Paper presented during the Seminario de Complejidad Sintáctica*, Universidad de Sonora, Hermosillo, México, 12–13 November.

Álvarez González, Albert. 2010. Predicación secundaria en yaqui: Resultativa y depictiva. In *Estudios de lenguas amerindias,* 2: *Contribuciones al estudio de las lenguas originarias de América*, Estrada Fernández Zarina & Ramón Arzápalo Marín (eds), 17–39. Hermosillo: Universidad de Sonora.

Álvarez González, Albert. 2018. Apuntes sobre el Manual para administrar a los indios del idioma cahita los Santos Sacramentos. In *Las lenguas de México: Diálogos historiográficos*, Bárbara Cifuentes & Rodrigo Mártinez Baracs (eds), 211–232. Mexico: IIB UNAM/ Somehil.

Bisang, Walter. 2009. On the evolution of complexity – Sometimes less is more in East and mainland Southeast Asia. In *Language Complexity as an Evolving Variable*, Geoffrey Sampson, David Gil & Peter Trudgill (eds), 34–49. Oxford: OUP.

Bisang, Walter. 2013. Grammaticalization in languages with a high degree of hidden complexity: Two types of maturation and two types of hidden complexity. *Paper presented during the Seminario de Complejidad Sintáctica*, Universidad de Sonora, Hermosillo, México, 11–12 November.

Bisang, Walter. 2014. Overt and hidden complexity – Two types of complexity and their implications. *Poznan Studies in Contemporary Linguistics* 50(2): 127–143. https://doi.org/10.1515/psicl-2014-0009

Brody, Jill. 1987. Particles borrowed from Spanish as discourse markers in Mayan language. *Anthropological Linguistics* 29: 507–521.

Buelna, Eustaquio (ed.). 1890. *Arte de la Lengua Cahita*. México: Imprenta del Gobierno Federal.

Buitimea Valenzuela, Crescencio. 2007. *Peesio betana nottiwame. Regreso de Hermosillo*. Hermosillo: Unison.

Carlson, Robert. 1987. Narrative connectives in Supyire. In *Coherence and Grounding in Discourse* [Typological Studies in Language 11], Russell Tomlin (ed.), 1–19. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.11.02car

Clayman, Steven E. 2012. Address terms in the organization of turns at talk: The case of pivotal turn extensions. *Journal of Pragmatics* 44: 1853–1867. https://doi.org/10.1016/j.pragma.2012.08.001

Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Chicago IL: Chicago University Press.

Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity* [Studies in Language Companion Series 71]. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.71

Dedrick, John & Casad, Eugene H. 1999. *Sonora Yaqui language structure*. Tucson AZ: The University of Arizona Press.

Dedrick, John. 1985. Las cartas en Yaqui de Juan Bandera. *Tlalocan* X: 119–187

Diessel, Holger. 1996. Processing factors of pre- and postposed adverbial clauses. *Berkeley Linguistics Society* 23: 71–82. https://doi.org/10.3765/bls.v22i1.1344

Diessel, Holger. 2001. The ordering distribution of main and adverbial clauses. *Language* 77(2): 433–455. https://doi.org/10.1353/lan.2001.0152

Estrada Fernández, Zarina & Álvarez González, Albert. 2008. *Parlons Yaqui*. Paris: L'Harmattan.

Estrada Fernández, Zarina, Buitimea Valenzuela, Crescensio, Gurrola Camacho Adriana E., Castillo Celaya, María Elena & Carlón Flores, Anabel. 2004. *Diccionario yaqui-español y textos: obra de preservación lingüística*. México: Plaza y Valdés.

Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and Empirical Foundations*, Irina Nikolaeva (ed.), 366–431. Oxford: OUP.

Evans, Nicholas. 2009. Insubordination and the grammaticalisation of interactive presuppositions. *Paper presented during Methodologies in determining morphosyntactic change*, Osaka, Japan, March 2009. <http://www.r.minpaku.ac.jp/ritsuko/english/symposium/pdf/symposium_0903/Evans_handout.pdf>

Ford, Cecilia E., & Thompson, Sandra A. 1986. *Conditionals in discourse: A text-based study from English*. In *On Conditionals*, Elizabeth Closs Traugott, Alice ter Meulen, Judy Snitzer Reilly, & Charles A. Ferguson (eds), 353–373. Cambridge: CUP. https://doi.org/10.1017/CBO9780511753466.019

Frajzyngier, Zygmunt. 1988. Polish conjunctions and switch reference. Ms, University of Colorado at Boulder.

Fraser, Bruce. 1990. An approach to discourse markers. *Journal of Pragmatics* 14: 383–395. https://doi.org/10.1016/0378-2166(90)90096-V

Giacalone Ramat, Ana & Mauri, Caterina. 2011. The grammaticalization of coordinating interclausal connectives. In *The Oxford Handbook of Grammaticalization*, Heiko Narrog & Bernd Heine (eds), 656–667. Oxford: OUP.

Givón, T. 2001. *Syntax. An introduction*. Amsterdam: John Benjamins.

Givón, T. 2009. Discourse-structure constraints on the grammaticalization of clausal conjunctions. Paper presented during the *Seminario de Complejidad Sintáctica,* Universidad de Sonora, Hermosillo, México, 12–13 November.

Guerrero, Lilián. 2019. Adverbial subordinators in Yaqui. In *Diverse Scenarios of Syntactic Complexity* [Typological Studies in Language 126], Albert Álvarez González, Zarina Estrada Fernández & Claudine Chamoreau (eds). Amsterdam: John Benjamins. (this volume)

Hayashi, L. 1989. Conjunctions and referential continuity. Ms, University of Oregon at Eugene.

Hernández Doode, Gabriela. 2002. Marcadores discursivos: Una aproximación en la narrativa yaqui. MA Thesis, Universidad de Sonora, Hermosillo.

Heine, Bernd & Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511613463

Hopper Paul J. & Traugott, Elizabeth. Closs. 1993. *Grammaticalization*. Cambridge: CUP.

Johnson, Jean. 1962. *El idioma yaqui*. México: INAH.

Kortmann, Bernd. 1997. *Adverbial subordination: A Typology and History of Adverbial Subordinators Based on European Languages*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110812428

Kortmann, Bernd. 2001. Adverbial conjunctions. In *Language Typology and Language Universals*, Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, & Wolfgang Raible (eds), 842–854. Berlin: Mouton de Gruyter.

Kuteva, Tania. 2017. The "mirror" of insubordination. Paper presented at the *Linguistics Departamental Seminar*, SOAS, University of London, 13 June 2017. <https://www.youtube.com/watch?v=Jfq1KcoRens&t=7s> (6 November 2017).

Li, Charles. 1988. Conjunctions as switch-reference devices in Green Hmong. Ms, University of California at Santa Barbara.

Lindenfeld, Jacqueline. 1973. *Yaqui Syntax.* Berkeley CA: University of California Press.

Lionnet, Andrés. 1977. *Los elementos de la lengua cahita*. México: UNAM.

Martin Zorraquino Maria Antonia & Portalés Lázaro, José. 1999. Los marcadores discursivos. In *Gramática descriptiva de la lengua española*, Vol. 3, Ignacio Bosque & Violeta Demonte (eds), 4051–4213. Madrid: Espasa Calpe.

Meillet, Antoine. 1958[1921]. *Linguistique historique et linguistique générale*. Paris: Champion.

Miller, Jim. 2006. Influence of literacy on language development. In *Encyclopedia of Language & Linguistics*, 2nd edn, Keith Brown (ed.), 661–665. Oxford: Elsevier. https://doi.org/10.1016/Bo-08-044854-2/00853-1

Miller, Wick R. 1984. The Classification of the Uto-Aztecan Languages Based on Lexical Evidence. *International Journal of American Linguistics* 50(1): 1–24.

Mithun, Marianne. 1988. The grammaticization of coordination. In *Clause Combining in Grammar and Discourse* [Typological Studies in Language 18], John Haiman & Sandra A. Tompson (eds), 331–359. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.18.13mit

Myers-Scotton, Carol. 1993. *Duelling Languages: Grammatical Structure in Code-switching*. Oxford: Clarendon Press.

Onodera, Noriko. 2011. The grammaticalization of discourse markers. In *The Oxford Handbook of Grammaticalization*, Heiko Narrog & Bernd Heine (eds), 614–624. Oxford: OUP.

Pfaff, Carol. 1982. Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. In *Spanish in the United States*, Jon Amastae & Lucía Elías-Olivares (eds), 264–297. Cambridge: CUP.

Rojas-Nieto, Cecilia. 2019. Constructional grounding in emerging complexity: Early COMP-*que* constructions in Spanish acquisition. In *Diverse Scenarios of Syntactic Complexity* [Typological Studies in Language 126], Albert Álvarez González, Zarina Estrada Fernández & Claudine Chamoreau (eds). Amsterdam: John Benjamins. (this volume)

Schiffrin, Deborah. 1987. *Discourse markers.* Cambridge: CUP. https://doi.org/10.1017/CBO9780511611841

Silva Encinas, Carlos, Álvarez Romero, Pablo, & Buitimea Valenzuela, Crescencio (eds). 1998. *Jiak nikpoe'tejoim, pláticas en Yaqui*. Hermosillo: Universidad de Sonora.

Stassen, Leon. 1985. *Comparison and Universal Grammar: An Essay in Universal Grammar*. Oxford: Blackwell.

Thompson, Sandra A. & Longacre, Robert E. 1985. Adverbial clauses. In *Language Typology and Syntactic Description*, Vol. II: *Complex Constructions*, Timothy Shopen (ed.), 171–234. Cambridge: CUP.

Torres, Lourdes. 2002. Bilingual discourse markers in Puerto Rican Spanish. *Language in Society* 31: 65–83. https://doi.org/10.1017/S0047404502001033

Traugott, Elizabeth Closs. 1988. Pragmatic strengthening and grammaticalization. In *Proceedings of the Fourteenth Annual Meeting of he Berkeley Linguistics Society*, Shelley Axmaker, Annie Jaisser & Helen Singmaster (eds), 406–416. Berkeley CA: BLS.

Traugott, Elizabeth Closs. 1995. Subjectification in grammaticalisation. In *Subjectivity and Subjectivisation*, Susan Wright & Dieter Stein (eds), 31–54. Cambridge: CUP. https://doi.org/10.1017/CBO9780511554469.003

Traugott, Elizabeth Closs. 2004. Le rôle de l'évolution des marqueurs discursifs dans une théorie de la grammaticalisation. In *Structure informationnelle et Particules énonciatives. Essai de typologie*, M. M. Jocelyne Fernandez-Vest & Shirley Carter-Thomas (eds), 295–334. Paris: L'Harmattan.

Waltereit, Richard & Detges, Ulrich. 2007. Different functions, different histories: Modal particles and discourse markers from a diachronic point of view. *Catalan Journal of Linguistics* 6: 61–80. https://doi.org/10.5565/rev/catjl.124

Authors index

Language index

Subject index

This volume surveys the phenomenon of syntactic complexity in a diversity of languages and from a diversity of theoretical perspectives. The topics include clause combining strategies such as relative, complement, and adverbial clauses, serialization, clausal nominalizations, but also the switch reference systems involved in clause chains, the role of insubordination and the influence of language contact in the development of syntactic complexity as well as the acquisition of complex clauses in child language and the grammaticalization processes leading to syntactic complexity. These studies illustrate the varied aspects involved in clause combining and help to understand how syntactic complexity works and evolves in the world's languages, how it varies across languages, how it is influenced by language contact, how it is acquired. As such, this book gives the opportunity for readers to expand both their typological and their theoretical knowledge about syntactic complexity in a variety of languages.

ISBN 978 90 272 0343 4

John Benjamins Publishing Company