

STUDIES IN
LANGUAGE
COMPANION
SERIES 219

Studies at the Grammar-Discourse Interface

Discourse markers and discourse-related
grammatical phenomena

Edited by
Alexander Haselow
Sylvie Hancil

JOHN BENJAMINS PUBLISHING COMPANY

Copyright © 2021, John Benjamins Publishing Company. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair use permitted under U.S. or applicable copyright law.

Studies at the Grammar-Discourse Interface

Studies in Language Companion Series (SLCS)

ISSN 0165-7763

This series has been established as a companion series to the periodical *Studies in Language*.

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

Founding Editor

Werner Abraham
University of Vienna / University of Munich

Editors

Werner Abraham
University of Vienna / University of Munich

Elly van Gelderen
Arizona State University

Editorial Board

Bernard Comrie
University of California, Santa Barbara

William Croft
University of New Mexico

Östen Dahl
University of Stockholm

Gerrit J. Dimmendaal
University of Cologne

Ekkehard König
Free University of Berlin

Christian Lehmann
University of Erfurt

Elisabeth Leiss
University of Munich

Marianne Mithun
University of California, Santa Barbara

Heiko Narrog
Tohoku University

Johanna L. Wood
University of Aarhus

Debra Ziegeler
University of Paris III

Volume 219

Studies at the Grammar-Discourse Interface
Discourse markers and discourse-related grammatical phenomena
Edited by Alexander Haselow and Sylvie Hancil

Studies at the Grammar-Discourse Interface

Discourse markers and discourse-related
grammatical phenomena

Edited by

Alexander Haselow

University of Münster

Sylvie Hancil

University of Rouen

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/slcs.219

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2021005956 (PRINT) / 2021005957 (E-BOOK)

ISBN 978 90 272 0869 9 (HB)

ISBN 978 90 272 5989 9 (E-BOOK)

© 2021 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

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

Table of contents

Grammar, discourse, and the grammar-discourse interface <i>Alexander Haselow and Sylvie Hancil</i>	1
Part I. Discourse markers: The grammar-discourse interaction from a developmental perspective	
CHAPTER 1	
On the rise of discourse markers <i>Bernd Heine, Gunther Kaltenböck, Tania Kuteva and Haiping Long</i>	23
CHAPTER 2	
On the pragmatic development of modal particles in Navarrese-Lapuradian Basque: <i>Hori emain ote nauzu?</i> <i>Sergio Monforte</i>	57
CHAPTER 3	
On divergent paths and functions of ‘background’-based discourse markers in Korean <i>Seongha Rhee and Hyun Jung Koo</i>	77
CHAPTER 4	
Reanalysis and the emergence of adverbial connectors in the history of Japanese <i>Reijirou Shibasaki</i>	101
Part II. Discourse markers: The grammar-discourse interaction from a synchronic, usage-based perspective	
CHAPTER 5	
The meaning and functions of French <i>je pense (que)</i> : A constructionalist and interactional account <i>Juliette Angot and Maj-Britt Mosegaard Hansen</i>	127
CHAPTER 6	
Discourse markers and brain lateralization: Evidence for dual language processing from neurological disorders <i>Alexander Haselow</i>	157

CHAPTER 7		
Vietnamese expletive between grammatical subject and subjectivity marker: <i>Nó</i> at the syntax-pragmatics (discourse) interface		195
	<i>Huy Linh Dao</i>	
CHAPTER 8		
The final particle <i>like</i> in Northern English: A particle of reformulation in the context of interenunciative readjustment		229
	<i>Sylvie Hancil</i>	
CHAPTER 9		
On pragma-semantics of expressives: Between words and actions		245
	<i>Suren Zolyan</i>	
Part III. Discourse-related grammatical phenomena		
CHAPTER 10		
A just amazing marker in French: “Juste”: “Juste excellentissime!” “Juste super heureux!” “Juste irréal!” “Juste pas possible!”		275
	<i>Danh-Thành Do-Hurinville</i>	
CHAPTER 11		
On how the distinction between reciprocal and collective verbs affects (anti-)control		299
	<i>Anna Snarska</i>	
CHAPTER 12		
The rise of cause/reason adverbial markers in Yaqui (Uto-Aztecan)		313
	<i>Albert Álvarez González</i>	
Index		353

Grammar, discourse, and the grammar-discourse interface

Alexander Haselow and Sylvie Hancil
University of Münster / University of Rouen

1. Introduction

‘Discourse’ and ‘grammar’ are distinct, but not strictly separate levels of language structure and cognitive linguistic activity, given that they interact and complement each other in various ways. Traditionally, each of these levels has been associated with a different set of principles that guide a speaker’s verbal behavior: while ‘grammar’ is typically conceived of as the domain responsible for establishing relations on a local (e.g. phrasal or clausal) level, providing the machinery required for creating “well-formed” morphosyntactic and semantic structures, ‘discourse’ refers to the set of principles according to which these grammatical structures are integrated under a more global, coherent piece of text, that is, a meaningful contribution in a specific communicative context, including the different ways in which pieces of text (e.g. turns, utterances) serve the management of interpersonal relationships. However, the simple fact that speakers need to satisfy structural constraints on both levels at every moment in a communicative encounter in order to be comprehensible (Ariel 2009) implies that at some point both need to converge into one structure in a single communicative act. Hence, the principles and constraints holding on each of the two levels “meet” or “overlap” at various points of intersection. Under this view, the idea that at some point grammatical relationships end (e.g. at sentence boundaries) and discourse relationships take over and that such a conceptual dividing line would block principles from one domain to be productive in the other is difficult to maintain.

The relatively large – and still growing – body of literature on phenomena at the grammar-discourse intersection has expanded our understanding of the various ways in which grammar and discourse, or grammar and ‘pragmatics’ (a term that is often used to refer to aspects of discourse structure) interact (see e.g. the important studies by Du Bois [2003] and Ariel [2008, 2009]). One the most important findings in this respect is certainly the close relation between discourse goals, grammatical

principles, and frequency: discourse affects grammar in that speakers frequently choose those grammatical forms and structures that serve their specific discourse goals best, that is, discourse goals guide speakers in selecting specific grammatical forms and structures at different points in communicative interaction; grammar, in turn, interacts with or reflects discourse in that grammars often conventionalize what speakers do most often in discourse (e.g. Du Bois 1987; Bybee 2006), i.e. a subset of discourse patterns may (and often does) turn grammatical (see Ariel 2009). Hence, this interaction is also an important driving force in language change, given that such changes, which may proceed in or extend into two directions ('grammar->discourse' or 'discourse->grammar'), typically affect patterns or constructions that serve specific discourse purposes. In this sense, both systems closely interact in widely encompassing ways, including potentially all aspects of the language system in use.

This volume seeks to contribute to research on the interaction between grammar and discourse, providing 12 case studies that look at different phenomena at the grammar-discourse interface from different (e.g. historical, (neuro)cognitive, or interactional) perspectives. Most of these studies deal with linguistic expressions that can be subsumed under the umbrella term 'discourse marker' (DM) in the widest sense (e.g. Schiffrin 1987; Fischer 2006; Fraser 1999; Heine 2013), which refers to linguistic signs and structural units whose use is conditioned by the discourse situation rather than by morphosyntactic and semantic aspects and which encompasses elements that are characterized by a catalogue of formal and functional features, the prototypes of these expressions showing e.g. (a) morphological invariability, (b) syntactic independence from their environment, (c) relational or linking functions "outside" (sentence-)grammatical relationships, and (d) expression of procedural rather than lexical-conceptual meaning in the sense of Blakemore (2002), i.e. they serve as processing cues for utterance interpretation. Functionally, they relate an utterance to the situation of discourse, that is, they contribute to the organization of (spoken or written) text, to different aspects of speaker-hearer interaction, and to the expression of speaker attitudes. The focus on discourse markers in a volume on the grammar-discourse interface is natural, given that these elements are a showcase example of the intersection of grammar and discourse: from a historical-developmental perspective, DMs show a strong tendency to derive from elements that once served as sentence-internal constituents (e.g. adverbs) and that were transferred from the sentence-grammatical to the discourse-structural domain (e.g. Heine 2013; Brinton 2017); some of them represent borderline cases in that they are relevant for sentence-internal relations, but in many uses also cross the sentence boundary and serve relations in discourse (e.g. English *and*, *but*, *so*); from a structural perspective the question is how much

grammar is involved in their use (e.g. in terms of positional constraints), and from a functional point of view it is not entirely clear whether their (discourse-)structuring function is part of the grammar of a language or a (discourse-)pragmatic phenomenon. Nine of the thirteen contributions to this volume include a discussion of one or more of these questions and related issues, looking at different types of DMs (including e.g. final particles, expletives, adverbial connectors, or expressives) from different perspectives and based on data from various languages such as English, Basque, French, Korean or Japanese, thus offering a broader picture by including grammatical systems that differ fundamentally from each other. Some of these contributions deal with DMs from a historical-developmental perspective, providing detailed descriptions of the transition of particular linguistic signs from grammar “into” discourse, others deal with the borderline status of particular linguistic units that may be used as either clause-internal constituents or as discourse markers “outside” the clause (see Section 3).

In addition to the study of DMs, this volume offers a discussion of phenomena that are more closely associated with the grammar of a language (thus addressing morphosyntactic and semantic phenomena), but that cannot be plausibly explained without taking specific discourse constraints into consideration and which are therefore equally interesting for a better understanding of the grammar-discourse interface. The three chapters that are devoted to such phenomena focus on (i) the uses of an adverb (French *juste*), which may either modify or modalize an adjective on the sentence-grammatical level or an utterance as a whole on the discourse level, (ii) the phenomenon of ‘Partial Control’ in the context of collective predicates and reciprocal verbs, and on (iii) the development and uses of three adverbial markers of cause/reason (*betchi’ibo*, *bwe’ituk*, and *po(r)ke*) in Yaqui (Uto-Aztecan), whose use cannot be explained by referring to sentence structure alone. Looking at these studies, it appears that many grammatically available options or grammatical relations receive their concrete interpretation in a particular discourse context. This way, the meaning or function of certain structural patterns is not just dictated by morphosyntactic and semantic principles, but (also) the product or by-product of discourse-specific strategies pursued by a speaker.

We will now take a closer look at the main research questions dealt with in the three main parts of this volume and integrate these questions in a broader context by linking them to current discussions in relevant fields of research.

2. Current issues in research on the grammar-discourse interface

2.1 Part I – Discourse markers: The interaction of grammar and discourse from a historical-developmental perspective

Despite extensive research on the function and uses of discourse markers (DMs) (e.g. Schourup 1985; Schiffrin 1987; Lenk 1998; Jucker & Smith 1998; Fraser 1999, 2015; Aijmer 2002; Fischer 2006; Heine 2013; Degand, Cornillie & Pietrandrea 2013; Crible 2018; Haselow 2019) and the enormous progress that has been made in this research line during the past four decades as regards our understanding of the nature of these linguistic elements, many aspects are still waiting to be explored. One of these concerns the processes by which elements of grammar can be functionalized for discourse purposes and thus come to cross the border from one domain to the other over time. The problem is discussed in the first four studies in this volume, which address two major questions: (a) how can we explain the transition of particular linguistic forms “out” of grammar and “into” discourse, and (b) how exactly does this transition occur in different languages?

A. **Explaining the development or “rise” of DMs.** DMs are often said to be the result of grammaticalization, that is, of a unidirectional development of a linguistic sign from a lexical to a grammatical form involving a set of changes on different levels of the language system, above all semantic change (loss or “bleaching” of semantic features), phonological reduction (loss of phonetic substance), loss of the morphosyntactic constraints of the original word class, host-class expansion, syntactic expansion, and sometimes also loss of formal independence. These changes have been described in terms of different, but comparable sets of parameters (e.g. Heine & Kuteva 2007; Lehmann 2015). While many of these parameters can be observed in the development of DMs in various languages, the grammaticalization approach poses a serious problem: DMs are typically outside the confines of sentence grammar and thus syntactically as well as semantically independent of their (morpho)syntactic environment. Hence, the question arises how we can account for the fact that the processes behind the development or “evolution” of DMs did not lead into grammar under the traditional conceptualization of the term as referring to sentence-internal structuring principles, but “out of” grammar in the sense that the respective linguistic signs that are no longer bound to the constraints that determine the grammaticality of linguistic units (e.g. sentences), but are subjected to usage-conditions that hold on the discourse level. A proposal for answering this question is made in Heine, Kuteva, Kaltenböck & Long’s contribution (this vol.), whose concept of *cooptation* seeks to account for the manifold cases in which what was once a sentential constituent (e.g. an adverb

or a clausal fragment) is reanalyzed as an element with relative, or even full, structural-semantic independence of a host structure. *Cooptation* means that, similar to parentheticals in the classical sense, units of language can be and are “taken out” of sentence grammar at any time in speech production and instantaneously come to serve functions on the discourse or the interactional level (see also Kaltenböck, Heine & Kuteva 2011: 874–875; Heine et al. 2017: 828). Thus, long-term processes such as grammaticalization are not required for the transition from (sentence) grammar to discourse *per se*, but may optionally follow cooptation and lead to later conventionalization as well as to formal and functional changes of a “coopted” unit.

- B. **Typological issues.** Much of what we know about the development and the use of DMs is based on the study of Indo-European languages, especially English. This makes it difficult to see whether geographically and/or typologically distinct languages show similarities or differences as regards the processes underlying the rise of particular types of DMs. A very basic question is, for instance, which syntactic categories typically serve as starting points for DM development in different languages. While in most Indo-European languages DMs are typically based on adverbs, adverbials and conjunctions, in Asian languages the closest equivalents to DMs tend to derive from clausal connectives (Thompson & Suzuki 2011; Rhee 2012) and from verbal particles, e.g. evidential markers (e.g. Strauss 2005) or quotative markers (Rhee 2016). However, while the sources or starting points of the development of DMs differ, this does not exclude the possibility that the processes leading to DM status are, in some way, comparable across languages. The contribution by **Sergio Monforte** (this vol.) on the development of Basque *ote* into a pragmatic marker, the one by **Seongha Rhee & Hyun Jung Koo** (this vol.) on the development of ‘background’-based discourse markers in Korean, and the study by **Reijirou Shibasaki** (this vol.) on the rise of syntactically independent (“stand-alone”) adverbial connectors with projecting functions in the history of Japanese are insightful in this respect and allow for new hypotheses on the nature of the processes underlying the emergence of more or less prototypical instances of DMs in distinct languages.

2.2 Part II – Discourse markers: The grammar-discourse interaction from a synchronic, usage-based perspective

The language system does not only provide principles for combining words into structured syntactic units on the sentence-grammatical level, in which these words function as clausal “constituents” and are, as such, formally linked to other constituents within the same structural unit, but also linguistic means for the creation of structures beyond such units on the discourse level, such as DMs. There are

various tests that can be used to demonstrate that these expressions do not form part of an “integrated” morphosyntactic unit with the surrounding segments and thus appear to be “outside” sentence grammar. They are, for instance, unable to be the focus of a cleft-sentence (1a), they cannot be questioned (1b), negated (1c) or coordinated with other expressions or constituents (1d), and they can be omitted without affecting the grammaticality and the meaning of the rest of the structure (1e), as illustrated for *well* in the utterance *well I started off applying for jobs* in (1).

- (1) *well* I started off applying for jobs [...]
- a. *It is **well** that I started off applying for jobs.
(compare to e.g. it is a **job** that I started off applying for)
- b. *How/What is it? – ***well**
(compare to: What is it (that you started off applying for)? – a **job**)
- c. *It is not **well** that I started off applying for, but...
(compare to: It is not a **JOB** that I started off applying for, but...)
- d. ***well** and **so** I started off applying for jobs
(compare to: I started off applying for **jobs** and **scholarships**.)
- e. \emptyset I started off applying for jobs.

In spite of these properties, DMs are an integral part of the unit they accompany as they integrate the structural unit with which they occur in surrounding discourse or the communicative situation in which it is produced in a broader sense, thus doing “metatextual work” (Traugott 1997). Their scope is not restricted to a clause, but usually over units of discourse, which may vary in size and thus allow for the establishment of more global structural relations between discourse units (Schiffrin 1987; Lenk 1998). They serve the creation of structure outside formal dependency relationships, linking several segments of discourse into a coherent whole.

The fact that DMs are outside sentence-grammatical relations does, however, not necessarily mean that they are outside the grammar of a language in general: they have been accounted for in many functional approaches to grammar, most of which solve the problem of integrating them into ‘grammar’ by distinguishing between a domain referring to clausal grammar, which basically encompasses clausal constituents, the syntactic machinery required for their linearization into clausal configurations, and semantic relationships, and a domain comprising linguistic material that is not part of this machinery and thus operating “outside” clausal structures. The distinction surfaces, for instance, in Dik’s (1997) *Functional Grammar*, which distinguishes between clausal constituents and so-called “extra-clausal constituents” (ECCs), the latter of which include e.g. interjections, discourse markers, backchannels, tag questions, forms of address, or formulae of social exchange (e.g. greetings, summonses) (Dik 1997: 379–401). It is also found in Huddleston & Pullum’s (2002: 1350–1362) academic grammar, in which a distinction is made

between clausal structures and *supplements*, the latter of which are morphosyntactically unintegrated into the host construction and thus unable to function as dependents to any head. This kind of grammatical dualism is also at the heart of Kaltenböck, Heine & Kuteva's (2011) model of *Discourse Grammar*, which distinguishes *Sentence Grammar* and *Thetical Grammar*, and Haselow's (2017) model of grammar, which distinguishes between *microgrammar* and *macrogrammar*. In all these models, grammar and discourse (or 'more grammatical' vs. 'more discourse-related' aspects of language structure) are conceptualized as separate, but closely interacting domains of language structure and language-related cognitive activity, and it is precisely this interaction that generates a number of questions, some of which are addressed in one or more of the four contributions in the second part of this volume.

- A. **The position of DMs.** An aspect that is immediately relevant for discussions of the grammar-discourse interface is the distribution of DMs within a discourse unit (e.g. a unit of talk or a turn). Since DMs are outside the constraints of sentence grammar, their position is relatively flexible, with a strong tendency for turn- or utterance-initial uses. However, initial use is far from being universal, not even within a single language such as English: while some DMs do occur only at the beginning of a discourse unit (e.g. *oh*), others are used both initially and finally (e.g. *anyway*, with subtle differences in meaning, though), and again others may be used initially, medially and finally, e.g. *actually, you know* or *I think*. The predominance of initial and final positioning suggests that the topological preferences of DMs are related to their communicative function and to discourse structure (see e.g. the contributions in Beeching & Detges 2014; also Traugott 2015), given that the two ends or "peripheries" of an utterance, a turn or a unit of talk are important transition points from one unit of discourse to another, where the risk of disruptions in the structural coherence of emergent discourse is relatively high. Correlations between specific DM functions and the position of a DM within and at the peripheries of an utterance are, for instance, addressed in **Juliette Angot & Maj-Britt Mosegaard Hansen's** study (this vol.) on French *je pense*. The final position as a structural slot for DMs has attracted much scholarly attention in the last decade (e.g. Haselow 2013; Beeching & Detges 2014; Hancil, Haselow & Post 2015; Traugott 2016), but there is still potential as regards the precise function of DMs at the end of a discourse unit, an aspect that is addressed in **Sylvie Hancil's** study on final *like*.
- B. **The alternation between sentence-grammatical vs. discourse-structural uses of linguistic expressions.** Often one and the same linguistic form can function both as a sentence-internal constituent or as a discourse-structuring element, thus moving on the boundary of grammar and discourse at a given time in

the development of a language. An interesting case in this respect is the optional expletive subject pronoun *nó* in Vietnamese, which is discussed in this volume by **Huy Linh Dao**. *Nó* may be used either as a clausal subject or as a marker with discourse-related functional properties. The observation that there appear to be overlaps in use and function between more grammatical uses (as syntactic constituents) and more discourse-pragmatic uses raises issues in relation to recent proposals concerning the interplay between *Sentence Grammar* and *Discourse Grammar* (Heine 2019) or between two types of cognitive activity called *Type 1 processing*, which concerns higher-level semantic and conceptual processing including discourse- and interaction-related structural relations and meanings, and *Type 2 processing*, which encompasses the structuration of sentences and the semantic content of linguistic units (e.g. phrases or clauses) (Heine, Kuteva & Long 2020). The distinction between grammatical/sentence-structural and discourse-structural aspects of language use and linguistic cognition is also a common thread in neurolinguistic research, where increasing evidence for a basic neuroanatomic difference between the processing of aspects of language related to sentence structure (i.e. morphosyntax and semantics) and those aspects that refer to discourse organization and interaction management has been brought forward over the past three decades. One of the central findings deriving from clinical neuropsychology and neuroimaging studies is the relation between brain lateralization and the processing of specific aspects of language: there is substantial neurological evidence suggesting that processing activities related to sentence structure involve strong neural activity in the left hemisphere (LH) of the human brain (e.g. Grodzinsky 2000), whereas discourse- and interaction-related aspects correspond to neural activation in the right hemisphere (RH) (e.g. Lehman Blake 2006; Myers 1999). This means that hemisphere-specific damage leads to selective impairment of specific linguistic abilities, with LH damage prototypically leading to different forms of agrammatism, i.e. impaired uses of grammatical and lexical resources, whereas RH damage involves impaired abilities as regards the overall structuration of discourse, pragmatic aspects of language use and comprehension (e.g. inferential abilities), and interactional aspects (e.g. paying attention to the listener's needs, interpreting emotive aspects of language use), often leaving sentence-grammatical abilities largely intact (e.g. Marini et al. 2005; Marini 2012; Prat, Long & Baynes 2007; Sherratt & Bryan 2012). This suggests that the processing of DMs as elements "outside" sentence structure is likely to involve neural substrates that are distinct from those involved in the processing of clausal constituents and (sentence-)grammatical relationships. Comparing the incidence of DMs in spontaneous speech produced by patients suffering from either LH or RH damage, **Alexander Haselow** (this vol.) finds evidence

for dual processing mechanisms, based on the empirical observation that DMs tend to be underrepresented in the speech produced by RH-impaired speakers compared to the frequency of use of DMs produced by speakers with diagnosed LH damage.

- C. **Semantic contribution vs. semantic non-restrictedness.** Most DMs or “pragmatic markers” are not essential to the meaning expressed in the structural unit with which they are associated, i.e. their omission does not alter the propositional content of this unit in any way, as illustrated in (1) for *well*. Yet, often the dividing line between semantic contribution (or logical, semantic characteristics) and lack of ability to add to the semantic content of an utterance (due to meanings that are anchored in the communicative situation) is very thin. An interesting discussion on such a case is provided by **Suren Zolyan** in this volume, who develops a theoretical approach to the meaning of expressives, a category that is notoriously difficult to analyze in semantic terms (see e.g. Foolen 2015) as these expressions are intermediate between truth-conditional semantics and communicative-pragmatic usage conditions.

2.3 Part III – Discourse-related grammatical phenomena

As argued above, grammar and discourse complement each other: discourse can only be based on grammar since the units it comprises must be based on a selective use of grammar, and grammar is integrated in emergent discourse since no grammatical unit is produced in a communicative vacuum, but serving specific discourse goals. The interaction between both domains has long been subject to intense research (e.g. Van Dijk 1972, 1977; Du Bois 2003; Ariel 2008, 2009), resulting in the formulation of a wide range of principles and theories such as *Preferred Argument Structure* (PAS, Du Bois 1987, 2003) and in more large-scale attempts at developing a grammar of text (e.g. Mann & Thompson 1988).

The three contributions in the third section of this volume focus on language-specific grammatical phenomena that, each in their own way, illustrate the association between grammatical conventions and discourse, showing that the analysis of particular grammatical phenomena often involves or requires more than the consideration of merely formal syntactic criteria, encompassing the wider discourse context in which they are embedded. **Danh Thành Do-Hurinville** (this vol.) shows that the semantic interpretation of more recent uses of the French adverb *juste* (as e.g. in *c'est juste sensationnel*) is often indeterminate as it may have scope over a single syntactic unit (e.g. an adjective) and over an utterance in general, creating a hyperbolic effect. However, this effect is bound to specific conditions, such as the type of co-occurring element and the discourse environment in which *juste* is used. **Anna Snarska** (this vol.) focuses on the phenomenon of *Partial Control*

and the distinction between reciprocal verbs and collective predicates, illustrating the complexity involved in identifying formally ‘silent’, but semantically implied subjects. Finally, **Albert Álvarez Gonzalez** (this vol.) deals with the development of the adverbial connectors *betchi’ibo*, *bwe’ituk*, and *po(r)ke* in Yaqui, which are used to introduce a clause that provides the cause/reason based on which the situation denoted in the associated clause is carried out. The meaning differences between these three connectors correlate with specific contextual features, i.e. their uses and their interpretation extends, to a certain degree, into the discourse domain.

3. Overview of the papers in this volume

In the opening paper of Part I (“On the rise of discourse markers”), **Heine, Kuteva, Kaltenböck** and **Long** argue that the rise of DMs cannot be described exhaustively in terms of parameters of grammaticalization, an important reason being that any approach that is restricted to grammaticalization cannot account for the problem that, unlike other “products” deriving from this process, DMs are only loosely integrated into the set of constraints that regulate grammaticality. For instance, they do not function as syntactic constituents and their semantic-pragmatic scope is beyond the syntactic unit they accompany. To facilitate comparability of the grammaticalization approach and the proposed alternative – a process called *cooptation* – the paper discusses precisely those data that are used by other authors to support the grammaticalization hypothesis (viz. Degand and Evers-Vermeul [2015] and Brinton [2017]), namely the English DMs *I admit* and *if you will* and the French DM *alors*. Heine, Kuteva, Kaltenböck and Long suggest that DMs are ‘theticals’ in the terminology of Discourse Grammar (Kaltenböck, Heine & Kuteva 2011), which means that following cooptation they serve a wide range of functions relating to the discourse situation.

Sergio Monforte’s paper (“On the pragmatic development of modal particles in Navarrese-Lapurcian Basque. *Hori emain ote nauzu?*”) provides an analysis of the ongoing pragmatic extension of the epistemic particle *ote* in the Navarrese-Lapurcian dialect. Basque has a range of modal particles like other European languages such as German or Italian. They are mainly related to evidentiality and epistemicity, such as evidential *ei* and *omen* or epistemic *ahal*, *bide* and *ote*, some of which have recently been studied from a semantic-pragmatic perspective (Garmendia 2014; Korta & Zubeldia 2016), a syntactic perspective (Etxepare 2010; Etxepare & Uria 2016; Alcázar 2017, Monforte 2018a; Trotzke & Monforte 2019) and from a historical point of view (Monforte 2018b). In his paper, Monforte discusses the emergence of new pragmatic uses of the particle *ote* based on spoken and written data. For instance, when used with information-seeking requests, *ote* serves

politeness in that it mitigates the illocutionary force of a question and allows the addressee to leave the question unanswered. As shown by the author, this development is reminiscent of similar processes going on in other languages, e.g. in Romanian, where the particle *oare* is used to form what Farkas (2017) calls “non-intrusive” questions, and suggests the existence of a cross-linguistic grammaticalization path to be found in typologically unrelated languages.

Drawing upon data taken from historical and contemporary corpora, **Seongha Rhee & Hyun Jung Koo**’s paper (“On divergent paths and functions of ‘background’-based discourse markers in Korean”) provides an analysis of the grammaticalization pathway of the Korean DMs *kulssey* and *kuntey*, both of which originate from adverbial connectives and acquired various discourse functions, *kulssey* serving as a stance marker indicating e.g. uncertainty, reluctance/hesitance, or discontent, *kuntey* being used e.g. as a topic shifter, a prompt for elaboration or as a preface to dispreferred information. The authors discuss the development of the two DMs, which share lexical and grammatical resources (e.g. the verbs *kule-* ‘be so’, *ha-* ‘do’ and the postpositional particle *-ey* ‘at’), including aspects such as (i) the role of the original semantics in the change toward DMs, (ii) prosodic factors (*kulssey* is more sensitive to prosodic variation with respect to functions than *kuntey*, which is prosodically more stable across functions), (iii) the functional differences between the two DMs, and (iv) the meaning differences of each single DM depending on its position. In their most frequent position, which is at the left periphery, their functions are vastly different, whereas in medial positions and at the right periphery they are (nearly) identical, with some subtle differences, though. Such position-dependent functional differences have been discussed for various DMs in recent publications (e.g. Beeching & Detges 2014; Haselow 2016) and appear to be a crucial feature of DMs.

Reijirou Shibasaki’s study (“Reanalysis and the emergence of adverbial connectors in the history of Japanese”) deals with the historical development of Japanese adverbial connectors such as *totan(-ni)*, *shunkan(-ni)*, or *hyoosi(-ni)*, all of which correspond to the meaning ‘at the moment’, with subtle differences. These connectors, which in present-day Japanese occur as independent items in sentence-initial position, have adopted a discourse organizing function in that they anticipate more to come in ongoing discourse, thus serving a projecting function. As Shibasaki shows, in all cases the development is based on reanalysis, more precisely, the reanalysis of temporal nouns that once served as the head of adnominal clauses, and thus of the sequential relation between the preceding clause, especially the form of the verb, and the following nominal head. The development can thus be interpreted as the rise of pragmatic markers from earlier clause-combining conjunctions (see Koyanagi 2018: 115–120; Shibasaki 2018 for related issues). Based on data from Modern and Present-day Japanese corpora, Shibasaki dates the process back to the first half of

the twentieth century, even though the cause of the change (the inflectional leveling of adnominal and conclusive forms) is considered to have started as early as in the Kamakura period (1185–1333). Historical documents are seen as a factor that is likely to have facilitated the reanalysis of adnominal forms as conclusive forms because the well-known modern punctuation style had not yet been established around the end of the nineteenth century (Watanabe et al. 1993), which means that the interpretation of sentence/clause boundaries was often ambiguous, depending on the language user's interpretation, and had gone unsolved for centuries.

Part II begins with a paper by **Juliette Angot** and **Maj-Britt Mosegaard Hansen** ("The meaning and functions of French *je pense (que)*: a Constructionalist and interactional account"), who examine the use of the French construction *je pense* ("I think") in social interaction, where it is frequently used as a pragmatic marker serving interpersonal, face-related as well as discourse-organizational functions. Adopting a Construction Grammar approach, the authors analyze *je pense* as a micro-construction with two basic elements of meaning (one epistemic, the other one subjective), and provide a detailed discussion of its functional value in conversational interaction and, given its syntactic variability, of the relation between the function of *je pense* and its syntactic position. The paper makes an interesting contribution to the discussion of the grammar-discourse divide since a clear distinction between instances where *je pense* functions as a pragmatic marker with discourse-organizational and interactional meanings and uses with propositional meaning often appears impossible, one important reason being that, as the authors show, the process of semantic bleaching has not advanced to such a degree that all of its original semantic content (or propositional meaning) has disappeared.

Alexander Haselow's paper ("Discourse markers and brain lateralization: Evidence for dual language processing from neurological disorders") contributes to ongoing research on the processing of discourse markers (DMs) in the brain, which provides important clues to language organization and the categorization of DMs in linguistic theory. Based on the relatively robust neurolinguistic evidence for left-hemispheric dominance in grammatical processing and right-hemispheric dominance in processing pragmatic, discourse- and interaction-related aspects of language, Haselow examines speech data produced by two groups of neurologically impaired speakers (with either left or right hemisphere damage) compared to unimpaired speakers. The empirical data show that the incidence of DMs is significantly higher in the speech produced by speakers with an impaired left hemisphere compared to their frequency with speakers suffering from right hemisphere damage, approximating the frequency of DMs measured for unimpaired speakers. The observation that DMs are underrepresented in the speech of patients who show relatively preserved syntactic skills (i.e. speakers with unimpaired left-hemispheric activity) suggests that their use is not bound to processing activities related to

sentence-internal structuration, and that discourse organization and interaction management form a separate (but not autonomous) processing domain that is strongly associated with right-hemispheric processing. The data thus provide further evidence for the assumption that linguistic cognition is not organized in a monolithic way, but exhibiting a dualistic organization (see also Heine 2019; Haselow & Kaltenböck 2020), with two interacting domains of language structure and cognitive activity, a microdomain in which local, clause-internal relations between single constituents are processed, and a macrodomain in which global relations outside the clause between segments of discourse and interactional moves are processed.

Huy Linh Dao's study ("Vietnamese expletive between grammatical subject and subjectivity marker: *nó* at the syntax-pragmatics (discourse) interface") provides an analysis of the expletive subject pronoun *nó* (*nó*_{EXPL} hereafter) in modern spoken Vietnamese. As Dao shows, in contrast to referential *nó*, *nó*_{EXPL} does not behave like a full-fledged pronoun in terms of its syntactic and semantic behavior, but has developed discourse-related uses and is thus strongly sensitive to contextual (i.e. pragmatic) factors. Drawing on a thorough analysis of *nó*_{EXPL}, Dao suggests that the expression is best treated as an ego-evidential (or egophoric, speaker-related) marker and that the referential and expletive uses of *nó* derive from the same set of features. The differences between referential *nó* and subjective-expressive *nó*_{EXP} are illustrated by contrasting the feature profile of the two expressions, which in the case of *nó* comprises the features [+Referential; ±Sentient; –Speaker; –Discourse participant], but for *nó*_{EXP} [–R(eferential), +Sen(tient), +Sp(eaker), +D(iscourse-Participant)]. The analysis thus provides interesting insights into the progressive move of particular uses of a grammatical form "out of" grammar into the direction of discourse and the subtle changes in the syntactic and semantic behavior of such forms.

Sylvie Hancil's contribution ("The final particle *like* in Northern English – a particle of reformulation in the context of interenunciative readjustment") studies the final particle *like* in Northern English dialects. Over the last twenty years, a growing number of linguists have contributed to an increased interest in Northern dialects of English (see the interesting discussion in Wales [2006]). Among the linguistic features characterizing the North-East of England are discourse markers used in sentence-final position (see Beal, Burbano Elizondo & Llamas 2012). Hancil's paper takes up this issue and sheds some more light on a typical feature of Northern English grammar, namely the use of adverbials in sentence-final position. Among the various possible adverbials that can be encountered in this position (e.g. *but*, *like*, *mind*), *like* represents a highly interesting case since corpus-based data provided by the NECTE corpus (Newcastle English Corpus of Tyneside English) show that it has undergone semantic and pragmatic changes in Northern English, especially in Geordie English, where it has developed into a discourse marker with

a great variety of meanings, extending from an anaphoric/cataphoric value to an intensifying value via a phatic value. The final position of the marker correlates with the development of a number of interactional meanings, e.g. in the sense that *like* indicates the speaker's discourse strategy relative to the hearer. The study of sentence-final *like* also addresses aspects referring to Brown & Levinson's (1987) theory of politeness.

In his paper "On pragma-semantics of expressives: between words and actions", **Suren Zolyan** develops a theoretical-descriptive approach to the linguistic category of *expressives*, one of the most problematic phenomena in semantics, given that expressives are a borderline case between action and speech (or, more precisely, between action and the expressive function of language), thus crossing the descriptive levels of semantics and pragmatics, which renders attempts at grasping the semantic features of expressives complicated. The analysis of expressives provided by Zolyan is based on the assumption that meaning is the result of a conjunction of linguistic and extra-linguistic systems in the process of social interaction and communication. This approach enables Zolyan to link logical and semantic characteristics with cognitive and pragmatic ones into a "pragmasemantic" approach, which means that the verbalization of mental representations as well as non-verbal and verbal actions are seen as interacting entities that mutually influence each other. The chapter discusses three main theoretical approaches to expressives and their interrelation: (1) expressivity as a characteristic concomitant with any linguistic phenomenon in the process of communication, (2) expressives as a special class of speech acts, and (3) expressives as a special class of lexical and quasi-lexical units expressing emotional states and attitudes. In the context of the present volume expressives are an interesting case as they represent one of the most immediate links between the language system (in terms of conventional meanings) and the discourse situation (in terms of their situation-specific interpretation and that motivates their use).

Part III begins with a paper by **Danh Thành Do-Hurinville** ("A just amazing marker in French: "Juste"), which deals with the recent adverbial uses of French *juste* as a modal adverb, as e.g. in *c'est juste merveilleux*. In these uses, which occur predominantly in everyday conversation and in the media, *juste* is anteposed to extreme adjectives (e.g. *génial*, *magnifique*, *merveilleux*) and, as shown by Do-Hurinville, fulfils a twofold modalizing function: it intensifies the semantics of the adjective it accompanies, but it also modalizes the utterance in which it occurs as a whole. Thus, as with other phenomena discussed in this volume, these more recent uses of *juste* can be seen as indeterminate between a sentence-internal modifier with a scope over a specific syntactic constituent and a stance marking element with metalinguistic, pragmatic uses on the discourse level. However, the second interpretation is bound to specific conditions, which are thoroughly studied in the contribution.

Anna Snarska's paper ("On how the distinction between reciprocal and collective verbs affects (anti-)control") focuses on the phenomenon of *Control*, which is a relation between a missing subject of a non-finite clause (usually marked as PRO in the linguistic literature) and a matrix argument that imposes its reference on the former. Various kinds of control have been identified; this paper contributes to a better understanding of the nature of so-called *Partial Control* (e.g. Landau 2000, 2003, 2004; Boeckx, Hornstein & Nunes 2010), as illustrated in (2).

- (2) John₁ told Mary₂ that he₁ wants [PRO₁₊ to meet in the morning].

The referent of the silent PRO seems to include *John* along with other individuals salient in the context (in this case it is *Mary*, hence the notation ₁₊ on PRO). Since the matrix subject only partially controls the reference of the lower subject, this phenomenon has come to be known as *Partial Control* (PC). The supporters of the existence of PC stress the fact that only when the non-finite clause contains a collective predicate such as *meet*, *congregate* or *assemble*, a 'group' interpretation is accessible. This view is challenged in Snarska's paper. Supported by evidence from Polish, Snarska argues that *spotykać się* 'meet' is, in fact, an inherently reciprocal verb (like, e.g. *kłócić się* 'argue') that involves two arguments playing the same role in the event. Only such verbs are able to generate the 'group' construal. By co-occurring with a null discontinuous phrase, they introduce an additional participant of the act denoted by the verb; hence the alleged 'PRO₊ effect' illustrated in (3).

- (3) Janek₁ czuł, że Maria₂ chciała [PRO₂₊ się spotkać z nim]
 John felt that Mary wanted REFL to-meet with him
 'John felt that Mary wanted to meet.'

Truly collective predicates such as *gromadzić się* 'congregate/gather' are incapable of participating in PC since they are incompatible with discontinuous phrases, let alone their null versions, as shown in (4).

- (4) *Janek₁ wiedział, że Maria₂ chciała [PRO₂₊ się zgromadzić w tym
 John knew that Mary wanted REFL to-gather in this
 kościele].
 church
 'John knew that Mary wanted to gather/congregate in this church.'

Snarska's approach provides a possible explanation for why the 'group' interpretation is accessed in anti-control constructions (cf. Brandt et al. 2016) in Polish, where a co-referential interpretation between the missing subject and the matrix argument is impossible. Among other things, the topic discussed by Snarska illustrates the context-bondedness of the semantic interpretation or recovery of "silent"

subjects in specific constructions, and thus the interaction between grammar and lexical semantics, on the one hand, and meaning aspects “outside” merely formal constraints on the sentence level, which listeners need to include for inferential processes involved in the identification “silent” referents, on the other hand.

Albert Álvarez González’s article (“The rise of cause/reason adverbial markers in Yaqui (Uto-Aztecan)”) deals with three cause/reason markers (*betchi’ibo*, *bwe’ituk*, *po(r)ke*) in Yaqui. The causal adverbial clauses introduced by these three different markers exhibit different internal structures, in particular regarding the coding of the subject of the adverbial clause and the position of the causal marker. Álvarez González traces the development of the markers, suggesting that they derive from three different origins: (i) extension in use from phrasal to clausal level (*betchi’ibo*); (ii) combination of a discourse marker (*bwe*) which, as also argued in Álvarez (2019), moved from discourse into syntax in order to participate in the creation of the cause/reason adverbial marker in Yaqui, and a linguistic element associated with an older strategy used for marking cause/reason clauses (*ituca*), and (iii) language contact (*po(r)ke* – a loanword from Spanish). The three different ways by which interclausal connectors of cause came into existence and the ways in which they are used offer interesting insights into the grammar-discourse interplay since, as is shown by Álvarez González, the identification of meaning differences between the three adverbial connectors, which are rather subtle, requires a detailed pragmatic account of the system of interclausal causality in Yaqui and an analysis of contextual features that correlate with the use of each of the three causal markers.

4. Conclusion

This volume makes no attempt at questioning or negating the differences between grammar and discourse, but seeks to contribute to a better understanding of the meeting points between the two. The studies in this volume concern the fluid boundary between both domains as is evidenced e.g. in the evolution of discourse markers, the interpretation of a grammatical construction as a discourse pattern (e.g. when complement-taking predicates develop discourse-bound uses, see Heine, Kaltenböck, Kuteva & Long, this vol., Angot & Mosegaard Hansen, this vol.), or in usage patterns that mediate between grammar and (spoken) discourse (Part 3). For instance, although French *juste* can, grammatically speaking, occur with a broad range of adjectives, the precise meaning and the scope of the adverb differ across different subsets of adjectives and, in the case of evaluative or “extreme” adjectives, carry over to discourse, where *juste* modifies the utterance as a whole. Such cases suggest that there are discourse principles that regulate specific selective processes in grammar and may open up alternative semantic interpretations. All in all, the

studies address various aspects of the grammar-discourse interaction, offering intriguing insights from historical, typological and predominantly usage-based perspectives, based on fresh empirical data from a broader set of languages.

References

- Aijmer, Karin. 2002. *English Discourse Particles. Evidence from a Corpus* [Studies in Corpus Linguistics 10]. Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.10>
- Alcázar, Asler. 2017. A syntactic analysis of rhetorical questions. In *Proceedings of the 34th West Coast Conference on Formal Linguistics*, Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel & Edward J. Rubin (eds), 32–41. Somerville MA: Cascadilla.
- Álvarez González, Albert. 2019. From discourse to syntax: The use of the discourse marker *bwe* in the creation of interclausal connectives in Yaqui. In *Diverse Scenarios of Syntactic Complexity* [Typological Studies in Language 126], Albert Álvarez González, Zarina Estrada-Fernández & Claudine Chamoreau (eds), 217–247. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.126.09gon>
- Ariel, Mira. 2008. *Pragmatics and Grammar*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511791314>
- Ariel, Mira. 2009. Discourse, grammar, discourse. *Discourse Studies* 1(1):5–36. <https://doi.org/10.1177/1461445608098496>
- Beal, Joan, Burbano Elizondo, Lourdes & Llamas, Carmen. 2012. *Urban North-Eastern English: Tyneside to Teeside*. Edinburgh: EUP.
- Beeching, Kate & Detges, Ulrich (eds). 2014. *Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Leiden: Brill. <https://doi.org/10.1163/9789004274822>
- Blakemore, Diane. 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.
- Boeckx, Cedric, Hornstein, Norbert & Nunes, Jairo. 2010. *Control as Movement*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511761997>
- Brandt, Patrick, Trawiński, Beata & Wollstein, Angelika. 2016. (Anti-)Control in German: Evidence from comparative, corpus- and psycholinguistic studies. In *Co- and Subordination in German and other Languages*, Ingo Reich & Augustin Speyer (eds), 77–98. Berlin: De Gruyter.
- Brinton, Laurel. 2017. *The Development of Pragmatic Markers in English: Pathways of Change*. Cambridge: CUP. <https://doi.org/10.1017/9781316416013>
- Brown, Penelope & Levinson, Stephen C. 1987. *Politeness*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511813085>
- Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language* 82: 711–733. <https://doi.org/10.1353/lan.2006.0186>
- Crible, Ludivine. 2018. *Discourse Markers and (Dis)fluency. Forms and Functions across Languages and Registers*. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.286>
- Degand, Liesbeth, Cornillie, Bert & Pietrandrea, Paola (eds). 2013. *Discourse Markers and Modal Particles. Categorization and Description* [Pragmatics & Beyond New Series 234] Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.234>

- Degand, Liesbeth & Evers-Vermeul, Jacqueline. 2015. Grammaticalization or pragmaticalization of discourse markers? More than a terminological issue. *Journal of Historical Pragmatics* 16(1): 59–85. <https://doi.org/10.1075/jhp.16.1.03deg>
- Dik, Simon C. 1997. *The Theory of Functional Grammar*, Part 2. Berlin: Mouton de Gruyter.
- Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63: 805–855. <https://doi.org/10.2307/415719>
- Du Bois, John W. 2003. Discourse and grammar. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Vol. 2, Michael Tomasello (ed.), 47–87. Mahwah NJ: Lawrence Erlbaum Associates.
- Etxepare, Ricardo. 2010. Omen bariazioan. In *Euskara eta euskarak: Aldakortasun sintaktikoa aztergai*, Beatriz Fernández, Pablo Albizu & Ricardo Etxepare (eds). Vitoria-Gasteiz: University of the Basque Country.
- Etxepare, Ricardo & Uriá, Larraitz. 2016. Microsyntactic variation in the Basque hearsay evidential. In *Microparameters in the Grammar of Basque* [Language Faculty and Beyond 13], Beatriz Fernández & Jon Ortiz de Urbina (eds), 265–288. Amsterdam: John Benjamins. <https://doi.org/10.1075/lfab.13.10etx>
- Farkas, Donka F. 2017. Going beyond the prototypical: Special interrogatives and special declaratives. Paper presented at the Stanford Colloquium, 8 December 2017.
- Fischer, Kerstin (ed.). 2006. *Approaches to Discourse Particles*. Amsterdam: Elsevier.
- Foolen, Ad. 2015. Expressives. In *The Routledge Handbook of Semantics*, Nick Riemer (ed.), 473–490. London: Routledge.
- Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 3(7): 931–952. [https://doi.org/10.1016/S0378-2166\(98\)00101-5](https://doi.org/10.1016/S0378-2166(98)00101-5)
- Fraser, Bruce. 2015. The combining of discourse markers – A beginning. *Journal of Pragmatics* 86: 48–53. <https://doi.org/10.1016/j.pragma.2015.06.007>
- Garmendia, Joana. 2014. A pragmatic analysis of the Basque particle *ote*. In *Cognitive and Pragmatic Aspects of Speech Actions*, Iwona Witzczak Plisiecka (ed.), 97–166. Bern: Peter Lang.
- Grodzinsky, Yosef. 2000. The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences* 23: 1–21. <https://doi.org/10.1017/S0140525X00002399>
- Hancil, Sylvie, Haselow, Alexander & Post, Margje (eds). 2015. *Final Particles*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110375572>
- Haselow, Alexander. 2013. Arguing for a wide conception of grammar: The case of final particles in spoken discourse. *Folia Linguistica* 47(2): 375–424. <https://doi.org/10.1515/flin.2013.015>
- Haselow, Alexander. 2016. A processual view on grammar: Macrogrammar and the final field in spoken syntax. *Language Sciences* 54: 77–101. <https://doi.org/10.1016/j.langsci.2015.12.001>
- Haselow, Alexander. 2017. *Spontaneous Spoken English. An Integrated Approach to the Emergent Grammar of Speech*. Cambridge: CUP. <https://doi.org/10.1017/9781108265089>
- Haselow, Alexander. 2019. Discourse marker sequences: Insights into the serial order of communicative tasks in real-time turn production. *Journal of Pragmatics* 146: 1–18. <https://doi.org/10.1016/j.pragma.2019.04.003>
- Haselow, Alexander & Kaltenböck, Gunther (eds). 2020. *Grammar and Cognition: Dualistic Models of Language Structure and Language Processing* [Human Cognitive Processing 70]. Amsterdam: John Benjamins. <https://doi.org/10.1075/hcp.70>
- Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6): 1205–1247. <https://doi.org/10.1515/ling-2013-0048>
- Heine, Bernd & Kuteva, Tania. 2007. *The Genesis of Grammar*. Oxford: OUP.

- Heine, B. 2019. Some observations on the dualistic nature of discourse processing. *Folia Linguistica* 53(2): 411–442. <https://doi.org/10.1515/flin-2019-2016>
- Heine, Bernd, Kuteva, Tania, Kaltenböck, Gunther & Long, Haiping. 2017. Cooptation as a discourse strategy. *Linguistics* 55(4): 813–855. <https://doi.org/10.1515/ling-2017-0012>
- Heine, Bernd, Kuteva, Tania, & Long, Haiping. 2020. Dual process frameworks on reasoning and linguistic discourse. A comparison. In Haselow & Kaltenböck (eds), 59–89.
- Huddleston, Rodney & Pullum, Geoffrey. 2002. *Language Description: The Cambridge Grammar of the English Language*. Cambridge: CUP. <https://doi.org/10.1017/9781316423530>
- Jucker, Andreas H. & Smith, Sara W. 1998. And people just you know like ‘wow’ – Discourse markers as negotiating strategies. In *Discourse Markers. Theory and Descriptions* [Pragmatics & Beyond New Series 57], Andreas Jucker & Yael Ziv (eds), 171–201. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.57.10juc>
- Kaltenböck, Gunther, Heine, Bernd & Kuteva, Tania. 2011. On thetical grammar. *Studies in Language* 35(4): 848–893. <https://doi.org/10.1075/sl.35.4.03kal>
- Korta Carrión, Kepa & Zubeldia Arruabarrena, Larraiz. 2016. Bide partikula: Eduki doxastikoa eta ebidentziala. *Gogoia* 15: 11–29.
- Koyanagi, Tomokazu. 2018. *Research on Grammatical Change* Tokyo: Kuroshio. (In Japanese).
- Landau, Idan. 2000. *Elements of Control. Structure and Meaning in Infinitival Constructions*. Dordrecht: Kluwer.
- Landau, Idan. 2003. Movement out of control. *Linguistic Inquiry* 34: 471–498. <https://doi.org/10.1162/002438903322247560>
- Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22: 811–877. <https://doi.org/10.1007/s11049-004-4265-5>
- Lehmann, Christian. 2015. *Thoughts on Grammaticalization*, 3rd edn [Classics in Linguistics 1]. Berlin: Language Science Press. https://doi.org/10.26530/OAPEN_603353
- Lehman Blake, Margaret. 2006. Clinical relevance of discourse characteristics after right hemisphere brain damage. *American Journal of Speech-Language Pathology* 15(3): 255–267. [https://doi.org/10.1044/1058-0360\(2006\)024](https://doi.org/10.1044/1058-0360(2006)024)
- Lenk, Uta. 1998. *Marking Discourse Coherence: Functions of Discourse Markers in Spoken English*. Tübingen: Narr.
- Mann, William C. & Thompson, Sandra A. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8(3): 243–281. <https://doi.org/10.1515/text.1.1988.8.3.243>
- Marini, Andrea. 2012. Characteristics of narrative discourse processing after damage to the right hemisphere. *Seminars in Speech and Language* 33(1): 68–78. <https://doi.org/10.1055/s-0031-1301164>
- Marini, Andrea, Carlomagno, Sergio, Caltagirone, Carlo & Nocentini, Ugo. 2005. The role played by the right hemisphere in the organization of complex textual structures. *Brain and Language* 93: 46–54. <https://doi.org/10.1016/j.bandl.2004.08.002>
- Monforte, Sergio. 2018a. About the particle *al* in basque. In *Novas Perspectivas na Lingüística Aplicada*, Marta Díaz Ferro, Jorge Diz Ferreira, Ania Pérez Pérez & Ana Varela Suárez (eds), 19–31. Lugo: Editorial Axac.
- Monforte, Sergio. 2018b. Question particles in basque. *Isogloss. A Journal on Variation of Romance and Iberian Languages* 4(1): 29–53. <https://doi.org/10.5565/rev/isogloss.48>
- Myers, Penelope S. 1999. *Right Hemisphere Damage: Disorders of Communication and Cognition*. San Diego CA: Singular.

- Prat, Chantel S., Long, Debra L. & Baynes, Kathleen. 2007. The representation of discourse in the two hemispheres: An individual differences investigation. *Brain and Language* 100(3): 283–294. <https://doi.org/10.1016/j.bandl.2006.11.002>
- Rhee, Seongha. 2012. Context-induced reinterpretation and (inter)subjectification: The case of grammaticalization of sentence-final particles. *Language Sciences* 34(3): 284–300. <https://doi.org/10.1016/j.langsci.2011.10.004>
- Rhee, Seongha. 2016. From quoting to reporting to stance-marking: Rhetorical strategies and intersubjectification of reportative. *Language Sciences* 55: 36–54. <https://doi.org/10.1016/j.langsci.2016.02.003>
- Schourup, Lawrence. 1985. *Common Discourse Particles in English Conversation*. New York NY: Garland.
- Schiffirin, Deborah. 1987. *Discourse Markers*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511611841>
- Sherratt, Sue & Bryan, Karen. 2012. Discourse production after right brain damage: Gaining a comprehensive picture using a multi-level processing model. *Journal of Neurolinguistics* 25(4): 213–239. <https://doi.org/10.1016/j.jneuroling.2012.01.001>
- Shibasaki, Reijirou. 2018. From the inside to the outside of the sentence: Forming a larger discourse unit with *jijitsu* ‘fact’ in Japanese. In *New Trends in Grammaticalization and Language Change* [Studies in Language Companion Series 202], Sylvie Hancil, Tine Breban & José V. Lozano (eds), 333–360. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.202.14shi>
- Strauss, Susan. 2005. Cognitive realization markers in Korean: A discourse-pragmatic study of the sentence-ending particles *-kwun*, *-ney* and *-tela*. *Language Sciences* 27(4): 437–480. <https://doi.org/10.1016/j.langsci.2004.09.014>
- Thompson, Sandra A. & Suzuki, Ryoko. 2011. The grammaticalization of final particles. In *The Oxford Handbook of Grammaticalization*, Bernd Heine & Heiko Narrog (eds), Oxford: OUP.
- Traugott, Elizabeth Closs. 1997. The role of the development of discourse markers in a theory of grammaticalization. Paper presented at the *ICHL XII*, Manchester, UK.
- Traugott, Elizabeth Closs. 2015. Investigating “periphery” from a functionalist perspective. *Linguistics Vanguard* 1. <https://doi.org/10.1515/lingvan-2014-1003>
- Traugott, Elizabeth Closs. 2016. On the rise of types of clause-final pragmatic markers in English. *Journal of Historical Pragmatics* 17(1): 26–54. <https://doi.org/10.1075/jhp.17.1.02tra>
- Trotzke, Andreas & Monforte, Sergio. 2019. Basque question particles: Implications for a syntax of discourse particles. *Linguistic Variation* 19(2): 352–385. <https://doi.org/10.1075/lv.18002.tro>
- van Dijk, Teun A. 1972. *Some Aspects of Text Grammars*. The Hague: Mouton.
- van Dijk, Teun A. 1977. *Text and Context: Explorations in the Semantics and Pragmatics of Discourse*. London: Longman.
- Wales, Katie. 2006. *Northern English. A Social and Cultural History*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511487071>
- Watanabe, Fumio, Muraishi, Shōzō & Kabe, Sasuke. 1993. *Nihongo kaishaku katuyo jiten* (A dictionary of reading and application of Japanese). Tokyo: Gyosei.

PART I

Discourse markers

The grammar-discourse interaction
from a developmental perspective

On the rise of discourse markers

Bernd Heine, Gunther Kaltenböck, Tania Kuteva
and Haiping Long

University of Cologne / University of Graz / Heinrich-Heine
University Düsseldorf / Sun Yat-sen University

Discourse markers exhibit a range of grammatical properties that set them apart from many other lexical and grammatical forms. A number of hypotheses have been proposed to account for these properties. Most commonly, such accounts have drawn on grammaticalization theory, less commonly also on the notion of pragmaticalization. As argued in this paper, however, the rise of discourse markers cannot be described exhaustively in terms of parameters of grammaticalization or pragmaticalization. Looking at the rise of two English and one French discourse marker, the paper argues that this development is also shaped by a general strategy of discourse processing called cooptation, which accounts for properties of discourse markers that are beyond the scope of grammaticalization theory. Thus, discourse markers are described best as grammaticalized (paren)theticals.

Keywords: cooptation, decategorialization, discourse grammar, discourse marker, grammaticalization, thetical

1. Introduction

1.1 The present paper

The rise and development of discourse markers (DMs) is widely understood to be the result of grammaticalization or pragmaticalization. But more recently it has been argued that this process also involves an operation called *cooptation*, whereby pieces of a text are deployed for use on the metatextual level, serving to monitor the production of texts and to provide instructions on how to interpret the texts (cf. Heine 2013, 2018; Heine et al. 2017). This view is challenged by Degand and Evers-Vermeul (2015: 72) and Brinton (2017: 37), who conclude that the development of DMs can largely or entirely be described in terms of grammaticalization (see Section 5 below).

The goal of the present paper is to argue that the latter position is in need of reconsideration, and that any approach that is restricted to grammaticalization is

unlikely to provide a comprehensive account of this development and to explain why DMs are structured the way they are. The goal thus is a highly limited one, which means that many of the observations that have been made on the subject matter of the paper will have to be ignored (see Section 6).

To ensure maximal comparability for evaluating the two positions just sketched, the paper is restricted to the data presented by the authors mentioned in support of their position. To this end, we will look at three DMs discussed by these authors, namely the English markers *I admit* (Section 2) and *if you will* (Section 3), and the French marker *alors* (Section 4). We will, however, not be able to do justice to the rich descriptions and reconstructions presented by the authors, on whose work our analysis rests; rather we will be confined to the goal mentioned and our concern is therefore only with evidence that is immediately relevant to this goal.

1.2 Discourse marker, grammaticalization, and cooptation

There are three concepts that are central for an understanding of the ensuing discussions, namely ‘discourse marker’, ‘grammaticalization’, and ‘cooptation’, and this section is devoted to defining these concepts.

The term *discourse marker* has been used in a wide range of senses and for quite a number of different phenomena, extending from monosyllabic interjection-like particles, formulae of social exchange, and hesitation fillers to clausal expressions. Also called discourse particles, pragmatic markers, discourse connectives, adverbials, connecting adverbials, or simply conjunctions, etc., DMs have been the subject of many studies (see, e.g., Aijmer 2002; Dér 2010 for convenient overviews; see also Furkó 2014). The most detailed characterization of DMs that we are aware of is provided by Brinton (2017: 8–9, Table 1.1). In the present paper DMs are defined as in (1) (cf. Heine (2013: 1211, 2018).

- (1) Discourse markers are (a) invariable expressions which are (b) syntactically independent from their environment, (c) typically set off prosodically from the rest of the utterance,¹ and (d) their function is metatextual, relating a text to the situation of discourse, that is, to the organization of texts, the attitudes of the speaker, and/or speaker-hearer interaction.²

Whether DMs constitute a distinct ‘category’ is an issue that is notoriously controversial. The definition in (1) seems to take care of most of the expressions that

1. The hedge ‘frequently’ indicates that prosodic separation is not an absolutely reliable criterion, being dependent on factors such as relative speech rate, complexity of the parenthetical expression, placement within an utterance, emphasis, and processing factors (see, e.g., Dehé 2014).

2. The components attitudes of the speaker and speaker-hearer interaction relate, respectively, to the terms ‘subjectivity’ and ‘intersubjectivity’ of Traugott (1982, 1989; see also Traugott and Dasher 2002: 22)

have crosslinguistically been classified as DMs. At the same time, it differs from definitions proposed in the tradition of Schiffrin (1987: 31), like that of Fraser (1999: 938) or Traugott (2018: 27), which stipulate that DMs signal some kind of relationship between clauses, utterances, units of talk, or discourse segments. For Traugott (2018: 27), for example, a DM is “a metatextual marker that signals some kind of relationship between clauses/utterances”.

The reason for proposing (1) instead is that in a number of their uses, DMs do not signal such a relationship. For example, expressions such as *I think* and *you know* are commonly considered to be paradigm instances of English DMs in many of their uses. Yet if someone says, *Istanbul is, I think/you know, not the capital of Turkey*, one may wonder whether the function of *I think* and *you know* is fully in accordance with Traugott’s definition. We will rather take the properties listed in (1) to be criterial for what is and what is not a DM, being aware that these properties are not the only ones associated with DMs.

One may wish to also mention a number of other features associated with DMs, e.g., that they normally can be omitted without loss of grammaticality or propositional content or that their use is optional (but see Dér 2010), that they do not contribute to the truth conditions of the clause, that their meaning is procedural rather than conceptual, that they cannot be negated or become the focus of a cleft sentence, that they are semantically transparent (cf. van Bogaert 2011: 298–9), and that their functions are restricted to the here and now of the situation of discourse in which they are used – accordingly, they have been classified as indexicals (Aijmer 2002; Furkó 2014: 292) or discourse deictics (Levinson 2006).

DMs have been classified, on the one hand, as a sub-type of pragmatic markers (Fraser 2009) and, on the other hand, as a sub-type of parentheticals (Kaltenböck 2007: 31; Fischer 2007a, 2007b) or, as we will say here, of theticals (Kaltenböck et al. 2011; Heine 2013, 2018; Heine et al. 2013: 165, 2017).³ What distinguishes them from many other theticals is that they are largely or entirely invariable, that is, they typically do not allow internal modification (see also Fischer 2007a, 2007b). Not everybody might agree to call all the expressions discussed in this paper DMs. But to the extent that they conform to the definition in (1) they will be referred to as ‘discourse markers’ – a term corresponding to what Brinton (1996, 2008, 2017) calls ‘pragmatic markers.’⁴

3. The term ‘parenthetical’ has been used for a wide range of features or categories. In the present paper, it is restricted to its use in paradigm studies such as Kaltenböck (2007) and Dehé (2014). In this sense, ‘parentheticals’ and ‘theticals’ have the features to be presented in (9) of Section 2.2. Since we are relying here on the framework of Discourse Grammar, the term ‘thetical’ stands for pieces of discourse belonging to thetical grammar rather than to sentence grammar (Heine et al. 2013).

4. Note that this use of ‘pragmatic marker’, which is largely equivalent to what we call here a DM, is not the same as that of Fraser (2009) alluded to above.

The second concept is *grammaticalization*, which is defined here in short as the development from lexical to grammatical forms and from grammatical to even more grammatical forms (Heine & Kuteva 2002: 2). On this definition, which is some way or other in accordance with most of the over fifty definitions that there are, grammatical developments that do not conform to the definition are not strictly within the scope of grammaticalization theory (Kuteva et al. 2019).⁵ Note that the definition is concerned with change in ‘forms’, including the constructions of which they are a part, that is, with linguistic expressions having a semantic, a morpho-syntactic, and a phonological component. Thus, it differs from some other definitions that are restricted essentially to change in meaning (or function), like that by Hopper and Traugott (2003: 18), even if this difference will not affect the content of the present paper.⁶

Grammaticalization thus is an essentially unidirectional process, and it is based on the context-induced manipulation of expressions in discourse (see the context extension model of Heine 2002). It is most commonly described in terms of five principles by Hopper (1991: 25–9),⁷ or of nine parameters by Lehmann ([1982] 2015: 132), or of four parameters by Heine and Kuteva (2007: 33–44). With reference to the subject matter of the present paper it is most of all the effects listed in (2) that expressions undergoing grammaticalization exhibit.

- (2) Common effects of grammaticalization (in parentheses: the relevant parameters of Heine and Kuteva 2007: 33–44; see also Lehmann [1982] 2015: 132)
 - a. Meaning: The expression loses concrete semantic features in favor of more abstract grammatical features (desemanticization).
 - b. Function: It becomes functionally dependent on its host constituent (desemanticization leading to bonding).
 - c. Syntax: It loses morphosyntactic features, including part or all of its syntactic independence (decategorialization).
 - d. Placement: It tends to become fixed in the position next to or close to its host constituent (decategorialization).
 - e. Phonology: It tends to lose phonetic substance, including part or all of its prosodic distinctiveness (erosion).

5. Grammaticalization theory is neither a theory of language nor of language change; its goal is to describe grammaticalization, that is, the way grammatical forms arise and develop through space and time, and to explain why they are structured the way they are (Heine 2003: 575).

6. Hopper and Traugott (2003: 18) define grammaticalization as “the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions.”

7. Notice, however, that Hopper (1991: 32) points out that the five principles of grammaticalization he identifies “are not the exclusive domain of grammaticalization, but are common to change in general”.

We may illustrate these effects with the following crosslinguistically common pathway of grammaticalization (see Kuteva et al. 2019 for more examples). In many languages, verbs of volition have grammaticalized into future tense markers, and these processes are usually in accordance with the features in (2). Thus, when the Old English verb *willan* ‘to want, to wish, to desire’ or the Swahili verb *-taka* ‘want’ were grammaticalized into the future tense markers *will* and *-ta-*, respectively, (a) they lost most of their lexical semantic features, and (b) they became dependent on the new main verb, that is, their host constituent. (c) They also lost most of their verbal properties, such as the ability to be inflected, and they lost their morphosyntactic independence that they had as full verbs, with *will* now serving as a verbal auxiliary and Swahili *-taka* as a verbal prefix (*-ta-*). (d) They also lost phonetic substance, *will* optionally being reduced to *’ll* and *-taka* losing its final syllable and becoming *-ta-*, and both lost much of the prosodic distinctiveness they had as lexical verbs. (e) Finally, the erstwhile verbs also lost their positional flexibility, being fixed in the position to the left of the new main verb.

In accordance with the tradition which has been prevalent in grammaticalization studies over roughly the last fifty years, grammatical changes that contradict (2) will not be considered here to be instances of grammaticalization.

In some lines of research a distinction is made between two perspectives of grammaticalization, namely between a narrow or restricted and a wider or expanded view of grammaticalization (e.g., Degand & Simon-Vandenberg 2011: 290; Degand & Evers-Vermeul 2015: 77), or between grammaticalization viewed as reduction and increase in dependency vs. grammaticalization viewed as expansion (Traugott & Trousdale 2013: 99–112). We will ignore such distinctions in the remainder of the present paper, for the following reasons: First, none of the studies cited argues that the generalizations in (2) are not covered by grammaticalization theory, second, none of those studies seems to provide a plausible account of why the development of DMs results in changes such as the ones listed in (4) below, and, third, there is reason to question the relevance of this distinction in the first place (Heine 2018).

The third concept to be discussed in the paper is that of *cooptation*, which is defined in (3) (cf. Kaltenböck et al. 2011: 874–875; Heine et al. 2017: 828).

- (3) Cooptation is a fully productive operation whereby a chunk of sentence grammar, such as a word, a phrase, a reduced clause, a full clause, or some other piece of text is deployed for use on the metatextual level of discourse monitoring,⁸ thereby turning into a thetical. Its functions are determined by the situation

8. The term ‘metatextual’ refers to the function of a text piece as being anchored in and relating the meaning of a sentence to the situation of discourse (cf. Traugott 1995: 6; Heine 2018: 34). In more general terms, a ‘metatextual unit’ presents a statement whose topic is the text itself (Witosz 2017: 108; see also Genette 1982).

of discourse, serving (a) to overcome constraints imposed by linearization in structuring texts, (b) to package text segments, (c) to place a text in a wider perspective, e.g. by elaborating, providing an explanation, a comment or supplementary information, (e) to describe the attitudes of the speaker, and/or (e) to interact with the hearer.

Coopted units (called theticals) are – depending on their specific discourse function – placed either within a sentence or at its left or right periphery without being grammatically integrated in the sentence,⁹ and some can even form utterances of their own. Cooptation entails most or all of the grammatical changes in (4) (see also Kaltenböck et al. 2011: 853; Heine et al. 2017: 817; Heine 2018: 32–35). All evidence that there is suggests that these changes are not the result of a long, gradual process but rather take place jointly in a fairly abrupt manner, even if it takes some time before they are conventionalized. The criteria in (4) are based on a larger set of data, including DMs, and in the remainder of the paper the usefulness of the criteria is tested by confronting them with specific data from English and French. Note, however, that in a given case not all of the criteria may clearly apply, or may apply only in some modified form.

- (4) Grammatical changes typically associated with cooptation
- a. Meaning: From meaning as part of the sentence to meaning outside the sentence
 - b. Function: From sentence-structuring function to metatextual function
 - c. Syntax: From syntactic constituent of the sentence to syntactically detached status¹⁰
 - d. Prosody: From prosodically integrated to unintegrated or less integrated status¹¹
 - e. Semantic-pragmatic scope: From more restricted scope to scope beyond the sentence¹²

9. We are ignoring here the fact that ‘left periphery’ and ‘right periphery’ can refer to quite different concepts depending, e.g., on which kind of adjacency pair is involved (see Beeching 2016). Furthermore, we are aware that the terms are not really fortunate considering that they invoke a visual view of language based on written structure in languages such as English (Haselow 2016: 83). Nevertheless, we are using the terms loosely for linguistic material located outside the ‘core’ of a sentence, clause, or phrase. See also Traugott (2015).

10. But even when being syntactically unattached it may retain cohesive relations with the sentence, such as coreference.

11. In written documents, prosodic non-integration tends to be signaled by punctuation marks like commas.

12. Note that the notion ‘semantic-pragmatic scope’ differs from that proposed by Traugott (1995) and, most importantly, from notions of ‘scope’ based on a syntactic perspective, like that of

- f. Placement: From positionally constrained to less constrained placement, most commonly at the left periphery of a sentence.

The notion cooptation corresponds largely to what Waltereit (2006) calls discursive reanalysis, whereby a lexical or grammatical item is used at the discursive level, or what Dostie (2009: 203) describes as category change, whereby a lexical or grammatical item “becomes a pragmatic item.”

Cooptation is a ubiquitous operation that can be used any time and anywhere, both in speech and in writing, even if less commonly in writing. A typical example is provided in (5), where the chunk *I would suggest*, printed in bold, is classified as a coopted unit, that is, a thetical inserted in the relative clause in (5): (a) its meaning is not part of the sentence meaning, (b) its function can be said to be metatextual, (c) it is syntactically unattached, not being a constituent of the sentence, (d) it is likely to be set off by means of intonation and/or pause features, (e) its semantic-pragmatic scope extends over a larger piece of discourse in which it occurs, and (f) it can be moved to various slots of the utterance in (5).

- (5) *It's an issue which **I would suggest** will not go away.* (ICE-GB: s2b-046-30)

Which chunks or ‘text pieces’ qualify for cooptation is an issue that is in need of more research but at the present stage, the following generalizations seem possible: First, the piece must be such that it can be accepted by the hearer as being a meaningful contribution to the structure and the content of the discourse. Second, it seems that it is text pieces used recurrently that are favored in cooptation,¹³ but more information is needed on this issue. Third, the piece may not be, and frequently is not syntactically complete as long as the meaning of the missing part is recoverable from the situation of discourse (the ‘context’). For example, the text piece *I would suggest* in (5) can be said to lack a complement, but the meaning of the complement is recoverable from the context. And fourth, the piece must honor word boundaries, that is, it may not consist of or contain some incomplete part of a word such as an affix.¹⁴

Tabor & Traugott (1998), of Role and Reference Grammar (Matosević 2008: 49), or of Functional Discourse Grammar (Hengeveld & Mackenzie 2008, 2011). Having metatextual functions (see (4b)), semantic-pragmatic scope of DMs extends beyond the sentence over the situation of discourse, that is, it relates the use of a DM directly to the organization of texts, the attitudes of the speaker, and/or speaker-hearer interaction (see Heine et al. 2017).

13. We are grateful to Alexander Haselow (p.c. of 20-05-2019) for having drawn our attention to this observation.

14. There are, however, exceptions. In Korean, the plural suffix *tul* has provided the source for a DM having a wide range of functions, most of all expressing mirativity, irritation, and friendliness (Rhee 2018).

1.3 Conclusions

As we saw in Section 1.2, there are two contrasting mechanisms that need to be distinguished for the reconstruction of DMs, and these mechanisms can be characterized thus: Grammaticalization enables speakers to develop new patterns and forms for constructing sentences, whereas cooptation makes it possible to transfer linguistic expressions to the metatextual level of discourse processing. Cooptation is responsible for salient grammatical properties of DMs, such as the ones listed in (4) above, but it marks only one stage in the development of DMs. Grammaticalization is a more general mechanism, accompanying this development from its beginning to its end – that is, it continues to work on linguistic expressions even after their cooptation (see Heine 2013, 2018). The main contrasting effects of the two mechanisms are summarized in Table 1.

Table 1. Contrasting changes typically observed in expressions undergoing cooptation and grammaticalization

	Grammaticalization	Cooptation
Meaning	Change of meaning within the sentence	From meaning as part of the sentence to meaning outside the sentence
Function	Increasing functional dependence within the sentence	From function within the sentence to metatextual function
Syntax	Increasing syntactic integration within the sentence	From integrated to detached syntax
Prosody	Decrease in prosodic distinctiveness	Increase in prosodic distinctiveness
Semantic-pragmatic scope	Scope may change within the sentence	From scope within the sentence to wider scope
Placement	Decrease in freedom of placement	Increase in freedom of placement

2. The marker *I admit*

In English, like in many other languages, there are linguistic expressions exhibiting two contrasting uses where one is a constituent of the clause and contributes to the meaning of that clause while the other lacks these properties, being outside the clause and serving metatextual functions. One example of this kind is the subject of the present section, namely the expression *I admit*. The data are taken from the seminal study of Brinton (2017) and like Brinton, we are also aiming at a diachronic account to understand why such doublets exist.

2.1 Introduction

The English speech act verb *admit*, also used as a performative verb, is associated with a range of different uses. Our interest here is restricted to the two kinds of uses illustrated in (6).

- (6) a. *I admit that I did this, mea culpa, let me try to help figure this out.*
 (2009 CBS_FacNat [COCA]; Brinton 2017: 168)
- b. *Having to mix a new batch of this laundry formula from time to time is, I admit, a bit inconvenient.*
 (1990 Mother Earth News [COCA]; Brinton 2017: 168)

Brinton (2017: 168) describes both uses of the expression *I admit* with the formula [*I/you* (modal) *admit*]. This type has either a first or a second person subject pronoun and optionally a modal auxiliary; accordingly, it may take a range of forms, like *I admit*, *I will admit*, *I must admit*, *I have to admit*, *you admit*, *you will admit*, *you must admit*, and *you have to admit*.¹⁵ We follow Brinton in assuming that the various members of this construction behave essentially the same as *I admit* with reference to what we have to say in the remainder of this section.

Our concern here is exclusively with differences between the two uses of *I admit* in (6), which will be referred to, respectively, as *I admit*₁ (6a) and *I admit*₂ (6b). The main difference can be sketched as follows. *I admit*₁ has all the features of a verb of sentence grammar, taking a pronominal subject and a phrasal or a clausal complement typically introduced by *that*. It is fully productive in that it can be negated (*I don't admit*), questioned (*do I admit?*), etc. We will call *I admit*₁, as in (6a), a sentence-grammar unit.

*I admit*₂, by contrast, does not take complements, it is not a syntactic constituent of the sentence in which it occurs, and it is likely to be set off prosodically, frequently signaled in writing by commas or other punctuation marks. *I admit*₂ is called a comment clause (Quirk et al. 1985: 1112ff.; Brinton 2008), a parenthetical disjunct or, more generally, a pragmatic marker (Brinton (2017: 168), belonging to the class of what Thompson and Mulac (1991) call epistemic parentheticals; note that Urmson (1952) had called *admit* a parenthetical verb (see also Schneider 2007; Schneider, Glikman & Avanzi 2015). Expressions like *I admit*₂ are in fact theticals in the framework of Discourse Grammar (Kaltenböck et al. 2011; Heine et al. 2013).

15. *I must admit* and *I have to admit* are referred to by Fraser (1996) as hedged performatives and are somewhat different from the other variants in that they are generally not addressable (cf. Boye & Harder's (2007) addressability test for complement-taking predicate clauses).

We will refer to *I admit*₂ as a thetical and, to the extent that it corresponds to the definition (1), as a DM.¹⁶

The question that we are concerned with in this section is: What accounts for the structural differences between *I admit*₁ and *I admit*₂ mentioned above? To this end, we will first look at the diachronic data presented by Brinton (2017) and subsequently look for an answer to this question.

2.2 Reconstruction

*I admit*₁, having a speech act meaning, can be traced back to Middle English, even if it was not very common. Example (7) is an early attestation of it.

- (7) *If eny man aske whi y wole not admytte [...] þat all menal moral vertues [...] be in þe iiije table oonli ...*
 ‘if any man asks why I will not admit [...] that all ancillary moral virtues [...] are in the fourth table only ...’ (c1454 Pecoock, *The follower to The Donet* (Roy 17.D.9) 200/ 17 [MED]; Brinton 2017: 179)

Unambiguous examples of *I admit*₂, that is, thetical *I admit* in medial or final position (Brinton 2010: 295), do not appear until the 18th century, cf. (8).¹⁷

- (8) *for the three former qualities, I admit, make you worthy of happiness.*
 (1749 Fielding, *Tom Jones* [CLMET3.0]; Brinton 2017: 182)

Thus, there seems to be fairly clear evidence to the effect that *I admit*₁, including the variants of it mentioned above, preceded the first occurrence of *I admit*₂ in time and we follow Brinton in assuming that *I admit*₂ is historically derived from *I admit*₁. But how does one account for the properties in (9) that distinguish the two expressions and, hence, are likely to have evolved in the transition from *I admit*₁ to *I admit*₂? We found no conclusive evidence to suggest that the properties in (9) evolved gradually over a longer period of time; rather, *admit*₂ appears to have acquired them fairly suddenly in the 18th century or earlier. Note that these properties

16. *I admit*₂ corresponds to the definitional features of DMs in (1) except for (1a). In view of the formula [*I/you* (modal) *admit*] that Brinton proposes for it (see above), one may describe it as a weakly grammaticalized DM, having features of a partially filled, or partially substantive construction on the basis of a construction grammar perspective.

17. Brinton (2017: 182) says: “Unambiguous examples of parenthetical *admit* (those occurring in medial or final position) do not appear until the eighteenth century or later”. Since an example like (8) seems to be unambiguous, the phrase “or later” is omitted above.

are not restricted to *I admit*₂ but have more generally been proposed to characterize theticals (or parentheticals) (see, e.g., Heine 2018: 37, (7)).¹⁸

- (9) Thetical properties of *I admit*₂ distinguishing it from *I admit*₁
 - a. Meaning: Its meaning is outside of the sentence it is associated with.
 - b. Function: It's function is metatextual.
 - c. Syntax: It is not a syntactic constituent of the sentence in which it occurs.
 - d. Prosody: It is likely to be set off prosodically from the rest of the sentence.
 - e. Semantic-pragmatic scope: It has scope beyond the sentence.
 - f. Placement: Depending on its function it can be placed within the sentence or at its left or right periphery. (cf. Brinton 2017, Table 6.2)

Brinton (2017) invokes the grammaticalization hypothesis of Thompson and Mulac (1991), commonly known as 'the matrix clause hypothesis', to answer this question.¹⁹ According to this hypothesis, epistemic parentheticals like *I think* or *I suppose* arose as matrix clauses taking *that*-complement clauses. Deletion of the complementizer *that* created an indeterminate structure where the original matrix clause was 'reanalyzed' as parenthetical, as illustrated in the following hypothesized development: *I think that the world is flat* > *I think* \emptyset *the world is flat* > *The world is flat, I think* (Brinton 2017: 157). Brinton also suggests that *admit* lost the complementizer *that* and that this led to indeterminate structures where the original main clause could be reanalyzed as parenthetical (Brinton 2017: 187–188, 289). According to Olga Fischer (p.c. of 24-06-2019), however, the DM *I think* already occurred before *that* got lost.

For a critical appraisal of the grammaticalization hypothesis of Thompson and Mulac (1991), see Fischer (2007a, 2007b). Irrespective of how one may evaluate this hypothesis, *I admit*₂ can at best be interpreted as a case of low degree of grammaticalization. But what about the 'reanalysis' of a sentence grammar unit like *I admit*₁ as a parenthetical or, as we say here, its cooptation as the thetical *I admit*₂: How could *I admit*₂ possibly acquire the properties that define it as a thetical, such as the ones in (9)?

It seems that grammaticalization theory is hard-pressed to answer this question, for the reasons given in (2) of Section 1: First, in accordance with the parameters of desemanticization and bonding (Heine and Kuteva 2007: 33–44; Lehmann

18. As we will see in Section 3.2, there is one important constraint on the generalizations in (4e) and (4f).

19. We are restricted here to the development from *I admit*₁ to *I admit*₂. Accordingly, earlier developments of *admit*, like that from event verb ('allow to enter') to speech act verb ('acknowledge/concede as true'; Brinton 2017: 188, Figure 6.1), must be ignored.

2015: 132), grammaticalizing nouns or verbs will gradually become functionally dependent on their host constituent (cf. (2a)). Second, in accordance with the parameter of decategorialization, grammaticalization leads as a rule to reduction or loss of morphosyntactic independence, rather than to an increase in independence (2b). And third, in accordance with the parameter of erosion, grammaticalizing nouns, verbs, and other units are likely to lose phonetic substance, including prosodic features; jointly with their host constituent they tend to become part of one and the same intonation unit (2c).

In short, the development from *I admit*₁ to *I admit*₂ appears to be at variance with the parameters in (2) and the unidirectionality principle of grammaticalization, which would predict grammatical change from less dependent to more dependent linguistic expressions rather than the other way round.²⁰

2.3 Cooptation

In view of problems such as the ones just outlined, an alternative account of discourse markers like *I admit*₂ is proposed by Heine (2013, 2018), Furkó (2014), and Heine et al. (2016, 2017). This account is overall in agreement with the detailed reconstruction work of (Brinton (2017) but adds one consideration that appears to be relevant for understanding the diachronic development of the units looked at in the preceding section. On this account, developments like that from *I admit*₁ to *I admit*₂ involve cooptation (see (3) of Section 1 for a definition).

Cooptation is a highly productive operation and the text pieces that are coopted for specific discourse functions can take a wide range of forms. In the case of the English verb *admit* it appears to have been used extensively, and some of these uses have given rise to a range of more or less formulaic coopted units,²¹ referred to as parentheticals, comment clauses, etc. (see Brinton 2017: 168–190).

The question then is: Which of the many possible expressions or constructions involving the verb *admit* provided the source for the cooptation of *I admit*₂? In spite of detailed reconstructions carried out by Brinton (2017), no conclusive answer seems possible since there is no information on the particular situation of discourse that induced earlier speakers of English to coopt *I admit* as a thetical for the first time. However, given the limited data that are available, the hypothesis proposed by Brinton, based on meticulous reconstruction work, is the most plausible one,

20. This lack of syntactic integration of *I admit*₂ is also difficult to reconcile with Boye and Harder's (2007) account of grammaticalizing complement-taking predicates (CTPs), which equally seems to suggest a tighter link to the host clause, at least in initial position.

21. Concerning the term 'formulaic coopted unit', see Heine et al. (2017: 827).

namely that that source was provided by a sentence grammar use of *I admit* which did not involve a complementizer (*that*), as in Example (10). Note, however that *I admit* in (10) is conceivably not a sentence grammar unit but already a thetical placed at the left periphery of the sentence. In the absence of prosodic information, this issue must remain unresolved and we follow Brinton's analysis.

- (10) *I admit the Northwest point standeth due North, and my course is to go due West,*
(1574 Bourne, *A regiment for the sea* [EEBO]; Brinton 2017: 181)

Once coopted or, in the terminology of Brinton (2017: 187), 'reanalyzed as parenthetical' in the 18th century, it appears to have acquired the properties characteristic of theticals in (9), as the description by Brinton suggests. In short, with its cooptation, *I admit*₂ must have undergone the changes in (11), accounting for the properties it has in Present-Day English.

- (11) Hypothesized changes of *I admit*₂ following cooptation
- a. Meaning: It was no longer part of the meaning of the sentence.
 - b. Function: It now serves a metatextual function.
 - c. Syntax: It lost its status as a syntactic constituent of the sentence.
 - d. Prosody: It appears to have lost in prosodic integration, as is suggested by the punctuation marks in (8).²²
 - e. Placement: It lost constraints on placement, now occurring in initial, medial, or final position (Brinton 2017: 182)

Similar observations were made by Fischer (2007a, 2007b) with reference to the development of the English theticals *I guess* and *I think*: "Once the parenthetical has evolved, it can be put in any position in the sentence". And once coopted as a thetical, *I admit*₂ was ready to grammaticalize in the direction of a DM, gradually losing the ability to take subject arguments other than the first person pronoun *I* and to be inflected for tense (Laurel Brinton, p.c. of 23-05-2019). Since, as we argue, *I admit* was coopted without a complementizer, we see no reason to assume that there was a grammaticalization process leading to the loss of the complementizer *that*.

To conclude, the development of the expression *I admit* appears to have been largely shaped by cooptation in accordance with the properties listed in (4) of Section 1, showing hardly any traces of pronounced grammaticalization following cooptation. As we will see in the next section, however, cooptation is frequently followed by pronounced grammaticalization.

22. It is possible, however, that the punctuation marks were added only later (see Traugott 2015).

3. The marker *if you will*

3.1 Grammaticalization

Once an expression has been coopted, this affects its external and its internal structure in different ways: Its external status changes from sentence constituent to sentence-external thetical which, depending on its specific discourse function, may be placed at the left or the right periphery of the sentence or else be inserted somewhere within the sentence.

In its internal status, both the form and the meaning of the expression are still largely what they used to be prior to cooptation (see Section 5.1). But when used recurrently, the expression may grammaticalize in accordance with the parameters listed in (12) (cf. (2) of Section 1). However, since the external structure of the coopted unit is not essentially affected by this process, grammaticalization will now take place on what may be called the metatextual level of discourse.

- (12) Grammaticalization following cooptation (Heine & Kuteva 2007: 33–44)
- a. Context extension: The expression is extended to new contexts inviting new meanings.
 - b. Desemantization: It loses semantic features in favor of the new functions evoked or supported by new contexts.²³
 - c. Internal decategorialization: The morphological components of the expression lose their morphosyntactic identity and coalesce into an invariable, particle-like unit.

This process can be illustrated with the English expression *if you will*, called a comment clause (Brinton 2008: 162–183), a discourse marker, a disjunct expressing meta-linguistic comments (Quirk et al. 1985: 618), or a formulaic in subordinate clause (Heine, Kaltenböck and Kuteva 2016: 56–58). It is first attested in the 10th and 11th centuries, as exemplified in (13).²⁴

- (13) *ealswa þu miht nu, gif þu wilt, æt God gebiddon þæt þes man gehæled wurdæ*
 ‘also you might now, if you will, entreat of God that this man be healed’
 (c1000 *The Old English Life of St. Giles* 218 [DOEC]; Brinton 2017: 169, (8a))

The expression, like *gif þu wilt* in (13), still had referential or propositional meaning in Old English but it was already a thetical. In fact, as the description by Brinton

23. An anonymous reviewer asks whether this change can be taken to reflect what in some studies is called ‘pragmatic strengthening’. The answer is in the affirmative.

24. The following account, though not the interpretation proffered, is based entirely on Brinton (2008: 168–80).

(2017: 168–171) suggests, it already exhibited the grammatical properties of a coopted unit listed in (4) of Section 1: (a) Its meaning was no longer part of the sentence meaning, (b) it appears to have had a metatextual function, (c) it was no longer a syntactic constituent of the sentence of (13), (d) it seems to have been prosodically separated from the rest of the sentence, as the commas to its left and its right might suggest, (e) it presumably had semantic-pragmatic scope beyond the sentence, and (f) it was no longer positionally constrained, occurring in initial, medial, or final position of the sentence (Brinton 2017: 168).

The existing evidence suggests that grammaticalization neither preceded nor coincided with cooptation; rather, it must have set in subsequently – that is, after *if you will* had turned into a thetical. The way this happened is described in detail by Brinton (2008: 177–180); the following is a short summary in accordance with the parameters in (12):

Context extension. By the modern period, *if you will* was extended from directive contexts, where its meaning was ‘if you are willing [to do so]’, to non-directive contexts inviting the meaning ‘if you are willing [to say so]’, and eventually to contexts where it assumed the function of a general hedge on the speaker’s claim.

Desemantization. *If you will* gradually lost its propositional meaning in favor of new meanings evoked or supported by the new contexts, such as expressing what Brinton (2017: 179) calls negative politeness. And it also seems to have lost its significance as an expression that is sensitive to referentiality and truth conditions.

Internal decategorialization. The three components of *if you will* gradually lost their morphological distinctiveness and coalesced into an invariable DM, undergoing internal, morphological fixation into a single, unalterable form (van Bogaert 2011: 308), and acquiring, as Fischer (2007b) suggests, features of a ‘lexicalized item’ – a process commonly referred to in grammaticalization theory as univerbation.²⁵

To conclude, the internal structure of *if you will* exhibits the main hallmarks of grammaticalization, as described by Brinton (2008: 162–183). But in its external structure, it has properties that are incompatible with grammaticalization, being the result of an earlier operation of cooptation, which accounts, e.g., for the fact that, like other theticals, *if you will* in Present-Day English is neither part of the syntax nor of the meaning of the sentence in which it occurs. This peculiar behavior, to be

25. Univerbation obtains when a collocation of two or more words or morphemes loses its internal morphological boundaries and turns into a new, invariable word (cf. Lehmann [1982] 2015: 161).

observed in a number of English DMs, is described by van Bogaert (2011: 302) as the ‘decategorialization paradox’.²⁶ On the view expressed in this paper, the paradox can be explained thus: Decategorialization applies only to the internal structure of DMs whereas their external structure is shaped by cooptation, which somehow has the opposite effects of decategorialization, namely those in (4) of Section 1.

3.2 On constituent anchored *if you will*

Among the grammatical changes typically associated with cooptation, as listed in (4) of Section 1, there are two that are seemingly problematic, namely that cooptation leads from restricted semantic-pragmatic scope to scope beyond the sentence (see (4e)) and from positionally constrained to largely unconstrained placement (4f). It would seem that this generalization is in need of qualification. In Heine et al. (2017: 823), a distinction is made between utterance anchored and constituent anchored coopted units (theticals).²⁷ Our interest in this paper is mainly with the former: They constitute the majority of theticals and the generalizations in (4) apply exclusively to them.

Constituent anchored theticals have one particular constituent of the host utterance as their anchor rather than the utterance as a whole, and this constituent is typically a noun phrase or an adverbial phrase. Unlike utterance anchored theticals, they have semantic-pragmatic scope only over that constituent and are placed immediately before or after that constituent – accordingly, (4e) and (4f) do not apply to them. What they share with utterance anchored DMs, however, is that their functions are metatextual (see (4b)) and, hence, their scope also relates to the situation of discourse.

Some utterance anchored theticals can also be used as constituent anchored theticals, with the effect that in such uses, both their semantic-pragmatic scope and their placement are restricted to the constituent concerned. This also applies to *if you will*. The data provided by Brinton (2017: 168–71, (8), (9)) suggest that *if you will* might have been essentially an utterance anchored thetical in Old and Middle English, cf. Example (13) above. But in the modern period from the mid-16th

26. “As a sign becomes more grammaticalized, it tends towards a fixed position in a given construction. This loss of syntagmatic variability is known as fixation. It can be seen that the evidence in favor of one grammaticalization criterion (decategorialization) flouts another (fixation). Thus, it would seem that the grammaticalization of CTMPs [complement taking mental predicates; a.n.] is plagued by a “decategorialization paradox”” (van Bogaert 2011: 302).

27. In addition, there are two further types which need not concern us here, namely context anchored coopted units having a host, and context anchored coopted units without a host (Heine et al. 2017: 823).

century onward there are also instances of constituent anchored *if you will*. Thus, Example (14) of a 17th century text suggests that *if you will* is anchored in and has scope over the noun *Quatrumvirate* rather than a larger piece of discourse, in that the ‘landing site’ of the thetical seems to be provided not by a sentence but rather by a noun phrase.

- (14) *The whole Triumvirate, or if you will, Quatrumvirate are included*
 (1684 Goddard, *Plato’s Demon; or the State-physician*
Unmaskt 53 [OED]; Brinton 2017: 171)

This situation suggests that constituent anchored *if you will* was a later innovation. In fact, this would not seem to be an isolated development. As the reconstruction by Davidse, De Wolf and Van linden (2015) shows, the DM *no doubt* arose in Middle English as an utterance anchored thetical with scope over the sentence as a whole. But in Late Modern English (1710–1920), the use of *no doubt* was extended to contexts where its semantic-pragmatic scope was restricted to one constituent of the sentence – that is, where it acquired uses as a constituent anchored DM.

Both cases are best interpreted as instances of grammaticalization following cooptation, taking place in specific contexts and leading to a reduction in semantic-pragmatic scope from sentence to sentence constituent – having the effect that the DM concerned is also restricted to placement immediately before or after that constituent. Assuming that this hypothesis is correct it would be in accordance with the scope restriction hypothesis of Lehmann ([1982] 2015), according to which grammaticalization leads or may lead to a shrinking in ‘structural scope’.

4. The French marker *alors*

The observations made in the preceding sections suggest that the cooptation hypothesis proposed in Heine (2013, 2018) to account for the development of DMs entails that a number of previous analyses of discourse markers and other theticals should be looked at from a somewhat different perspective. This perspective does not question the reconstructions proposed by the authors concerned but suggests that in order to understand salient properties of DMs, in particular those listed in (9), cooptation is needed as an explanatory notion.

This perspective does not only concern the development of English DMs, as can be illustrated with an example from a language other than English. In their discussion of the development of DMs, Degand and Evers-Vermeul (2015: 74–75) assert that the development of discourse markers “falls within the scope of grammaticalization” and that DMs “are a part of the grammar”. Neither of these conclusions is questioned here. Where we differ, however, is in arguing, first, that that

development cannot be reduced to grammaticalization and, second, that their account does not explain why DMs have the properties they have, such as the ones in (9) (see Heine 2018a: 42).

Both points can be illustrated with the very example that Degand and Evers-Vermeul (2015) and Degand and Fagard (2011) present in support of their grammaticalization hypothesis, namely the French marker *alors*, glossed in dictionaries as an adverb meaning ‘then, now’. The following sketch rests on the data provided by them (Degand and Evers-Vermeul 2015: 75–77; Degand and Fagard 2011) but proffers an alternative interpretation of these data.²⁸

Going back to the Latin prepositional phrase *ad illam horam* (‘at that hour’, or possibly the ablative form *illa hora*; Maj-Britt Mosegaard Hansen, p.c. of 25-06-2019), *alors* was grammaticalized to a sentence adverbial in Old French of the 12th century, an integrated adjunct of the sentence restricted to a clause-internal position and having temporal usages. This situation strongly suggests that it was a constituent of sentence grammar. And this situation does not appear to have changed when *alors* underwent further grammaticalization at the end of the 13th century, being extended to contexts where causal meanings evolved (cf. the parameter of context extension in (12a)), and from the 14th century on also occasionally conditional meanings. Note that changes from temporal to causal and conditional meanings are paradigm instances of grammaticalization (see Kuteva et al. 2019 for examples).

The data at our disposal do not allow establishing when exactly cooptation took place for the first time.²⁹ But beyond any reasonable doubt this must have happened at the latest in the 15th or 16th century in Middle French. Thus, Degand and Fagard (2011, Section 5.2) note that “*alors* jumps from medial position in Old French to initial position from Middle French onwards.” From this time on, *alors* exhibits the hallmarks of a coopted unit (see (9) of Section 2.2), like those in (15).³⁰

28. Note, however, that the interpretation offered in this section must be taken with care, for the following reason: With their emphasis on semantic reconstruction, Degand and Fagard (2011) employ a framework of analysis that does not always allow relating the categories used by them unambiguously to the notions used in this paper (see Section 1 above for the latter).

29. This does not apply to non-coopted uses of *alors*, which continued to be employed as a sentence grammar unit. Conceivably, uses of coopted *alors* can be traced back already to the 14th century, where *alors* is found in initial position set off by a comma from the following sentence (see Degand & Fagard 2011, ex. (18)).

30. Degand and Evers-Vermeul (2015: 76) argue that it was movement to the clause-initial position that was responsible for changes such as the ones in (4). However, there are also cases of cooptation where movement to the clause-initial position does not play any role (see Heine et al. 2017; Heine 2018a).

- (15) Properties suggestive of cooptation in the development of French *alors* (Degand & Fagard 2011; Degand & Evers-Vermeul 2015)
- a. Meaning: It is no longer part of the meaning of the sentence, it can be omitted without changing the semantic content, and it no longer establishes a temporal relation (Degand & Fagard 2011, Section 3.2).
 - b. Function: It acquired ‘metadiscursive’ functions typical of theticals, such as assuming a connective function not restricted to connecting adjacent clauses.³¹ And it was used to express coherence relations between adjacent clauses and to work as a discourse-structuring device.
 - c. Syntax: It turned from a syntactically integrated clause-internal adverbial into a non-integrated conjunct “outside the core syntactic clause” (Degand & Evers-Vermeul 2015: 75).
 - d. Prosody: There is no information on the prosodic form of coopted *alors*.
 - e. Semantic-pragmatic scope: Its scope was extended beyond the clause it was associated with.
 - f. Placement: It acquired positional flexibility, now being preferably placed clause-initially and, distinctly less commonly, also clause-finally.³²

There are unfortunately no appropriate text examples in the data at our disposal to show when and how exactly cooptation took place, our interpretation therefore rests entirely on the description provided by Degand and Fagard (2011) and Degand and Evers-Vermeul (2015). But it would seem that the undated example in (16) reflects an early use of *alors* suggestive of cooptation in the sense of (15).

- (16) French (Earlier French of unknown date; Degand & Fagard 2011, (18))
Or actendez, monseigneur, ce dit elle. Et maintenant vous me voiez bien, faictes pas? – Par Dieu! m’amy, nenny, dit monseigneur, comment vous verroie je? vous avez bouché mon dextre oeil, et l’autre est crevé passé a dix ans. – Alors, dist elle, or voy je bien que c’estoit songe voirement qui ce rapport m’a fait.
 ‘Wait a minute now, my Lord, she said. Now you can see me well, can’t you? – By God! My dear, no, said his Lordship, how could I see you? You have blocked my right eye, and the other one has been dead for ten years now. – *Alors*, she said, now I can see that it was really all a dream.’

31. Parameters used by Degand and Fagard (2011, Section 3.2) to define *alors* as a metadiscursive marker are: (a) It does not establish a temporal or argumentative relation, (b) it can be left out without changing the semantic content, and (c) it can be glossed by other French topic shifters, such as *bon* ‘well’, or transition markers, such as *et puis* ‘and then’.

32. Degand and Fagard (2011) note that from Middle French on, a majority of cases of *alors* is found in initial position, with a slow rise of occurrences in final position from Classical French onwards. See also Beeching (2016: 224).

With the transition to metadiscursive functions following cooptation, the earlier temporal meaning of *alors* as a sentence grammar adverb was not lost but was retained up to Present-Day French.³³

Subsequent to cooptation, *alors* underwent another round of grammaticalization from conjunct to DM, perhaps most dramatically in the 20th century when it acquired metadiscursive meanings as a topic shifter and transition marker. Thus, the overall development of *alors* seems to be compatible with the scenario in (17).

- (17) Main stages in the development of discourse markers (Heine 2013: 1238–1239)
(Grammaticalization >) cooptation > grammaticalization

In accordance with (17), cooptation can come in at any stage in the development of a grammaticalizing expression, with one exception: On account of the word constraint pointed out in Section 1, the expression must not have been grammaticalized to the extent that it has lost its status as an independent word or word collocation (see Section 1). Conceivably it is expressions that are used recurrently and are in the initial stage of grammaticalization that are particularly well suited for cooptation leading to DMs, but this is an issue that is in need of further analysis.³⁴ Note that after cooptation, the earlier expression will not disappear; rather, it is likely to continue to be used, as it appears to have happened with *alors*, whose use as a temporal adverb was continued up to Present-Day French, even if only in written use (Beeching 2016: 213). Not uncommonly therefore, the coopted and the non-coopted units co-occur in the synchronic state of the language concerned, being distinguishable on the basis of the properties listed in (9) of Section 2.2. A paradigm example was provided in Section 2.1, where we saw that non-coopted *I admit*₁ co-occurs with its coopted counterpart *I admit*₂ in Present-Day English.

To sum up, the development of *alors* from an adverbial, clause-internal use to discourse-structuring, peripheral (clause-initial and clause-final) use in spoken Present-Day French, or from ‘core grammar’ to ‘discourse grammar’ (Degand & Evers-Vermeul (2015: 77), is hard to reconcile with parameters of grammaticalization (Heine & Kuteva 2007: 33–44; Lehmann [1982] 2015: 132), even if one were to adopt a wider notion of grammaticalization.

On the view of Degand & Evers-Vermeul (2015) it was word order change, rather than cooptation, that was responsible for the rise of DMs like French *alors*, cf. (18), and Degand and Fagard (2011, Section 6.2) conclude: “Put more strongly, syntactic change would be a prerequisite for semantic change”. Extension

33. Beeching (2016: 213) points out, however, that the temporal reading of *alors* is only found in written, not in spoken Present-Day French.

34. We are grateful to Alexander Haselow (p.c. of 20-05-2019) for having made this suggestion.

of placement to the left or the right periphery of a sentence, which would be in accordance with the cooptation hypothesis (see (4f)), is well motivated, serving in particular information structuring functions (Detges & Beeching 2014: 2). But to establish that syntactic change is a prerequisite for semantic change is a hypothesis that would seem to be in need of further evidence.

- (18) Rather, a new syntactic position (clause-initial) gave rise to new meanings, requiring syntactic scope extension over the host clause in the case of the clausal conjunct, and over potentially more than the host clause in the case of the DM. (Degand & Evers-Vermeul 2015: 77)

There is no doubt that the left periphery of a sentence or utterance is the favored position of English DMs (see, e.g., Schiffrin 1987: 31–2, 328; Traugott 1995: 5–6; Jucker & Ziv 1998; Hansen 1998: 156; Brinton 1996: 33–35; Schourup 1999: 233; Maschler 2009: 44; Rysová 2017: 13), and the same applies to some other languages (Onodera 2011); see (4f) of Section 1. Still, there are also languages where this does not generally seem to be the case. For example, Dér and Markó (2010: 144–5) found that of the 14 most frequently used Hungarian DMs, only 39.6% occur in a turn-initial position while 40.2% occurred turn-medially, 8.3% turn-finally, and 11.9% were stand-alones.

Ignoring such observations, there are problems with the claim in (18). First, French *alors* is not and *was* not restricted to the initial position; already in its use as what Degand and Evers-Vermeul (2015) call a conjunct did it also occur occasionally in final position (see Beeching 2016 on right periphery *alors*). And second, there are DMs that do not and never did occur in clause-initial position. This appears to apply, for example, to the English DM *as it were*, as the reconstruction work by Brinton (2008: 171–5) suggests: Neither in its rise nor in its subsequent development or in its present use was this marker associated noticeably with the initial position, and one may wonder how the hypothesis in (18) would account for such cases.

To conclude, we surmise that movement of French *alors* to clause-initial position more likely was an epiphenomenal effect of cooptation, as predicted by the generalization in (4f): Once deployed as a thetical, *alors* could be placed not only within the sentence but also at its left and right periphery. And since DMs “prototypically introduce the discourse segments they mark” (Hansen 1997: 156), it comes as no surprise that the left periphery of an utterance tends to be a favored position of DMs.

5. Discussion

The framework proposed in this paper raises a number of questions, which are looked into in the present section.

5.1 Problems

The observations made in the preceding sections do not question the careful reconstruction work by Brinton (2017) and Degand and Evers-Vermeul (2015). Rather, they were meant to argue that these reconstructions can be enriched by also taking findings on cooptation into account. However, the cooptation hypothesis proposed in this paper has not gone unchallenged, it was criticized by these authors, cf. (19) and (20).

- (19) We cannot agree with this “spontaneous character” [of cooptation; a.n.] to account for the development of discourse markers, because many different studies in many different languages have exposed the gradual rise of linguistic expressions as discourse markers [...]. (Degand & Evers-Vermeul 2015: 72)
- (20) [G]rammaticalization [...] *motivates* the changes that we see in a way that the idea of cooptation, with a sudden change from sentence to discourse meaning and function and no reference to how semantic change works in particular cases, does not. That is, cooptation would not seem to explain why any particular form might be suited (and hence coopted) to serve a certain pragmatic function. Problems of dating, which Heine points out, are likely the result of insufficient or inadequate data. (Brinton 2017: 37)

A major concern of both authors of (19) and (20) is with the ‘spontaneous character’ or the ‘sudden change’ characterizing cooptation, we will return to this point in the next Section 5.2. Brinton (2017: 37) raises two further issues in the quotation of (20). First, she argues that cooptation “does not seem to explain why any particular form might be suited (and hence coopted) to serve a certain pragmatic function”. This problem applies in much the same way to grammaticalization, where there is so far no general agreement on how exactly its functional motivations are to be defined and how exactly these motivations determine the selection of forms or constructions serving as the input of the process.

The second issue concerns “problems of dating, which Heine points out, are likely the result of insufficient or inadequate data” (Brinton 2017: 37). Both the datings and the data used by Heine (2013: 1236, Table 1) in support of his cooptation hypothesis are taken from Brinton’s (2008) work. What Heine argues for is that grammaticalization takes a very long time to materialize. Now, the data provided by

Brinton (2008) suggest that the time span between the first attestation of a sentence grammar unit and its attestation as a DM was extremely short in a number of cases. Accordingly, an interpretation in terms of grammaticalization seems problematic in such cases while an interpretation in terms of cooptation is not, given that the latter does not need an extended time span to materialize (see Section 5.2).³⁵

5.2 Instantaneous vs. gradual change

As was observed in the preceding section, Degand and Evers-Vermeul (2015: 72) and Brinton (2017: 37) have a problem with regard to the ‘spontaneous character’ or ‘sudden change’ involved in cooptation (see (18) and (19)). It would seem, however, that this is not really a problem for an analysis of the development of DMs. Take the following text example of a paradigm instance of thetical use: At the moment when the text piece *we can't hide the fact* was coopted and inserted into the sentence of (21) in some specific situation of discourse, it instantly acquired all the properties defining it as a thetical, namely those in (9).

- (21) *And of course during the nineteenth century we can't hide the fact Egypt became very much of a hunting ground for agents on behalf of museums in Europe.*
(ICE-GB: s2a-052-81)

Cooptation thus is in fact an instantaneous cognitive-communicative operation at the disposal of the individual speaker, which helps him or her to respond more flexibly to the immediate requirements of a particular situation of discourse. It can be observed daily when we enrich our speech or writing with (paren)thetical expressions;³⁶ both its nature and its grammatical effects can therefore easily be tested without having to draw on historical text corpora.

Cooptation therefore contrasts with grammaticalization, whose outcome is the result of a gradual and a long process, as also pointed out by Degand and Evers-Vermeul (2015) and Brinton (2017). To be sure, there is also an alternative perspective: Grammaticalization can also be said to have a ‘spontaneous’ component when analyzed as consisting of a series of instantaneous acts of innovation, and it is only when viewed from the bird’s eye perspective of the historical linguist that this series of acts, taken together, has the appearance of a gradual or a continuous

35. Thus, Heine (2013: 1237) says: “In none of these examples do we find clear evidence that the rise of these early theticals was due to a prior process of grammaticalization.”

36. There are, however, clear differences between speech and writing in that theticals are used more commonly in the former than in the latter.

process, commonly described in terms of a ‘chain’ (Claudi & Heine 1986; Craig 1991: 455–6; Heine 1992; Heine, Claudi & Hünemeyer 1991: 220–228), a ‘continuum’ (Heine & Reh 1984: 15), a ‘scale’ (Lehmann 2015: 22), or a ‘cline’ (Hopper & Traugott 2003: 6–7).

Cooptation entails a fairly dramatic change in the external status of the expression concerned: The expression no longer belongs to sentence grammar. But its conceptual meaning, while no longer part of the sentence meaning, is not dramatically affected. For example, with its cooptation, the thetical *we can't hide the fact* in (21) essentially retained the semantic content it had prior to cooptation. This also means that cooptation can cut across an ongoing process of grammaticalization – with the effect that that process may now continue on a metatextual level, as appears to have happened in the case of French *alors* (see Section 4; Maj-Britt Mosegaard Hansen, p.c. of 28-06-2019).

But when used repeatedly, the unit coopted can turn into a constructional thetical (Heine et al. 2017: 819–20), that is, into what following Construction Grammar one may call a partially filled construction. In this case, its meaning is likely to change,³⁷ as seems to have happened with *I admit* (see Section 2.1), described by Brinton (2017: 168) with the formula [*I/you (modal) admit*], where *admit* is the substantive component while the rest is variable in limits.

Finally, in some very rare cases such as *if you will* and French *alors*, the coopted unit is adopted by a wider discourse community and may become conventionalized and, if also adopted by subsequent generations, it may grammaticalize, turning into what Hilpert (2015: 23) calls a grammaticalizing construction. In this case, its meaning will change in appropriate contexts in accordance with principles of grammaticalization. As a rule, however, grammaticalization does not affect its external status in that its thetical properties will be preserved (see (9) of Section 2.2).

In sum, the development of DMs has in fact the appearance of a gradual process but, in accordance with (17), this appearance is not the result of cooptation but rather of subsequent grammaticalization.

5.3 Alternative views

The question of how to explain the development of DMs has generated quite some research activity – an issue that we had to ignore in the previous sections. This development includes accounts in terms of lexicalization (Wischer 2000; Fischer 2007a, 2007b) but, perhaps most importantly, in terms of pragmaticalization

37. We are grateful to Laurel Brinton (p.c. of 23-05-2019) for having pointed this out to us, even if our interpretations are slightly different.

(e.g., Erman & Kotsidas 1993; Aijmer 1997; Günthner 1999; Dostie 2004: 26–8; Waltereit 2002, 2006; Frank-Job 2006; Ocampo 2006;³⁸ Hansen 2008: 64; Claridge & Arnovick 2010: 21–23; Arroyo 2011; Haselow 2011; Wiese 2011: 1017–1019; Beijering 2012: 56–59; Claridge 2013). In accordance with the approach proposed here, students of pragmaticalization maintain that the development exhibits properties that are beyond the scope of grammaticalization theory and, in fact, these are for the most part also the properties that define discourse markers as theticals.

As pointed out in Heine (2013: 1217–1219, 2018: 32), there are only two minor differences between the notions of pragmaticalization and the present approach: First, the latter makes a distinction between two mechanisms, that is, cooptation and grammaticalization. And second, cooptation is a more general mechanism that accounts for a wide range of ‘parenthetical’ phenomena of discourse (see Heine et al. 2017). And within this range, DMs form only a small subset of all the theticals to be found in the grammar of English or French.

In other approaches again, the development of DMs is classified as a special case or a subtype, or a subspecies of grammaticalization (Prévost 2011; Degand & Evers-Vermeul 2015: 65–67; see also Brinton 2017: 35). And some authors propose a wider or expanded notion of grammaticalization, or a view of grammaticalization as expansion rather than as reduction (e.g., Degand & Simon-Vandenberg 2011; Traugott & Trousdale 2013; Degand & Evers-Vermeul 2015). Considering that some of the properties characterizing this development, such as the ones listed in (4) of Section 1, seem to be incompatible with grammaticalization, however, one may hesitate to adopt such a view, unless one is prepared to run the risk of turning ‘grammaticalization’ into a largely vacuous term for any change in grammar (Heine 2018). This is also the position to be observed in other research traditions devoted to the study of grammaticalization, such as Cognitive Grammar (e.g., Langacker 2011).

But do DMs belong to ‘grammar’? Or do they only relate to a ‘wider’ or ‘expanded’ notion of grammar, as has been suggested in some studies (e.g., Traugott 1995; Hansen 1998; Lenker 2000; Traugott & Dasher 2002; Brinton & Traugott 2005: 136–140; Brinton 2008; Prévost 2011; Diewald 2011a, 2011b; van Bogaert 2011)? Since the present paper is restricted to diachronic regularities in the development of DMs, this question would need a separate treatment – one that should more generally be concerned with the foundations of grammatical theory building.

Nevertheless, for the time being we see no intrinsic reason to question the position taken in most mainstream models of contemporary linguistics as well as in reference grammars of English, such as Quirk et al. (1985: 631–646), Biber et al.

38. Rather than ‘pragmaticalization’, Ocampo (2006: 317) uses the term ‘discoursivization’.

(1999: 1086–1088), and Huddleston & Pullum (2002: 1356ff.),³⁹ according to which DMs indubitably belong to grammar (see Heine (2018: 42–3), or that taken within the framework of Discourse Grammar (Kaltenböck et al. 2011; Heine et al. 2013), where two components of grammar are distinguished, namely sentence grammar and thetical grammar, with DMs being part of the latter component.

6. Conclusions

In the *World Lexicon of Grammaticalization* of Kuteva et al. (2019), over 500 processes of grammaticalization have been identified based on data from more than a thousand languages. A predominant pattern of change to be observed in these processes is one where more or less free forms are gradually integrated semantically, morphosyntactically, and phonetically within a sentence, or a phrase. Of the 500-plus processes discussed, ten concern the development of discourse markers, and this development seems to be in accordance with the account proposed in this paper: Rather than being gradually integrated, the expressions giving rise to discourse markers move out of the sentence, they turn into semantically, morphosyntactically, and frequently also prosodically detached units serving the monitoring of discourse.

The reasons for this behavior were the subject of the present paper. In our daily linguistic interaction we constantly co-opt text pieces and insert them within a sentence or at its left or right periphery for metatextual functions of discourse processing. Once such a text piece is transferred via cooptation to the level of discourse processing, it turns into a parenthetical in the terminology of Kaltenböck (2007: 31) and Fischer (2007b), or a thetical in the terminology of Discourse Grammar (Heine et al. 2013), acquiring the properties that were listed in (9) of Section 2.2. On account of these properties, the piece will be elusive to the rules or conventions used for forming sentences; nevertheless, interlocutors will readily accept them as appropriate and meaningful contributions to linguistic discourse.

Most of these pieces are used once and never again, but some are used regularly, and a few are transmitted to subsequent generations of speakers, undergoing processes of grammaticalization, and turning into discourse markers, that is, invariable particles used recurrently and serving a range of discourse functions – we have looked at three of such particles in the preceding sections, namely English *I admit* and *if you will*, and French *alors*. Thus, like cooptation, grammaticalization is part

39. The term ‘discourse marker’ is used only in one of the three works mentioned, namely Biber et al. (1999: 1086–1088), whereas in Quirk et al. (1985: 631–646) DMs are mainly referred to as conjuncts and in Huddleston & Pullum (2002: 1356ff.) as supplements.

of the development of discourse markers, as aptly pointed out in many studies on this subject matter.

The paper was restricted to one single factor that can be identified in the rise of discourse markers. Accordingly, we had to ignore many other factors that also have, or may have a bearing on this development. Two factors in particular deserve mention. One concerns analogy (see especially Fischer 2007a). There is reason to assume, first, that changes as they can be observed in expressions like *I admit, if you will, or alors* were triggered or influenced by analogical processes as they took place in other linguistic material moving from sentence grammar to thetical grammar. And second, there is also reason to assume that analogy is a driving force that is at least as important in cooptation as it is in grammaticalization. The second factor is constructional change, as it has been described in a wide range of studies (e.g., Hilpert 2013; Traugott & Trousdale 2013). As such studies suggest, a perspective based on constructions rather than on single expressions allows for a distinctly more fine-grained analysis of grammatical change and, in fact, cooptation can be viewed as mechanism leading to the creation of a new kind of construction (see, e.g., Shibasaki 2018). Whether or to what extent such a perspective might affect the cooptation hypothesis expounded in the present paper is an issue to be left for future research.

The title of the paper may be seen as promising a comprehensive account of how discourse markers arise; yet, we were concerned only with one particular line of development. While this line is presumably the most common one in the languages surveyed there is at least one further line in addition, namely via borrowing. Crosslinguistically, discourse markers belong to the linguistic material that is fairly frequently transferred from one language to another in situations of language contact (see Heine 2016 for an overview). It would seem, however, that the outcome of such transfers does not differ essentially from that observed in the preceding sections, in that a borrowed discourse marker is likely to retain the properties defining it as a thetical, such as those listed in (9) for *I admit*₂.

To conclude, the observations made in this paper are in support of Ocampo's (2006: 317) view that grammaticalization is movement towards syntax and morphology whereas DMs move precisely to the opposite end, namely outside of syntax and towards discourse, and of Norde's (2009: 23) conclusion that "movement towards discourse is genuinely different from movement towards grammar, and the two are therefore best kept separate". Grammaticalization theory provides one of the most powerful tools available to the historical linguist but, as is argued in the paper, there are limits on where this tool can be applied.

Acknowledgements

The paper has benefitted greatly from valuable comments made by Mikyun Ahn, Kate Beeching, Laurel Brinton, Ulrike Claudi, Liesbeth Degand, Olga Fischer, Maj-Britt Mosegaard Hansen, Alexander Haselow, Diana Lewis, Seongha Rhee, and Elizabeth Traugott, to whom we wish to express our deeply-felt gratitude. Furthermore, our thanks go to two anonymous reviewers for highly constructive suggestions made on an earlier version of the paper.

References

- Aijmer, Karin. 1997. *I think* – An English modal particle. In *Modality in Germanic Languages: Historical and Comparative Perspectives*, Toril Swan & Olaf Jansen-Westvik (eds.), 1–47. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110889932.1>
- Aijmer, Karen. 2002. *English Discourse Particles: Evidence from a Corpus* [Studies in Corpus Linguistics 10]. Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.10>
- Arroyo, José Luis Blas. 2011. From politeness to discourse marking: The process of pragmaticalization of *muy bien* in vernacular Spanish. *Journal of Pragmatics* 43: 855–874. <https://doi.org/10.1016/j.pragma.2010.10.002>
- Beeching, Kate. 2016. *Alors/donc/then* at the right periphery: Seeking confirmation of an inference. *Journal of Historical Pragmatics* 17(2): 208–230. <https://doi.org/10.1075/jhp.17.2.03bee>
- Beijering, Karin. 2012. Expressions of Epistemic Modality in Mainland Scandinavian: A Study into the Lexicalization-grammaticalization-pragmaticalization Interface. PhD dissertation, Rijksuniversiteit Groningen.
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Boye, Kasper & Harder, Peter. 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in Language* 31(3): 569–606. <https://doi.org/10.1075/sl.31.3.03boy>
- Brinton, Laurel J. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110907582>
- Brinton, Laurel J. 2008. *The Comment Clause in English: Syntactic Origins and Pragmatic Development*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511551789>
- Brinton, Laurel J. 2010. From performative to concessive disjunct: *I/you admit* and *admittedly*. In *Language Change and Variation from Old English to Late Modern English: A Festschrift for Minoji Akimoto* [Studies in Language and Communication 14], 279–302. Bern: Peter Lang.
- Brinton, Laurel J. 2017. *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge: CUP. <https://doi.org/10.1017/9781316416013>
- Brinton, Laurel J. & Traugott, Elizabeth Closs. 2005. *Lexicalization and Language Change*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511615962>
- Claridge, Claudia. 2013. The evolution of three pragmatic markers: *As it were*, *so to speak/say* and *if you like*. *Journal of Historical Pragmatics* 14(2): 161–184. <https://doi.org/10.1075/jhp.14.2.01cla>
- Claridge, Claudia & Arnovick, Leslie. 2010. Pragmaticalisation and discursisation. In *Historical Pragmatics* [Handbook of Pragmatics 8], Andreas H. Jucker & Irma Taavitsainen (eds), 165–192. Berlin: De Gruyter Mouton.

- Claudi, Ulrike & Heine, Bernd. 1986. On the metaphorical base of grammar. *Studies in Language* 10(2): 297–335. <https://doi.org/10.1075/sl.10.2.03cla>
- Craig, Colette G. 1991. Ways to go in Rama: A case study in polygrammaticalization. In *Approaches to Grammaticalization*, Vol. 2 [Typological Studies in Language 19], Elizabeth Closs Traugott & Bernd Heine (eds), 455–92. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.19.2.20cra>
- Davidse, Kristin, De Wolf, Simon & Van linden, An. 2015. The development of the modal and discourse uses of *there/it is/I have no doubt* expressing modal and interactional meaning. *Journal of Historical Pragmatics* 16(1): 25–58. <https://doi.org/10.1075/jhp.16.1.02dav>
- Degand, Liesbeth & Evers-Vermeul, Jacqueline. 2015. Grammaticalization or pragmaticalization of discourse markers? More than a terminological issue. *Journal of Historical Pragmatics* 16(1): 59–85. <https://doi.org/10.1075/jhp.16.1.03deg>
- Degand, Liesbeth & Fagard, Benjamin. 2011. *Alors* between discourse and grammar: The role of syntactic position. *Functions of Language* 18(1): 29–56. <https://doi.org/10.1075/fol.18.1.02deg>
- Degand, Liesbeth & Simon-Vandenberghe. 2011. Introduction: Grammaticalization and (inter) subjectification of discourse markers. *Linguistics* 49(2): 287–294. <https://doi.org/10.1515/ling.2011.008>
- Dehé, Nicole. 2014. *Parentheticals in Spoken English: The Syntax-Prosody Relation* [Studies in English Language]. Cambridge: CUP. <https://doi.org/10.1017/CBO9781139032391>
- Dér, Csilla Ilona. 2010. On the status of discourse markers. *Acta Linguistica Hungarica* 57(1): 3–28. <https://doi.org/10.1556/ALing.57.2010.1.1>
- Dér, Csilla Ilona & Markó, Alexandra. 2010. A pilot study of Hungarian discourse markers. *Language and Speech* 53(2): 135–80. <https://doi.org/10.1177/0023830909357162>
- Detges, Ulrich & Beeching, Kate. 2014. *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Leiden: Brill.
- Diewald, Gabriele. 2011a. Grammaticalization and pragmaticalization. In *The Oxford Handbook of Grammaticalization*. Heiko Narrog & Bernd Heine (eds), 450–461. Oxford: OUP.
- Diewald, Gabriele. 2011b. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49(2): 365–390. <https://doi.org/10.1515/ling.2011.011>
- Dostie, Gaetane. 2004. *Pragmaticalisation et marqueurs discursifs: Analyse sémantique et traitement lexicographique*. Brussels: De Boeck & Larcier. <https://doi.org/10.3917/dbu.dosti.2004.01>
- Dostie, Gaétane. 2009. Discourse markers and regional variation in French: A lexico-semantic approach. In *Sociolinguistic Variation in Contemporary French* [Impact: Studies in Language, Culture and Society 26], Kate Beeching, Nigel Armstrong & Françoise Gadet (eds), 201–214. Amsterdam: John Benjamins. <https://doi.org/10.1075/impact.26.15dos>
- Erman, Britt & Kotsinas, Ulla-Britt. 1993. Pragmaticalization: The case of *ba* and *you know*. *Studier i Modern Sprakvetenskap* 10: 76–92.
- Fischer, Olga. 2007a. *Morphosyntactic Change*. Oxford: OUP.
- Fischer, Olga. 2007b. The development of English parentheticals: A case of grammaticalization? In *Tracing English through Time: Explorations in Language Variation: A Festschrift for Herbert Schendl on the Occasion of his 65th Birthday*, Ute Smit et al. (eds), 103–118. Vienna: Braumüller.
- Frank-Job, Barbara. 2006. A dynamic-interactional approach to discourse markers. In *Approaches to Discourse Particles*, Kerstin Fischer (ed.), 359–374. Amsterdam: Elsevier.
- Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics* 6 (2), 167–190.

- Fraser, Bruce. 1999. What are discourse markers. *Journal of Pragmatics* 31: 931–952.
[https://doi.org/10.1016/S0378-2166\(98\)00101-5](https://doi.org/10.1016/S0378-2166(98)00101-5)
- Fraser, Bruce. 2009. Topic orientation markers. *Journal of Pragmatics* 41: 892–898.
<https://doi.org/10.1016/j.pragma.2008.08.006>
- Furkó, Bálint Péter. 2014. Cooptation over grammaticalization. *Argumentum* 10: 289–300.
- Genette, Gérard. 1982. *Palimpsestes: La littérature au second degré*. Paris: Seuil.
- Günthner, Susanne. 1999. Entwickelt sich der Konzessivkonnektor *obwohl* zum Diskursmarker? Grammatikalisierungstendenzen im gesprochenen Deutsch. *Linguistische Berichte* 180: 409–446.
- Hansen, Maj-Britt Mosegaard. 1997. *Alors* and *donc* in spoken French: A reanalysis. *Journal of Pragmatics* 28: 153–187. [https://doi.org/10.1016/S0378-2166\(96\)00086-0](https://doi.org/10.1016/S0378-2166(96)00086-0)
- Hansen, Maj-Britt Mosegaard. 1998. The semantic status of discourse markers. *Lingua* 104(3–4): 235–260. [https://doi.org/10.1016/S0024-3841\(98\)00003-5](https://doi.org/10.1016/S0024-3841(98)00003-5)
- Hansen, Maj-Britt Mosegaard. 2008. *Particles at the Semantics/Pragmatics Interface: Synchronic and Diachronic Issues*. Oxford: Elsevier.
- Haselow, Alexander. 2011. Discourse marker and modal particle: The functions of utterance final *then* in spoken English. *Journal of Pragmatics* 43(14): 3603–3623.
<https://doi.org/10.1016/j.pragma.2011.09.002>
- Haselow, Alexander. 2016. A processual view on grammar: Macrogrammar and the final field in spoken syntax. *Language Sciences* 54: 77–101. <https://doi.org/10.1016/j.langsci.2015.12.001>
- Heine, Bernd. 1992. Grammaticalization chains. *Studies in Language* 16(2): 335–368.
<https://doi.org/10.1075/sl.16.2.05hei>
- Heine, Bernd. 2002. On the role of context in grammaticalization. In *New Reflections on Grammaticalization* [Typological Studies in Language 49], Ilse Wischer & Gabriele Diewald (eds), 83–101. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.49.08hei>
- Heine, Bernd. 2003. Grammaticalization. In *The Handbook of Historical Linguistics*, Brian D. Joseph & Richard D. Janda (eds), 575–601. Oxford: Blackwell.
<https://doi.org/10.1002/9780470756393.ch18>
- Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6): 1205–1247. <https://doi.org/10.1515/ling-2013-0048>
- Heine, Bernd. 2016. Language contact and extra-clausal constituents: The case of discourse markers. In Kaltenböck, Keizer & Lohmann (eds), 243–272.
- Heine, Bernd. 2018. Are there two different ways of approaching grammaticalization? In *New Trends on Grammaticalization and Language Change* [Studies in Language Companion Series 202], Sylvie Hancil, Tine Breban & José Vicente Lozano (eds), 23–54. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.202.02hei>
- Heine, Bernd, Claudi, Ulrike & Hünemeyer, Friederike. 1991. *Grammaticalization: A Conceptual Framework*. Chicago IL: University of Chicago Press.
- Heine, Bernd, Kaltenböck, Gunther & Kuteva, Tania. 2016. On insubordination and cooptation. In *Insubordination* [Typological Studies in Language 115], Nicholas Evans & Honoré Watanabe (eds), 39–63. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.115.02hei>
- Heine, Bernd, Kaltenböck, Gunther, Kuteva, Tania & Long, Haiping. 2013. An outline of discourse grammar. In *Functional Approaches to Language*, Shannon Bischoff & Carmen Jany (eds), 175–233. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110285321.155>
- Heine, Bernd, Kaltenböck, Gunther, Kuteva, Tania & Long, Haiping. 2017. Cooptation as a discourse strategy. *Linguistics* 55: 1–43. <https://doi.org/10.1515/ling-2017-0012>

- 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* [Studies in the Development of Language 9]. Oxford: OUP.
- Heine, Bernd & Reh, Mechthild. 1984. *Grammaticalization and Reanalysis in African languages*. Hamburg: Buske.
- Hengeveld, Kees & Mackenzie, J. Lachlan. 2008. *Functional Discourse Grammar: A Typological-Based Theory of Language Structure*. Oxford: OUP.
<https://doi.org/10.1093/acprof:oso/9780199278107.001.0001>
- Hengeveld, Kees & Mackenzie, J. Lachlan. 2011. *Functional Discourse Grammar*. In *The Oxford Handbook of Linguistic Analysis*, Bernd Heine & Heiko Narrog (eds), 367–400. Oxford: OUP.
- Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word-formation, and Syntax*. Cambridge: CUP. <https://doi.org/10.1017/CBO9781139004206>
- Hilpert, Martin. 2015. From hand-carved to computer-based: Noun-participle compounding and the upward strengthening hypothesis. *Cognitive Linguistics* 26(1): 23–36.
<https://doi.org/10.1515/cog-2014-0001>
- Hopper, Paul J. 1991. On some principles of grammaticization. In: Elizabeth Closs Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization*. Volume 1. Amsterdam: John Benjamins, 17–35.
- Hopper, Paul J. & Traugott, Elizabeth Closs. 2003. *Grammaticalization* (2nd rev. edn). Cambridge: CUP. <https://doi.org/10.1017/CBO9781139165525>
- Huddleston, Rodney & Pullum, Geoffrey K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP. <https://doi.org/10.1017/9781316423530>
- Jucker, Andreas H. & Ziv, Yael. 1998. Discourse markers: Introduction. In *Discourse Markers: Descriptions and Theory* [Pragmatics & Beyond New Series 57], Andreas H. Jucker & Yael Ziv (eds), 1–12. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.57.03juc>
- Kaltenböck, Gunther. 2007. Spoken parenthetical clauses in English. In *Parentheticals*. [Linguistik Aktuell/Linguistics Today 106], Nicole Dehé & Yordanka Kavalova (eds), 25–52. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.106.05kal>
- Kaltenböck, Gunther, Heine, Bernd & Kuteva, Tania. 2011. On thetical grammar. *Studies in Language* 35(4): 848–893. <https://doi.org/10.1075/sl.35.4.03kal>
- Kuteva, Tania, Heine, Bernd, Hong, Bo, Long, Haiping, Narrog, Heiko & Rhee, Seongha. 2019. *World Lexicon of Grammaticalization*, 2nd extensively revised and updated edn. Cambridge: CUP. <https://doi.org/10.1017/9781316479704>
- Langacker, Ronald W. 2011. Grammaticalization and Cognitive Grammar. In *The Oxford Handbook of Grammaticalization*, Heiko Narrog & Bernd Heine (eds), 79–91. Oxford: OUP.
- Lehmann, Christian. [1982] 2015. *Thoughts on Grammaticalization*, 3rd edn [Classics in Linguistics 1]. Berlin: Language Science Press.
- Lenker, Ursula. 2000. *Soþlice* and *witodlice*: Discourse markers in Old English. In *Pathways of Change. Grammaticalization in English* [Studies in Language Companion Series 53], Olga Fischer, Anette Rosenbach & Dieter Stein (eds), 229–249. Amsterdam: John Benjamins.
<https://doi.org/10.1075/slcs.53.12len>
- Levinson, Stephen C. 2006. Deixis. In: Laurence R. Horn & Gregory L. Ward (eds.), *The Handbook of Pragmatics*. London: Blackwell, 87–120.

- Maschler, Yael. 2009. *Metalanguage in Interaction: Hebrew Discourse Markers* [Pragmatics & Beyond New Series 181]. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.181>
- Matasović, Rolf. 2008. Patterns of grammaticalization and the layered structure of the clause. In *New applications of Role and Reference Grammar: Diachrony, Grammaticalization, Romance Languages*, Rolf Kailuweit, Björn Wiemer, Eva Staudinger & Rolf Matasović (eds), 45–57. Newcastle upon Tyne: Cambridge Scholars.
- Norde, Muriel. 2009. *Degrammaticalization*. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780199207923.001.0001>
- Ocampo, Francisco. 2006. Movement towards discourse is not grammaticalization: The development of /claro/ from adjective to discourse particle in spoken Spanish. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, Nura Sagarra & Almeida Jacqueline Toribio (eds), 308–319. Somerville MA: Cascadilla Proceedings Project.
- Onodera, Noriko Okada. 2011. The grammaticalization of discourse markers. In *The Oxford Handbook of Grammaticalization*, Heiko Narrog & Bernd Heine (eds), 611–620. Oxford: OUP.
- Prévost, Sophie. 2011. *A propos* from verbal complement to discourse marker: A case of grammaticalization? *Linguistics* 49(2): 391–413. <https://doi.org/10.1515/ling.2011.012>
- Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Rhee, Seongha. 2018. Grammaticalization of the plural marker in Korean: From object to text to stance. *Journal of Language Sciences* 25(4): 221–249. <https://doi.org/10.14384/kals.2018.25.4.221>
- Rysová, Magdaléna. 2017. Discourse connectives: From historical origin to presentday development. In *New Perspectives on Cohesion and Coherence*, Katrin Menzel, Ekaterina Lapshinova-Koltunski & Kerstin Kunz (eds), 11–35. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.814460>
- Schiffirin, Deborah. 1987. *Discourse Markers* [Studies in Interactional Sociolinguistics 5]. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511611841>
- Schneider, Stefan. 2007. *Reduced Parenthetical Clauses as Mitigators: A Corpus Study of Spoken French, Italian and Spanish* [Studies in Corpus Linguistics 27]. Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.27>
- Schourup, Lawrence Clifford. 1999. Discourse markers. *Lingua* 107: 227–265. [https://doi.org/10.1016/S0024-3841\(96\)90026-1](https://doi.org/10.1016/S0024-3841(96)90026-1)
- Shibasaki, Reijirou. 2018. From the inside to the outside of the sentence: Forming a larger discourse unit with *jijitsu* ‘fact’. In *New Trends on Grammaticalization and Language Change* [Studies in Language Companion Series 202], Sylvie Hancil, Tine Breban & José Vicente Lozano (eds), 333–360. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.202.14shi>
- Schneider, Stefan, Glikman, Julie & Mathieu Avanzi (eds). 2015. *Parenthetical Verbs*. Berlin: De Gruyter.
- Tabor, Whitney & Traugott, Elizabeth C. 1998. Structural scope expansion and grammaticalization. In *The Limits of Grammaticalization* [Typological Studies in Language 37], Anna Giacalone Ramat & Paul J. Hopper (eds), 229–272. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.37.11tab>
- Thompson, Sandra A. & Mulac, Anthony. 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In *Approaches to Grammaticalization*, Vol. 2 [Typological Studies in Language 19], Eliabeth C. Traugott & Bernd Heine (eds), 313–329. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.19.2.16tho>

- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In *Directions for Historical Linguistics*, [Current Issues in Linguistic Theory 24], Winfred P. Lehmann & Yakov Malkiel (eds), 245–271. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.24.09clo>
- Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65(1): 31–55. <https://doi.org/10.2307/414841>
- Traugott, Elizabeth Closs. 1995. The role of the development of discourse markers in a theory of grammaticalization. Paper presented at the International Conference of Historical Linguistics XII, Manchester 1995.
- Traugott, Elizabeth Closs. 2015. Investigating “periphery” from a functionalist perspective. *Linguistics Vanguard* 1(1): 119–130. <https://doi.org/10.1515/lingvan-2014-1003>
- Traugott, Elizabeth Closs. 2018. Modeling language change with constructional networks. In *Beyond Grammaticalization and Discourse Markers: New Issues in the Study of Language Change* [Studies in Pragmatics 18], Salvador Pons Bordería & Óscar Loureda Lamas (eds), 17–50. Leiden: Brill.
- Traugott, Elizabeth Closs & Dasher, Richard B. 2002. *Regularity in Semantic Change* [Cambridge Studies in Linguistics 96]. Cambridge: CUP.
- Traugott, Elizabeth Closs and Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780199679898.001.0001>
- Urmson, James Opie. 1952. Parenthetical verbs. *Mind* (New Series) 61(244): 480–496. <https://doi.org/10.1093/mind/LXI.244.480>
- van Bogaert, Julie. 2011. *I think* and other complement-taking mental predicates: A case of and for constructional grammaticalization. *Linguistics* 49(2): 295–332.
- Waltereit, Richard. 2002. Imperatives, interruption in conversation and the rise of discourse markers: A study of Italian *guarda*. *Linguistics* 40: 987–1010. <https://doi.org/10.1515/ling.2002.041>
- Waltereit, Richard. 2006. The rise of discourse markers in Italian: A specific type of language change. In *Approaches to Discourse Markers* [Studies in Pragmatics 1], Kerstin Fischer (ed.), 61–76. Amsterdam: Elsevier.
- Wiese, Heike. 2011. *So* as an emerging focus marker in German. *Linguistics* 49(5): 991–1039. <https://doi.org/10.1515/ling.2011.028>
- Wischer, Ilse. 2000. Grammaticalization versus lexicalization: ‘Methinks’ there is some confusion. In *Pathways of Change. Grammaticalization in English* [Studies in Language Companion Series 53], Olga Fischer, Anette Rosenbach & Dieter Stein (eds), 355–370. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.53.17wis>
- Witosz, Bożena. 2017. Metatext in the discourse of the theory of text, stylistics and pragmalinguistics. *Forum Lingwistyczne* 4: 107–112.

On the pragmatic development of modal particles in Navarrese-Lapuridian Basque

Hori emain ote nauzu?*

Sergio Monforte

University of the Basque Country

The grammaticalisation of modal particles has mainly been focused on their source and historical development. This article deals with the ongoing pragmatic extension of the epistemic particle *ote* in the Navarrese-Lapuridian dialect. When used in questions, this particle has traditionally expressed that the speaker believes that none of the potential denoted responses will have the necessary strength of commitment and, therefore, no answer is expected. Based on the fact that questions containing *ote* are felt as less direct and that the particle minimises the interrogative force, speakers of such dialect have reinterpreted the use of *ote* as a resource to reduce the expectancy and directiveness of information seeking questions where the speakers hope to receive a response.

Keywords: modal particles, pragmatic extension, grammaticalisation, Basque

1. Introduction

In this paper an attempt to account for the new use *ote* found in Navarrese-Lapuridian Basque is put forth.¹ I propose that *ote* has experienced (a maybe first step of) the grammaticalization into a question particle in the sense of Heine (2003). Typological evidence (Bayer 2012; Kinuhata 2012; Hack 2014; Yang & Wiltschko 2016; Farkas 2017) favours this analysis; even intralanguage data put forward this path ‘modal particle > question particle’ as the relation between the modal particle *ahal* and the question particle *al* suggests (Monforte 2018, 2020a).² Nevertheless,

* Translated into English as ‘Will you give me that?’

1. I follow the classification of the current Basque dialects proposed by Zuazo (2014) along this article.

2. Modal particles in Basque are claimed to derive from verbs (Lakarra 2019; Monforte 2020a) except the question particle *al* which would have developed from the epistemic particle *ahal*.

the use of *ote* in such contexts is still restricted to questions conveying a request or invitation and those which may be considered as intrusive, hence, it is pragmatically restricted, unlike *al* which can occur in a wide range of question types. The aim of the speaker to weaken the strength of the interrogative justifies the use of *ote* in those kinds of questions since the general *ote* used in all dialects modifies the illocutionary force turning an information seeking question into a conjectural one, i.e. blocking the conversational obligation to provide an answer to the uttered question. This effect was reinterpreted by speakers of those varieties as a source to reduce inquiring nature of questions and, therefore, it may function as a politeness marker.

The modal particle *ote*, used historically and in the present day in all Basque dialects, has been described to add the speaker's doubts to the proposition it appears, mostly questions (Larramendi 1729; Harriet 1741; Lécluse 1826; Gèze 1873; Astarloa 1883; Azkue 1891; Ithurry 1895; Chaho 1836; Altube 1930; Villasante 1980; de Rijk 2008; Garmendia 2014; Monforte 2020b). Also, some grammarians as de Rijk (2008) claim that *ote* contributes to the proposition by adding an 'I wonder' effect, for instance:

- (1) Nora joan ote da oporretan?
 where.ADL go PART AUX.3SG.ABS holidays.IN
 '[I wonder] where did s/he go on holidays?'

This particle occurs not only in *wh*-questions but also in polar and disjunctive questions conveying the same interpretation:

- (2) a. Egia esan ote du?
 truth.ABS say PART AUX
 '[I wonder] Did s/he say the truth?'
 b. Aita ala semea izan ote da?
 Father.ABS or son.ABS be PART AUX
 '[I wonder] Was it the father or the son?'

Nevertheless, this is not the only clause type in which it can be used: *ote* can also appear in non-interrogative contexts under attitudinal predicates (Döring 2007) and belief-, knowledge- and conjectural-predicates (Dor 2005). For instance:

- (3) Samar bat irintxi ote dian iruditu zatan.
 trace one swallow PART AUX.C seem AUX
 'It seems to me that I may have swallowed some (cider) dregs.'

(Barandiaran 1972)

This behaviour is reasonably consistent along all dialects; however, speakers of varieties from the French territory have developed a new pragmatic use during the last century.³ Let us briefly exemplify this:

- (4) Nahi ote duzu eskual bixkotxa?
 want PART AUX basque cake.ABS
 ‘May you like a piece of Basque cake?’

This use apparently shows no syntactic difference compared to the previous one,⁴ as *ote* behaves as expected for all modal particles. Nevertheless, the use of *ote* in such contexts is pragmatically deviant for speakers from other varieties. Along this article I study this new pragmatic use emerged, at least,⁵ in the Navarrese-Lapurdian dialect based on spoken data and written sources. In fact, in addition to literary works written in the local variety (e.g. Larzabal, Etxamendi and Thikoipe), 40 interviews were conducted to elicitate data regarding the use of the particle *ote* and other particles following the methodology put forth by Matthewson (2004) and Burton and Matthewson (2015). The sample was representative of the following dialects: Western dialect, (4 interviews) Central dialect (12 interviews), Navarrese (8 interviews) Basque and Navarrese-Lapurdian Basque, included Amikuze’s area (16 interviews). Interviewees were chosen following the standards in the traditional dialectology (Camino 2009); therefore, they were native speakers of the dialect spoken in each area and the influence of the standard language was none or very little. Two groups were differentiated regarding their age and gender: 10 people were men and 22–42 years old, 10 people were women and 22–42 years old, 10 people were men and 52–80 years old and 10 people were women and 52–80 years old. During the interviews speakers were exposed to some situations and asked to take part in the development of the situation by producing the clauses they found suitable in such situations. Since the conversations were guided, all speakers agree on the same clauses in almost all contexts; for instance, all interviewees produced the following

3. First testimonies of this use are found in theater plays written by Larzabal in the XXth century.

4. The particle *ote* has a high degree of microvariation in Eastern Basque, namely that concerning its syntactic behaviour. However, the separate syntactic pattern found in those varieties does not correlate with the distinct interpretational use dealt with in this paper. Indeed, both developments do not share the same geographical area, since the syntactically differentiated *ote* does not occur and is not accepted in the area extended along the Lapurdian coast where the use of *ote* as in Example (4) seems to be stronger and have spread from.

5. The varieties spoken in French territory are grouped in two dialects following Zuazo’s (2014) classification: the Navarrese-Lapurdian Basque and Souletin Basque. Data from the latter were not included in this work.

clauses in the same contexts, something expected: ‘Are you ok?’, ‘[I wonder] where did I leave the newspaper?’, ‘Would you like a piece of cake?’ and so on.

This paper is organised as follows: first, some aspects of the Basque grammar dealt with below are briefly described in Section (2) and, also, the role of syntax in this separate interpretation is ruled out; in Section (3) a look is taken into the modal particle *ote* and its canonical contribution to the proposition; Section (4) discusses typological evidence of two types of change in modal particles, on the one hand, their grammaticalization into question particles and, on the other hand, their secondary use as politeness markers; following this in Section (5), I return to the Basque particle *ote* and it is proposed that speakers from Navarrese-Lapurdian dialect has reinterpreted it and spread its use to information seeking questions; finally, some conclusions are presented in Section (6).

2. Some basic notions on question formation and modal particles in Basque

The neutral sentential order has been claimed to be SOV (de Rijk 1969; Villasante 1980; Hualde & Ortiz de Urbina 2003; Pastor 2019);⁶ however, this may vary due to the movement of constituents related to the information structure such as topics or foci:

- (5) a. Maitena Ziburun bizi da.
Maitena Ziburu.IN live AUX
‘Maitena lives in Ziburu.’
- b. MAITENA [bizi da] Ziburun **bizi da**?
Maitena live AUX Ziburu.IN
‘It is Maitena who lives in Ziburu?’

Since the modal particle *ote* is mainly used in questions, let us look into their syntactic derivation. Root questions in Basque, whether they contain modal particles or not, are marked in syntactic terms by movement of the verbal complex (i.e. the lexical and auxiliary verbs) to the CP-domain and, if there were a focal constituent, those would be also fronted to such domain (see Example (6a) and (b)):

- (6) a. Non [bizi da] Maitena ~~non~~ bizi da?
where live AUX Maitena
‘Where does Maitena live?’

6. This topic has long discussed in Basque linguistics and, although most researchers agree that the order SOV is the neutral one in Basque, we can also find other opinions such as those of Hidalgo (1999) and Esnal (2012).

- b. ZIBURUN [bizi da] Maitena **bizi da?**
 Ziburu.IN live AUX Maitena
 ‘Is it Ziburu where Maitena lives?’

As stated above, the presence of modal particles does not modify such derivation described in Examples ((6a) and (b)); therefore, the use of *ote* does not alter the word order arrangement in questions:

- (7) a. Non [bizi ote da] Maitena ~~non bizi ote da?~~
 where live PART AUX Maitena
 ‘Where does Maitena live? (I’m wondering)’
 b. ZIBURUN [bizi ote da] Maitena **bizi ote da?**
 Ziburu.IN live PART AUX Maitena
 ‘Is it Ziburu where Maitena lives? (I’m wondering)’

In what concerns this paper, questions containing the modal particle *ote* and those lacking it are syntactically derived in the same way and the interpretational difference between both cases is due to the contribution of the particle to the proposition and its interaction with the illocutionary force. The function of modal particles seems to happen covertly since modal particle in Basque occupy a position in the TP-domain and there is no clear evidence of the relation between modal particles and illocutionary force. Let us briefly describe their syntactic behaviour.

Traditionally, grammarians (de Rijk 1969; Euskaltzaindia 1987; Albizu 1991; Elordieta, G. 1997; Elordieta, A. 2001) have grouped some modifiers of the illocutionary force together based on their similar syntactic behaviour, namely *modal particles*. Although the number of modal particles in Basque is not as high as in Asian and Germanic languages, they can be classified into different groups concerning their interpretation or function: those conveying evidentiality (*ei, omen*), those expressing epistemicity (*ahal, bide, ote*) and the interrogative particle (*al*). Let us briefly exemplify this:

- (8) Barakaldon hasi omen da lanean.
 Barakaldo.IN begin PART AUX work.IN
 ‘Apparently, s/he began working in Barakaldo.’
 (9) Itoko ahal haiz!
 drown.FUT PART AUX
 ‘I wish/ I do hope you drown!’
 (10) Bazkaldu al duzu etxean?
 have-lunch PART AUX house.IN
 ‘Did you have lunch at home?’

The main shared property by these particles is that they all form a morphological word with the inflected verb. Evidence of this behaviour is found in (a) negative contexts (11a) and (b) focus contexts in eastern dialects (11b) since the particle moves along with the finite verb to the left; and also in (c) non-inflected clauses where modal particles cannot occur (11c and d) since there is no finite verb:

- (11) a. Ez al duzu egunkaria erosi ez al duzu?
not PART AUX newspaper.ABS buy
'Didn't you buy the newspaper?'
- b. Jonek ote dua Jonek hori erran ote du?
Jon.ERG PART AUX.PART that.ABS say
'Was it Jon who said that? (I'm wondering)'
- c. Ea hori erosi (*al) eta bestea bota al duen ea galdetu dut.
this.ABS buy PART and other.ABS throw PART AUX.C ask AUX
'I asked whether s/he bought this one and threw away the other one.'
- d. Ez dakit nora joan (*ote)⁷
not know where.ADL go PART
'I don't know where to go.'

As said in Section (1), the use of *ote* under exam in this paper shows no apparent distinct syntactic behaviour; this can be noted in the following example, since speakers from other varieties interviewed for this work (speakers of western, central and Navarrese Basque) accept it but their interpretation is different:

- (12) Erretzen ahal ote dut hemen?⁸
burn.IPFV can PART AUX here
'[I wonder] Can I can smoke here? [standard interpretation]
'May I smoke here?' [interpretation limited to the Navarrese-Lapurdián Basque]
(Thikoipe 2009: 113)

In the first interpretation no answer is expected, since it is assumed that neither the speaker, nor the addressee knows the answer. However, in the second interpretation

7. The symbol * indicates that the sentence would be ungrammatical if the constituent between brackets were included in the sentence.

8. An anonymous reviewer wonders about the syntactic category the modal verb *ahal* belongs to and whether it should be considered an auxiliary. Although I cannot go into greater detail, auxiliary and modal verbs have their own particular syntactic behaviours (auxiliaries are fronted along with the negation particle, whereas modal verbs are not), as well modal verbs need auxiliaries to create a grammatically correct finite sentence.

Also this anonymous reviewer wonders whether the sequence *ahal ote* is a collocation in Basque. Particles in Basque always occur preceding the finite verb; that means that it appears following the lexical or modal verb in those contexts where the adjacency between lexical and finite verbs is not broken.

the speaker awaits a reply from the addressee. I will turn back to both interpretations in Sections 3 and 5.

To sum up, the new use presented here is not attributed to a differentiated syntactic way of the modal particle *ote* to modify the illocutionary force. In fact, syntax is discarded as the driving force of this separate use and focus on an ongoing change regarding the interpretation of the particle. Below, I examine the common interpretation of *ote* in standard Basque, also common to all dialects.

3. The modal particle *ote*: Its general contribution

The occurrence of this particle is limited to some kind of clauses, leading to two separate interpretations, as shown in the following lines. First, its prototypical use occurs in questions; pragmatically *ote* functions as an indicator for an ‘I wonder’ interpretation in questions:

- (13) a. Etxera heldu da?
house.ADL arrive AUX
‘Did s/he get home?’
b. Etxera heldu ote da?
house.ADL arrive PART AUX
‘[I wonder] Did s/he get home?’

It appears not only in polarity question, but also in *wh*-questions and disjunctive questions as seen in (1) and (2b). Based on that interpretation, it seems that the use of the modal particle *ote* is intrinsically linked to conjectural questions (Littell *et al.* 2010).⁹ This type of question conveys that the speaker does not know the answer and does not expect to receive it from the addressee. Therefore, the use of *ote* in an information seeking question is not pragmatically correct, as the following example proves since the information data asked by the speakers is expected to be known by the addressee:

- (14) #Nora joan ote zara oporretan?¹⁰
where.ADL go PART AUX.2SG.ABS holidays.IN
‘[I wonder] where did you go on holidays?’

9. Rhetorical questions can also contain the particle *ote*; however, their use in such contexts is not obligatory and it is not clear which function and contribution they realise in that kind of questions. This situation is not only restricted to the field of modal particles in Basque since German modal particles also display a similar situation (Gödeke 2019).

10. I use the symbol # to mark pragmatic deviance following the general convention.

Additionally, it can also appear in non-interrogative contexts but only if it is embedded under attitudinal predicates (Döring 2007) and belief-, knowledge- and conjectural-predicates (Dor 2005):

- (15) Geixko esan ote dezun iruditzen zait.
 bad say PART have.C seem AUX
 ‘It seems to me that you may have said it wrong.’ (Altzaga 1994: 101)

The function of *ote* in these contexts seems to be related to the epistemic evaluation of the speaker to the proposition. Hence, I conclude that *ote* is an epistemic particle which conveys the speaker’s judgment about the proposition not only in declarative sentences, but also in questions. Indeed, Littell et al. (2010) claim that epistemic particles in Lillooet Salish, Thompson Salish and Giksan languages turn a *bona fide* question into a conjectural one by producing a set of weakened possible answers. This weakening of the strength of the answers is obtained by using an epistemic particle. A similar analysis can be proposed in the case of *ote*; consider the following example:

- (16) Non ote dago semea? (Etxean ote? Eskolan ote?)
 where PART be SON.ABS house.IN PART school.IN PART
 ‘I’m wondering where s/he went on holidays. (Maybe at home? Maybe at the school?...)’

Therefore, in the case of *wh*-questions containing *ote*, those produce a set of possible answers lacking enough commitment to be determined as the answer to the *wh*-word.

The fact that *ote* expresses that lack of commitment to the possible answers implies the sense that no reply is needed in such questions (since no possible answer can satisfy the interrogative requirements) and, thus, the interrogative force and the directiveness of the question are minimised. This would explain the kind of questions *ote* is spreading to in Navarrese-Lapurcian varieties, namely certain information seeking questions which can be seen as costly or intrusive for the addressee. In what follows I will focus on this innovation concerning the pragmatic use of *ote* in Navarrese-Lapurcian varieties; nevertheless, I will first present similar grammaticalisation processes or analyses put forth for other languages.

4. On the literature of pragmatic change regarding modal particles

Previously in Section (1) I have hypothesised that speakers from Navarrese-Lapurdian may have reinterpreted the contribution of the modal particle *ote* and expanded its use to information seeking questions, a context banned for the rest of dialects. Before examining this piece of microvariation concerning the particle *ote* in Navarrese-Lapurdian varieties, I will bring up similar cross-linguistic hypotheses analysing similar uses and development of modal particles.

Below, I will present two different approaches to the variation regarding some modal particles which occur in interrogative contexts. First, hypotheses postulating that modal particles have become question particles will be briefly introduced; then, analyses discussing a secondary use of modal particles as markers of politeness will be put forward.

4.1 The grammaticalization of modal particles into question particles

Some works such as Bayer (2012) and Hack (2014) suggest that the origin of the question particles in Bavarian German and Dolomitic Ladin is found in some epistemic particles. With regard to the German language, the modal particle *denn* (related to English ‘then’ (Bayer & Obenauer 2011)) has been described as a marker of the epistemic attitude conveyed by the speaker and it can occur in polarity and *wh*-questions in the standard language when speakers are concerned about the answer their questions will elicit (Bayer & Obenauer 2011; Bayer 2012):

- (17) Wo hast du denn meine Schlüssel hingelegt?
 where have you PART my keys put-down
 ‘Where did you put my keys? (I’m wondering)’
 (Bayer & Obenauer 2011: 454–455)

In the Bavarian dialect, however, this particle, which has been phonologically reduced from *denn* to *-n*, has turned into a question particle; in fact, its occurrence is obligatory in *wh*-questions and it does not express any semantic cue.¹¹ For instance:

- (18) Wann hod-a-s-da-n zoagt?
 when has-he-it-you-PART shown
 ‘When did he show it to you?’
 (Bayer 2012: 13)

11. Speakers of the Bavarian dialect use another modal particle, i.e. *no(u)*, in order to express what speakers from other areas convey by adding *denn* to the question (Bayer 2012: 13–14).

Similar processes have been claimed for particles of other languages such as the question particle *po/pa* (similar to English ‘then’) in Dolomitic Ladin and the interrogative particle *ka* in Japanese.¹² In the case of the particle *po/pa* Hack (2014) provides evidence of this phenomenon in different varieties of Dolomitic Ladin; as we can see in the following examples, the same sentence can be interpreted uniquely in different varieties:

- (19) a. Olà vas=to pa? (Fodom variety)
 where go.2SG=CL PART
 ‘Where are you going now (given that x has happened)?’ (Hack 2014: 54)
- b. Ul’a vas=te *(pa)? (Badiot variety)
 where go.2SG=CL PART
 ‘Where are you going now?’ (ibidem: 55)

In the Example (19a) *po/pa* functions as a modal particle expressing that the speaker and the addressee share a common ground regarding that question; on the other hand, in the Example (19b), however, it indicates that the proposition is a question. Indeed, *po/pa* is obligatory in questions in those varieties where it is used as noted in (19b). Based on the data displayed by the different varieties of Dolomitic Ladin, Hack (2014) proposes a grammaticalization path from modal particle towards general question marker.

To sum up, cross-linguistically some modal particles have been claimed to produce question particles by bleaching their semantic-pragmatic contribution. Nevertheless, some hypotheses do not contemplate a completed grammaticalization process as argued in the next subsection.

4.2 The use of modal particles to minimize the interrogative strength

As can be noted, the hypotheses developed by Bayer (2012) and Hack (2014) focus on the diachronical development of modal particles into question particles without taking into account how the interpretation of modal particles leaked and became more functional step by step. Now I will introduce a different perspective accounting for a similar use in Romanian and Mandarin Chinese.¹³ In both cases speakers

12. Kinuhata (2012) looks into the historical development of the question particle *ka* from Old Japanese to nowadays and concludes that it has evolved from subjective to objective meaning by losing its speaker-orientedness. In other words, it changed from a modal particle to a question particle, similarly to the cases described for German (Bayer 2012) and Dolomitic Ladin (Hack 2014).

13. Vanrell *et al.* (2014: 12) claim that the particle *-a* in Sardinian can also be used for such purposes, i.e. to minimize the strength of the question.

are claimed to use particles in cases where they are not expected to appear and they are examined in pragmatic terms, i.e. their use is linked to the effect of minimising the strength of the question they occur in. Let us briefly discuss both hypotheses.

On the one hand, Farkas (2017) examines the Romanian particle *oare* (eng. ‘really’); this particle is used in questions and adds an ‘I wonder’ effect marking conjectural questions, similar to Basque *ote*:

- (20) Oare Ioana e aici?
 PART Ioana is here
 ‘Is Joana here, I wonder?’ (Farkas 2017: 1)

Nevertheless, this is not the only use this particle shows, since Farkas (2017) notices *oare* can also be used in true information seeking questions, for instance:

- (21) Oare mai ai migrene?
 PART still AUX migraine
 ‘Do you still have migraines, I wonder?’ (Farkas 2017: 15)
- (22) Oare ai avea chef sa vii cu mine la piață?
 PART AUX joera SUBJ etorri COM lps ADL ART merkatu
 ‘Would you feel like coming with me to the market, I wonder?’ (ibidem)

Based on that, Farkas claims that *oare* functions in a way that it makes questions less intrusive. In other words, although the speaker expects to receive an answer, s/he does not want the addressee to feel obliged to; this reduction of the inquisitiveness of the question is caused by the use of the modal particle *oare*.

On the other hand, we have the work of Yang and Wiltschko (2016) on Mandarin Chinese. These authors present data regarding the use of the particle *ha* and associate its use in some contexts with an effect in politeness. Mandarin *ha* occurs at the end of questions and expresses that speakers expect addressees to confirm the proposition and add it to their common ground; for instance:

- (23) Nimen shi jiu dianzhong kai men de ha?
 zuek izan bederatzi orduetan zabaldu ate NOMLZ PART
 ‘You opened at nine o’clock, right?’

These authors propose that *ha* weakens the strength of an assertion and, therefore, it gives the addressee the feeling that s/he can accept or reject it. Having this in mind, they hypothesise that *ha* occurs in contexts where speakers feel a high degree of commitment to the proposition but this may be considered as threatening for addressees and, therefore, speakers may use this particle to minimise this high commitment and provide addressees a way out, if they do not feel in position to accept it.

In conclusion, these hypotheses argue that speakers based on the basic interpretation of some modal particles may have reinterpreted them and use them now as markers weakening the interrogative strength. After having presented these works dealing with variation concerning modal particles, now I will turn back to the Basque data and the analysis of the pragmatically non-canonical *ote*.

5. Pragmatic microvariation on Navarrese-Lapurdian

Returning to the topic of this paper, it has previously been stated that the modal particle *ote* mainly occurs in questions but only in those in which no answer is requested since neither the speaker, nor the addressee is supposed to know it (these contexts are also known as ‘can’t find the value’ questions following Obenauer (2004, 2005, 2006) or conjectural questions following Littell et al. (2010)); that means that *ote* is not felicitous in information seeking questions (see Example (14)). However, it is in Navarrese-Lapurdian varieties where speakers accept its use under certain circumstances discussed below. The following examples can be taken as a sample of such use gathered in literary works and interviews:

- (24) Nahi ote dun hik ere?
 want PART AUX you.ERG too
 ‘I wonder whether you would like it too.’ (Thikoipe 2007: 52)
- (25) Eskualde hauetakoa ote zira?
 region these.GEN.ABS PART be
 ‘I wonder whether you are from this area. (asked by a stranger in the train)’
 (Etxamendi 2011: 9)
- (26) Esku-ukaldi bat emaiten ahal ote dautazu?
 hand one give.IPFV can PART AUX
 ‘Could you give me a hand, please?’ (Landart 1983)
- (27) Jakiten ahal ote dut neskatxa hori nor den?
 know.IPFV can PART AUX girl that who be.C
 ‘I wonder if you could explain to me who that girl is.’ (Larzabal 1992)
- (28) Hortxeko oihan hortarik ekartzen ahal ote zinuke izai bat?
 there.GEN forest that.ABL bring.IPFV can PART AUX fir one
 ‘May you (please) carry a fir from that forest over there?’ (Minaberri 1983)
- (29) Zerbait gisaz laguntzen ahal ote gituzu izai baten
 something mode.INS help.IPFV can PART AUX fir one.GEN
 apaintzen?
 decorate.NOM
 ‘May you (please) help us decorating one fir somehow?’ (ibidem)

These questions cannot be considered conjectural questions, since they expect some kind of answer from the addressee: for instance, speakers want to know whether the addressee would like some food too (24), whether she is from that region (25) or whether the addressee may help bringing the Christmas tree or some decorations in (28) and (29). Nevertheless, they do not all look only for information, but also for confirmation or rejection of an action to carry on. If we look upon Escandell's (1987) work on questions, she distinguishes the following ones:

- a. Non-biased questions such as information seeking questions and directives.
- b. Biased questions, for instance, negative polar questions, rhetorical questions, hypothetical and exclamatives.

This classification may shed some light on the use of *ote* in those contexts. If we consider the questions listed above (4, 12, 24–29) again from the perspective of Escandell's (1987), they can be grouped as information seeking questions and directives. As Escandell (1999) describes them, these questions do not constitute distinct classes, since, after all, the petition of information implies also an action but in this case verbally; hence, information seeking and directive questions both form a single group. However, some of those questions can be considered as petitions or offerings according to the aim of the speaker. Escandell (1999) offers an interesting perspective on those questions based on how costly the acceptance of the action under question may be considered and whether it is beneficial for the speaker or the addressee. For instance:

- (30) a. Lagunduko didazu altzariak etxe berrira eramaten?
 help.FUT AUX furniture.ABS.PL house new.ADL carry.NOM
 'Will you help me carrying the furniture to the new house?'
 b. Zerbait edan nahi duzu?
 something drink want AUX
 'Would you like to have something to drink?'

The acceptance of (30a) supposes a cost for the addressee but a benefit for the speaker; on the other hand, the acceptance of (30b) presumes a cost for the speaker but a benefit for the addressee. Escandell adds that the presence of politeness markers is not unusual in such contexts, namely in requests which suppose a cost for the addressee in the sense of the works of Yang and Wiltschko (2016) and Farkas (2017). Considering the contribution of *ote*, it can be hypothesised that, if *ote* developed into a question particle, it would be first reinterpreted and used in contexts containing a petition to do something costly for the addressee, since the contribution of *ote* affects the illocutionary force of questions by reducing its strength and directiveness. This prediction is born out since *ote* in Navarrese-Lapurdian occurs

in questions requesting to perform an action or to accept a request which may be costly for the addressee, for instance Examples ((26), (28) and (29)).

Furthermore, this particle occurs in contexts which may be considered as inquisitive or intrusive such as Example (25) since it is the beginning of a conversation between two strangers on a train or Example (27) since the speaker knows that the information asked is private and, therefore, addressees may not intend to give a reply if they feel that their privacy is threatened.

Finally, *ote* can also appear in offerings as Examples (4) and (24). There is no cost for addressees or threat to their privacy in those contexts; in fact, speakers expressed during the interviews that, although it is possible to use it, the question sounded too polite, even deceptive. Nevertheless, this does not contradict the hypothesis put forward above; in fact, those contexts mostly occur with no particle and those few cases containing *ote* could be explained in terms that the speaker would not want to impose the addressee to accept the offer. Also, the sense of deceptiveness would be caused because the speaker would actually expect the offer to be rejected, since it would be costly for her/him.

This use of *ote* agrees with the one described for the Romanian particle *oare* and even for the Mandarin particle *ha*; both particles are used to present a request in a politer way. This effect is achieved because those particles modify the strength of the question by minimising its directness. As Blum-Kulka (1987) states, petitions may be interpreted as costly actions for addressees and it is usual to minimise their directiveness to make them less threatening and more appealing for those who will perform the requested action.

Moreover, Van Eys (1879), Larrasquet (1939) and Orpustan (2019) offer evidence in favour of such analysis in the descriptions of this particle regarding its use in Navarrese-Lapurdian Basque. On the one hand, Van Eys (1879) has the following entry in his dictionary:

othe: *Expletive word, signaling doubt, interrogation.* Joan ote da? Did s/he leave? Egin ote duzu? Did you do it? Ez ote duzu egin? Didn't you do it? Bai ote? Is it true? Ezta otherik. There's no doubt. *It seems to us that othe and othoi* (eng. 'please') *have a common origin, or more accurately othe is the syncope of othoi; othoi je vous prie, je vous demande please I beg you, I ask you. Hence, Joan othe da really means: Je vous demande, je vous prie, est-il parti? I ask you, I beg you, is s/he gone?*¹⁴

14. This is the original text in French:

othe: *Mot explétif, de doute, d'interrogation.* Joan ote da? Est-il parti? Egin ote duzu? Avez-vous fait? Ez ote duzu egin? N'avez-vous pas fait? Bai othe? Est-il vrai? Ezta otherik. Il n'y a pas de doute. *Il nous semble que othe, et othoi ont une origine commune, ou plutôt que othe est la syncope de othoi* othoy, je vous prie, je vous demande. *Ainsi: Joan othe da, signifie au fond: Je vous demande, je vous prie, est-il parti?* (Van Eys 1879)

of their illocutionary and epistemic functions, as stated by Wegener (2002: 382). Second, the apparent loss of the particle *ote* has brought gains too: its interpretation has been bleached, but this has caused that the particle can be used not only in conjectural and rhetorical questions, but also in information seeking questions; thus, its contexts have been generalised. This situation resembles the term *extension* or *context generalisation* in its pragmatic manifestation (Harris & Campbell apud Heine 2003: 579–580): a linguistic term is considered to have extended when it can be used in more contexts than it could originally.

A parallelism to this change can be found in the central dialect regarding the question particle *al* (Monforte 2020a); this particle has been claimed to be diachronically related to the modal particle *a(ha)l* (which may be translated to English as ‘surely’) and, as in the case of *ote*, it appear only in polarity questions. Nevertheless, its grammaticalization seems to be more developed since (a) it appears in a wider range of yes/no questions and (b) the use of *a(ha)l* is almost lost. Hence, this particle would have fulfilled the entire process in (34).

6. Conclusion

The Basque modal particle *ote* occurring in interrogative contexts turns questions into conjectural ones. This behaviour is found consistently in all dialects (Euskaltzaindia 1987, 2016). Nevertheless, in this work it has been proved that the particle *ote* has expanded its contextual domain in Navarrese-Lapurdián Basque and, therefore, it occurs also in information seeking questions, unlike the rest of dialects examined for this work (all of them except the Souletin one). By using *ote* speakers aim to minimise the feeling of inquiring and the strength of questions such as in petitions allowing the addressee to avoid replying it; this also causes the sense of politeness. Hence, I propose that speakers of that dialect have reinterpreted *ote* as a marker to reduce the expectancy and directiveness of questions, since the context where it generally occurs does not expect any answer, i.e. conjectural questions. This pattern has also been found in other languages such as Romanian (Farkas 2017) and Mandarin Chinese (Yang & Wiltschko 2016); the case of the Romanian particle *oare* deserves to be remarked since, apart from occurred in conjectural questions, it is also used to minimise the strength of the interrogative force. Finally, the grammaticalisation path experienced by the particle *ote* in Navarrese-Lapurdián Basque fits the second stage of the model proposed by Heine (2003: 590), i.e. $A > A, B > B$. In fact, the ungrammaticalised and grammaticalised *otes* coexist in this dialect. Thus, *ote* is currently an ambiguous particle in this dialect and only its context clarifies it. This has been caused because the desemantization of the general *ote* in that dialect

has increased the number of contexts where this particle can appear (extension or context generalisation in Heine's (2003: 579) terms) and, consequently, it has created concrete uses in those specific contexts.

Acknowledgements

I am grateful to Xabier Artiagoitia and two anonymous reviewers for their valuable comments, as well Maitena Duhalde-De Serra for the discussion of the data regarding this topic. I am also thankful to the audience of the workshop "Euskal Filologiaren adarrak" held in Vitoria-Gasteiz for their questions. All errors are mine.

Funding

This study has been made possible thanks to the research project PGC2018-100686-B-I00 from the Spanish Ministry of Science, Innovation and Universities.

Consulted corpora

- Altzaga, Toribio. 1994. *Neskazar; Amantxi; Aterako gera; Biozberak* [Auspoa liburutegia 223]. Oiartzun: Sendoa Argitaldaria.
- Barandiaran, José M. 1972. *Obras completas* (Eusko-folklore). Bilbao: La Gran Enciclopedia Vasca.
- Etxamendi, Eñaut. 2011. *Idazlanen bilduma / Eñaut Etxamendi*. Baiona: Maiatz.
- Landart, Daniel. 1983. Bai ala ez. In *Egan 1–6*. Donostia: Euskalerrriaren adiskideen elkarte.
- Larrasquet, Jean. 1939. *Le basque de la Basse-Soule orientale*. Paris: C. Klincksieck.
- Larzabal, Piarres. 1992. *Piarres Larzabal idazlanak III*. San Sebastian: Elkar.
- Minaberri, Marijane. 1983. *Haur Antzerki*. Bilbao: Bizkaiko Aurrezki Kutxa/ Euskaltzaindia.
- Thikoipe, Mixel. 2007. *Hotz bero: ixtorioak*. Baiona: Maiatz.
- Thikoipe, Mixel. 2009. *Jin bezala*. Baiona: Maiatz.

References

- Albizu, Pablo. 1991. *Sobre la existencia del movimiento largo de núcleos en euskera*. Ms, Instituto Universitario Ortega y Gasset, Madrid.
- Altube, Seber. 1930. *Erderismos*. Bilbao: Euskaltzaindia.
- Astarloa, Pablo P. 1883. *Discursos filosóficos sobre la lengua primitiva, ó Gramática y análisis razonada de la Euskara ó Bascuence*. Bilbao: Pedro Velasco.
- Azkue, Resurrección M. 1891. *Azkuero Resurrección Mariak eratu ta biurturiko euskal-izkindea*. Bilbao: Astuiko Josén Moldegintza.

- Bayer, Josef. 2012. From modal particle to interrogative marker: A study of German *denn*. In *Functional Heads. The Cartography of Syntactic Structures*, Laura Brugé (ed.), 13–28. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780199746736.003.0001>
- Bayer, Josef & Obenauer, Hans-Georg. 2011. Discourse particles, clause structure, and question types. *Linguistic Review* 28(4): 449–491. <https://doi.org/10.1515/tlir.2011.013>
- Blum-Kulka, Shoshana. 1987. Indirectness and politeness in requests: Same or different? *Journal of Pragmatics* 11(2): 131–146. [https://doi.org/10.1016/0378-2166\(87\)90192-5](https://doi.org/10.1016/0378-2166(87)90192-5)
- Burton, Strang & Matthewson, Lisa. 2015. Targeted construction storyboards in semantic fieldwork. In *Methodologies in Semantic Fieldwork*, M. Ryan Bochnak & Lisa Matthewson (eds), 135–156. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780190212339.003.0006>
- Camino, Iñaki. 2009. *Dialektologiatik euskalkietara tradizioan gaindi*. San Sebastian: Elkarlanean.
- Chaho, Augustin. 1836. *Études grammaticales sur la langue euskarienne*. Paris: Arthus Bertrand.
- de Rijk, Rudolf. 1969. Is Basque an SOV language? *FLV* 1(3): 319–352.
- de Rijk, Rudolf. 2008. *Standard Basque: A Progressive Grammar*. Cambridge MA: The MIT Press.
- Dor, Daniel. 2005. Toward a semantic account of that-deletion in English. *Linguistics* 43(2): 345–382. <https://doi.org/10.1515/ling.2005.43.2.345>
- Döring, Sophia. 2007. Zur Kontextverschiebung bei deutschen Diskurspartikeln. MA thesis, Humboldt-Universität zu Berlin.
- Elordieta, Arantzazu. 2001. Verb Movement and Constituent Permutation in Basque. PhD dissertation, Leiden University, LOT.
- Elordieta, Gorka. 1997. Morphosyntactic Feature Chains and Phonological Domains. PhD dissertation, University of Southern California.
- Escandell, María V. 1987. La interrogación en español: Semántica y pragmática. PhD dissertation, Universidad Complutense de Madrid.
- Escandell, María V. 1999. Los enunciados interrogativos. Aspectos semánticos y pragmáticos. In *Gramática descriptiva de la lengua española: Entre la oración y el discurso-morfología*, M. Ryan Bochnak & Lisa Matthewson, 3929–3992. Madrid: Espasa Calpe.
- Esnal, Pello. 2012. Hitz-ordena. Erabilera estrategikoa. *Euskera* 57(3): 471–494.
- Euskaltzaindia [The Academy of the Basque Language]. 1987. *Euskal gramatika. Lehen urratsak II*. Bilbao: Euskaltzaindia.
- Euskaltzaindia [The Academy of the Basque Language]. 2016. *Euskararen herri hizkeren atlasa, VII: Lexikoa, Izen morfologia, sintaxia*. Bilbao: Euskaltzaindia.
- Farkas, Donka F. 2017. Going beyond the prototypical: Special interrogatives and special declaratives. In *Stanford Colloquium*. Stanford CA: Stanford University.
- Garmendia, Joana. 2014. Ote: hiztunaren ziurtasunik eza. *Gogoia* 12–13: 7–26.
- Gèze, Louis. 1873. *Éléments de grammaire basque, dialecte souletin: Suivis d'un vocabulaire basque-français et français-basque*. Baiona: Veuve Lamaignère.
- Hack, Franziska M. 2014. The particle *Po* in the varieties of Dolomitic Ladin-grammaticalisation from a temporal adverb into an interrogative marker. *Studia Linguistica* 68(1): 49–76. <https://doi.org/10.1111/stul.12022>
- Harriet, Bernard M. 1741. *Gramatica escuaraz eta francesez, composatua frances hitzcunça ikhasi nahi dutenen faboretan*. Baiona: Fauvet alarguna eta J. Fauvet.
- Heine, Bernd. 2003. Grammaticalization. In *The Handbook of Historical Linguistics*, Briand D. Joseph & Richard, Janda D. (eds.), 575–601. Malden MA: Blackwell. <https://doi.org/10.1002/9780470756393.ch18>
- Hidalgo, Victor. 1999. Hitz ordenaren estatistikak euskaraz. *ASJU* 33(2): 393–452.

- Hualde, José I. & Ortiz de Urbina, Jon. 2003. *A Grammar of Basque*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110895285>
- Ithurry, Jean. 1895. *Grammaire basque: Dialecte labourdin*. Baiona: Lamaignère.
- Kinuhata, Tomohide. 2012. Historical development from subjective to objective meaning: Evidence from the Japanese question particle *ka*. *Journal of Pragmatics* 44(6–7): 798–814. <https://doi.org/10.1016/j.pragma.2012.03.004>
- Larramendi, Manuel. 1729. *El imposible [sic] vencido. Arte de la lengua Bascongada*. Salamanca: Antonio Joseph Villargordo.
- Lécluse, Fleury. 1826. *Grammaire basque*. Toulouse/Baiona: J.n-M.eu Douladoure/L. M. Cluzeau.
- Littell, Patrick, Matthewson, Lisa & Peterson, Tyler. 2010. On the semantics of conjectural questions. In *Evidence from Evidentials* [University of British Columbia Working Papers in Linguistics], Tyler Peterson & Uli Sauerland (eds), 89–104. Vancouver BC: University of British Columbia.
- Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70(4): 369–415. <https://doi.org/10.1086/429207>
- Monforte, Sergio. 2018. About the *al* particle in Basque. In *Novas perspectivas na lingüística aplicada*, Marta Díaz Ferrom Jorge Diz Ferreira, Ania Pérez Pérez & Ana Varela Suárez (eds), 19–31. Vigo: Axac.
- Monforte, Sergio. 2020a. The development of modal particles and its implication in the syntax: The case of Basque *ahal*. Ms submitted for publication.
- Monforte, Sergio. 2020b. Galderetako -a, al and ote partikulak euskaraz: Sintaxia, mikroaldakortasuna eta interpretazioa. PhD dissertation, UPV/EHU.
- Obenauer, Hans-Georg. 2004. Nonstandard *wh*-questions and alternative checkers in Pagotto. In *Syntax and Semantics of the Left Periphery, Interface Explorations*, Horst Lohnstein & Susanne Trissler (eds), 343–384. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110912111.343>
- Obenauer, Hans-Georg. 2005. La syntaxe des questions non standard: Les questions de surprise-désapprobation en bellunese. *Recherches Linguistiques de Vincennes* 33: 55–90. <https://doi.org/10.4000/rlv.1267>
- Obenauer, Hans-Georg. 2006. Special interrogatives-Left Periphery, *wh*-doubling, and (apparently) optional elements. In *Romance Languages and Linguistic Theory 2004* [Current Issues in Linguistic Theory 278], Jenny Doetjes & Paz González (eds), 247–273. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.278.12obe>
- Orpustan, Jean-Baptiste. 2019. *Basque et Français. Méthode abrégée de traduction Navarrolabourdin classique*. Baigorri: Izpegi.
- Pastor, Luis. 2019. Estrategias facilitadoras del procesamiento en lenguas OV-VO. Estudio comparativo de corpus. PhD dissertation, UPV/EHU.
- Van Eys, Willem J. 1879. *Grammaire comparée des dialectes basques*. Paris: Maisonneuve.
- Vanrell, Maria d. M., Ballone, Francesc, Schirru, Carlo & Prieto Vives, Pilar. 2014. Intonation and its interfaces in Sardinian polar questions. *Loquens* 1(2): e1–e16.
- Villasante, Luis. 1980. *Sintaxis de la oración simple*. Oñati: Editorial Franciscanas Aránzazu.
- Wegener, Heide. 2002. The evolution of the modal particle *denn*. In *New Reflections on Grammaticalization* [Typological Studies in Language 49], Ilse Wischer & Gabriele Diewald (eds), 379–394. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.49.24weg>
- Yang, Xiaodong & Martina Wiltschko. 2016. The confirmational marker *ha* in Northern Mandarin. *Journal of Pragmatics* 104: 67–82. <https://doi.org/10.1016/j.pragma.2016.09.004>
- Zuazo, Koldo. 2014. *Euskalkiak*. Donostia: Elkarlanean.

On divergent paths and functions of ‘background’-based discourse markers in Korean

Seongha Rhee and Hyun Jung Koo

Hankuk University of Foreign Studies / Sangmyung University

This paper analyzes two polyfunctional discourse markers (DMs) in Korean, *kulssey* and *kuntey*, which share lexical and grammatical sources, with the meaning of ‘at the time’ for *kulssey* and ‘at the place’ for *kuntey*. They denote the background of an action/event (‘in such circumstances’), which is their primary meaning when used as connective adverbs. Despite these source commonalities, when they developed into DMs, their functional differences are pronounced, only with a few shared functions. They show a certain degree of positional preferences by function, but at a more comprehensive level, left- and right-periphery do not show rigid functional dichotomy with respect to subjectification and intersubjectification, often hypothesized in the literature. Instead, the role of prosody is more crucial.

Keywords: discourse marker, pragmatic inference, divergent functions, functional network, Korean

1. Introduction

Korean has a large inventory of discourse markers (DMs) with diverse discourse-organizing and interpersonal functions. On the synchronic plane, discourse markers are typically polyfunctional (Fischer 2006 and works therein, Koo 2018), and there are multiple determinants of these functions (Rhee 2017, 2018; Koo & Rhee 2018, 2019). The function of a particular DM largely depends on the context in which it occurs and the prosodic features that accompany it (Hancil 2013 and works in Hancil & Hirst 2013). In other words, despite the multiplicity of functions of any given DM, the function of a situated DM is largely unambiguous because diverse interpretive cues are available in the context.

On the diachronic plane, it has been widely accepted in grammaticalization studies that the meaning of the source construction and pragmatic inferences are

among the most important determinants of the function of a DM (Rhee 2017; Rhee & Koo 2020; Rhee 2020, cf. Traugott 2014). In view of these synchronic and diachronic characteristics, this paper analyzes two polyfunctional DMs *kulssey* and *kuntey* that share lexical and grammatical sources and yet carry different functions in addition to certain shared functions. Both of them began their life as lexical items with a similar meaning indexically referring to ‘background’ (i.e., ‘in such circumstances, while it is so, etc.’). They both share certain lexical and grammatical sources in their construction, i.e., the verb *kuleha*- ‘be so’, an adnominalizer, and the postpositional particle *-ey* ‘at’ (see Section 2 for more detail).

The objectives of this paper are threefold. We will describe the grammaticalization paths of the two DMs *kulssey* and *kuntey* in Section 2; illustrate their semantic-pragmatic functions in Section 3; and discuss various factors contributing to their functional development in Section 4. Section 5 summarizes the findings and concludes the paper. The data for historical development are largely taken from the 21st Century Sejong Corpus, a 200-million word corpus, developed by the Korean Ministry of Culture and Tourism. This corpus encompasses data from the 15th to the early 20th century (1446–1912). The contemporary data are taken from the Drama-Cinema Corpus, a 24 million word corpus consisting of 7,424 drama and cinema scenarios, compiled by Min Li of Seoul National University. Scenarios are dated from 1992 through 2015.

2. The development of *kulssey* and *kuntey*

The DMs *kulssey* and *kuntey* originated from adverbial connectives which are polymorphic in formal makeup, sharing some key morphemes, as shown in (1):

- (1) a. *kulssey* (Rhee 2015; Song & Rhee 2017; cf. Ryu 1962)
kuleha ‘be so’ + *-l-* ‘adnominal (Prospective)’ + *sA* ‘time/occasion’ +
-Ay ‘at’
 b. *kuntey* (Jeong 2003)
kuleha ‘be so’ + *-n-* ‘adnominal (Anterior)’ + *tA* ‘place’ + *-ey* ‘at’

As shown in (1), the two forms have two distinct yet similar lexical nouns *sA* and *tA*, which as dependent nouns participated in the formation of a number of grammatical forms. According to Jeong (2003: 61), the nominal *sA* originally denoted ‘the sun’ from which the notion ‘temporal/spatial expanse or interval’ was derived. The nominal *tA* originally denoted ‘the place’ and the meaning still persists in Modern Korean.

When used as adverbial connectives, *kulssey* has the meaning of 'then, therefore, meanwhile' and *kuntey* 'then, but, however'. Considering their morphological construction shown in (1), their semantic developments can be characterized as in (2):

- (2) a. *kulssey*: 'at the time' (lit. 'at the time when it would be so')
 > 'when it is so' > 'then; therefore; meanwhile'
 b. *kuntey*: 'at the place' (lit. 'at the place where it is/was so')
 > 'where it is so; whereas' > 'then; but; however'

Even though the two forms are based on the lexical items that were productively used in Middle Korean, i.e. *sA* 'time' and *tA* 'place', the composite forms have differential temporal depth and have undergone much formal and functional change as shown in (3):

- (3) a. *kulssey* (Song & Rhee 2017):
kulehAlsAy, kulehAlssAy (15c.) 'at the time when it would be so'
 > *kulelsAy, kulelssAy* (15c.) 'when it is/was so; while; then; therefore'
 > *kulssey* (PDK) DM
 b. *kuntey* (cf. Chung 1996: 122–124 for *-ntAy*):
kulehAntAy (19c.) 'at the place where it was/is so'
 > *kulentey* (20c.) 'then; but; however'
 > *kuntey* (PDK) 'then; but; however; DM'

As shown in (3), *kulssey* has a long history but its contemporary usage is restricted to DM usage with diverse meanings (see 3.1), whereas *kuntey* has a much shorter history as a fully unverbated form and is used as a connective adverbial or as a DM with diverse meanings (see 3.2).

3. Multiple DM functions of *kulssey* and *kuntey*

3.1 Kulssey

After *kulssey* became a DM, it acquired various discourse functions, marking the speaker's diverse epistemic, emotional, and interactional stances, e.g. uncertainty, pause-filling, reluctance/hesitance, (feigned) surprise, discontent, protest, emphatic reassertion, disregard for emphasis, and agreement. We exemplify the usage of each function in turn indicating their preferred positions: left-periphery (LP), right-periphery (RP) or medial (MED).

3.1.3 Reluctance/Hesitance (“Well, I need time to reveal my stance”)

The reluctance-marking function is closely related to the pause-filler function. Whereas pause-filling involves the speaker’s intrapersonal need, reluctance/hesitance involves interpersonal, interactional, i.e., intersubjective, need. Reluctance may be with reference to acceptance or refusal, motivated by social needs of politeness (cf. ‘delay device,’ ‘face-threat mitigator’ Jucker 1993; ‘temporizing, slowing down effect’ Wierzbicka 1976; ‘polite hedge’ Lee 1993: 153–154). This function of marking reluctance and hesitance is among the primary functions of *kulssey* in contemporary Korean, which usually occurs at LP, exemplified in (6):

- (6) (B wonders why A wanted to see him and A suggests having lunch together)
 A: [I just wanted to have lunch together. I know a great place serving mudfish soup. Do you like mudfish soup?]
 B: ...*kulssey*... *mwe*... *chwuethang-un na-n pyellwu-y-a, sasil*
 DM DM mudfish.soup-TOP I-TOP not.great-be-END fact
 ‘[DM: Well, wait a minute]... that is... Mudfish soup, I’m not particularly fond of it, to say the truth.’ (1997 Drama *Kutay kuliko na* Episode #17)

In (6) B does not like mudfish soup which A intends to treat him to, but in order to save A’s face, B uses the reluctance/hesitance DM before he informs him that he, contrary to A’s expectation, does not like it.

3.1.4 (Feigned) surprise (“Guess what happened!; Why on earth”)

Kulssey can signal the surprising quality of upcoming information. This function is often exploited to dramatize a narrated story and carries the attention-attracting function. The DM usually occurs at MED, as shown in (7):

- (7) (A woman, speaking of her childhood memory of her mother’s chance encounter with a woman who married her ex-husband,)
 [The story doesn’t end there. At that moment the woman caught sight of us.]
kuleteni wuli emma-ka kulssey, kongsonhi... ku pwuin-hanthey
 then our mom-NOM DM politely that woman-to
cangmallwu 90.to cel-ul ha-nun-ke-y-ess-e-yo
 truly 90.degree bow-ACC do-ADN-NOMZ-be-PST-END-POL
 ‘Then, my mom, [DM: guess what happened!], bowed a deep bow (i.e. bending 90 degrees) politely to the woman – that’s what happened.’
 (1998 Drama *Kutay kuliko na* Episode #40)

In (7) the speaker is describing a surprising scene in which her mother showed exceptional politeness to a woman who married her own ex-husband. The speaker intends to draw the addressee’s attention to this rather surprising display of extreme politeness to a woman toward whom her mother would have harbored hostility and

hatred at heart for stealing her husband. The speaker in the course of the narration, may be genuinely re-experiencing the surprise she felt long ago, but, as is often the case in story-telling from memory about the past experience, the surprise may be ‘feigned’ for a strategic reason for a dramatic effect (see Koo 2008, Koo & Rhee 2013 for strategic language use leading to grammaticalization).

3.1.5 *Discontent (“I cannot understand; I’m not happy”)*

The DM *kulssey* can signal the speaker’s discontent about the situation. The discontent may be with respect to the preceding proposition or the addressee. Discontent thus marked is often associated with an undesirable turn of events or uncooperativeness of the addressee. The DM with this function usually occurs at RP, as illustrated in (8):

- (8) (Two aunts, A and B, are visiting C. C just came home after he was badly beaten up by a group of bully boys at a beach. A is speaking to B and C.)
 A: [Of all days to pick, my good-natured nephew has been beaten up today (when we are visiting). Oh, no! You were scared, weren’t you!]
cyay-ka wenakey mal-twu-epsi swunha-n ay-ntey
 that.child-NOM by.nature word-even-without gentle-ADN child-CONN
haphilimyen onul mac-kwu tuleo-ney kulssey!
 of.what.necessity today be.beaten-and come.in-EXCL DM
 ‘That child is quiet and gentle by nature but today, of all days to pick, he came home all beaten up! [DM: I cannot understand.]’
 (1997 Drama *Kutay kuliko na* Episode #3)

In (8) an aunt is expressing her anger and discontent, signaled by the DM *kulssey*, toward the situation that her nephew returned home all beaten up.

3.1.6 *Protest (“What are you doing?”)*

The DM *kulssey* is also used to signal protest. The protest-marking function is closely related to the discontent-marking function, to the extent that the two are often indistinguishable in terms of interpretation. The major difference, however, is that whereas discontent is mild and self-directed (thus often used in monologues), protest is aggressive and other-directed (thus in dialogues). The DM with this function usually occurs at RP, as exemplified in (9):

- (9) (A is infuriated to see his son badly beaten by someone and grabs a pickaxe saying he would go retaliate. A’s younger son, B, tries to restrain him.)
 B: [(while forcibly taking the pickaxe from him) Where do you mean you are going, Father? Do you mean you want to hurt yourself? You will break your bone. Give it to me.]

- A: [Are you thinking that I'm a feeble old man? I'm still strong and powerful enough. Give it back to me. (tries to take it back)]
- B: (struggling not to let it go) *apeci-n kaman kyeys-ey-yo, kulssey.*
 father-TOP quietly exist-END-POL DM
cey-ka ka-yo! cey-ka ka-yo!
 I.HUM-NOM go-POL I.HUM-NOM go-POL
 'Father, you remain still here [DM: Don't be silly!]. I will go. I will go.'
 (1997 Drama *Kkokci* Episode #14)

In (9) the speaker is alarmed by his father who is about to harm the one who hurt his son (the speaker's older brother), and is protesting to his father while pleading to stop. This DM, typically occurring at RP, is provocative and thus it is often used in confrontational contexts.

3.1.7 *Emphatic reassertion ("I'm telling you!; For sure!")*

The DM *kulssey* signals the speaker's emphatic reassertion of what has already been indicated. Since this involves repetition of what the speaker has already said or, if not said, what has been situationally obvious, the DM in this function has the overtone of irritation as well (thus resembling the discontent-function). The DM with this function usually occurs at RP, as shown in (10):

- (10) (Upon returning home, A feels with her cold hand the forehead of her daughter sleeping in bed and thinks that she has fever.)
- A: [Mother, Hyerim is running fever. Feel her. She's running fever.]
- B: [(after feeling Hyerim's forehead) No, she doesn't have fever.]
- A: [(trying to feel her forehead again) Yes, she does. I mean it.]
- B: (snatching A's cold hand away from Hyerim's forehead)
aikwu any-a kulssey
 INTJ be.NOT-END DM
 'Come on, she does not, [DM: I'm telling you!]
 (1999 Drama *Chengchwunuy tech* Episode #1)

In (10) the speaker is reasserting emphatically her previous claim that her granddaughter has no fever. She is partly irritated by her daughter touching her granddaughter with a cold hand, and tries to stop her doing that by assuring that what she said is true.

3.1.8 *Disregard for Emphasis ("No matter what!; Whatever!")*

The DM *kulssey* can also signal the speaker's belittling stance, i.e. disregard for other possibilities or circumstances. The disregarding stance is to emphasize the message in order to say that his or her message holds true. *Kulssey* used for this

function typically signals discontent or protest as well. The DM occurs usually at LP, as shown in (11), but sometimes it occurs at RP.

- (11) (A's daughter-in-law has kept a secret bankbook, the content of which her husband now forcibly tries to check. The struggle develops into a conflict and she begins to cry. A asks her son what kind of bankbook it is, but he says it's none of her business. Deeply offended and troubled, she is asking about it.)
 A: *a kulssey totaycheyka kuke-y mwusun thongcang-i-ntey*
 INTJ DM why.on.earth that-NOM what.kind bankbook-be-CSL
eymi-ka cele-nyakwu!
 daughter.in.law-NOM do.that-Q
 'Ah, [DM: whatever!] What kind of bankbook is it and why on earth is she acting like that?' (2002 Drama *Hwangkummacha* Episode #36)

In (11) the speaker is troubled by the situation, particularly her son's saying that the issue is none of her business. The speaker, in turn, disregards what he said, by means of the DM *kulssey*, effectively saying 'I don't care what you say; no matter what' and asks again about the bankbook.

3.1.9 Agreement ("You're right!")

The final function of the DM *kulssey* is that of marking agreement. As compared to other usages, this usage is relatively infrequent (Song & Rhee 2017). The DM with this function invariably occurs at LP, as exemplified in (12):

- (12) (A, speaking to her uncle B, about a young actress trying to seduce her father)
 A: *eyis! ku yeca kipwunnappu-e!*
 INTJ that woman be.unpleasant-END
 'Sh**! That woman! I don't like her.'
 B: *kulssey maliya! sakopangsik-i thulli-emwuk-ess-e!*
 DM DM thought.pattern-NOM be.wrong-PEJ-PST-END
 '[DM: Right], that's what I'm saying! Her way of thinking is despicably wicked.' (1994 Drama *Kimkaika* Episode #21)

In (12) the speaker B is agreeing with his niece's evaluation of the woman. The DM *kulssey* used in this function often accompanies *malya* 'it's what (I'm) saying' as in (12), and is often interchangeable with another agreement DM *kulekey* (Rhee 2015).

3.2 Kuntey

The DM *kuntey* has diverse functions as well, and is among the most frequent DMs in contemporary Korean (Jeon & Nam 2005; Choi 2007; Sohn 2016). We will look at the examples for each of the seven functions of *kuntey*.

3.2.1 Topic presentation ("So")

The DM *kuntey* is one of the most frequently-used topic initiators and attention-attractors in Present-day Korean. Unlike its less reduced counterpart, *kulentey* 'but, then', the DM *kuntey* rarely functions as an adversative or consecutive connector, a clear instance of 'divergence' (Hopper 1991). The DM has become reinterpreted as a signal of the speaker's desire to secure the addressee's attention for the verbal action that is to follow. The DM with this function invariably occurs at LP, as exemplified in (13):

- (13) (Two men are visiting A at her office; after an exchange of brief self-introduction,)
 A: (studies two visitors and emits a light sigh)
kuntey mwusun il-lo ce-l chac-usi-nun
 DM what.kind matter-with I.HUM-ACC look.for-HON-ADN
ke-ci-yo?
 NOMZ-END-POL
 '[DM: So,] for what business are you here to see me?'
 (2003 Drama *1%-uy etten kes*, Episode #1)

In (13) the speaker, after an exchange of brief formalities, signals her desire to learn about the visitors' purpose of the visit and proceed to the next phase of the discourse by opening the discourse with a question.

3.2.2 Topic shift ("By the way")

Kuntey is one of the most frequently-used topic-shifters as well. Even though the lexical meaning of *kuntey* is contrastive or adversative 'whereas, but, however', the DM *kuntey* does not prominently make a contrastive reference to the previous discourse; the contrast effect only resides in the cessation of, or departure from, the previous topic. In actual usage, the departure is only minimal since the newly introduced topic tends to be only branching out of the previous topic. For an obvious reason, the DM with this function occurs at LP, as shown in (14):

- (14) (A and B are having lunch on the SAT day after the morning test session. A laments and whines that she performed badly on the morning tests.)
 A: [All I could figure out was black parts were printed letters and yellow parts were the paper. My tuition has been wasted. Truly wasted.]
 B: [Why are you complaining already? The results are not out yet.]
 A: [You did well on the test, right? You, betrayer!] (then studies B's face)
kuntey ne cincca sinpangkwa ka-lkey-a?
 DM you really communication.department go-INTN-END?
 '[DM: By the way], are you really planning to go to a communication studies department?'
 (2009 Drama *Oyinkwutan* Episode #10)

In (14) Speaker A changes the topic from the performance on the SAT test to intended college majors. Since both involve college entrance, the topic shift is only minimal.

3.2.3 *Elaboration request* (“So?; Then what?”)

The elaboration request function is closest to its literal meaning for direct connection with the previous discourse. The speaker is prompting the previous speaker to elaborate a particular part of the utterance. Since this DM functions as a prompt it always occurs at LP, as illustrated in (15):

- (15) (A’s would-be in-laws are not happy with him and are not willing to let their daughter marry A soon. A and his friend B are talking.)

A: [It’s driving me crazy.]

B: [Why is that? They still haven’t consented to your marriage?]

A: *ani helak-un ha-sy-ess-e.*

no consent-TOP do-HON-PST-END

‘Yes. As long as consent is concerned, they did consent.’

B: *kuntey(?)*

DM

‘[DM: Then (what?)]’

A: [The problem is the time. They insist that we marry in the fall next year.]

(2003 Drama *1%-uy etten kes* Episode #18)

In (15) Speaker B requests A to elaborate on the reason for the worry because, as he thinks, if the would-be in-laws have already consented, there would be no conceivable reasons to worry. It is interesting to note that when the DM is used for this function it normally is not spoken with the rising intonation typical of a question. This usage resembles co-construction of an on-going utterance, i.e., the DM is uttered on behalf of the previous speaker, prompting with a connective ‘but’ (see 4.3 for more). For this reason the DM used for this function typically stands alone, inviting the interlocutor to continue and elaborate on what has just been said.

3.2.4 *Preface to dispreferred information* (“Well.. but”)

The DM *kuntey* may signal that dispreferred information is to follow. This function bears the semantic trace of the lexical meaning of the adversative connective, i.e., ‘however’. *Kuntey* used for this function occurs at LP, and tends to be followed by a pause which indicates hesitance, as exemplified in (16):

- (16) (A, drinking hot water B offered, notices that B has been knitting something; B hides it.)

B: [You shouldn’t see this now. You should see it later... I might give it to you as a present when I’m done.]

A: [Is it (going to be) mine?]

B: [(nods)] *kuntey-yo* (A looks at her puzzled) *eccemyen mos*
 DM-POL possibly cannot

tuli-l-ci-to *moll-a-yo*
 give-FUT-NOMZ-also not.know-END-POL

‘[DM: Well.. but], as it may turn out, I may not be able to give it to you.’
 (2008 Drama *Kamwunuy yengkwang* Episode #24)

In (16) Speaker B, after admitting that what she is working on is going to be a gift for A, says that it may not be realized (because she has failed to make a number of things for him already), contrary to A’s expectation. She uses the DM *kuntey* as a preface to this dispreferred information. The DM tends to be followed by a pause, which indicates the speaker’s reluctance to disclose the information of undesirable nature.

3.2.5 (*Feigned*) Surprise (“Guess what happened!; Why on earth”)

Still another function of the DM *kuntey* is to signal that the speaker is surprised or puzzled. It typically occupies the MED position, but not in the slot for traditional clausal connective. This usage resembles that of the DM *kulssey* in MED position signaling (feigned) surprise for a dramatic effect (see 3.1.4 above). The use of *kuntey* for this function is exemplified in (17):

(17) (A woman, noticing her husband is already in bed,)

i yangpan-i kuntey onul way ilehkey ilccik ca-ø.
 this man-NOM DM today why like.this early sleep-END

‘This man, [DM: why on earth], why has he gone to bed so early today?’

(2009 Drama *Solyakkwukcip atultul* Episode #21)

In (17) the speaker is signaling her surprise at her husband who has gone to bed very early, which is exceptional considering his habit of going to bed very late. Since her (monologal) speech coincides with her realization of the situation, the surprise may be spontaneous and genuine. However, as we noted in the discussion of *kulssey* in 3.1.4, the surprise, especially in narratives, may be only rhetorical in order to create such a feeling on the part of the interlocutor.

3.2.6 *Discontent* (“I don’t understand!”)

Kuntey may also signal the speaker’s discontent. The discontent-marking function is closely related to the surprise-marking function, described above, for their shared semantic feature of counter-expectation. The DM with this function usually occurs at RP, as exemplified in (18):

(18) (A, a social studies teacher, is talking with another teacher, B, and a student, C, about the school festival.)

B: [What is this program “School Bags in the Good Old Days” all about?]

C: [That’s a joint program with the student government and the film club. It is to exhibit the teachers’ photos of their high-school days. As the title of the festival is Dongkwang Festival with Teachers, that should fit quite well.]

A: *haksaynghoy hayngsa-ey way tto yenghwapan nom-tul-i*
 student.government event-at why again film.club guy-PL-NOM
natay-nun-ke-nya kuntey?

act.up-ADN-NOMZ-Q DM

‘Why on earth are those film club guys acting up in a student government event, [DM: I don’t understand]?’ (1999 Drama *Hakkyo-2*, Episode #21)

In (18) the teacher, A, asks a rhetorical question which obviously signals his displeasure about the film-club students playing an active role in the school festival. The utterance-final DM *kuntey* signals that the speaker is not happy about the situation and thus the question is only rhetorical (“they shouldn’t”), not seeking information on the reason.

3.2.7 Protest (“What are you doing?”)

The final function of *kuntey* is to signal the speaker’s protesting attitude. The protest-marking function is closely related to the discontent-marking function, and the two functions are often indistinguishable. As noted with respect to the similar function of *kulssey* (see 3.1.6 above), discontent is largely mild and self-directed in nature, whereas protest is aggressive and other-directed. The DM *kuntey* with the protest function usually occurs at RP, as exemplified in (19):

(19) (A is agonized over the imminent death of his secret love. His brother, B, unaware of the situation and only noticing that he is acting strangely, playfully charges him.)

B: [Big Brother, you look strange these days.]

A: [What do you mean I look strange, you dude?]

B: [Your mind seems to be somewhere else... Hey, Big Brother, isn’t it that you have hidden a secretly born baby somewhere?]

A: (raising voice) *mwe-y-a! imma! i casik-i malha-nun*
 what-be-END this.guy this fella-NOM speak-ADN
ke-y... kuntey...

NOMZ-NOM DM

‘What? Look at this guy! The way this wretched one says is... [DM: what are you doing?]...’ (2009 Drama *Solyakkwukcip atultul* Episode #13)

In (19) the speaker protests his brother's playful, absurd charge of having an illicit baby without others' knowledge, by means of the DM *kuntēy* at the utterance-final RP position. *Kuntēy* used in this function tends to be followed by a pause or ellipsis as in (19), which indicates that the speaker is emotionally overwhelmed and is unable to complete the sentence.

4. Functional determinants

4.1 Functions compared

The diversity of the DM functions of *kulssey* and *kuntēy* is partly due to their source meaning, use context, and linguistic realization patterns, e.g., semantics, prosody, syntagmatic position, turn position, etc. Since the meaning of the main lexeme *kuleha*- 'be so' contains indefiniteness and vagueness as shown by the English translation including 'so', it is applicable to a wide context. It has been noted that the two DMs emerged from very similar source constructions, i.e. those both involving *kuleha*- 'be so', an adnominal, a dependent noun, and a locative particle (see (1) above), and as a natural consequence, it is expected that they have shared functions. Despite the similarity of their sources, however, the two DMs show more differences than commonalities. The functions can be summarized as in Table 1, in which most functions at MED and RP positions are similar, as marked by shaded cell (see, however, 4.2 for more discussion), whereas all LP functions are dissimilar.

It is also noteworthy that LP is the most favored position for both DMs, nine out of eighteen. This may be due to the fact that LP is where the original connectives occurred before their development into DMs, i.e. LP preference is the retention of the original position.

Table 1. Functions and positions of *kulssey* and *kuntēy*

DM	LP	MED	RP
<i>kulssey</i>	Uncertainty	(Feigned) Surprise	Discontent
	Reluctance/Hesitance	Pause-Filling	Protest
	Agreement		Emphatic Reassertion
	Pause-Filling		Disregard for Emphasis
	Disregard for Emphasis		
<i>kuntēy</i>	Topic Presentation	(Feigned) Surprise	Discontent
	Topic Shift		Protest
	Elaboration Request		
	Preface to Dispreferred Info		

4.2 Syntagmatic independence

Brinton (1996), Heine (2013: 1209) and Heine et al. (forthcoming) list the properties that constitute the defining characteristics of DMs. At the syntactic level, DMs are not a syntactic constituent of the sentence in which they occur. Since DMs arise from a clause-internal position, usually at a syntactically licensed position, the source forms necessarily need to acquire syntactic freedom in order to develop into DMs.

In Korean, the connective function of many complex connectives is usually carried by bound morphemes involving the particle *-ey* ‘at’ (or its historical variant *-Ay*) that occurs at the final position of the construction (note that Korean is a head-final language). As we noted in (1), the DMs under the present analysis *kulssey* and *kuntey* still carry the trace of the connective. These DMs gained positional freedom by becoming morphosyntactically independent connective words (i.e. adverbs) when the bound connective *-ey/-Ay* hosted the dependent nouns, *sA* ‘time’ and *tA* ‘place’, which in turn are modified by the verb *kuleha-* ‘be so’ with the help of an adnominalizer (i.e. prospective *-l* or anterior *-n*).

The complex source forms *kuleha-l-sA-ey* and *kuleha-n-tA-ey*, underwent morphosyntactic ‘compacting’ (Hopper & Traugott 2003[1993]: 140) or ‘coalescence’ (Haspelmath 2011) and increased conceptual cohesion, and became fully ‘univerbated’ (Lehmann 2015[1982]: 160) as *kulssey* and *kuntey*. This syntagmatic process seems to have been facilitated by the anaphoric meaning of *kuleha-* ‘be so’ (note ‘so’ which is ultimately from the speaker-distal deixis *ku* ‘that’, which is semantically vague). After *kulssey* and *kuntey* fully developed as non-bound, lexical connectives, freely occurring at non-argument positions, their further development into DMs is an easy process, as is well-known with the English DMs, *and*, *so*, *but*, *though*, etc.

4.3 Semantic persistence and periphery

Since it is widely accepted that grammatical functions are closely related to, or, in a strong position, determined by the meaning of the source lexemes and constructions (cf. ‘persistence’, Hopper 1991; ‘the source determination hypothesis’, Bybee et al. 1994), the relationship between the source meanings and the DM functions of *kulssey* and *kuntey* merits discussion. This issue is closely related to positionality, i.e. periphery (as)symmetry

An analysis of the relationship between the source meanings and the DM functions bears special significance when we compare *kulssey* and *kuntey* because they share much in source characteristics as indicated above. In other words, considering the similarity of their sources, we would expect them to have developed into DMs of similar functions. However, as noted in 4.1, the overlap in their functions

is only partial. Immediately relevant factors to functional differences are variable degrees of semantic change and positionality (other factors are also discussed in 4.4 and 4.5). Semantic change largely involves bleaching, i.e. divergence from, and loss of the previous meanings. Positionality involves (as)symmetry of functions in peripheral positions, a point of much discussion in recent literature (Beeching & Detges 2014 and works therein, and numerous others including the articles in *Journal of Historical Pragmatics* 17(2), 2016). With respect to positionality, *kulssey* and *kuntsey* occur either at LP or RP, but their functions seem to favor certain positions, as indicated in Table 1 above.

The functions of the DMs have variable degrees of semantic-functional bleaching of the major lexemes and grammatical markers. For contemporary Korean speakers, the diverse DM functions currently available can be grouped according to their perceived differential degrees of bleaching and their preferred positions, as shown in Table 2.²

Table 2. Degrees of perceived connection between the Source Semantics and DM Functions

DM	Strongly perceivable connection	Weakly perceivable connection	No perceivable connection
<i>kulssey</i>	[none]	Uncertainty (L) Reluctance/ Hesitance (L) Pause-Filling (L/M)	Agreement (L) Disregard for Emphasis (L/R) (Feigned) Surprise (M) Discontent (R) Protest (R) Emphatic Reassertion (R)
<i>kuntsey</i>	Topic Presentation (L) Topic Shift (L) Elaboration Request (L) Preface to Dispreferred Info (L)	[none]	(Feigned) Surprise (M) Discontent (R) Protest (R)

The grouping of the functions in Table 2 shows a number of interesting facts. First, the source meaning is more strongly visible in *kuntsey* than in *kulssey*. This may be due to the fact that the development of *kuntsey* is a recent phenomenon (from a 19th c. adverbial), whereas *kulssey* has a much longer history (from a 15th c. adverbial). The longer historical depth of *kulssey* seems to be correlated with the more extensive semantic bleaching. As a matter of fact, in Modern Korean there are few people who are aware that *kulssey* involves the old defective noun *sA* ‘time’.

2. The classification was initially prepared by the authors and was confirmed by a group of native speakers, who were asked if the source meanings of the DMs, i.e. *kulentey* ‘but’, *kulelssay* ‘meanwhile’ as connectives, are strongly visible, weakly visible, or entirely opaque in each of the DM functions. The classification received unanimous agreement.

Another aspect is that the source connection is most opaque in RP functions for both DMs. This may be due to the fact that RP functions have undergone the greatest extent of change. RP is not where the original connectives could occur before their development into DMs. Thus, the distance between the meanings can be coextensive with the extent of grammaticalization.

Also noteworthy is that, in their most frequent position, i.e. at LP, the functions of the two DMs are vastly different, whereas their functions at MED and RP show more similarity (see also Table 1). This can be explained by the tendency hypothesized to be universal (though contested) that RP functions are largely interactional, intersubjective, and interpersonal, thus convergence is relatively expected. However, even the RP functions have different levels of intensity in meaning, i.e. *kulssey* signals stronger emotion than *kuntsey*, a state of affairs possibly reflecting the fact that *kulssey* has a longer history (and has undergone more extensive grammaticalization) than *kuntsey*.

We have seen that the role of source semantics in grammaticalization is not deterministic. The developmental paths may be modified by other factors, such as the degree of grammaticalization which may affect the extent of semantic bleaching and differential positionality which may affect the degree of (inter)subjectification.

In this context, there is another issue that warrants discussion, which concerns subjectivity and intersubjectivity with respect to the positionality. Certain functions such as Uncertainty, Disregard for Emphasis, Pause-filling, Surprise, Topic Shift, etc. are speaker-internally motivated functions. These functions tend to be at LP and MED. Some other functions such as Discontent, Protest, Emphatic Reassertion, Feigned Surprise, Elaboration Request, Agreement, etc. are more interactional and intersubjective by nature. These functions tend to occur at RP, but are not always so. This state of affairs suggests that there is a general tendency of correlation between subjective functions and LP on the one hand and intersubjective functions and RP on the other, but the correlation is not exclusive but only a tendency.

4.4 Prosody

The DMs *kulssey* and *kuntsey* are realized with different characteristic prosodies. Even though prosodic qualities could not be quantified, an impressionistic generalization is that *kulssey* has a higher level of sensitivity to prosodic variation with respect to functions, whereas *kuntsey* is prosodically more stable across functions. Prosodic differences are also variable with sentence-types. The functions of the two DMs according to their prosodic characteristics with respect to duration and speed are as summarized in (20):

(20) Duration and Speed

- a. Short & Fast: Agreement, Elaboration Request, Disregard for Emphasis, Emphatic Reassertion, Topic Presentation, Topic Shift, Surprise, Discontent, Protest
- b. Long & Slow: Uncertainty, Pause-filling, Reluctance/Hesitance, Preface to Dispreferred Information

The characteristics listed in (20) show that the functions associated with the determinate attitude tend to be short, whereas those with the indeterminate attitude tend to be long. This is a general tendency in language use involving affective attitudes, not restricted to DMs.

The prosodic characteristics with respect to intonation as associated with the DM functions are as shown in (21):

(21) Intonation

- a. Falling: Agreement, Elaboration Request
- b. Rising: Disregard for Emphasis, Emphatic Reassertion, Topic Presentation, Topic Shift
- c. Falling-Rising: Surprise, Discontent, Protest
- d. Level-Elongated: Uncertainty, Pause-filling, Reluctance/Hesitance, Preface to Dispreferred Information

The observed characteristics listed in (21) can be interpreted in the following way. First, the rising (and falling-rising) intonation is more closely related to the functions that solicit the addressee's attention. This seems to be a general tendency in language use since rising intonation increases perceptual saliency on the part of the addressee. Second, the falling intonation is associated with compliance and acceptance. This also seems to be a general tendency since falling intonation signals propositional and attitudinal alignment, which is conceptually and interactionally unmarked. Third, the function of Elaboration Request, which demands the addressee's attention to the reason or cause of the described situation or event, occurs with a falling intonation, unlike a regular question which carries a rising intonation. As briefly alluded to in 3.2 with Example (15), this usage resembles co-construction of an utterance, thus the DM is simply inserted on behalf of the current speaker rather than taking up a full speaker-turn. Fourth, the level intonation is associated with the functions of marking indeterminacy. This also seems to be a general tendency of language use, i.e. the level intonation is usually combined with elongation often suggesting reluctance for cognitive reasons (such as word search) or social, interactional reasons (such as hesitation for politeness).

The prosodic characteristics with respect to pause, either preceding or following the DM, are as follows:

(22) Pause

- a. With a pause: Uncertainty, Pause-filling, Reluctance/Hesitation, Preface to Dispreferred Information, Topic Initiation, Topic Shift
- b. Without a pause: Agreement, Disregard for Emphasis, Emphatic Reassertion, (Feigned) Surprise, Elaboration Request, Discontent, Protest

The observed characteristics listed in (22) show that the functions associated with the determinate attitude tend to occur without a pause, whereas those with the indeterminate attitude tend to occur with a pause. The characteristics of pause are largely parallel with those of duration, i.e. the functions associated with the presence of a pause overlap with those associated with the long and slow prosody, and those associated with the absence of a pause largely overlap with those associated with short and fast prosody. The described situation is a general tendency in language use involving affective attitudes, not restricted to DM prosody, i.e. the language involving determinate attitude tends to occur without a pause and realized with short and fast prosody.

A related issue is whether there is a correlation between prosody and the position of the DMs. The distribution of the prosodic features by the DM positions is shown in Table 3.

Table 3. Prosody and positionality of the DMs

Prosodic category	Prosodic feature	LP (9 functions)	MED (3 functions)	RP (total 6 functions)
Duration & Speed	Short & Fast	5 (55.6%)	2 (66.7%)	6 (100%)
	Long & Slow	4 (44.4%)	1 (33.3%)	0
Intonation	Rising	3 (33.3%)	0	2 (33.3%)
	Falling	2 (22.2%)	0	0
	Falling-Rising	0	2 (66.7%)	4 (66.7%)
	Level-Elongated	4 (44.4%)	1 (33.3%)	0
Pause	With Pause	6 (66.7%)	1 (33.3%)	0
	Without Pause	3 (33.3%)	2 (66.7%)	6 (100%)

Table 3 shows that the two DMs occurring at RP are invariably short and fast and without accompanying a pause, and further they tend to be of a falling-rising intonation. At the MED position, the DMs tend to be short and fast, with a falling-rising intonation and without a pause. At LP, they tend to occur with a pause (typical of ‘comma intonation’, Hansen 1997: 156), but there is no other notably strong tendency. The strong tendency of the DMs occurring at RP is largely due to the interactional, emotive nature of the functions at the position as alluded to in the foregoing exposition.

4.5 Pragmatic inference and functional network

A large body of literature shows that pragmatic inferences are actively involved in grammaticalization (cf. 'context-induced reinterpretation' Heine et al. 1991; Rhee 2015; 'invited inference' Traugott & König 1991; Hopper & Traugott 2003[1993]; Traugott & Dasher 2002, among others). Different inferences from an identical form may lead to the emergence of divergent, seemingly contradictory, functions. This is well illustrated in Rhee (2015), in which different inferential paths in the development of the three major functions of *kulssey*, i.e. Hesitance, Agreement, and Disregard for Emphasis, are presented, as in (23), modified from Rhee (2015: 23–24):³

- (23) a. Hesitance
 A: [Can you lend me some money?]
 B: *kulssey* ('Well, I don't know.')
- Inferential extension: lit. 'at it being so' > 'while that is the case' > 'while I acknowledge what you say' > 'I'm not sure, while I acknowledge what you say' > 'I need more time to answer.'
- b. Agreement
 A: [Mr. Kim is tardy again today.]
 B: *kulssey* ('You're right.')
- Inferential extension: lit. 'at it being so' > 'because that is so' > 'because he is tardy again today' > 'We are saying this because he is tardy again' > 'You're right!/You can say that again!'
- c. Disregard for Emphasis
 A: [Please let me go out to play.]
 B: *kulssey antway* ('Everything notwithstanding, no!')
- Inferential extension: lit. 'At it being so, you can't' > 'While that is so, you can't.' > 'While I know what you want, you can't.' > 'Everything notwithstanding, you can't.'

Even though the function of Disregard for Emphasis is illustrated with an explicit negation *antway* 'no' in (23c), such function does not require such negative language form as shown in 3.1 with Example (11). The inferences involved in the development of *kulssey* and *kuntey* can be diagrammatically presented as in Figures 1 and 2, respectively, along the three broad conceptual domains of ACKNOWLEDGMENT, CONCESSION, and UNEXPECTEDNESS.

3. Rhee (2015) interprets the source construction of *kulssey* as 'at it being so' rather than the more literal 'at the time that it would be so' or the more concise and natural 'while it is so'.

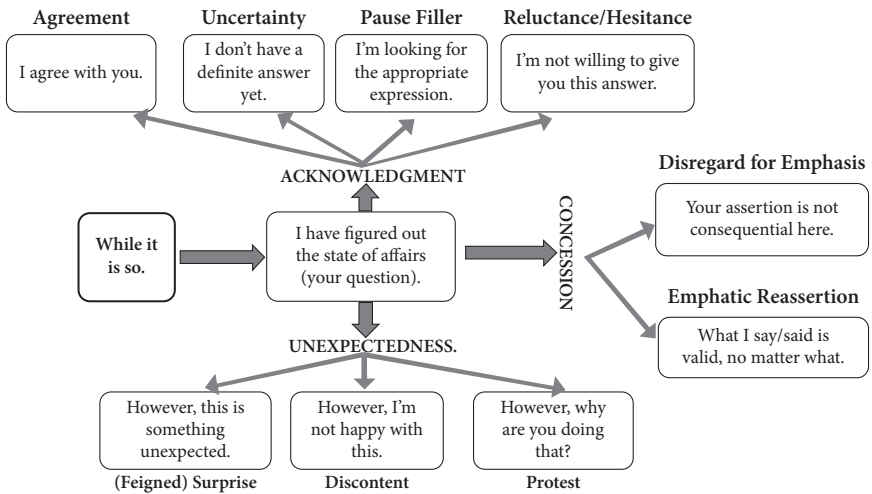


Figure 1. Inferences in the development of *kulssey*

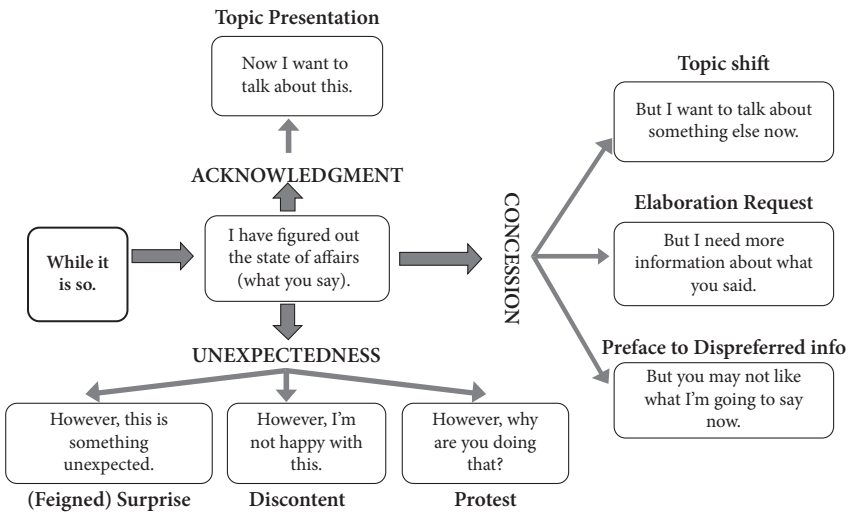


Figure 2. Inferences in the development of *kuntey*

As we have seen in the previous discussion, the development of the DMs *kulssey* and *kuntey* have proceeded along a series of morpho-syntactic and semantic changes, beginning from the dependent nouns *sA* 'time' and *tA* 'place' in Middle Korean. The functional development across the levels of grammar can be diagrammatically presented for *kulssey* and *kuntey* as in Figure 3. The diagram shows the historical and conceptual progression of grammaticalization across lexical (dependent noun),

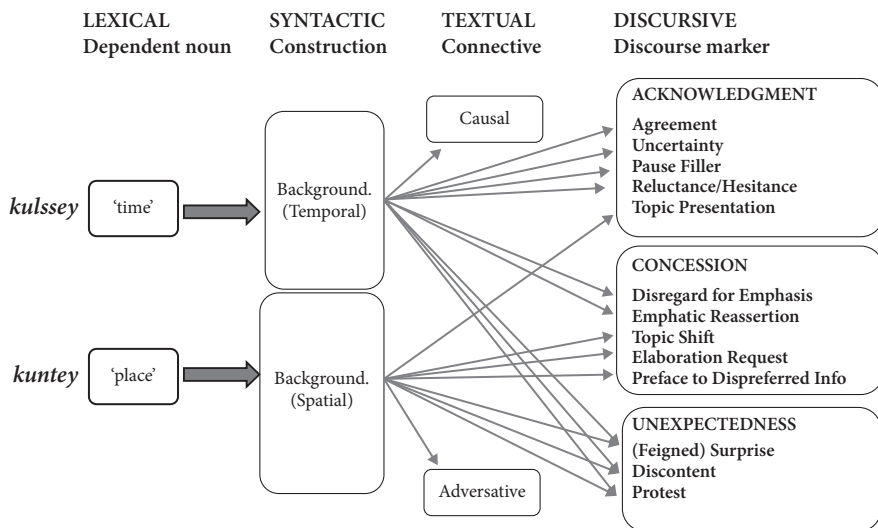


Figure 3. Developmental paths of the DMs *kulssey* and *kuntesy*

syntactic (phrasal construction), textual (adverbial connective), and discursive (discourse marker) levels, a pattern consonant with most reported instances of grammaticalization across languages.

5. Summary and conclusion

Korean DMs *kulssey* and *kuntesy* developed from similar source constructions. *Kulssey* developed into a marker of Uncertainty, Pause-filling, Reluctance/Hesitance, Agreement, Disregard for Emphasis, Emphatic Reassertion, Surprise, Discontent, and Protest. *Kuntesy* developed into a marker of Topic Presentation, Topic Shift, Elaboration Request, Preface to Dispreferred Information, Surprise, Discontent, and Protest. The functional similarities are evident in the affect-related functions mostly at the RP, which emerged most recently. We have seen that LP and RP do not show strict functional dichotomy with respect to subjectification and intersubjectification, but there are evident positional preferences by the functions. We further argued that the functional extension has been enabled by pragmatic inferences from the sources and conceptual relatedness among the notions that emerged in the course of grammaticalization.

Acknowledgements

An earlier version of this paper was presented at the International Conference on Current Trends in Linguistics, Université de Rouen, Mont-Saint-Aignan, France, March 28–29, 2019. We are thankful for the comments made by the audience. Also special thanks go to Shaun Manning for kindly reading an earlier version of the paper.

Funding

This research was supported by the Hankuk University of Foreign Studies research fund, the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A5A2A01035042) for the first, corresponding author, and by the Sangmyung University research fund for the second author.

Abbreviations

ACC	accusative	INTJ	interjection
ADN	adnominal	L(P)	left-periphery
ATT	attemptive	M(ED)	medial
CONN	connective	NOM	nominative
CSL	causal	NOMZ	nominalizer
DM	discourse marker	PEJ	pejorative
END	sentence-ender	PL	plural
EXCL	exclamative	POL	polite
FUT	future	PST	past
HON	honorific	Q	interrogative
HUM	humiliative	R(P)	right-periphery
INTN	intentional	TOP	topic.

References

- Beeching, Kate & Detges, Ulrich (eds). 2014. *Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Leiden: Brill. <https://doi.org/10.1163/9789004274822>
- Brinton, Laurel J. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110907582>
- Bybee, Joan L., Perkins, Revere D. & Pagliuca, William. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago IL: University of Chicago Press.
- Choi, Boyun Jane. 2007. A Corpus-based Discourse Analysis of Korean Discourse Markers: An Analysis of Spoken and Written Use. PhD dissertation, University of California, Los Angeles.

- Chung, Jaeyoung. 1996. *A Study on the Grammaticalization of dA Syntagma in Middle Korean*. Seoul: Kwukehakhoy.
- Fischer, Kerstin. 2006. Toward an understanding of the spectrum of approaches to discourse particles: Introduction to the volume. In *Approaches to Discourse Particles*, Kerstin Fischer (ed.), 1–20. Amsterdam: Elsevier.
- Hancil, Sylvie. 2013. Introduction. In Hancil & Hirst (eds), 1–31.
- Hancil, Sylvie & Hirst, Daniel (eds). 2013. *Prosody and Iconicity* [Iconicity in Language and Literature 13]. Amsterdam: John Benjamins. <https://doi.org/10.1075/ill.13>
- Hansen, Maj-Britt Mosegaard. 1997. *Alors and donc* in spoken French: A reanalysis. *Journal of Pragmatics* 28(2): 153–187. [https://doi.org/10.1016/S0378-2166\(96\)00086-0](https://doi.org/10.1016/S0378-2166(96)00086-0)
- Haspelmath, Martin. 2011. The gradual coalescence into 'words' in grammaticalization. In *The Oxford Handbook of Grammaticalization*, Heiko Narrog & Bernd Heine (eds), 342–355. Oxford: OUP.
- Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6): 1205–1247. <https://doi.org/10.1515/ling-2013-0048>
- Heine, Bernd, Claudi, Ulrike & Hünemeyer, Friederike. 1991. *Grammaticalization: A Conceptual Framework*. Chicago IL: The University of Chicago Press.
- Heine, Bernd, Kaltenböck, Gunther, Kuteva, Tania & Long, Haiping. Forthcoming. On the rise of discourse markers.
- Hopper, Paul J. 1991. On some principles of grammaticalization. In *Approaches to Grammaticalization*, Vol. 1 [Typological Studies in Language 19], Elizabeth Closs Traugott & Bernd Heine (eds), 17–35. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.19.1.04hop>
- Hopper, Paul J. & Traugott, Elizabeth Closs. 2003 [1993]. *Grammaticalization*, 2nd edn. Cambridge: CUP. <https://doi.org/10.1017/CBO9781139165525>
- Jeon, Young-ok & Nam, Killim. 2005. Kwuewa mwuneuy cepsok phyohyen pikyo yenkwu: "kulentey, -nuntey"-lul cwungsimulo (A comparative study of conjunctions in spoken and written Korean: Focusing on *kulentey*, *-nuntey*). *Korean Language Research* 17: 169–194.
- Jeong, Howan. 2003. Hankwukeuy Paltalkwa Uyconmyengsa (The development of Korean and dependent nouns). Seoul: Ihoymwunhawa.
- Jucker, Andreas H. 1993. The discourse marker *well*: A relevance-theoretical account. *Journal of Pragmatics* 19(5): 435–452. [https://doi.org/10.1016/0378-2166\(93\)90004-9](https://doi.org/10.1016/0378-2166(93)90004-9)
- Koo, Hyun Jung. 2008. Grammaticalization of negation markers in Korean. *Discourse and Cognition* 15(3): 1–27. <https://doi.org/10.15718/discog.2008.15.3.1>
- Koo, Hyun Jung. 2018. Hankwuke tamhwaphyoci yenkwuuy tonghyangkwa cenmang (Research trends and prospect of Korean discourse markers). In *Hankwuke Yenkwuuy Saylowun Hulum* (New Research Trends of Korean), Yong-Kwon Shin (ed.), 221–255. Seoul: Pakijong Publishing.
- Koo, Hyun Jung & Rhee, Seongha. 2013. "I will do it... but I'm asking you to do it": On the emergence of polite imperative from promissive. *Procedia-Social and Behavioral Sciences* 97: 487–494. <https://doi.org/10.1016/j.sbspro.2013.10.263>
- Koo, Hyun Jung & Rhee, Seongha. 2018. On the emergence of polyfunctionality of discourse markers: The case of *kulay* 'it is so' in Korean. Paper presented at the 36th AESLA Conference, University of Cádiz, Spain, 19–21 April.
- Koo, Hyun Jung & Rhee, Seongha. 2019. From self-talk to grammar: The emergence of multiple paradigms from self-quoted questions in Korean. *Lingue e Linguaggi* 31: 255–278. <https://doi.org/10.1285/i22390359v31p255>

- Lee, Won-pyo. 1993. Discourse analysis of three 'volitional interjections,' 'yey,' 'kulssey,' and 'ani' in Korean. *Inmwunkwahak* (Yonsei University) 69–70: 139–185.
- Lehmann, Christian. 2015 [1982]. *Thoughts on Grammaticalization*. Berlin: Language Science Press. https://doi.org/10.26530/OAPEN_603353
- Nam, Ki-Shim & Ko, Yong-Kun. 1985. *Phyocwun Kwukemwunpeplon* [Standard Korean Grammar]. Seoul: Tower Press.
- Rhee, Seongha. 2015. On the emergence of Korean markers of agreement. *Journal of Pragmatics* 83: 10–26. <https://doi.org/10.1016/j.pragma.2015.04.005>
- Rhee, Seongha. 2017. Periphery and prosody as determinants of discourse marker functions: A case in Korean. Invited lecture at the Meiji International Symposium 2017: New Directions in Pragmatic Research: Synchronic and Diachronic Perspectives, Meiji University, Nakano, Japan, 20 March.
- Rhee, Seongha. 2018. On multiple determinants of discourse marker functions: Peripheral asymmetry revisited. Paper presented at the 20th International Congress of Linguists, Cape Town, South Africa, 2–6 July.
- Rhee, Seongha. 2020. Pseudo-hortative and the development of the discourse marker *eti poca* 'well, let's see' in Korean. *Journal of Historical Pragmatics* 21(1): 53–82. <https://doi.org/10.1075/jhp.00036.rhe>
- Rhee, Seongha & Koo, Hyun Jung. 2020. From quotation to surprise: The case in Korean. *Journal of Pragmatics* 155: 83–100. <https://doi.org/10.1016/j.pragma.2019.10.006>
- Ryu, Chang-Don. 1962. Hesahwa kokwu [From content word to function word]. *Inmwunkwahak* (Yonsei University) 7: 1–25.
- Sohn, Sung-Ock S. 2016. Development of the discourse marker *kulentey* ('but, by the way') in Korean: A diachronic and synchronic perspective. *Journal of Historical Pragmatics* 17(2): 231–254. <https://doi.org/10.1075/jhp.17.2.045oh>
- Song, Inseong & Rhee, Seongha. 2017. On the role of prosody in the development of discourse markers: The case of *kulssey*. Paper presented at the 2017 Winter Conference of Korean Association of Linguistic Science, Pusan National University, Busan, 14 February.
- Traugott, Elizabeth Closs. 2014. On the function of the epistemic adverbs surely and no doubt at the left and right peripheries of the Clause. In *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*, Kate Beeching & Ulrich Detges (eds), 72–91. Leiden: Brill.
- Traugott, Elizabeth Closs & König, Ekkehard. 1991. The semantics-pragmatics of grammaticalization revisited. In *Approaches to Grammaticalization*, Vol. 1 [Typological Studies in Language 19], Elizabeth C. Traugott & Bernd Heine (eds), 189–218. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.19.1.10c1o>
- Traugott, Elizabeth Closs & Dasher, Richard B. 2002. *Regularity in Semantic Change*. Cambridge: CUP.
- Wierzbicka, Anna. 1976. Particles and linguistic relativity. *International Review of Slavic Linguistics* 1(2–3): 327–367.

Reanalysis and the emergence of adverbial connectors in the history of Japanese

Reijirou Shibasaki

Meiji University

This study deals with issues related to the development of stand-alone adverbial connectors such as *totan(-ni)*, *shunkan(-ni)*, *hyoosi(-ni)*, and *yasaki(-ni)* in sentence-initial position in the history of Japanese. These adverbial connectors, which can be witnessed from the early twentieth century, developed from the erstwhile head of adnominal clauses through a series of steps, i.e. reanalysis. The adverbial connectors under investigation here serve to anticipate more to come in the following discourse, labeled *projectors* by Hopper and Thompson (2008). In a nutshell, the emergence of adverbial connectors involves an increase in structural scope over a stretch of discourse.

Keywords: reanalysis, adverbial connectors, projector, grammaticalization

1. Introduction

This study deals with issues related to the development of stand-alone adverbial connectors such as *totan(-ni)*, *shunkan(-ni)*, *hyoosi(-ni)*, and *yasaki(-ni)* in sentence-initial position in the history of Japanese. These expressions all mean ‘(at the) moment’, albeit written in different Chinese characters with subtle shades of meaning. As shown in the following examples, these adverbial connectors of nominal origin have (a) a clause-combining use in addition to (b) an adverbial use in contemporary Japanese; other uses are addressed in later sections on an as-needed basis. The elements in focus are underlined for clarity; corpora and dictionaries used for this study are summarized in Section 3.

- (1) *totan* (途端 [‘road’ + ‘edge’ = (lit.) ‘(on the) edge of the road’])
- a. Head of adnominal clause use
- Doa o aketa totan, neko ga tobi-konde kita.*
 door ACC open.ADN moment cat NOM leap-lean.into.and come.PST
 ‘Just as (I) opened the door, the cat leapt in.’ (2015 Group Jammassy: 370)

- b. Sentence-initial adverbial use
Ocha o irete modotte kita. Totan, kanojo wa
 (green)tea ACC pour.and return.and came moment she TOP
angurito kuti o aketa.
 agape mouth ACC open.PST
 ‘(She) made tea and came back there (with the tea). Just then, her jaw
 dropped in amazement.’ (2005 *Meikyuu no Kioku*: BCCWJ)
- (2) *shunkan* (瞬間 [‘blink’ + ‘interval’ = (lit) ‘in a twinkling’])
- a. Head of adnominal clause use
Tati-agatta shunkan ni ie ga guratto ookiku yureta
 stood-rise.ADN moment PT house NOM violently widely shake.PST
 ‘The moment I stood up, the house shook violently from side to side.’
 (2015 Group Jammassy: 155)
- b. Sentence-initial adverbial use
Amon no midori no me ni wa, ai de wa naku, awaremi
 name GEN green GEN eye PT TOP love COP TOP not compassion
ga atta. Shunkan, Kyoko wa zotto-sita.
 NOM was moment name TOP shiveringly-do.PST
 ‘(Kyoko found that) in Amon’s green eyes, there were “poor you” looks but
 no love. (At that) moment, Kyoko felt a shiver.’
 (1992 *Kazenotenpyoorin*: BCCWJ)
- (3) *hyoosi* (拍子 [‘beat’ + suffix? = ‘in the act of beating/doing’])
- a. Head of adnominal clause use
Sikashi, sakki tumazuita hyoosi ni, baggu o yuka e
 but a.moment.ago tumble.ADN moment PT bag ACC floor PT
oite kite simatta.
 put.and come.and put.away.PST
 ‘The moment (I) stumbled just now, (I accidentally) left (behind my) bag
 on the floor.’ (2003 *Mikeneko Hoomuzu*: BCCWJ)
- b. Sentence-initial adverbial use
Usui natumono no sukaato ga hirogaru. Hyoosi ni, Karen
 thin summer.wear GEN skirt NOM stretch.out moment PT name
no ryoohiza ni kuroi ana no yoona kizu ga aru no o
 POSS both.knees PT black hole GEN like scar NOM be NML ACC
Yuko wa mita.
 Yuko TOP see.PST
 ‘(Karen’s) thin summer skirt got hiked. At that moment, Yuko got to see
 scars on her knees, which looked like black holes.’
 (2004 *Nepentesu*: BCCWJ)

(4) *yasaki* (矢先 ['arrow' + 'tip' = (lit.)'arrowhead'])

a. Head of adnominal clause use

“*Nantoka sinoide ikeru ka to zonzite ori-masita*
 somehow manage go PT COMP think.HUM be.HUM-POL.ADN
yasaki, koo nari-masite...”
 moment this.way become-POL.and

‘The moment (I) thought that (I would) make it through the day, this happened...’
 (1988 *Oyosi no Chawan*: BCCWJ)

b. Sentence-initial adverbial use

Sorezore ga “ima o ikiru” koto o jikkoo ni utusi-hajimeru.
 each NOM now ACC live NML ACC carry.out PT shift-begin
Sono yasaki ni higeki ga okotta.

DEM.MED moment PT tragedy NOM happen.PST

‘Each (member) put into practice “Live in the present.” At that moment, (such a) tragedy happened.’
 (2004 *Eiga de Eisi Nyuumon*: BCCWJ)

There are some similarities and differences in the usage of these expressions. Despite the combination of different Chinese characters shown in (1) through (4), they come to share one core meaning, ‘moment’, which seems to be metaphorically (or metonymically) derived and conventionalized over time; however, their morphosyntactic behaviors are not exactly uniform. In (1) and (2), for example, *totan* and *shunkan* are found to be used sentence-initially as bare nominal forms serving as adverbial connectors. On the other hand, *hyoosi* needs the particle *ni* ‘at, in’ as an adverbial connector, as in (3),¹ while *yasaki* needs the medial demonstrative *sono* ‘that’ or the proximal demonstrative *kono* ‘this’ as well as the particle *ni* to be an adverbial phrase, at least in my database. Of course, there may soon be innovative bare nominal uses of *hyoosi* and *yasaki* on the Internet. However, the varying pace of change in the development of these adverbial connectors deserves a closer look, in order to see whether they emerged historically in a similar manner or whether some of them developed under the influence of the other more frequently occurring expressions (see Section 4). Another point that should be mentioned is that Group Jammassy (2015), the handbook of Japanese grammar patterns for teachers and learners used nation-wide, supplies definitions for and gives detailed explanations of *totan* and *shunkan*, while making no mention of *hyoosi* and *yasaki*. Phrases or words that occur frequently appear to be selected for efficiency in practice (Section 2).

1. In the BCCWJ (*Shoonagon*) corpus, this is the only example of sentence-initial *hyoosi ni* ‘at (the) moment’ without any demonstrative. *Sono hyoosi ni* ‘at that moment’ is a little more frequent albeit moderate in numbers (17 tokens).

In addition to the role of frequency or preference in language pedagogy, the morphological characteristic of the verb that precedes and modifies these temporal nouns is worthy of attention: they all end with the *ta*-form in (1) through (4) in contemporary Japanese.² Historically, *ta* derived from the adnominal form *taru* of the perfect auxiliary *tari* usually used in written/literary language. The inflectional ending *ru* of *taru* began to be eroded in the Kamakura period (1185–1333), with conclusive and adnominal forms being leveled. As a result, the remaining *ta* played a predominant role as both conclusive and adnominal forms from the Edo period (1603–1867) onwards, with the meaning of past tense or perfect (e.g. Yamada 1936: 62, 68; Yuzawa 1955: 54; Noguchi 1994: 236–237; Nomura 2013: 26, 71–72; Kurashima 2019: 112–113; also see Aoki 2016 for a useful study). These are summarized in Table 1.

Table 1. The inflectional paradigm of *ta* ‘past or perfect’ in the Edo period (Yuzawa 1955: 364)¹

Mizen (irrealis)	Ren’yoo (adverbial)	Shuushi (conclusive)	Rentai (adnominal)	Izen (realis)
<i>tara</i>	<i>tari</i>	<i>ta (tari)</i>	<i>ta, taru</i>	<i>tare</i>

¹ The parentheses have been added by the present author: (*tari*), because *tari* might have survived in written styles such as *gikobun* (i.e. a (pseudo-)classical style) and *kanbun kundoku* (i.e. a Japanese reading of a Chinese passage) but not in spoken genres. Note in passing that the merger of adnominal and conclusive forms *in general* is often described as taking place from the eleventh through fourteenth centuries. I am grateful to one reviewer for these comments.

Noguchi (1994: 237) states that the conclusive form *tari* became obsolete in the first half of the Meiji period (1868–1912); a similar observation is provided in Nomura (2013: 72). Furthermore, Noguchi (1994: 237) points out that the adnominal *taru* is not integrated in the inflectional paradigm of *ta*, as mentioned in Yuzawa (1954: 417), which implies that *ta* appears to have been the preferred form for both conclusive and adnominal inflections, at least in spoken language, in the Edo period.

These observations all suggest that (1a), for example, may have been interpreted in the following two ways (5).

- (5) a. [clause 1 ... V-ta] *totan*, [clause 2 ...]
 ‘Just as (I) opened the door, the cat leapt in.’
 b. [sentence 1 ... V-ta]. *Totan*, [sentence 2 ...]
 ‘(I) opened the door. (At that very) moment, the cat leapt in.’
 (Shibasaki in press)

2. According to Tanaka (2010: 242), *V-ru-totan(-ni)* is also attested, albeit infrequently.

That is, how language users interpret the sequential relation between the form of the verb and *totan*, in either speaking or writing, depends on the immediate context.

The point of relevance here is that, as in (5b), the *ta*-form of the auxiliary verb does not necessarily need any nominal head to modify to complete the utterance, as long as language users interpret it as a grammatical sequence, i.e. an example of the rise of adverbial connectors through reanalysis (see Shibasaki 2018b, 2019 and in press for related issues; cf. Koyanagi 2018: 115–120). As discussed in Section 5, since syntactic boundaries can also be blurred and reinterpreted in other languages (Burridge & Bergs 2017: 148–151), such a reanalysis of (5a) as (5b) is not unusual, especially in spoken discourse. Reanalysis is defined here as “a mechanism which changes the underlying structure of a syntactic pattern and which does not involve any immediate or intrinsic modification of its surface manifestation” (Harris & Campbell 1995: 61). It is worth noting that punctuation marks, especially the period (.), had not been fully established even around the turn of the twentieth century in Japanese (Sugimoto 1979: 139–171), which may have enabled or facilitated the creation of new usages in this specific period or presumably in earlier periods.

For the following descriptive purposes, I will set forth one working hypothesis in (6) and test it through the discussion below.

(6) Working hypothesis

Adverbial connectors of temporal nominal origin such as *totan* and *shunkan*, as in (5b), all derived from the head of adnominal or noun-modifying constructions, as in (5a). In the course of this development, one (or some) frequently used construction(s) created through reanalysis played a leading role in facilitating the growth of related (but less frequent) constructions in a cluster.

This study is structured as follows. In Section 2, I will introduce the preceding studies directly related to this present study, with the addition of a follow-up survey, while in Section 3, I will introduce corpora and dictionaries used for this study. In Section 4, I will provide an overview of the historical development of the four expressions and analyze the survey results to verify the validity of the hypothesis in (6) from a language-internal perspective. Section 5 is devoted to the discussion of how reanalysis works to offer a possible explanation of innovative constructions in reference to similar phenomena in other languages; Section 6 is the concluding remarks. In addition, some socio-cultural aspects of this phenomenon, i.e. language contact, are briefly noted in the Appendix.

2. Preceding studies with corroborating information

In the history of Japanese, there are many other temporal nouns that can serve as the head of temporal adverbial clauses, for example, *tabi(-ni)* ‘(occasion >) every time’, *tuide(-ni)* ‘(occasion >) when/while’, *hazumi(-ni)* ‘(bounce >) when’, *saichuu(-ni)* ‘in the midst (of)’, *sanaka(-ni)* ‘in the midst (of)’, *ori(-ni)* ‘(occasion >) when/while’, and *sai(-ni)* ‘(time >) when’ (e.g. Muraki 2012: 327–340).³ Muraki (2012) makes a detailed study of voice, aspect, mood, tense, affirmative/negative, and types of adnominal forms in these adverbial clauses. Tanaka (2010: 237–277) delves into the semantic similarities and differences in temporal adverbial clauses including *totan(-ni)* and *shunkan(-ni)*. Despite their thorough studies in morphosyntax and semantics, the uses of free-standing *totan(-ni)* and *shunkan(-ni)* in discourse are not addressed.⁴

Group Jammassy (2015), which was cited in Section 1, is a handbook that illustrates a variety of useful expressions especially for Japanese learners, while Nitta (2002) enumerates a wide range of adverbial expressions under the rubric of semantic categories. However, they only briefly touch on *totan* and *shunkan*, while paying no attention to *hyoosi* and *yasaki*. Hida and Asada (2018) is a well-organized reference guide with a large array of examples with clear explanations. Under the heading of *totan*, one can see examples accompanied by phraseological comments and synonyms such as *tatimachi* ‘in a moment’ and *totuzen* ‘suddenly’. Nevertheless, none of the others, i.e. *shunkan*, *hyoosi*, and *yasaki*, are mentioned in the book. Of course, it is not that the selection process for each adverb is flawed in these books but that the basis for selection varies from one research study to another.

On the other hand, Saegusa (2015) takes a fine-toothed approach to the analysis of a cluster of related expressions, albeit from a synchronic perspective. Saegusa (2015: 204–210) focuses on the sentence-initial uses of *ikioi* ‘(force >) continuously’, *jissai* ‘(truth >) actually’, *jijitu* ‘(fact >) in fact’, and *kekka* ‘(result >) as a result’ in

3. The number of morae seems to have a bearing on the creation of stand-alone adverbial expressions. For example, expressions of three or four morae, such as *totan* (3 morae) and *shunkan* (4 morae), are found to serve as independent adverbs, albeit not frequently, while those of two morae, such as *tabi*, *ori*, and *sai*, usually cannot be used as independent adverbs, even with the addition of the particle *-ni* (see Note 5 below too). This phenomenon is the subject of part of my future studies (but see Figure 1 in Section 4.2).

4. Another set of sentence-initial adverbs, i.e. *reason* adverbs such as *yuwe ni* ‘hence’ can be attested in the Heian Period, which presumably arose under the influence of Chinese; I am grateful to one anonymous reviewer for this comment.

relation to their nominal-oriented uses.⁵ Consider the expressions in Table 2, to which I add *totan*, *shunkan*, *hyoosi*, and *yasaki* for the sake of discussion. If given usages are found in both her and my databases, the corresponding sections are check-marked (✓). Nom-pred=nominal predicate; topical=topicalization.

Table 2. Expressions of nominal origin: Their forms and functions
(based on Saegusa 2015: 210)¹

Form	-no	-wa	-da	-ni	-∅
Example	(genitive)	(topical)	(nom-pred)	(adverbial)	(adverb)
<i>jissai</i> 'in fact'	✓	✓	✓	✓	✓
<i>jijitu</i> 'in fact'	✓	✓	✓		✓
<i>ikioi</i> 'continuously'	✓	✓	✓		✓
<i>kekka</i> 'as a result' ²	✓	✓	✓		✓
<i>totan</i> 'at the moment'	✓			✓	✓
<i>shunkan</i> 'at the moment'	✓			✓ ³	✓
<i>hyoosi</i> 'at the moment'	(✓) ⁴			✓	
<i>yasaki</i> 'at the moment'				(✓) ⁵	

¹ One reviewer suggests that the tests for nounhood in Table 2 should include demonstratives such as *kono* 'this (proximal)', *sono* 'that (medial)', and *ano* 'that (distal)' as well as other case particles; I am grateful for the reviewer's suggestion. Such usage of demonstratives is useful as addressed in Figure 1 (Section 4.2). Also see Note 11.

² See Takahashi & Higashiizumi (2014) for both variant forms and functions of *kekka* 'as a result'.

³ In my database, *shunkan* is often found with the medial demonstrative as in *sono shunkani ni* 'at that moment'.

⁴ The archaic diction *hyoosi ga* (GEN) *warusa* ['poorness'] 'out of time/tune' is witnessed in 1477 (Kitahara 2006b, s.v. *hyoosi*). However, the genitive usage seems to be obsolete in contemporary Japanese.

⁵ Forms with both proximal and medial demonstratives *kono/sono yasaki ni* 'at this/that moment' can be attested; *kakaru* 'such (as this)' can also be used as a prenominal modifier, although all of these modifiers are very small in number.

Saegusa's approach to the discourse-pragmatic patterns of the first four expressions tells us that they retain their erstwhile nominal properties because they appear in genitive forms (-*no*) and because they can be topicalized as nominals with the addition of the topic marker *wa*. Overall, these four expressions are versatile in morphosyntax, expanding the functions of the older nominal properties to the newer adverbial ones (see Shibasaki 2018b for the diachrony of *jijitu*). On the other

5. In addition to these four expressions, Saegusa (2015: 210) includes *jitu* 'substance, truth' in her analysis and points out that *jitu* cannot grammatically act as a bare adverb but needs the particle *-ni* to perform as an independent adverbial phrase (typically as a verbal adverb not as a sentence-initial adverbial connector as treated in this study). For this reason, I will not discuss *jitu* in this study. But see Note 3 for a relevant comment.

hand, the latter four expressions are functionally restricted to genitive and adverbial uses, although the genitive use appears to be very infrequent in comparison to the adverbial use. For example, a quick survey of each use in the *BCCWJ* corpus provides telling evidence, as in Table 3. The numbers represent the raw frequencies, while the percentages in parentheses stand for their relative frequencies.

Table 3. A set of temporal nouns on the noun-adverb continuum (based on *BCCWJ*)

Form	-no (genitive)	-ni (adverbial)	Total
Example			
<i>totan</i> 'at the moment'	17 (1.2%)	1410 (98.8%)	1427 (100%)
<i>shunkan</i> 'at the moment'	237 (17.8%)	1091 (82.2%)	1328 (100%)
<i>hyoosi</i> 'at the moment' ¹	4 (1.4%)	281 (98.6%)	285 (100%)
<i>yasaki</i> 'at the moment'	58 (36.9%)	99 (63.1%)	157 (100%)
Total	316 (9.9%)	2881 (90.1%)	3197 (100%)

¹ All the examples of *hyoosi* meaning 'beat, rhythm' have been sifted out.

In contrast to the first set of four (*jissai*, *jijitu*, *ikioi*, *kekka*) in Table 2, the second set of four of nominal origin repeated in Table 3 is skewed toward adverbial use across the board. Importantly, each expression shows a marked difference in the frequency of use in this corpus: *totan* and *shunkan* are used much more frequently than *hyoosi* and *yasaki*, which serves as justification for why many preceding works pay attention only to *totan* (and *shunkan* at best). Another finding from this additional research is that the low-frequency item *yasaki* has a relatively higher tendency to retain its nominal property, which might suggest that high-frequency items such as *totan* have advanced in their degree of grammaticalization.⁶

Based on these preceding works and supplemental remarks, I will probe the historical development of these four temporal adverbs of nominal origin in the

6. Concerning the degree of grammaticalization, the three-way writing system of Japanese should be mentioned: *kanji*, *hiragana*, and *katakana*. In general, content words are written in *kanji*, while function words are written in *hiragana*; loan words (especially from Western languages) and onomatopoeic words are written in *katakana* (Iwasaki 2002: 11). For example, about 65% of *totan* (927 out of 1427 tokens) are written in *hiragana*, while about 91% of *yasaki* (143 out of 157 tokens) and 100% of *hyoosi* (285 out of 285 tokens) are written in *kanji*. In this respect, *shunkan* is quite exceptional, because despite its higher frequency, it has more than a 99% chance of being written in *kanji* (1323 out of 1328). One possible reason for this is that *shunkan* may have some connection with Buddhism; however, the earliest example of *shunkan* in Kitahara (2006a, vol. 16: 1458) is attested in *Seimikaisou* (1837–1847), i.e. the translation of the Dutch version of *Elements of Experimental Chemistry* (1799) by the British chemist William Henry (1775–1836). The translator was Yooan Udagawa (1798–1846), a famous Dutch scholar of the time. This issue goes beyond the realm of this study.

sections that follow. Unlike the aforementioned preceding studies, the emphasis/focus of this current study is the investigation of a set of synonymous expressions together from a diachronic perspective, to follow the processes of change based on genuine data, and to test the hypothesis in (6) that constructional clustering is a local phenomenon, with one (or some) high-frequency construction(s) accelerating the expansion of related low-frequency constructions in a network but not in isolation (see Note 13; see Hopper and Traugott 2003: xv).

3. Corpora and dictionaries

In this study, I conduct a qualitative analysis of some triggers for the rise of adverbial connectors, instead of the fully-fledged quantitative approach taken in my previous studies (e.g. Shibasaki 2018b, 2019, in press), because this study is to be a synopsis of how this particular phenomenon is derived through reanalysis. It is, however, advantageous to make good use of the widely used dictionaries in (7) and corpora in (8) and (9) to give a broad overview of the development of each expression at the ‘quantitative turn of the tide’ (Janda 2013).

(7) Dictionaries

- a. *Edogo no Jiten* (= Maeda 1979)
- b. *Edogo Jiten* (= Okubo & Kinoshita 1991)
- c. *Nihon Kokugo Daiijiten* (= Kitahara 2006a)
- d. *Seisenban Nihon Kokugo Daiijiten* (= Kitahara 2006b)

(8) Main corpora⁷

- a. *Taiyo Corpus* (*Taiyo*, 1895, 1901, 1909, 1917, 1925) himawari ver. 1.5
- b. *Balanced Corpus of Contemporary Written Japanese* (*BCCWJ*, 1971–2008) the online search engine *Shonagon*
- c. *The Corpus of Historical Japanese* (*CHJ*, from Nara to Taisho periods) the online search engine *Chuunagon*

(9) Supplemental corpora

- a. *Senjichuu no Hanasikotoba* (1937–1945) (= Endo et al. 2004)
- b. *Sharebon* (18th century) (= Ichimura 2015)
- c. *Ninjoobon* (19th century) (= Fujimoto & Takada 2015)

7. The *Taiyo Corpus* consists of the following numbers of words: 1.3 million words (1885), 1.2 million words (1901), 1.0 million words (1909), 0.97 million words (1917), and 0.87 million words (1925), according to Tanaka (2012). *BCCWJ* includes approximately 105 million words consisting of 11 genres, such as books, magazines, newspapers, blogs, etc.

Kitahara (2006a) can be considered the most thorough dictionary of Japanese, by which one can confirm the approximate dates for the earliest uses of each respective expression; Kitahara (2006b) includes some different examples from those cited in Kitahara (2006a). Maeda (1979) and Okubo and Kinoshita (1991) are two well-known dictionaries for the Edo period. I will utilize the two sets of corpora (8) and (9) to conduct additional research on their development and distributional patterns up to the present. (9a) is a corpus of radio drama scripts broadcast from 1937 to 1945. The two historical corpora, *Sharebon* (gay-quarter novelettes) in (9b) and *Ninjoobon* (love stories) in (9c), may reflect spoken language in the eighteenth and nineteenth centuries, respectively.

4. The survey results and discussion

As explained in the previous section, I will first use Kitahara (2006a,b) as a reference for the possible earliest records for both nominal and adverbial uses of *totan*, *shunkan*, *hyoosi*, and *yasaki*, and then conduct an additional survey of each respective form by using *Taiyo Corpus*, *CHJ* and *BCCWJ*. While other relatively frequent forms such as *sono yasaki ni* (DEM.MED + moment + PT) ‘at that moment’ in (4b) and *hyoosi ni* (Ø + moment + PT) ‘at (that) moment’ in (3b) are included in the following discussion, the rare forms are not addressed due to space limitations.

4.1 Boundary shifts and grammatical renewal through reanalysis

The following pairs show the earliest examples in the materials or corpora of the (a) head of adnominal clause use and (b) sentence-initial adverbial use for each expression, as found in the albeit limited datasets in (7) through (9).⁸

8. One reviewer was surprised to see the gap between the first use of the adnominal-head *hyoosi* (1447) and that of the sentence-initial *hyoosi* (1917), while stating “the adnominal/finite [i.e. conclusive; by the present author] merger is apparently irrelevant to the development”; I express my appreciation for his/her comment. However, as I explained in Section 1, rules for punctuation (especially period (.)) had not been conventionalized even around the late nineteenth century: language change is not just morphosyntactic but closely related to the development of punctuation conventions, if there, as in the case of Japanese. Therefore, the ‘bridging context’ discussed in Section 4.2 may not necessarily be as short as in the cases of (10) and (11) but may sometimes be much longer, as in (12) and (13), than usually thought. See Notes 12 and 15 for related discussions.

(10) *totan*a. Head of adnominal clause use⁹

... *imaimasii to haratati-magire ni tuki-tobasu totan ni hedate*
 annoying QP anger-confusion PT push-fly moment PT partition
no fusuma e atarite battari kokeru to...
 GEN sliding.door to hit.and plump fall and
 ‘...the moment (one) pushed away (the other person) in vexation in a
 moment of anger, (he) bumped into the fusuma and fell flat and...’ or ‘...
 (one) pushed away (the other person) in vexation in a moment of anger.
 At that very moment, (he) bumped into the fusuma and fell flat and...’

(1826 *Irofukamisoranenoyume*; CHJ)

b. Sentence-initial adverbial use

omotta kotoba mo dekaneta. Totan ni mata...
 thought word PT difficult.to.come.PST moment PT again
 ‘(Otaki was trying to make an excuse but good) excuses did not occur (to
 her mind). At that moment, again, (Katukoo shook his fist at Otaki).’

(1901 *Hazekooyo: Taiyo*)(11) *shunkan*

a. Head of adnominal clause use

... *keikoo-seraruru o okuru sono shunkan ni mo kimi*
 guidance-do.HON ACC send DEM.MED moment PT even lord
wa nao nagori no osim-ar-uru omomoti nite...
 TOP still reluctance GEN mourn-HON-HON look COP.and

‘... even at the moment (when) the Duchess began a journey (i.e. funeral
 ceremony), the Lord still had a wistful look and...’

(1895 *Otto von Bismarck-koo Zoku; Taiyo*)

b. Sentence-initial adverbial use

... *hito mo uma mo hitotukito boti ni tatimukafu. Sono*
 people too horse too at.a.dash graveyard PT head.over DEM.MED
shunkan ni kisi wa mottaru yari de...
 moment PT knight TOP carrying spear by

‘... both people and horses headed over to the graveyard. Just then, the
 knight speared (a bull)...’ (1909 *Meisi no Supein Kan: Taiyo*)

9. Note in passing that, as in the (a) examples in (10) through (13), the noun-modifying (adnominal) form was not necessarily the *ta*-form but often the *ru*-form (i.e. non-past tense) in these stages (see Note 2). Thus the *ta*-form seems to have been generalized over the twentieth century to the present.

(12) *hyoosi*

- a. Head of adnominal clause use

katto mae e hanete-saurahu e wo hanekiru hyausi
 vigorously front PT jump.and-AUX.HUM grip ACC spray moment
ni hidarite wo motte...
 PT left.hand ACC by.use.of

‘... (someone) stepped forward with renewed vigor. The moment (he) raised a spray of water with the grip of his sword (or spear?)...’

(1477 *Shikishoo, Khouhongi*; Kitahara 2006a, vol.11: 528)

- b. Sentence-initial adverbial use

Takagi wa jibun no koto no yooni kao o sikameta. Sono
 name TOP self GEN thing GEN like face ACC frown DEM.MED
hyoosi ni kare no megane ga kunya-kunya nami-utta.
 moment PT he GEN glasses NOM being.soft.and.pliable wave-hit

‘Takagi made a face as if (he were to blame). At that very moment, his glasses waved nervously.’

(1917 *Shundoo; CHJ*)

(13) *yasaki*¹⁰

- a. Head of adnominal clause use

ano simasaki ni oriyare ba koso ote fa kakure
 DEM.DIST tip.of.island PT be.HON if PT HON.hand TOP hold
watara nu sono yasaki ni, ote wo kakuru demo nasi
 cross NEG DEM.MED moment PT HON.hand ACC hold PT NEG

‘If (the lady) is at the tip of the island, (I [the demon]) would not hold (her) hand to cross the water; in such occasions [*yasaki*], I [the demon] would not even hold (her) hand’

(Late Muromachi *Toraakirabon, Setubun*;
 Kitahara 2006a, Vol.13: 92; cf. Fujita 2019: 67)

- b. Sentence-initial adverbial use

kotosi wa... kootoo site iru. Kono yasaki ni atarite
 this.year TOP fit do exist DEM.PROX moment PT at.the.occasion
ryookoo no hoogyo no hen ga okita no
 imperial.couple GEN demise.HON GEN mishap NOM happen NML
de aru.
 COP be

‘This year is ideal for (such an event). At this very moment, Their Imperial Majesties passed away.’

(1909 *Sinkoku Tanan no Aki; Taiyo*)

The most important finding from this survey is that in all cases, the nominal use of the adnominal clausal head is attested earlier than the sentence-initial adverbial

10. For the historical development of the temporal *yasaki*, albeit viewed from a different perspective, see Fujita (2019).

use, i.e. as adverbial connector. This holds true in general of other cases briefly summarized in Shibasaki (2019). Therefore, the working hypothesis (6) is partially justified by the historical documentation; however, the pathway of change is not uniform. The two frequent nominals, *totan* and *shunkan*, are forms of recent date in the Meiji period (1868–1912), while the other two, *hyoosi* and *yasaki*, are much older, tracing back to the Muromachi period (1336–1573). Nevertheless, it should be noted that the sentence-initial adverbial connectors emerged in the first half of the twentieth century, and *totan*, *shunkan*, and *yasaki* manifested their adverbial connector use in almost the same period. On this point, Kitahara (2006a, vol.9: 1237) comments that *totan* is often used adverbially when accompanied by the particle *in*, and that *hyoosi* and *yasaki* are rephrased synonymously as *totan* in their entries. These descriptions imply that *hyoosi* and *yasaki* as well as *totan* might have been adverbial-oriented from the very beginning. Notice that in (10) through (13), all eight examples are accompanied by *ni*, regardless of whether they are adverbial phrases or adverbial clauses.

The other important part of hypothesis (6) is which of these four adverbial connectors play(s) a leading role in terms of frequency in producing the constructional expansion. Although their adverbial connector usage has a relatively short history, I have made doubly certain to verify the frequency of each expression in the eighteenth and nineteenth centuries by use of the corpora in (9). The result is that in the corpora, only one example of *shunkan* (瞬間) was found around 1836; however, it is accompanied by phonetic transcriptions in hiragana, *matatakuma* ‘in a twinkling.’¹¹ In addition, Okubo and Kinoshita (1991) has only the entry for *yasaki ga warui* (moment + NOM + ‘wrong’) ‘badly timed’, which is placed under the heading of *yasaki* in Maeda (1979) with the same phrase, *yasaki ga warui*. Maeda (1979) also includes *totan no yasaki* (moment + GEN + ‘beat’) ‘(by) mere chance’; Fujita (2019) makes no clear mention of any frequency of *yasaki* in the Edo period. The available evidence suggests that adverbial connectors of each respective expression were not frequently used in the Edo period, even though the reanalysis of (5a) as (5b) had been under way language-internally. Therefore, the second half of the hypothesis (6) needs to be revised on the basis of a larger set of data in my future study.

On reconsidering the matter, one possible factor that facilitates the reanalysis may have been continuous exposure to Western languages in the next synchronic stage, i.e. the Meiji period. However, this is only briefly addressed in the Appendix for reasons of space.

11. The expression *matatakuma* is usually written 瞬く間, while *shunkan* is written 瞬間. Such different realizations come across occasionally due either to the Chinese-derived reading or its native Japanese equivalent reading.

4.2 Bridging context and constructional clustering

I introduce here the following Example (14) that can be considered to show a ‘bridging context’ (Heine 2002) from the older structure (5a) to the newer structure (5b).

- (14) ... *Oo wa nanno to hajikikaesita sono hyoosi ni,*
 name TOP damn.me QP hit.back.ADN DEM.MED moment PT
hidarizasi ni natte...
 a.left.handed.belt.grip (i.e. a wrestling trick) PT become.and
 ‘The moment *Oo* pushed back in a blind effort (saying,) “Damn me,” (he) pushed (his) left arm (under the opponent)...’ or ‘*Oo* pushed back in a blind effort (saying,) “Damn me”’. At the moment, (he) pushed (his) left arm (under the opponent)...’
 (1901 *Tokyobyoo; Taiyo*)

The function of the medial demonstrative *sono* ‘that’ is anaphoric reference to the preceding discourse, while the whole phrase *sono hyoosi ni* ‘at that moment’ directs the interlocutor’s or reader’s attention cataphorically to the following discourse. The forward-looking function of *sono hyoosi ni* may have served to grease the wheels of information flow between preceding and following clauses/sentences, resulting in the creation of the stand-alone adverbial connectors in the course of the change, i.e. reanalysis, from the head of adnominal clause use (12a) to the sentence-initial adverbial use (12b), namely in the ‘bridging’ context.¹² This process of change is also witnessed in the cases of *totan*, *shunkan*, and *yasaki*. Since the four adverbial connectors under investigation here serve to anticipate more to come in the following discourse, they can be regarded as an example of ‘projectors’ in the sense of Hopper & Thompson (2008). To put it another way, these projector constructions were created structurally through reanalysis and have become discourse-based pragmatically over time.

These observations can be summarized as in Figure 1. Note that I will use ‘preceding stage’ instead of ‘initial stage’ (Heine 2002: 86) simply because some earlier (or prehistoric) stages than (i) may exist. S-initial=sentence-initial. (.) means that period is implied but not always tangible in the material around the time (see Section 1).

I would like to emphasize the following findings. Firstly, the whole process of reanalysis in these cases was made in almost the same sequential order as in Figure 1, which is characteristic of reanalysis (Harris & Campbell 1995; see Section 5

12. Tanizaki (1975: 129–132) clearly states that novelists take full advantage of such an ambiguous structure as in (14), which obfuscates the grammatical sequence and enables potential readers to interpret it as they will.

(i) Preceding stage (annum 1447):

Head of adverbial clause (12a)

[clause 1 ... V-*ta* (/ru)]
 ↑
 adnominal
 (/conclusive) form

{ *toki*
sai
ori
totan
shunkan
hyoosi
yasaki
etc. } , [clause 2]

Reanalysis

(ii) Bridging context (annum 1901):

(In)dependent adverbial phrase/clause (14)

[clause/sentence 1 V-*ta* (/ru) (.)
 ↑
 adnominal/
 conclusive form

{ *toki*
sai
ori
totan
shunkan
hyoosi
yasaki
etc. } , [clause/sentence 2]

Showing
 formal
 reduction but
 discourse-dep
 endent
 cohesion

(iii) Switch context (app. 2004):

S-initial adverbial connector (3b)

[sentence 1 ... V-*ta* (/ru)]
 ↑
 (adnominal/
 conclusive form

∅ { **toki*
 **sai*
 **ori*
totan
shunkan
hyoosi
 ?*yasaki*
etc. } , [sentence 2]

Figure 1. Grammatical renewal through the reorganization of sequential relations

below). In the course of the change, the inflectional leveling between adnominal and conclusive forms, as in Table 1, lends impetus to the structural reorganization. In other words, the reanalysis basis of grammatical renewal pushes the boundary of morphosyntax toward discourse, i.e. from the inside of the clause/sentence (i) to the outside of the clause/sentence (iii) in Figure 1.

Secondly, the demonstrative *sono* ‘that’, appearing between preceding and following clauses or sentences, i.e. in the bridging context (ii), serves to lubricate the flow of information even through the reanalysis of (i) as (ii) and through the reorganization of (ii) as (iii). One point to notice here is that the demonstrative *sono* is still used as part of the adverbial connectors *sono yasaki ni* as in (4), which suggests that these adverbial connectors have not reached the next stage, i.e. ‘conventionalization,’ in which new meanings (and maybe new forms too) are “freed from the contextual constraints” (Heine 2002: 86).¹³

Lastly, as pointed out in Note 3, temporal nouns of two morae such as *toki*, *sai*, and *ori* can be used when they are modified by the preceding clauses or by the demonstrative, as in (14) and (ii) in Figure 1. However, they cannot appear without such modifications. On the other hand, *totan*, *shunkan*, *hyoosi*, and *yasaki* are versatile in this respect: they occur as stand-alone adverbial connectors in sentence-initial position. Although I could not find any free-standing *yasaki* in the database, it might have developed a similar usage by analogy with *totan*, *shunkan*, and *hyoosi*, all of which have already created (or established) the usage as in (1) through (3). If my expectation is fully realized, the hypothesis (6) turns out to be valid. Of course, it is well known that analogy often cannot be explained by a rule-based reasoning, as Traugott states: “much analogical thinking never results in change” (Traugott 2011: 55). On the other hand, if we focus on the narrower range of changes dealt with in this study, analogy seems to work well (see De Smet et al. 2017).¹⁴

13. For the nominalization basis of grammatical renewal, see Shibasaki (2018b), which takes a constructionalization approach to the phenomenon.

14. One reviewer points out the possibility that constructions in (i) through (iii) in Figure 1 may be different from each other but not positioned at some points of a continuum. However, as shown in Notes 8 and 12, the adnominal-conclusive merger is crucial; it may take language users less time to conventionalize one morphosyntactic change while taking longer to conventionalize another. Data from earlier stages are limited (Traugott 2007: 358); therefore, it would be safer to conduct a survey of related constructions together, while making the best use of the historical data available to us. Also see Aoki (1986) for a similar discussion.

5. (Re)creating grammar through reanalysis in other languages

We have thus far discussed the reanalysis of one particular sequential relation, i.e. the ambiguous inflectional form of the auxiliary *ta* and the following temporal noun, in creating the new grammar of Japanese.¹⁵ Yet renewing the grammar of a language through reanalysis is not restricted to Japanese. Burridge and Bergs (2017: 148–151) provide a couple of changes in European languages similar to those found in the history of Japanese. While they regard the (re)creation of grammar as a case of grammaticalization – e.g. lexical sources that create (new) grammatical forms or constructions, as well as changes of a given item that become more grammatical – they lay emphasis on reanalysis in the explanation of renewing the grammar. In what follows, I will illustrate two similar cases, one from Harris & Campbell (1995) and the other from Burridge & Bergs (2017).

Through the analysis of a number of relevant examples from many languages such as Nuclear Micronesian and Georgian, Harris and Campbell (1995: 61–65) state that reanalysis may affect the following aspects of structure in a language: (i) constituency, (ii) hierarchical structure, (iii) category shifts, (vi) grammatical relations, and (v) cohesion; semantic change can be witnessed hand-in-hand with reanalysis. For example, in the case of the well-known complementizer construction *for.. to* in English and *fer de... zu* ‘in order to’ in Pennsylvania German, the preposition + NP sequence was originally part of the main clause as noted in brackets in (15) and (16).

(15) Middle English

[*it is bet for me*] [*to sleen my self than ben defouled thus*]

‘It is better for me to slay myself than to be violated thus.’

(Chaucer; cited from Ebert 1978: 22; Harris & Campbell 1995: 62)

(16) Pennsylvania German

[*Es is ungewenlich* [*fer de John*]] [*harti Bicher zu lese*]

it is unusual for the John hard books to read

‘It’s unusual for John to read difficult books.’ (Burridge & Bergs 2017: 150)

In (15), while the constituent *for me* looks to be attached onto the main clause structurally as a prepositional phrase, it serves semantically as a logical subject of the following *to*-infinitive. By Present-day English, the *for* + NP + *to*-infinitive sequence has been reanalyzed as another constituent order, which now can be

15. Denison (2017: 317) states that “the structural change (or changes, if analysed in terms of micro-steps) would seem to arise from vagueness rather than ambiguity” showing a variety of cases from the history of English. In the case of this specific Japanese phenomenon, however, ambiguity rather than vagueness appears to be a driving force for the change.

used as the subject of the whole clause, as in (17).¹⁶ Pennsylvania German in (16) has undergone a similar reanalysis in which the preposition *fer* is reorganized “as a conjunction (or complementizer) heading a new subordinate clause” (Burridge & Bergs 2017: 150) and thus can be preposed, as in (18), where *fer* functions as the subject marker of the infinitive clause.

- (17) Present-day English
[For me to slay myself] [would be better than to be violated thus]
 (Harris & Campbell 1995: 62)
- (18) Pennsylvania German
Fer de John harti Bicher zu lese is ungewenlich
 “For John to read difficult books is unusual.” (Burridge & Bergs 2017: 150)

In both cases, the four aspects (i) through (iv) set forth by Harris and Campbell (1995) can be attested. Since the boundary shifts in these languages, both (i) constituency and (ii) hierarchical structure are affected and reanalyzed. The prepositions *for* and *fer* take on the functional characteristic of indicating the subject of the following infinitives, i.e. (iii) category shifts, and (iv) the grammatical relations can be realized in a different way (cf. Shibasaki 2018a).

These four aspects pertaining to reanalysis can be witnessed in the development of the adverbial connectors discussed in the previous sections. Furthermore, (v) cohesion may be strengthened and realized in a more discourse-oriented way, as discussed in the reanalysis in Figure 1 of Construction (ii) as Construction (iii); then, this case may be labeled as ‘coherence’ (or ‘co-text’; Halliday 1994: 271).

Burridge and Bergs (2017: 148) point out that although the majority of language users may be aware of changes happening to pronunciation and lexicon, “few realize that languages are also constantly renewing their grammar,” as shown in this study. Reanalysis changes underlying structure, so we find ourselves obliged to be aware of an unseen presence to keep up with the change.

6. Concluding remarks

In this study, I have presented the importance of reanalysis in the (re)creation of new grammatical expressions from those already available in a given language. In the case of Japanese, those temporal nouns – i.e. *totan*, *shunkan*, *hyoosi*, and *yasaki* – that were used as the head of adnominal clauses were reanalyzed as sentence-initial adverbial connectors in the first half of the twentieth century. The cause of the

16. See Fisher et al. (2017: 166–172) for a brief history of *to*-infinitives.

change has roots in the inflectional leveling of adnominal and conclusive forms that is considered to have started generally in the Kamakura period (1185–1333). Since a punctuation system was not fully established at that time, the interpretation of an ambiguous structural sequence such as that in (5) and (14) had gone unsolved for centuries. It was not until the increase in formal education and the gradual establishment of punctuation in the Meiji period (1868–1912) that the syntactic boundary began to be manifested. The survey result tells us that users of Japanese started to reanalyze the sequence of (5a) as (5b), unconsciously producing the sentence-initial adverbial connector in the early twentieth century. However, the reanalysis-based change did not occur in one single leap but through a series of steps, as illustrated in Figure 1. In addition, the pace of change seems to have been relatively slow, because a set of corpora in (8) and (9) did not provide many key examples even in the first half of the twentieth century. It is true that in the second half of the twentieth century, we can see clear distributional patterns of each expression, as in Table 3, from which one can deduce that the most frequent form, *totan*, may have played the central role in the development of related constructions with *shunkan*, *hyoosi*, and *yasaki*, as proposed in the hypothesis in (6). However, I need to concede that a larger set of data is required to back up this assumption (Section 4.1).

In addition to the language-internal issue, I briefly touch on the language-external issue, especially concerning the influence of English on the (re)structuring of Japanese grammar from the Meiji period onwards (Appendix). Of course, more research is necessary to tackle this issue; however, since one piece of crucial evidence to support the possibility of contact-induced grammaticalization in Japanese is proposed in Shibasaki (forthcoming), this line of reasoning would not be wide of the mark.

The change from clause-final element to sentence-initial *projector*, while seemingly unlikely especially to those who address language change within the purview of sentence grammar, happens in real life (cf. ‘cooptation’ in Heine 2013). Such a radical structural change would come within range if one takes a broader view and reconsiders how to combine and reanalyze clauses or sentences in a stretch of discourse, hopefully from a diachronic perspective.

Acknowledgements

I would like to express my sincere thanks to the following people: Heather Oumounabidji and Sue Browning for their proofreading of the earlier version of this paper, three anonymous reviewers for their constructive criticism and invaluable suggestions, and Sylvie Hancil for her patience. I need to concede that I could not fully revise the earlier version due to restrictions on outings in the face of a perceived unprecedented crisis, i.e. Covid-19.

Funding

This study is financially supported by the following research projects: Grant-in-Aid for Scientific Research (C) project (PI: Reijirou Shibasaki, No. 16K02781), Grant-in-Aid for Scientific Research (B) project (PI: Reijirou Shibasaki, No. 19K00693), and Grant-in-Aid for Scientific Research (B) project (PI: Ryoko Suzuki, No. 17KT0061), all under the supervision of the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Abbreviations

The glossing conventions are as follows

ACC	accusative	NEG	negative
ADN	adnominal	NML	nominalizer
AUX	auxiliary	NOM	nominative
COMP	complementizer	NP	noun phrase
COP	copulative verb	PASS	passive
DEM	demonstrative	POL	polite
DIST	distal	POSS	possessive
GEN	genitive	PST	past tense
HON	honorific	PT	particle
HUM	humble	QP	quotative particle
INT	interjection	TOP	topic
MED	medial	V	verb

The division of the history of Japanese in this study is: Nara period: 710–784; Heian period: 794–1192; Kamakura period: 1185–1333; Muromachi period: 1336–1573; Azuchimomoyama period: 1573–1600; Edo period: 1603–1867; Meiji period: 1868–1912; Taisho period: 1912–1926; Shoowa period: 1926–1989; Heisei period: 1989–2019; Reiwa period: 2019 – to the present.

Corpora

Balanced Corpus of Contemporary Written Japanese (BCCWJ). <<http://www.kotonoha.gr.jp/shonagon/>>

Ninjobon, Akari Fujimoto & Tomokazu Takada (eds). 2015. Tokyo: National Institute for Japanese Language and Linguistics.

Senjichuu no Hanashikotoba, Orié Endo, Taku Kimura, Takashi Sakura, Chieko Suzuki, Haruko Hayakawa & Toshiaki Yasuda (eds), 2004. Tokyo: Hituzi Syobo Publishing.

Sharebon, Taro Ichimura (ed.) 2015. Tokyo: National Institute for Japanese Language and Linguistics.

Taiyo Corpus (CD-ROM). 2005. Tokyo. Hakubunkan Shinsha.

The Corpus of Historical Japanese (CHJ) <https://pj.ninjal.ac.jp/corpus_center/chj/>

Dictionaries

- Edogo no Jiten* (A dictionary of the Edo language), Isamu Maeda (ed.). 1979. Tokyo: Kodansha.
- Edogo Jiten* (A dictionary of the Edo language), Tadakuni Okubo & Kazuko Kinoshita (eds). 1991. Tokyo: Tokyo-do.
- Nihon Kokugo Daijiten* (The dictionary of the Japanese language), vols. 1-13, 2nd edn, Yasuo Kitahara (ed.). 2006a. Tokyo: Shogakkan.
- Seisenban Nihon Kokugo Daijiten* (A carefully selected edition of *Nihon Kokugo Daijiten*), Yasuo Kitahara (ed.). 2006b. Tokyo: Shogakkan.

References

- Aoki, Haruo. 1986. Evidentials in Japanese. In *Evidentiality: The Linguistic Coding of Epistemology*, Wallace Chafe & Johanna Nichols (eds), 223–238. Norwood NJ: Ablex.
- Aoki, Hirofumi. 2016. Shuushikei, rentaikai no gooryuu ni tuite. In *Grammaticalization and Constructionalization in Japanese and English*, Minoji Akimoto, Hirofumi Aoki & Mitsuru Maeda (eds), 271–298. Tokyo: Hituzi Syobo. (In Japanese).
- Burridge, Kate & Bergs, Alexander. 2017. *Understanding Language Change*. London: Routledge.
- De Smet, Hendrik & Fischer, Olga. 2017. The role of analogy in language change: Supporting constructions. In *The Changing English Language: Psycholinguistic Perspectives*, Marianne Hundt, Sandra Mollin & Simone E. Pfenniger (eds), 240–268. Cambridge: CUP. <https://doi.org/10.1017/9781316091746.011>
- Denison, David. 2017. Ambiguity and vagueness in historical change. In *The Changing English Language: Psycholinguistic Perspectives*, Marianne Hundt, Sandra Mollin & Simone E. Pfenniger (eds), 292–318. Cambridge: CUP. <https://doi.org/10.1017/9781316091746.013>
- Du Bois, John W. 2003. Discourse and grammar. In *The New Psychology of Language*, Vol. 2, Michael Tomasello (ed.), 47–88. Mahwah NJ: Lawrence Erlbaum Associates.
- Ebert, Robert P. 1978. *Historische Syntax des Deutschen*. Stuttgart: Sammlung Metzler. <https://doi.org/10.1007/978-3-476-03873-9>
- Fisher, Olga, De Smet, Hendrik & van der Wurff, Wim. 2017. *A Brief History of English Syntax*. Cambridge: CUP. <https://doi.org/10.1017/9781139049559>
- Fujii, Noriko. 1991. *Historical Discourse Analysis: Grammatical Subject in Japanese*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110860528>
- Fujita, Yuko. 2019. Jikan kankei o arawasu yasaki ni tuite (On the temporal use of yasaki). *Proceedings of the Society of Japanese Linguistics Fall 2019*, 65–72.
- Group Jammassy (Yuriko Sunakawa & Priscilla Ishida (Supervising eds)). 2015. *A Handbook of Japanese Grammar Patterns for Teachers and Learners*. Tokyo: Kuro시오.
- Halliday, Michael A. K. 2004. *An Introduction to Functional Grammar*. London: Arnold.
- Harris, Alice C. & Campbell, Lyle. 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511620553>
- Heine, Bernd. 2002. On the role of context in grammaticalization. In *New Reflections of Grammaticalization* [Typological Studies in Language 49], Ilse Wischer & Gabriele Diewald (eds), 83–101. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.49.08hei>
- Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6): 1205–1247. <https://doi.org/10.1515/ling-2013-0048>

- Heine, Bernd & Kuteva, Tania. 2005. *Language Contact and Grammatical Change*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511614132>
- Hida, Yoshifumi. 2019. *Meiji umare no Nihongo* (Japanese words created in the Meiji period). Tokyo: Kadokawa.
- Hida, Yoshifumi & Asada, Hideko. 2018. *Gendai fukushi yooahoo jiten* (A usage dictionary of contemporary Japanese adverbs), rev. edn. Tokyo: Tokyodo.
- Hopper, Paul J. & Thompson, Sandra A. 2008. Projectability and clause combining in interaction. In *Crosslinguistic Studies of Clause Combining. The Multifunctionality of Conjunctions* [Typological Studies in Language 80], Ritva Laury (ed.), 99–123. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.80.06hop>
- Hopper, Paul J. & Traugott, Elizabeth Closs. 2003. *Grammaticalization*, 2nd edn. Cambridge: CUP. <https://doi.org/10.1017/CBO9781139165525>
- Ishikawa, Toshio. 1986. English borrowings in Japanese. In *English in Contact with Other Languages*, Wolfgang Viereck & Wolf-Dietrich Bald (eds), 457–471. Budapest: Akadémia Kiado.
- Iwasaki, Shoichi. 2002. *Japanese*, 1st edn [London Oriental and African Language Library 5]. Amsterdam: John Benjamins. <https://doi.org/10.1075/loall.5>
- Janda, Laura A. (ed.) 2013. *Cognitive Linguistics: The Quantitative Turn*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110335255>
- Kaneshige, Munekazu. 1979. Meijichuuki no joshi kyooiku ni tuite (On the education of women in the middle of the Meiji period). *Papers from Tokuyama University Economics Society* 13: 75–90.
- Koyanagi, Tomokazu. 2018. *Studies on Grammatical Changes in Japanese*. Tokyo: Kurosio. (In Japanese).
- Kurashima, Tokihisa. 2019. *Chugakusei karano Nihongo no rekishi* (History of Japanese for junior high school students). Tokyo: Chikuma Primer Shinsho.
- Loveday, Leo J. 1996. *Language Contact in Japan: A Socio-linguistic History*. Oxford: Clarendon Press.
- Muraki, Shinjiro. 2012. *Nihongo no hinshi taikai to sono shuhen* (The system of word classes in Japanese and related issues). Tokyo: Hituzi Syobo.
- Nitta, Yoshio. 2002. *Fukushiteki hyoogen no shosoo* (Aspects of adverbial expressions). Tokyo: Kurosio.
- Noguchi, Takehiko. 1994. *Sanninsho no hakken made* (Until the discovery of the third person). Tokyo: Chikumashobo.
- Nomura, Takashi. 2013. *Nihongo 'sutandaado' no rekishi* (History of Standard Japanese). Tokyo: Kodansha.
- Saegusa, Reiko. 2015. *Gokei kara imi e: Kinoushugi eno anchiteeze* (From word form to meaning: Antithesis to functionalism). Tokyo: Kurosio.
- Shibasaki, Reijirou. 2005. *Personal Pronouns and Argument Structure in Japanese: Discourse Frequency, Diachrony and Typology*. PhD dissertation, University of California, Santa Barbara.
- Shibasaki, Reijirou. 2018a. *Sequentiality and the emergence of new constructions*. In *Explorations in English Historical Syntax* [Studies in Language Companion Series 198], Hubert Cuyckens, Hendrik De Smet, Liesbet Heyvaert & Charlotte Maekelbergh (eds), 285–308. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.198.12shi>
- Shibasaki, Reijirou. 2018b. *From the inside to the outside of the sentence: Forming a larger discourse unit with jittu 'fact' in Japanese*. In *New Trends in Grammaticalization and Language Change* [Studies in Language Companion Series 202], Sylvie Hancil, Tine Breban & José Vicente Lozano (eds), 333–360. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.202.14shi>

- Shibasaki, Reijirou. 2019. On the rise of *douride* ‘no wonder’ as a projector and the reformulation of discourse sequential relations in Japanese. In *Japanese/Korean Linguistics 25*, Shin Fukuda, Mary Shin Kim, Mee-Jeong Park & Haruko Minegishi Cook (eds), 383–395. Stanford CA: CSLI.
- Shibasaki, Reijirou. In press. From a clause combining conjunction to a sentence-initial adverbial connector in the history of Japanese: With special attention to *totan(ni)* ‘at that moment’. In *Japanese/Korean Linguistics 26*, Shin Fukuda, Shoichi Iwasaki, Sun-ah Jun, Sung-Ock Sohn, Susan Strauss & Kie Zuraw (eds). Stanford CA: CSLI.
- Shibasaki, Reijirou. Forthcoming. From comparative standard marker to comparative adverb through language contact: On the grammaticalization of *yori* in Modern through Present-day Japanese. Manuscript. Meiji University.
- Sugimoto, Tsutomu. 1979. *Goi to kutoohoo* (Vocabularies and punctuation). Tokyo: Ohfusha.
- Takahashi, Keiko & Higashiizumi, Yuko. 2014. Usage of *kekka* in Modern Japanese corpora. In *Proceedings of the Sixth Corpus Japanese Linguistics Workshop*, 103–112. Tokyo: National Institute for Japanese Language and Linguistics (NINJAL). (In Japanese).
- Tanaka, Hiroshi. 2010. *Fukugooji kara mita Nihongo bumpoo no kenkyuu* (A study of the Japanese language from the perspective of compound particles). Tokyo: Hituzi Syobo.
- Tanaka, Makiro. 2012. *Meijikooki kara taishooki no goi no reberu to goshu* (Levels and types of words from late Meiji to Taisho eras). Tokyo: National Institute for Japanese Language and Linguistics.
- Tanizaki, Junichiro. 1975. *Bunsho dokuhon* (Text reader). Tokyo: Chuokoron-Shinsha.
- Toyoda, Minoru. 1939. *Nihon eigakushi no kenkyu* (A historical study of English literature in Japan). Tokyo: Iwanami.
- Traugott, Elizabeth Closs. 2007. Interview with Elizabeth Closs Traugott (Interviewer: Scott A. Schwenter). *Journal of English Linguistics* 25(4): 353–364.
- Traugott, Elizabeth Closs. 2011. Grammaticalization and mechanisms of change. In *The Oxford Handbook of Grammaticalization*, Heiko Narrog & Bernd Heine (eds), 19–30. Oxford: OUP.
- Yamada, Masanori. 1936. *Edokotoba no kenkyu* (A study of Japanese in the Edo era). Tokyo: Musashino-shobo.
- Yuzawa, Kokichiro. 1954. *Edokotoba no kenkyu* (A study of Japanese in the Edo era). Tokyo: Meiji-shoin.
- Yuzawa, Kokichiro. 1955. *Tokugawajidai gengo no kenkyu* (A study of Japanese in the Tokugawa period). Tokyo: Kazama-shobo.

Appendix. Language contact as a facilitator for reanalysis

In addition to some language-internal aspects of reanalysis discussed thus far, language-external aspects, i.e. language contact issues, cannot be ignored, although the following is just a preliminary sketch at the present stage. For convenience, let me start with educational reform in Japan. One important socio-cultural issue that cannot be overlooked in this study is the dissemination of education in the Meiji period (1868–1912). Until around the end of the Edo period, the study of the Chinese classics had been restricted to members of the Samurai class. However, the spread of public as well as private education in the Meiji period served to fill the educational gap between the social classes of those days. For example, the school enrollment rate steadily increased: 28.13% in 1873, 45% in 1887, 66.65% in 1897, 94.43% in 1904 (Hida 2019: 14); female education was also promoted and improved in tandem with the increase in formal education (Kaneshige 1979).

The evidence presented indicates that an increasingly large number of people from different social classes could acquire education, raising their cultural level and studying Western science as well as the Chinese classics.

Ishikawa (1986: 457) claims that in the late nineteenth through early twentieth centuries, a massive influx of loan words from English can be witnessed to meet the urgent need for modernization in areas such as science, technology, and higher education in order to learn about Western ideas, especially scientific and military technologies. Therefore, texts published just after this period can be regarded as reflecting the early grammatical influence of Western languages on Japanese; see Loveday (1996: 41) for American influence on Japanese vocabulary from the 1910s to the 1930s. Presumably, decades of exposure to Western languages and cultures are responsible not just for extending the lexical inventory of Japanese, but for influencing the purview of grammatical categories such as subject, person, number, gender (see Fujii 1991 on the development of grammatical subject in Japanese in the same vein; see Shibasaki 2005: Chapter 1).

The evidence in support of the position that European languages, especially English, carried increasing weight in the reorganization and expansion of Japanese grammar in this period is as follows. Toyoda (1939: 365–382, 449–489) chronologizes translated English novels in the first half of the Meiji period up until 1891: *The Pilgrim's Progress* (John Bunyan, translated and published in 1879), *Gulliver's Travels* (Jonathan Swift, translated and published in 1880), *Utopia* (Sir Thomas More, translated and published in 1882), to name but a few (Toyoda 1939: 455–457). One of the most popular translated materials was *Robinson Crusoe* (Daniel Defoe, translated and published in 1883), which ran into several editions, with the help of good translation by Tsutomu Inoue (ibid.: 457). Although these translated works may not include any crucial examples of *totan*, *shunkan*, *hyoosi*, or *yasaki* that clearly correspond to (*at*) *the moment* in English, a good number of English learners in those times may have read both source texts and translations, as mentioned above. Therefore, Japanese learners of English of the time appear to have had an opportunity to reinvent their grammar through translation. Additionally, the number of printed publications such as English dictionaries, textbooks, and grammar books had been on the rise (Toyoda 1939).

Heine and Kuteva (2005: 239–241) state that the length of contact necessary to undergo and complete a sequence of changes, i.e. replication, varies according to the degree of bilingualism among the communities concerned. Since the case of Japanese in contact with English and other European languages is not a situation involving large-scale bilingualism such as that proposed in Heine and Kuteva (2005), we cannot readily assess whether Japanese can provide another example of contact-induced grammaticalization. In fact, Japanese is not included in the index of languages addressed in Heine and Kuteva (2005), while Shibasaki (forthcoming) points out one potential case of contact-induced grammaticalization in the late nineteenth through the early twentieth century Japanese. Contact-induced grammaticalization or language change may be related to the following statement: “Speakers exploit available grammatical structures to realize their goals in speaking” (Du Bois 2003: 49).

PART II

Discourse markers

The grammar-discourse interaction
from a synchronic, usage-based perspective

The meaning and functions of French *je pense (que)*

A constructionalist and interactional account

Juliette Angot and Maj-Britt Mosegaard Hansen

The University of Manchester

This paper examines the meanings of the French construction *je pense (que)* ('I think (that)', henceforth *JP*), as well as its uses in social interaction. We propose an account of *JP* as a micro-construction with two basic elements of coded meaning, either one of which may be pragmatically foregrounded in a given context of use. As a direct result of its constructional meaning, which we argue is situated at the context level rather than at the content level of discourse, the expression frequently functions as a pragmatic marker in conversation. We show that uses of *JP* may fulfil both interpersonal, face-related functions and discourse-organizational ones. Our analysis has implications for the way *JP* and similar expressions in French and other languages are categorized and how their pragmatic functions may be seen to have developed. It also raises interesting questions with regard to the psycholinguistic processing of such constructions.

Keywords: pragmatic markers, epistemic verbs, parentheticals, Construction Grammar, Interactional Linguistics

1. Introduction

This paper examines the meanings of the French construction *je pense (que)* ('I think (that)', henceforth *JP*), as well as its uses in social interaction. Previous literature on *JP* and similar constructions (e.g. *je crois (que)*, 'I believe (that)', henceforth *JC*) has focused mainly on its role in expressing modality and on its parenthetical status (Apothéloz 2003; Dendale & Van Bogaert 2007; Schneider 2007; Gosselin 2018). The role of *JP* in interaction, on the other hand, has thus far been granted limited attention, although a few studies have investigated its role as a discourse marker (DM) (Andersen 2007; Mullan 2010). Below, we propose an account of *JP* as a micro-construction with two basic elements of coded meaning, either one of which may be pragmatically foregrounded in a given context of use. As a direct result of

its constructional meaning, which we argue is situated at the context level rather than at the content level of discourse (Hansen 2008), the expression frequently functions as a pragmatic marker (PM) in conversation. We show that uses of *JP* may fulfil both interpersonal, face-related functions and discourse-organizational ones.

The structure of the paper is as follows. Section 2 summarizes earlier studies of the marker. Section 3 presents our database and gives an overview of the occurrences of *JP*. In Section 4, we analyze *JP* as a micro-construction that has two basic elements of meaning – one epistemic, the other subjective – , and we argue for the usefulness of comparing the construction to other related ones in order to arrive at a more fine-grained analysis of its contribution. Section 5 examines interactional uses of *JP*, showing that it may on the one hand function as a face-saving device and, on the other, perform discourse-organizational functions. In Section 6 we look into correlations between the function and syntactic position of *JP*, and we discuss alternative ways of accounting for its positional variability. Finally, in Section 7, we draw conclusions and discuss the linguistic status of *JP*.

2. Previous studies

In previous literature, *JP* has been studied from various perspectives. The construction is often examined along with other expressions of opinion that exhibit similar syntactic behavior, such as *JC*, *je trouve (que)* ('I find (that)', henceforth *JT*) or *je considère (que)* ('I consider (that)'). From a syntactic and pragmatic perspective, the parenthetical use of such expressions has been the chief focus of research (Blanche-Benveniste 1989; Apothéloz 2003; Avanzi & Glikman 2009; Schneider, Glikman & Avanzi 2015).

Adopting a pragmatic perspective, Schneider (2007) describes the mitigating functions of a set of clauses that he refers to as "reduced parenthetical clauses" (RPCs). His aim is to understand how the speaker's claim to truth is weakened and what the consequences of this communicative strategy are. *JP*, *JC*, and *JT* are all included in Schneider's list of RPCs as mitigators. He concludes that they are all mitigating RPCs in that they downgrade speaker commitment. They do so by alleviating responsibility, i.e. the speaker directly reduces their burden of responsibility (Schneider 2007: 109).

In studies of modality, *JP* is generally categorized as an epistemic marker that weakens the speaker's commitment. Dendale and Van Bogaert (2007), however, show how it can be described in certain contexts as an evidential marker signalling the source of the information, rather than an epistemic modal marker. According to these authors, the contrast between *JT*, on the one hand, and *JP/JC*, on the other, can in part be explained in terms of direct vs non-direct evidence. Thus, *JT* can only be used when the speaker has direct experience of the state of affairs in its scope. *JP* and

JC, in contrast, are compatible with non-direct experience, in which case they signal that the speaker's opinion is based on an inference. As for the contrast between the latter markers, Dendale & Van Bogaert (2007: 72–75) explain that *JP* involves more intellectual processing than *JC*. Within the semantic domain of modality, Gosselin (2018) describes the constraints imposed by a small group of French expressions of personal opinion, including *JP*, *JC* and *JT*, on their embedded predicates. His analysis stresses that *JP* can serve a dual function: depending on the predicate it occurs with, it may thus function either as an epistemic or as a subjective marker. Moreover, he highlights key differences between *JP* and *JC*, whereby *JC* always expresses epistemic modality, whereas in certain contexts, *JP* marks the speaker's personal opinion without marking epistemic modality. When they are epistemic markers, Gosselin (2018) analyzes *JC* as based on knowledge about the situation, while *JP* is based on general knowledge. This distinction has overtones of evidentiality and can thereby be related to Dendale and Van Bogaert's (2007) study.

Andersen (2007) focuses on a group of DMs which from a morpho-syntactic viewpoint resemble clauses insofar as they contain a finite verb. However, as DMs, they grammaticalized into fixed formulae and cannot scope over any other elements in the sentence. In this group of DMs, she includes *JP*.

Finally, Mullan (2010) is to date the most in-depth study detailing the meaning and functions of *JP* in interaction. Working in a cross-cultural perspective, her main goal is to investigate and compare how social interactants linguistically mark their opinions, through the use of *JP*, *JC* and *JT* in French, *vs I think (that)* in Australian English. She shows how the inherent semantic content of each expression differs, and assigns to each of them a predominant function in discourse: organizational (such as marking a boundary or a different perspective from the prior turn), semantic (either expressing opinion or doubt), or pragmatic (face-saving function). Her findings show that all of the four expressions are grammaticalizing into DMs, on the basis of their organizational role. However, none of the French expressions is assigned a pragmatic function as its predominant role, their role as markers of opinion remaining important.

Taking a Constructionalist and corpus-based interactional approach, the present study seeks to unify these different strands of research and to bring new insights both on the interplay between the meaning and interactional function of *JP* in particular contexts and on the interplay between its interpretation and its position within the host clause. A particular aim is to highlight, alongside discourse-organizational functions, functions that are politeness-induced, and which have been ignored or underestimated in previous studies. In addition, our analysis has implications for the way *JP* and similar expressions in French and other languages are categorized and how their pragmatic functions may be seen to have developed. It also raises interesting questions with regard to the psycholinguistic processing of such constructions.

3. Presentation of the data

The data consist in three hours of informal conversation, involving exchanges of approximately ten to fifteen minutes' duration. They come from two corpora, recorded in France by Angot, the first in 2013, and the second in 2019. Each corpus involves four native speakers of French who were recruited on the basis of several criteria: age (20–30), gender (equally represented), and relationship to each other (either good friends who have known each other for several years or “friends of friends” meeting for the first time). The two corpora differ from each other with respect to the last variable: the first corpus involves both kinds of relationship, while the second only involves good friends. Participants were grouped together in pairs, or in groups of three or four. They were given topics of discussion prior to the recordings and were asked to talk about them for approximately ten or fifteen minutes. Topics were given in the form of assertions, so that participants would feel free to choose the perspective from which they wished to broach the topics (especially topics encouraging debate).

This methodology is inspired by Cosnier and Kerbrat-Orecchioni (1987: 7), who describe the conversations obtained as a hybrid genre, between debate and naturally-occurring conversation. The interactions clearly are artefacts and not spontaneous insofar as they were pre-arranged and constrained in terms of topics, time, and number of participants. The methodology used means that our data is characterized by two features that are not common in spontaneous verbal interaction, but specific to the kind of encounters constituting this corpus: First, it was a tacit condition that the participation of all participants was required, which led some speakers to explicitly designate another for the next turn; secondly, because participants were given topics to talk about for a specific duration, those topics are sometimes shifted more abruptly than would be the norm in non-elicited conversation (Levinson 1983: 312–316; Jefferson 1984; Sacks 1992), and conversational closings are achieved explicitly rather than through production of the standard sequences of closing-implicative utterances, passing turns, and terminal greetings (Schegloff & Sacks 1973; Levinson 1983: 316–318). These two characteristics must be kept in mind as they may be responsible for triggering the use of the target construction: whenever a current speaker chooses to designate someone as the next speaker by asking for their opinion, the use of a marker of opinion may become relevant; moreover, sudden shifts of topics and conversation endings can be seen as actions that require mitigation. We will see in §§ 5.1.1–5.1.2 that *JP* is in fact frequently used in both situations. Nevertheless, the fact that turn-taking was essentially free and that the chosen topics reflect everyday ones, the recordings arguably approximate naturalistic conversation sufficiently for our purposes.

The corpus has been transcribed following Conversation Analytic transcription conventions (see Appendix), and a total of 66 occurrences of *JP* has been found. Table 1 details these 66 occurrences by sentence structure.

Table 1. Occurrences of *je pense* by sentence structure

	Corpus 1	Corpus 2	Total
Matrix clause (<i>je pense que</i>)	12	18	30 (45%)
<i>je pense</i> Ø in matrix position	1	3	4 (6%)
Parenthetical constructions	8	7	15 (23%)
Standalone constructions	0	5	5 (8%)
Other	6	6	12 (18%)
Total	27 (41%)	39 (59%)	66 (100%)

Table 1 shows that 34 instances (51%) of *JP* occur in initial (matrix) position with respect to the proposition over which it takes scope. This can thus be considered to be its unmarked position. Four of these instances do not include the complementizer *que*. Nine instances are also turn-initial and in this position, *JP* always includes a complementizer. In six cases, however, we find *JP* (here again, always including *que*) as the second token of the turn. In such contexts, the linguistic form prefacing *JP* and opening the turn may be a DM (*donc* ‘so’, *mais* ‘but’, *ben/bah/beh* ‘well’, e.g. Hansen 1998), a positive or negative response particle (*nan* ‘no’, *ouais* ‘yeah’, cf. Hansen 2020), or the strong pronoun *moi* (‘me’). Moreover, *JP* takes the form of a reduced parenthetical in fifteen instances (23%). In this case, it occurs without a following complementizer in either medial or final position *vis-à-vis* the proposition in its scope. Five instances of *JP* (8%) are used as standalone expressions, i.e. its complement is omitted, but can be recovered from the immediately preceding context (usually the previous speaker’s turn). In such cases, the marker in fact never occurs completely on its own, but is in most cases accompanied by *ouais* (‘yeah’), an informal version of the positive response particle *oui* (‘yes’). Finally, twelve instances of *JP* (18%) have been classified in the category “Other”. In such cases, it is followed by the complementizer *que*, but either no embedded clause can be identified because the speaker self-interrupts before delivering a pragmatically complete message, or the structure of the turn is ambiguous. The former configuration is exemplified in (1), where *JP* occurs in a turn which is abandoned halfway through its production (1.1), possibly because Nicolas takes the floor in line 2:

(1) Corpus 1 – Nicolas / Paul¹

- 1 PAU j' pɛ̃sɛz kɛ []
I think that
- 2 NIC [j' kʁɔj]lɛs kɛ tu vɔlɛz klɛʃɛr sɛr sɛt
3 question ɛ̃hɛ=
I thought you wanted to argue about this question ɛ̃hɛ
- 4 PAU =ɛ̃hɛ ahɛ
5 (0.6)
- 6 PAU .h ouais j'aurais pu mais bon j' suis ɛ̃pas:: j' suis pas si
7 convaincu
.h yeah I could have but well I'm not I'm not that convinced.

4. The meaning(s) of *je pense (que)*

In this and the following section, we argue that the meaning and functions of JP can be better grasped:

1. if it is seen, in a Constructionalist perspective (Hoffman & Trousdale, eds. 2013), as a micro-construction consisting of two elements each contributing a central element of its semantics (§§ 4.1–4.2);
2. if it is analyzed in comparison with other related micro-constructions, viz. *JC* and *JT* (§ 4.3);
3. if we take into account the wider contexts in which the three constructions occur, that is, in interaction (§ 5).

4.1 *Je pense (que)* as a micro-construction

We adopt a Construction Grammar approach, which defines constructions as conventionalized pairings of form and meaning, which can be of any size and complexity. We propose that *JP* can be seen as a construction, and more particularly following Traugott (2008), as a “micro-construction”. To account for the development of constructions over time, Traugott (2008: 31–32) suggests the hierarchical distinction between “macro-constructions”, “meso-constructions”, and “micro-constructions”. At the highest level in the hierarchy are “macro-constructions”, defined as abstract, functional and schematic constructions; at an intermediary level, we find “meso-constructions”, which are sets of similar constructions; at the lowest level are

1. In the conversational excerpts, *je pense*, *je crois* and *je trouve* have been translated into English. However, their English translation does not necessarily reflect their individual functions in French. When the complementizer *que* follows, ‘that’ has been added.

“micro-constructions”, individual constructions entrenched in the mental grammar of speakers, which are representatives of meso-constructions, but which may have idiosyncratic properties. These three levels are abstractions and have to be distinguished from “constructs”, which are linguistic expressions as instantiated in situated language use.

Thus, at the highest level in the hierarchy, the fully schematic combination “Subject + Epistemic Verb” can be seen as a macro-construction, since neither element is lexically specified, but both are defined by function. The construction further includes an optional third element, “(complementizer) + p[roposition]”, which is syntactically the direct object of the former two and which constitutes their semantic scope. Importantly, this element will often have to be recovered from context. In other words, the macro-construction is specified for the possible null instantiation of its complement (Hilpert 2014: 44f), a property which is inherited by lower levels of the constructional hierarchy. At an intermediate level, we find “Subject + *penser*”. This partially schematic, partially substantive meso-construction is related to a set of other meso-constructions with broadly similar meanings, i.e. “Subject + *croire/trouver/considérer/...*”. Finally, at the micro-level, we find the individual construction *JP*, a lexically and morphologically fully specified construction, which may in at least some contexts compete with the micro-constructions *JC* and *JT* (see § 4.3 below).²

4.2 Constructional meaning of *je pense (que)*

As a construction, *JP* consists of two elements: the 1ps subject *je* and the epistemic verb *penser* in the present indicative. Each element of the construction contrasts paradigmatically with related linguistic items: Thus, the first element contrasts with other person/number combinations (*tu* ‘you’; *il/elle* ‘he/she’; etc.), while the second element contrasts, on the one hand, with other verbs denoting degrees of epistemic commitment, such as *sais* (‘know’) and *crois* (‘believe’), and on the other hand, with different tenses/moods of the same verb, e.g. *pensais* (‘thought’), *vais penser* (‘am going to think’), etc. In other words, *JP* as a micro-construction can be seen as composed of two elements of meaning, each encoded in a specific component of the construction: we refer to the semantic contribution of the subject *je* as the *subjective meaning*, and that of the verb form *pense* as the *epistemic meaning*.

2. Note that our proposal differs from the Constructionalist analysis of English *I think* and its congeners put forward by Van Bogaert (2011). The difference can largely be accounted for by the greater formal variability of the relevant English expressions.

To account for the epistemic meaning of *JP*, we build on the theory of Generalized Conversational Implicatures (GCIs, cf. Grice 1975; Gazdar 1979; Levinson 2000). The GCIs we are interested in here are inferences of epistemic uncertainty, and more specifically clausal and scalar implicatures, which fall under Grice's (1975: 26) first maxim of Quantity: "make your contribution as informative as is required". Neo-Gricean theories see such implicatures as being obtained on the basis of the existence of alternative expressions of different strengths. In the case of embedding constructions, many verbs, typically verbs of propositional attitude and verbs of saying, form contrastive pairs or scales (Levinson 2000: 110; Gazdar 1979: 61). For example, *believe/know*, *claim/reveal*, *think/realize* constitute such scales, involving a strong verb that entails or presupposes its complement, and a weak verb that does not, but instead implicates the negation of the stronger verb.

In French, *penser/savoir* ('think'/'know') form a contrastive pair of this type, *savoir* being a strong verb which entails its complement, while *penser* is a weak verb that gives rise to an inference of epistemic uncertainty. Thus, (2) and (3) below differ in that the embedding construction in (2) entails the embedded clause *il neige*, while the embedding construction in (3) fails to entail *il neige*:

- (2) *Jean sait qu'il neige.*
'Jean knows that it is snowing.'
- (3) *Jean pense qu'il neige.*
'Jean thinks that it is snowing.'

In addition to the clausal implicature that for all Jean knows, it may or it may not be snowing, (3) gives rise to the scalar implicature that it is not the case that John knows it is snowing. Note that in many contexts, (3) will additionally carry the implicature that the speaker knows or believes that it is, in fact, not snowing. Put differently, such sentences are often associated with the speaker's knowledge or belief that the embedded proposition is false. Analogously, (4) below, which features a 1ps subject followed by the past tense of *penser*, will often carry the implicature that the speaker now knows it was not snowing:

- (4) *Je pensais qu'il neigeait.*
'I thought it was snowing.'

When using the micro-construction *JP*, on the other hand, speakers always display their belief that the proposition denoted by the complement is true, as shown in (5) below. These observations in our view justify regarding *JP* as a micro-construction with idiosyncratic semantic/pragmatic properties that distinguish it from other uses of the meso-construction "Subj + *penser*".

- (5) *Je pense qu'il neige.*
 'I think that it is snowing.'

At the same time, by using the semantically weaker epistemic verb in (5), the speaker does implicate that they are not in a position to utter the stronger statement (*je sais qu'il neige* ('I know that) it is snowing'). JP therefore expresses an implicit acknowledgement that the speaker's beliefs may potentially be in conflict with another person's beliefs, including those of the addressee(s).

We propose that, while both the subjective and the epistemic meaning are present to some extent in all uses of JP, one of the two meanings is typically foregrounded at the expense of the other, which recedes into the background (cf. also Hooper 1975: 95). Thus, in some uses, JP mainly works to express a subjective stance towards the utterance, while in other cases it is chiefly an epistemic marker that reduces the speaker's commitment to the propositional content of their utterance. The context in which JP occurs determines which meaning will come to the foreground such that, when the construction occurs with objective facts, epistemicity is foregrounded insofar as the speaker flags a genuine uncertainty towards the propositional content of their utterance. When JP occurs with claims that are grounded in personal beliefs or opinion (i.e. where there are no objective facts), on the other hand, the subjective meaning is foregrounded and the propositional content reflects the speaker's own stance. Consider the examples in (6) and (7):

- (6) *Je pense que Laura a déménagé à Chicago.*
 'I think that Laura moved out to Chicago.'
- (7) *Je pense que le discours de Laura était remarquable.*
 'I think that Laura's speech was remarkable.'

In (6), the embedded proposition is an objective fact, whose accuracy can in principle be verified. The speaker is uncertain about the propositional content of the utterance, and the epistemic meaning, which rests on the existence of the <*penser / savoir*> scale, is in the foreground. Nevertheless, the subjective meaning, although backgrounded, remains present in (6), in that the speaker presents their own belief, in contrast with possible alternative beliefs held by other people. In contrast, the embedded proposition in (7) expresses a purely subjective stance, and it is possible that consensus may never be reached. In (7), it is the semantic contribution of the 1st subject that is more salient, and the subjective meaning is therefore in the foreground. The epistemic meaning recedes into the background but is still present, as the speaker cannot vouch for the intersubjective truth of the asserted proposition.

With objective facts as in (6), the omission of JP would modify the truth of the embedded proposition, which would be asserted as a factual statement, whereas

there would be no modification of the truth value of the embedded proposition when *JP* is omitted with subjective claims. In the latter case, *JP* is used to make explicit the subjective and potentially contestable nature of the utterance, and it commutes with *JT* (cf. § 4.3.2). An additional (rather more heavy-handed) test of the difference between the two meanings is to ask whether the sequence *je ne pense pas que p ; je sais que p* ('I don't think that p; I know that p') would be a possible alternative to *JP*, without markedly changing the contextual interpretation of the proposition designated as p. In that case, it is the epistemic meaning that is foregrounded.

This analysis implies that the meaning of the micro-construction *JP* is, in all of its uses, situated at what Hansen (2008: 16) calls the "context level" of discourse. Hansen defines context-level meanings as metadiscursive in nature, pertaining to relations between a described state-of-affairs and the discourse itself or the wider speech situation, including – among other aspects – the speaker's attitudes to the state of affairs in question. Context-level meanings stand in contrast to "content-level" meanings, which bear on the described state of affairs and/or its relation to other (real or imagined) states of affairs. Importantly, some types of linguistic constructions, such as PMs, express context-level meanings by definition. The fact that the micro-construction *JP* inherently expresses this type of meaning, too, in our view goes some way towards explaining the ease with which both this expression and its equivalents in other languages acquire parenthetical uses that afford them the type of positional variability that is characteristic of PMs. This may account for the apparent difficulty in documenting a diachronically gradual development towards parenthetical status where constructions like *JP* are concerned (cf. Heine et al. fc: § 1.2.1.2): semantically and pragmatically, such constructions convey meanings typically associated with PMs from the very beginning. It is therefore arguably natural for them to assume the syntactic behavior associated with PMs.

The examples below illustrate the uses of *je pense* as a primarily subjective vs primarily epistemic marker. In (8), *je pense* foregrounds a subjective meaning. Aurore and Paul are discussing one of the topics that were suggested to them, namely whether they agree with the fact that food and its preparation can be referred to as *culinary arts*. In what precedes the excerpt, Aurore takes a strong stance against this. Nevertheless, she acknowledges that there are some exceptions where food can be considered an art, and gives the example of chocolate sculptures. Her utterance is characterized by the use of *JP*, and the proposition expressed by the utterance hosting *JP* displays a personal opinion, which may vary from one person to another. The omission of *JP* would not modify the truth value of the proposition, but instead underscores a subjective claim.

(8) Corpus 1 – Aurore / Paul

- 1 AUR ... le côté - j' **pense que** oui on peut: avoir un côté artistique
 2 quand on fait des sculptures en chocola:t et cetera mais. .hh
 3 une assiette que tu vas dév- `fin dévorer en deux s`condes euh
 ... *the side - I think that yes it's possible to have an*
artistic side when somebody does chocolate sculptures and so on
but .hh a plate that you're going to eat up in two seconds um

In contrast, the excerpt in (9) displays a parenthetical use of *JP*, occurring clause-medially, and foregrounding the epistemic meaning. Bruno talks about a former colleague of his, and the proposition expressed by the host utterance (*il arrive bientôt à la quarantaine*) concerns an objective fact:

(9) Corpus 2 – Bruno / Céline

- 1 BRU le: seul collègue parisien avec qui j'ai gardé euh: des
 2 contacts,
 ... *the only colleague from Paris I kept in touch with,*
 3 (...)
 4 il arrive à: il arrive j' pense bientôt à la quarantaine, i'
 5 doit avoir entre trente-cinq et quarante.
 ... *he's soon I think going to turn forty, he must be between*
 ... *thirty-five and forty.*

Table 2 lists all the instances of *JP* in the database according to their predominant function and by corpus. The total number of instances of *JP* foregrounding an epistemic meaning (48%) is higher than those foregrounding a subjective function (35%), suggesting that *JP* is preferred as an epistemic marker. However, while a substantial difference can be observed in Corpus 2, Corpus 1 shows a somewhat more even distribution across both meanings, which suggests that the predominant meaning of *JP* may vary, either on the basis of the conversational topic or of individual speaker preference. A smaller proportion (17%) has been classified in a separate category ("Other"). It includes instances of *JP*, followed or not by the complementizer *que*, that cannot be assigned any predominant meaning. Most of the time, the turn in which they appear is abandoned halfway through its production, to be reformulated or not.

Table 2. Predominant function of *je pense*

Function	Corpus 1	Corpus 2	Total
Subjective	9	13	22 (35%)
Epistemic	12	21	33 (48%)
Other	6	5	11 (17%)
Total	27	39	66 (100%)

4.3 Fine-tuning the analysis of *je pense (que)*: Comparison with *je crois / trouve (que)*

As mentioned previously, the meaning of *JP* can be better grasped when it is compared to other related micro-constructions. Here, we focus on the two expressions already mentioned above, *JC* and *JT*, which similarly involve a 1ps subject and a verb of propositional attitude. In some (but not all) contexts, the three constructions are therefore interchangeable. Thus, in addition to forming a contrastive pair with *je sais*, *JP* may also contrast with either *JC* or *JT*. However, contrary to the pair *savoir/penser*, the pairs *croire/penser* and *trouver/penser* are not scalar in nature and no GCI arises from the use of one verb instead of the other.

Of the two alternative constructions, *JC* most closely resembles *JP*, in that it also forms a scalar contrastive pair with *je sais* ('I know'). Like *JP*, it can therefore in principle foreground either the epistemic or the subjective function. However, although *JC* is pragmatically felicitous with subjective claims, it always occurs with factual statements (in opposition to inherently subjective statements) in our database, and more specifically with states of affairs that directly relate to the speaker's past. In most cases, these states of affairs are either events that took place in the past, or they represent more or less exact information (e.g. figures) acquired previously. We take this as evidence that *JC* is used when the speaker already has some knowledge of the matter under discussion (cf. Angot In Progress).

JT, on the other hand, does not commute with *je sais* and is pragmatically felicitous only with inherently subjective statements. Accordingly, this construction is frequently combined with adjectives of subjective evaluation. To a lesser extent, it occurs with states of affairs (without any adjectives) where the topic of discussion is inherently subjective (cf. Angot In Progress).

In terms of frequency, *JC* occurs approximately as often as *JP*, with a total of 47 instances in the database (24 in Corpus 1 and 23 in Corpus 2). As for *JT*, we find 19 instances of this expression (10 in Corpus 1 and 9 in Corpus 2). In both cases, we observe regular associations between the constructions and their surrounding context, e.g. the state of affairs referred to in the utterance as well as any other linguistic or non-linguistic elements (hesitation markers, modals, etc.), which can be taken as clues to their respective semantic/pragmatic specificities.

4.3.1 *Je pense (que) vs je crois (que)*

In line with the analyses of Dendale and Van Bogaert (2007) and Gosselin (2018), our data suggest that when *JP* and *JC* are used to mitigate factual statements, the contrast between them relates to the source of information. In evidential terms, *JP* suggests a guess or a supposition given spontaneously, while *JC* indicates anterior knowledge or evidence. In the database, there is in fact no instance of *JC* as principally a marker of subjectivity. It invariably occurs in association with factual

propositions that rely on evidence from, or knowledge acquired in, the speaker's past. Thus, *JC* signals an inability to remember a fact accurately, as in (10) and (11). Before the beginning of the first excerpt, Aurore was speaking about a fast-food restaurant that she likes in the city where both participants live. Nicolas goes on to introduce a nearby brasserie, a type of French restaurant, which serves similar food (l. 1 to 4). Throughout his turn, Nicolas is trying to provide information about this brasserie, searching his memory in order to recover as much as possible.

(10) Corpus 1 – Aurore / Nicolas

1 NIC beh tu vois juste en face? au - 'fin au: (.) quand tu sors.
 2 sur la gauche j' **crois** un truc comme ça, .h t' as le:=mh::
 3 comment i' s'appelle t' as une brasserie qui s'appelle le -
 4 monsieur machin, j' **crois** .hh
*well you see just across the street? at - well at (.) when you
 go out. on the left I think something like that, .h there's
 the=mh how's it called there's a brasserie that's called the -
 "monsieur machin", I think .hh*

In addition to *JC*, various other linguistic features indicate that Nicolas' memory is faulty, namely the use of the general extender *un truc comme ça* 'something like that' (l. 2) and the self-directed question *comment il s'appelle* 'what's it called' (l. 3). To these linguistic features can be added the numerous cut-offs in ll. 1, 2, and 3.

The excerpt in (11) does not explicitly refer to an event in the speaker's past, but from the use of the past tense of the copula *être* (*c'était*, l.1), it seems clear that he is relying on something he learned previously to be able to assert the fairly exact figure of 170 million subscribers. In this excerpt, a range of interactional features reflect discontinuous speech and contribute to signaling that Bruno is searching his memory: the hesitation marker *euh*., the prolongation of the sound in *soixante-di:x*, the slower pace of *millions*, and the following repetition of this word, which is prefaced and followed by a 0.3 second pause.

(11) Corpus 2 – Adeline / Bruno

1 BRU j' **crois** qu' c'était euh: cent soixante-di:x <millions> (0.3)
 2 millions, (0.3) d'abonnés,
*I think that it was um a hundred and seventy million (0.3)
 million, (0.3) subscribers,*

In contrast, *JP* is preferred when the matter at hand is not one where the speaker enjoys privileged epistemic access. Before the excerpt in (12) begins, Nicolas has asked Charlotte whether or not she thinks that Asian restaurants in France actually serve traditional cuisine reflecting the food they would cook in their home country. Here, Charlotte's use of *JP* could be glossed as 'I'm guessing' or 'I imagine', as there is no indication that she is relying on memory.

(12) Corpus 1 – Charlotte / Nicolas

- 1 CHA nan. (.) j' pense [pas qu' i']=
no. (.) I don't think that they
- 2 NIC [(nan)]
(no)
- 3 CHA =peuvent - qui puissent pa'ce que:=mh: (0.5) ((hawks)) 'fin:
4 j' pense qu' i' adaptent=euh toujours aux: aux
5 goûts euh: (.) occidentaux.
could - can 'cause um (0.5) ((hawks)) well I think that they
always um adapt to to um (.) western tastes.

4.3.2 Je pense (que) vs je trouve (que)

As mentioned earlier, *JP* is predominantly subjective in meaning when the opposition between the 1ps and other grammatical persons is in the foreground. In such cases, the state of affairs is inherently subjective, and *JP* primarily forms a contrastive pair with *JT*. With either construction, the speaker presents their claim as a matter on which no-one has epistemic primacy, therefore acknowledging that others, including the recipient(s), might disagree with them. However, the meanings of the two markers differ in that the speaker must have direct experience of the state of affairs described in a proposition that is in the scope of *JT*, whereas *JP* can scope propositions that are the subject of either direct or non-direct experience (cf. Dendale & Van Bogaert 2007: 72–75). Thus, we get the contrasts in (13)–(14):

- (13) *J'ai vu ce film et je pense / je trouve qu'il est bien fait.*
'I've seen this film and I think / I find it's well made.'
- (14) *Je n'ai pas encore vu ce film, mais je pense / #je trouve qu'il doit quand même être bien fait.*
'I haven't seen this film yet, but I think / #I find it's probably well made.'

As a weak epistemic predicate, *JP* expresses a less assertive stance than *JT*. Accordingly, we observe that opinions are not open to debate and challenge with the latter construction to the same extent as with the former. In (15), the two participants are discussing the concept of 'traditional cuisine', and more specifically its meaning:

(15) Corpus 1 – Nicolas / Paul

- 1 NIC ouais cuisine traditionnelle () - c'est un peu un nom qui
2 veut un peu tout et rien dire quoi c'est comme euh un peu tout
3 genre - 'fin (0.4) dès qu'on parle de tradition j' trouve
4 qu' pour moi ça veut un peu rien dire..
yeah traditional cuisine () - it's a name that means like
everything and nothing it's a bit like everything like - well
(0.4) as soon as we speak about tradition I think that for me
it doesn't mean much I..

In line 3, *JT* is used to mark the speaker's opinion, which is stated with a certain degree of confidence in this case: in line 3, *dès que* ('as soon as') followed by the aspectually non-perfective present tense signals a repetitive action, namely every time Nicolas has heard of or discussed the topic under discussion, i.e. a form of direct experience. Nicolas does not employ any markers of tentativeness in stating his opinion; on the contrary, the *JT*-marked proposition restates – and thus serves to underline – what he has said earlier in the same turn in slightly different words. His subjective stance is further reinforced by the use of *pour moi* 'for me' (l. 4).

By comparison, *JP* is more tentative, which is probably attributable to the fact that semantically it conveys epistemic mitigation. Consider (16), where Adeline gives her opinion about a television show, *Young Sheldon*, using *JP* twice:

(16) Corpus 2 – Adeline / Bruno

- 1 BRU ah ouais c'est bien ça?
oh yeah is it good?
- 2 ADE .h (l.1) ça aurait pu êt` vachement pire.
it could've been far worse.
- 3 BRU ouais un spin-off quoi.
yeah like a spin-off.
- 4 (0.8)
- 5 ADE ça aurait pu ouais. [ça aurait pu êt`]=
it could have yeah. it could've been
- 6 BRU [beh ouais]
well yeah
- 7 ADE =vachement pire. j` m'attendais à pire que ça. franchement le:
8 le:: .h les: acteurs sont plutôt bien choisis.
far worse. I expected something far worse. to be honest the the the actors have been chosen well
- 9 BRU mouais
yeah
- 10 (1.0)
- 11 ADE c:'est p[↑]as euh:: phénomén[↑]al mais genre en série du mi[↑]di
12 euh:: t` sais pour manger [c'est rigolo. (.) j` pense]=
it's not um phenomenal but like during lunch time um you know while you're eating it's funny. I think
- 13 BRU [ça passe (ouais) (.) ouais]
it's alright (yeah) yeah
- 14 ADE =c'est euh aussi rigolo qu'un how I met your - 'fin (.) nan
15 c'est plus rigolo qu'un how I met your mother,
16 [c'est plus ri-]=
it's um as funny as a how I met your - well no it's funnier than a how I met your mother,
- 17 BRU [ah ouais quand même]
oh yeah really
- 18 ADE =.hhh j` **pense que** c'est l'équivalent de euh modern fami[↑]ly ou
19 des trucs comme ça
I think that it's the equivalent of um modern family or stuff like that

Throughout the excerpt, Adeline is hesitant when describing the show. First, her answer (l. 2) does not immediately follow Bruno's question (l.1). It is prefaced by a short inbreath, acknowledging the first pair-part of an adjacency pair to which she has to provide the second pair-part, as well as a 1.1 second silence which allows her to gain time before answering. She expresses reservation by repeating or paraphrasing herself, and by using downtoners. These features reflect Adeline's unfolding thought process, as she appears to be working out what she thinks about the show on the fly. This is evidenced further on by the introduction of two shows – *How I Met Your Mother* and *Modern Family* – as two elements of comparison she has just thought of. In ll. 11–12, she describes *Young Sheldon* as being reasonably funny and immediately produces another turn-constructional unit (TCU) where she compares it to *How I met your Mother* (l. 15). However, she self-interrupts before the end of her TCU and initiates self-repair, upgrading the show from *aussi rigolo* to *plus rigolo*, before finding a similarly funny show (*Modern Family*, l. 18). Throughout her description of the show, two utterances are modified by *JP*, one instance of which occurs in clause-initial position with the complementizer *que* (l. 18), and the other in clause-initial position (l. 12). The preference for *JP* instead of *JT* could be explained by the speaker's wish not to appear too assertive or conclusive, leaving the door open for revisions. Furthermore, in this excerpt *JP* arguably also assumes discourse-organizational functions, introducing new steps in the speaker's turn and punctuating her process of reflection.

The semantic difference between *JP* and *JT* can therefore be accounted for in epistemic terms: because *JP* conveys epistemic meaning even when it is used primarily as a marker of subjectivity, the force of its host utterance is toned down; in contrast, *JT* is not used to mitigate the speaker's commitment to the proposition, but to reinforce a subjective stance, while acknowledging that other stances are possible. In cooperative interaction, we would therefore expect *JP* to be preferred to *JT* when participants are discussing matters that are more open to challenge, and that is indeed what we mostly observe in our data (cf. Angot In Progress).

The contrastive approach has helped us highlight and refine the meaning of *JP*, which would be more difficult to capture if investigated on its own. We now turn to the different functions that the marker assumes in interaction.

5. Interactional functions of *je pense (que)*

In this section, our interest shifts to the way in which the coded meaning of *JP* allows the construction to be used by speakers to achieve a range of discourse-related functions. In particular, we discuss two functions that have thus far been granted limited attention: an interpersonal, politeness-driven function in which *JP* is used to soften a potential or actual face threat (Brown & Levinson 1987; cf. § 5.1); and various discourse-organizational functions that serve to achieve conversational goals (cf. section § 5.2).

5.1 *Je pense (que)* as a face-saving device

In many cases, face needs appear to motivate an instance of *JP*, as a strategy to maintain the social relationship between the participants. Uses of *JP* may be oriented towards saving either the speaker's face (cf. § 5.1.1) or the hearer's face (cf. section § 5.1.2), either by softening a (potential) threat or by weakening a claim to truth and thus avoiding presenting a statement as fact in case the speaker is proven wrong. We show that while the epistemic function of the marker is linked to the protection of the speaker's face, the subjective function is mostly linked to the hearer's face.

5.1.1 *Je pense (que)* as a redressive action for the speaker's face

When used with objective facts, i.e. when the epistemic function is foregrounded, *JP* can be interpreted as a device that protects the speaker's own face by acknowledging that they may be mistaken, thereby mitigating any loss of face effected by a potential subsequent correction. This is illustrated by Example (9), in § 4.2 above, and by (17) below. In (9), Bruno talks about a former colleague of his, and the proposition expressed by the host utterance (*il arrive bientôt à la quarantaine*) is toned down by *JP*. In addition to reducing the speaker's commitment towards their proposition, *JP* also acts to protect Bruno's face in case he is proven wrong about his colleague's age.

In (17), the marker is used with a similar function. Céline and Bruno are talking about the refurbishment of an old swimming pool in some mutual friends' garden. The context provides evidence that Bruno has never participated in the construction or refurbishment of a swimming pool. In line 5, he states that he would probably dislike it. By marking his statement with *JP*, he presents it as being based on a guess, thus protecting his own positive face in case it should turn out that he is wrong.

(17) Corpus 2 – Bruno / Céline

- 1 CEL c'est ça (.) boh après c'est sympa, (.) moi ça m' dérangerait
 2 [(pas)]
that's it (.) well actually it's nice, (.) I wouldn't mind
 3 BRU [mouais] (mo-) ouais [mais euh c'est pas le]:=
yeah (y-) yeah but um it's not the
 4 CEL [c'est comme les déménagements.]
it's like moving house.
 5 BRU =c'est pas l' bricolage que j'aime j' **pense**
it's not the kind of DIY I like I think.
 6 (0.3)
 7 CEL de nettoyer une piscine?
to clean up a swimming pool?
 8 BRU beh=euh ouais...
well um yeah

5.1.2 Je pense (que) as a redressive action for the hearer's face

As explained in § 4.2, when used with inherently subjective predicates (therefore supposing that no-one has epistemic primacy), *JP* does not so much seek to weaken the claim to truth as to highlight the subjectivity of the speaker's opinion, acknowledging the existence of alternative opinions. *JP* can thus be a strategy for speakers to avoid imposing their opinion. It is therefore unsurprising to find it in argumentative sequences, where speakers support their position while simultaneously showing openness, or where they display an alternative opinion as a second action in response to a first action. This is exemplified in (18), where *JP* is used as a device to mitigate a disaffiliative action (Lindström & Sorjonen 2013).

(18) Corpus 1 – Charlotte / Nicolas

- 1 NIC .h moi j' me re- j' remarque des fois j' sais pas s' tu fais
 2 ça mais .h fait je j'ai tendance à manger trop vite et du
 3 coup sans euh fait j' suis pas euh (0.3) en fait j' suis pas
 4 du tout dans l' truc besoin biologique mais des fois
 5 (t' sais;t' es) genre tu manges (0.4) là t' as fini ton
 6 plat °(et tu t' dis euh)° en fait j' (m'en) rappelle p'us du
 7 goût [fqu' ça avait .h du]=
I my- I notice sometimes I don't know if you do it but in fact
I - I tend to eat very fast and so without um in fact it's not
um (0.3) in fact it's not my thing at all biological need but
sometimes (you know;you're) like you eat (0.4) you're done
eating °(and you're telling yourself um)° in fact I don't
remember what it tasted like
 8 CHA [fhh ah ah ahf]
 9 NIC =coup ça fait un peu genre le mec qu' a bouffé vraiment
 10 [pour euh: .h sa]tisfaire=
so it's like the guy who really stuffed himself with food to
um satisfy
 11 CHA [((coughs)) ouais]
yeah
 12 NIC un besoin qui qu' a même pas fait attention quoi. (0.4) mais

- 13 euh j' sais pas ɛ'fin
a biological need who - who weren't even careful/who didn't
even paid attention. (0.4) but um I don't know well
- 14 CHA beh ça c'est quand on mange seul souvent ça fait ça.
well it happens when we eat alone it often happens.
- 15 NIC oui, [(bah ouais) ()]
yes, well yeah
- 16 CHA [pa'ce que - du coup quand t' es] avec des amis tu -
17 comme tu parles tu prends un peu plus le temps de:=mh:
'cause - so when you're with friends you -
'cause you speak you take a bit more time to um
18 (0.6)
- 19 NIC mouai:s, beh j' sais pas ɛmoi j- justement j' s'rai:s (.) plus
20 concentrée sur la conversation et .h b- pas faire gaffe à trop
21 c' que j' mange si j' fais pas 'fin si j' me pa- concentre pas
22 d'ssus à un moment [donné quoi]=
yeah, well I don't know for me a- actually I'd be (.) more
focused on the conversation and .h b- not be too careful with
what I'm eating if I don't well if I fo- don't focus on it at
some point then
- 23 CHA [ouais:]
yeah
- 24 NIC =.h ['fin c'est:]=
.h well it's
- 25 CHA [c'est vrai.]
that's right.
- 26 NIC c'est plus une étourderie j' **pense** plus qu(h)oi (après) que:=
it's more an absent-mindedness I think more then (but) than
- 27 CHA =et t' as remarqué qu' quand on est à table avec des potes on
28 finit toujours par parler d' bouffe?
and have you noticed that when eating with friends we always
end up talking about food?

At the beginning of the excerpt, Nicolas tells Charlotte in an extended turn (ll. 1–7, 9–10 and 12–13) that he is in the habit of eating fast, to the point that he quickly forgets the taste of what he just ate. In ll. 9–10 and 12, he is critical of himself because he feels that he eats without paying any attention to the food, just satisfying a biological need. In the next turn (ll. 14, 16–17), Charlotte offers a candidate explanation, suggesting that this is more likely to happen when people eat alone. Charlotte stops talking before producing a syntactically complete sentence (l.17) and after a 0.6 second pause (l. 18), Nicolas takes the floor. His turn is initiated by a slightly lengthened *mouais* ('yeah') and followed by *beh je sais pas* ('well I don't know'). Pekarek Doehler (2016) shows how the epistemic disclaimer *je sais pas* can be used to foreshadow a dispreferred action, and this is its role here, as Nicolas' turn in line 19–22 displays a different stance from Charlotte's. This is further supported by the use of the stressed pronoun *moi* ('as far as I am concerned', lit. 'me') as well as the adverb *justement* ('actually', lit. 'precisely'), both of which serve to mark a divergence between the two participants' viewpoints. In Nicolas' view, it is easier to pay attention to food when he eats alone, whereas he would be distracted while

eating with friends. In line 26, he adds a self-criticism (*c'est plus une étourderie*) that can be interpreted as an alternative account of his habit. This account is disaffiliative in that Nicolas is not endorsing Charlotte's viewpoint, from which he implicitly distances himself by his use of the comparative form *plus* ('more') in line 26. This utterance is marked by *JP*, which may be interpreted as mitigating Nicolas' disaffiliative action, by highlighting the assessment as subjective.

JP can also be used to mitigate a proposition in order to avoid a potential or actual conflict. As mentioned in Section 3, the data collection method affected certain aspects of the recordings, resulting, saliently, in more abrupt topic shifts and in more explicit and unilateral closings. Such actions can be seen as potentially threatening to the addressee's face, and hence as requiring mitigation. The use of *JP* in such contexts is a way for speakers to soften their actions, by seeking approval from their interlocutor(s) and not imposing themselves as leading the conversation. This is illustrated in (19) where Adeline invites Bruno, her interlocutor, to change the subject under discussion:

(19) Corpus 2 – Adeline / Bruno

- 1 ADE j- j' com|prends.
I- I understand.
2 (1.2)
3 ADE on peut p't-êt' switcher là [j' pense,]
we can perhaps switch now I think,
4 BRU [.hh] (0.3) ouai:s
.hh (0.3) yeah

By adding *JP* at the end of her turn, Adeline mitigates her utterance and signals that she is merely suggesting a topic shift rather than imposing it, showing that she is not leading the conversation. Here, *JP* occurs in collocation with the deontic modal verb *pouvoir* ('can'), in line with the idea that "mitigation affecting deontic modality reduces addressee's obligations" (Caffi 1999: 882). Note that this example also involves the use of the adverb *peut-être* ('perhaps') as a mitigating device.

JP also occurs in topic-shifting or conversation-ending sequences without being associated with deontic modality. In (20), due to the fixed time limit on the interactions in our corpus (cf. § 3 above), Céline's turn functions as a closing-implicative turn (Schegloff & Sacks 1973), and it is indeed the last turn of this conversation as Bruno does not provide any response. Although the context provides evidence that Céline genuinely does not know whether she and her interlocutor are in the position of ending the conversation (e.g. they have not yet looked at the time that has gone by), *JP* may be interpreted as a mitigator with similar functions to (19).

(20) Corpus 2 – Bruno / Céline

- 1 CEL .hh j' pense que c'est bon,
.hh I think that it's enough,

More generally, *JP* as a strategy of non-imposition on the addressee's face can co-occur with any form of action, i.e. not only in connection with topic shifts and closings. In such cases, it is always used in collocation with deontic modality in our database. Three times, it co-occurs with the deontic modal verb *falloir* ('to be necessary/must'), as exemplified in (21):

(21) Corpus 2 – Adeline / Céline / Quentin

1 CEL .hhh ben:::==mh: .tsk .h moi j' suis d'accord. et j' **pense**
 2 **qu'**i' faut consacrer plus de temps à l'éducation artistique à
 3 l'école,
 . hhh well mh .tsk .h as far as I'm concerned I agree. and I
 think that more time must be devoted to artistic education
 at school,

Interestingly, *falloir* only occurs when the subjective function is foregrounded. Thus, *JP* is used to mitigate not only the deontic modality, but the speaker's utterance as a whole, by acknowledging the existence of other opinions. It is also possible to consider subjective statements involving a deontic modal such as *falloir* to be more threatening and therefore requiring greater linguistic precaution. Note that, as the above instance of *JP* is used to initiate a new TCU within a current speaker's turn, it may also be interpreted as having a discourse-organizational function, namely that of introducing a new step in the discourse. Thus, this excerpt provides evidence that *JP* is a polyfunctional marker whose interactional functions result from several context-specific factors in any given case.

To sum up, *JP* can be a means to mitigate potential or actual face-threatening acts, thus reducing the risk of conflict between participants. Put another way, as a mitigation device it marks the content of the utterance as potentially threatening, while simultaneously reducing the risk of this threat. It can act as a strategy of non-imposition on either the speaker's or the addressee's face.

5.2 Discourse-organizational functions

In this section, we consider the discourse-organizational functions of *JP*, showing how the construction contributes to discourse coherence, particularly by marking transitions between discourse units.

5.2.1 *Indicating a transition in the speaker's turn-in-progress*

JP may occur in the middle of a speaker's turn for organizational purposes. In (22), it indicates a transition in a narrative. Before the beginning of the excerpt, Bruno and Quentin were talking about social media, and more specifically Facebook.

(22) Corpus 1 – Quentin / Bruno

- 1 QUE j'ai maté une vidéo hier qui durait dix minutes un truc comme
 2 ça, .h et=euh: c'est un type qui prend une euh::=mh:: .tsk une
 3 vieille clé à pipe
*I watched a video yesterday that lasted ten minutes something
 like that, and um it's a guy who takes a um .tsk an old
 socketwrench*
- 4 (29'')
- 5 QUE et i` t'en sors une espèce de:: de katana qui fait genre euh
 6 soixante centimètres de long, (0.2) mais qu` est ultra propre
 7 qu` est mais qu` est nickel quoi qu` est tranchant rasoir et
 8 tout euh .h qu` est superbe (0.4) euh i` t` fait - i` fait une
 9 trempe [sélective euh::]
*and he manages to make a kind of katana that's like um sixty
 centimeters long, (0.2) but that's really neat that's but
 that's great then that's sharp-edged and all um .h that's
 superb (0.4) um he does - he does a selective quenching um*
- 10 BRU [beh en même temps i` suf]fit de::
well at the same time you just have to
- 11 QUE i` t` fait une trempe sélecti:ve et tout euh `fin: voilà quoi
 12 [et euh:]
he does a selective quenching and all um well so yeah and um
- 13 BRU [i` suffit d` for]ger hein
you just have to forge
- 14 QUE et [c'est]=
and it's
- 15 BRU [ɛmhɛ]
- 16 QUE =c'est super propre, c'est - c'est impeccable et euh du coup
 17 bah j` pense que .h si j` m'étais pas euh inscrit euh:=au
 18 groupe de:: passion couteau sur facebook eh beh j'aurais
 19 jamais vu cette vidéo alors que je l'ai trouvée géniale quoi
*it's super neat, it's - it's impeccable and um and so well I
 think that .h if I hadn't um signed up um to the group "knife
 lovers" on facebook well I wouldn't have seen this video
 though I found it awesome*

In ll. 1–3, Quentin's first TCU projects a narrative, which lasts approximately one minute, during which Bruno's turns consist of helping Quentin identifying the name of a hand tool and providing receipt tokens.³ He describes the content of a video showing a man forging a knife from a hand tool. Up until ll. 16–19, which can be interpreted as the upshot of the story (Mandelbaum 2013: 495), there is no apparent connection between Quentin's story and the previous conversational topic – social media and Facebook, yet no explicit topical change has been effected. The topical coherence of Quentin's story can only be understood in the upshot. It is introduced in ll. 16–17 by the collocation of markers *et euh du coup bah* ('and um and so well'), which signal a new step in Quentin's turn. The telling of the story

3. For reasons of space, about thirty seconds of Quentin's story have been omitted (cf. l. 4). During this time, he is trying to recover the name of a hand tool, helped by Bruno who offers several suggestions.

was designed to establish a basis for further talk: Quentin aimed to highlight the usefulness of Facebook, without which he might not have discovered the video. Following the collocation of markers, *JP* marks the speaker's evaluation of the narrative (Labov & Waletzky 1967), which is clearly set apart by a break in the progressivity of the turn: *JP* is immediately followed by an audible inbreath. The interactional work achieved by *JP* is thus to mark the end of the story, allowing the resumption of the previous conversational topic. Because the turn-taking system is suspended while speakers are telling stories, the use of *JP* to mark the final evaluation segment can also be interpreted as contributing to the signposting of a return to normal turn-taking.

5.2.2 *Projecting a preferred answer: Je pense (que) in question-answer formats*

The two excerpts below show how prefacing a response by *JP* can be interpreted as a way to signal a preferred next action, i.e. one that conforms to the interlocutor's expectations (Pomerantz & Heritage 2013). Our focus of interest here is the second pair-parts of question-answer adjacency pairs (Sacks, Schegloff & Jefferson 1974), after a current speaker chooses to designate someone as the next speaker by asking for their opinion. In such contexts, the request for an opinion may be responsible for triggering the target construction, in that the use of a marker of opinion becomes relevant. Arguably, *JP* frames the subsequent turn as the second pair-part of the adjacency pair, and furthermore as a response that aligns with the preference of the first pair-part (Stivers 2008). The use of *JP* can be seen as a signal that the first pair-part has been interpreted as a request for an opinion. Thus, its function is simultaneously retrospective, showing how the previous turn has been interpreted, and prospective, signalling the direction in which the conversation is going. As such, it plays a bridging role between the two parts of the adjacency pair.

(23) below illustrates this organizational function. In ll. 1–2, the current speaker Quentin selects Céline as the next speaker by asking her opinion about a particular matter. The request for a personal opinion, as well as the verb chosen by Quentin (*considérer* 'consider'), may both be seen as triggering the use of *JP* in the response:

(23) Corpus 2 – Adeline / Céline / Quentin

- 1 QUE est-ce que tu considères que l' jardinage c'est plutôt
 2 [artistique, ou cultur]el.=
do you consider gardening to be more artistic, or cultural.
 3 ADE [ɛh eh eh (mh mh) ɛ]
 4 CEL =.tsk j' **pense que** c'est culturel. et j' trouv'rais ça trop
 5 bien que: [dans toutes]=
.tsk I think that it's cultural. and I would find it great if
in every
 6 QUE [.hh]
 7 CEL =les écoles y ait un p'tit potager:, ...
schools there was a little garden, ...

The concept of gardening is new to the discussion at this point and it is not clear why Quentin introduces it. Whatever his motives, there are several possible ways in which Céline could have designed her response. However, she aligns with Quentin's preference, choosing one of the two categories provided by him (in her opinion, gardening is a cultural matter) and volunteering further information (introduced in a second, truncated, TCU by *je trouverais*). The *JP*-preface may be interpreted as a clue to the aligning nature of the upcoming action.

In (24), the use of *JP* similarly projects the fundamentally aligning nature of Bruno's response, despite his exhibiting some difficulty in constructing that response:

(24) Corpus 2 – Adeline / Bruno / Céline / Quentin

- 1 ADE =quelle est ton opinion euh bruno.
what's your opinion um Bruno.
- 2 BRU bah euh:=f: (0.5) [j' (.)]=
well um (0.5) I (.)
- 3 ADE ((laughing silently)) [f.hɛ]
- 4 BRU =**pense** qu'i' faut do- laisser l' choix, j' sais pas en même
5 temps est-ce que faut laisser plus de place à l'éducation
6 artistique (0.2) et culturelle à l'école (.) .hh=
think that we should gi- leave the choice, I don't know at the
same time should we leave more space to artistic education
(0.2) and culture to school (.) .hh
- 7 ADE =on est d'accord c'est [l'école primaire hein]
we're on the same page it's primary school right
- 8 BRU [alors oui mais f]aut - ouais, (.)
9 mais euh::: i' faut qu' ce soit euh: .hh faut qu' ce soit
10 voulu,
then yes but we should yeah, (.) but um it
should be a choice,

Bruno's response (beginning in l. 2) is delayed by hesitation markers and a 0.5 second pause. His perceived difficulty in providing an immediate response is supported by other features. In his first TCU (*je pense qu'il faut do- laisser le choix*), Bruno self-initiates repair, replacing what could be heard as the first syllable of the verb *donner* ('give') by the verb *laisser* ('leave'). This is then followed by the epistemic disclaimer *je sais pas* ('I don't know'), introducing a second TCU where Bruno reformulates the suggested topic in the form of a question. Bruno's turn is thus designed in a way that signals tentativeness, and the use of *JP* can be seen as buying him time to construct a conforming action while pre-empting potential misinterpretation.

In other words, in response to a request for an opinion, *JP*-prefaces function as markers of alignment by explicitly framing the subsequent turn or TCU as constituting the relevant next action, thereby conveying the speaker's cooperative stance.

6. Correlations between the position and the function of *je pense (que)*

As we saw in Table 1, Sect. 3 above, the position of *JP* is highly variable. In this section, we examine whether there are meaningful correlations between the syntactic position of the marker and its pragmatic functions.

First, we observe a strong tendency for *JP* to occur in matrix position, representing 51% of the 66 instances in our database. In comparison, only 23% of the instances take the form of a reduced parenthetical, and 8% of standalone constructions. As a reduced parenthetical, *JP* preferentially occurs in clause-final (11 instances), as opposed to clause-medial (four instances), position.

Secondly, the data show that the marker tends to have discourse-organizational functions when it occurs in matrix position, while it mostly functions as an epistemic hedge when used parenthetically. Notice that this runs counter to what we might expect, if – as is usually assumed – parenthetical *JP* is further along in its development towards “PM-hood” than *JP* in matrix position.

Thompson and Mulac (1991) argue that *that*-deletion is what paves the way for the development of English *I think* and of epistemic constructions in general, into PMs. While in English, *I think* regularly occurs without the complementizer *that* (Thompson & Mulac 1991; Kärkkäinen 2003; Van Bogaert 2011), we find only four cases out of a total of 34 in our database where *JP* in matrix position does not include *que*. This modest number of examples makes it difficult to confirm with any degree of confidence whether *je pense* \emptyset in matrix position differs in meaning and/or function from *je pense que*. What we do observe is that, unlike *je pense que*, *je pense* \emptyset is never used in turn-initial position, i.e. as a strategy to take the floor. This, again, runs counter to what we would expect to find if *je pense que* represents the source construction and *je pense* \emptyset a stage of further grammaticalization. Moreover, all four instances of *je pense* \emptyset function as epistemic hedges, thus behaving similarly to parentheticals. This is exemplified by (21) in § 5.1.2 above.

Ultimately, it appears that *JP* is able to foreground both of its two coded elements of meaning and to fulfil any of its pragmatic functions, in any position and when occurring both with and without the complementizer. This raises the question of what is the most appropriate way to account for its variable syntax. We have already problematized Thompson and Mulac’s “classic” grammaticalization account. In contrast to that account, Heine et al. (2020) argue that parentheticals (or “theticals”, as they prefer to call them), and PMs more generally, arise through a process of “cooptation”. Cooptation takes a “chunk of sentence grammar, such as a word, a phrase, a reduced clause, or a full clause, [and deploys it] for use on the metatextual level of discourse monitoring” (Heine et al. 2020: § 2.2). It is instantaneous, but may be preceded, and/or (more commonly) followed, by the gradual

process of grammaticalization (Heine et al. 2020: § 2.4.2). However, as our analysis has shown, *JP* functions at the metatextual level whether or not it is syntactically parenthetical, and it is thus unclear what the explanatory value of cooptation might be in this case. Instead, we would argue that Hansen's (2008) distinction between content-level and context-level semantics, which is orthogonal to the distinction between propositional and non-propositional meaning, provides a better account of the peculiar syntactic, semantic, and pragmatic properties of *JP*.

7. Conclusion

In this paper, we have posited two components of coded meaning that coexist in any given use of the micro-construction *JP*, although one is typically foregrounded at the expense of the other. We have shown that, by virtue of these two coded elements of meaning, *JP* is able to fulfil interactional functions typically associated with the category of PMs. In one subset of its conversational uses, *JP* thus serves as a politeness marker, while in another subset it assumes discourse-organizational functions.

Unlike many other PMs, making a clear distinction between instances where *JP* functions as such *vs* those where it does not is not at all straightforward. This is because, as pointed out in § 4.2, *JP* inherently has context-level semantics, no matter what its use in any given instance. Further, it is unclear that the construction has undergone significant semantic bleaching of the kind that is usually seen as characteristic of PMs, inasmuch as its two elements of coded meaning are precisely what allows it to function successfully as a marker of politeness and to fulfil the specific discourse-organizational functions identified in this paper. Lastly, as we saw in sect. 6, there do not appear to be clear semantic and/or functional differences between uses of *JP* in matrix *vs* parenthetical position, and with or without a following complementizer. Although we do find occurrences where *JP* fulfills functions that are predominantly interactional, rather than propositional, and where its omission would not result in a change in the meaning of the host utterance, it retains a degree of propositional meaning in all its uses and seems to be relatively semantically constrained. In this respect, it is unlike prototypical pragmatic markers, e.g. *eh bien* ('well'), *alors* ('then'), or *donc* ('so') (cf. Hansen 1998), that are standardly defined as having no propositional meaning.

That said, like prototypical PMs, *JP* is a multifunctional expression, whose precise interactional function depends on its context of occurrence. Like PMs, its use reflects particular characteristics of conversation (cf. Beeching 2016). First, the mobility of *JP* with respect to the clause under its scope displays the spontaneous, moment-by-moment nature of conversation. As a syntactically free-standing device, *JP* also allows the speaker to carry out self-repair, hedging initial overly

strong assertions. Secondly, the uses of the construction reflects the interactional nature of conversation, insofar as it can serve the joint negotiation of conversational organization. Thirdly, by fulfilling politeness-driven functions and thereby playing a role in maintaining social relationships, *JP* reflects the normatively polite nature of conversation.

These observations seem to us to be of relevance to recent proposals concerning the dual nature of grammars. Thus, Heine et al. (2015, 2020) draw a distinction between “sentence grammar” and “thetical grammar”, the latter of which governs the behavior of PMs, fully or partially productive parenthetical expressions, interjections, and formulaic expressions that fulfil interpersonal functions, such as greetings.⁴ Importantly, they adduce evidence that constructions belonging to the two different types of grammar are processed differently by the human brain. As we have shown in this paper, *JP* instantiates a type of construction that in fact appears to occupy a middle ground between the two types of grammar. We therefore suggest that carrying out psycholinguistic experiments to test the processing of different uses of this expression (and/or cognate ones in other languages) ought to be of particular interest for the dual-grammar hypothesis.

Acknowledgements

The authors would like to thank the editors for having invited us to contribute and for the work they have done in putting together this volume. We also thank Bernd Heine and the two anonymous reviewers for their helpful comments on a prefinal version of this paper. Needless to say, they are not to blame for any remaining shortcomings.

References

- Andersen, Hanne Leth. 2007. Marqueurs discursifs propositionnels. *Langue française* 154(2): 13–28. <https://doi.org/10.3917/lf.154.0013>
- Angot, Juliette. In progress. *Je pense, je crois and je trouve* in French Conversation. PhD dissertation, University of Manchester.
- Apothélos, Denis. 2003. La rection dite ‘faible’: Grammaticalisation ou différentiel de grammaticalité? *Verbum* 25(3): 241–262.
- Avanzi, Mathieu & Glikman, Julia (eds). 2009. Entre rection et incidence: Des constructions verbales atypiques? Études sur *je crois, je pense* et autres parenthétiques. *Linx* 61. <https://doi.org/10.4000/linx.1322>
- Beeching, Kate. 2016. *Pragmatic Markers in British English. Meaning in Social Interaction*. Cambridge: CUP. <https://doi.org/10.1017/CBO9781139507110>

4. Haselow (2016) makes an analogous distinction between micro-grammar and macro-grammar.

- Blanche-Benveniste, Claire. 1989. Constructions verbales 'en incise' et rection faible des verbes. *Recherches sur le français parlé* 9: 53–73.
- Brown, Penelope & Levinson, Stephen C. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511813085>
- Caffi, Claudia. 1999. On mitigation. *Journal of Pragmatics* 31: 881–909. [https://doi.org/10.1016/S0378-2166\(98\)00098-8](https://doi.org/10.1016/S0378-2166(98)00098-8)
- Cosnier, Jacques & Kerbrat-Orecchioni, Catherine. 1987. *Décrire la conversation*. Lyon: Presses Universitaires de Lyon.
- Dendale, Patrick & Van Bogaert, Julie. 2007. A semantic description of French lexical evidential markers and the classification of evidentials. *Rivista di Linguistica* 19(1): 65–89.
- Gazdar, Gerald. 1979. *Pragmatics: Implicature, Presupposition, and Logical Form*. New York NY: Academic Press.
- Gosselin, Laurent. 2018. French expressions of personal opinion: *je crois / pense / trouve / estime / considère que p*. In *Epistemic Modalities and Evidentiality in Cross-Linguistic Perspective*, Zlatka Guentchéva (ed.), 179–195. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110572261-008>
- Grice, Herbert Paul. 1975. Logic and conversation. In *Syntax and Semantics, Vol. 3: Speech Acts*, Peter Cole & Jerry L. Morgan (eds), 41–58. New York NY: Academic Press.
- Hansen, Maj-Britt Mosegaard. 1998. *The Function of Discourse Particles. A Study with Special Reference to Spoken Standard French* [Pragmatics & Beyond New Series 53], Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.53>
- Hansen, Maj-Britt Mosegaard. 2008. *Particles at the Semantics / Pragmatics Interface: Synchronic and Diachronic Issues. A Study With Special Reference to the French Phasal Adverbs*. Oxford/Leiden: Elsevier/Brill.
- Hansen, Maj-Britt Mosegaard. 2020. The role of (historical) pragmatics in the use of response particles: The case of French. *Functions of Language* 27(3): 307–339. <https://doi.org/10.1075/fol.18024.mos>
- Haselow, Alexander. 2016. A processual view on grammar: Macrogrammar and the final field in spoken syntax. *Language Sciences* 54: 77–101. <https://doi.org/10.1016/j.langsci.2015.12.001>
- Heine, Bernd, Kuteva, Tania, Kaltenböck, Gunther & Long, Haiping. 2015. On some correlations between grammar and brain lateralization. *Oxford Handbooks Online*. Oxford: OUP. <https://doi.org/10.1093/oxfordhb/9780199935345.013.16>
- Heine, Bernd, Kaltenböck, Gunther, Kuteva, Tania & Long, Haiping. Forthcoming. *The Rise of Discourse Markers*. Cambridge: CUP.
- Hilpert, Martin. 2014. *Construction Grammar and Its Application to English*. Edinburgh: EUP.
- Hoffman, Thomas & Trousdale, Graeme (eds). 2013. *The Oxford Handbook of Construction Grammar*. Oxford: OUP. <https://doi.org/10.1093/oxfordhb/9780195396683.001.0001>
- Hooper, Joan B. 1975. On assertive predicates. In *Syntax and Semantics, Vol. 4*, John P. Kimball (ed.), 91–124. New York NY: Academic Press.
- Jefferson, Gail. 1984. On stepwise transition from talk about a trouble to inappropriately next positioned matters. In *Structures of Social Action: Studies in Conversation Analysis*, J. Maxwell Atkinson & John Heritage (eds), 191–222. Cambridge: CUP.
- Jefferson, Gail. 2004. Glossary of transcript symbols with an introduction. In *Conversation Analysis: Studies from the First Generation* [Pragmatics & Beyond New Series 125], Gene H. Lerner (ed.), 13–31. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.125.02jef>

- Kärkkäinen, Elise. 2003. *Epistemic Stance in English Conversation: A Description of its Interactional Functions, with a Focus on I think* [Pragmatics & Beyond New Series 115]. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.115>
- Labov, William & Waletzky, Joshua. 1967. Narrative Analysis. In *Essays on the Verbal and Visual Arts*, June Helm (ed.), 12–44. Seattle WA: University of Washington Press.
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511813313>
- Levinson, Stephen C. 2000. *Presumptive Meanings. The Theory of Generalized Conversational Implicature*. Boston MA: The MIT Press. <https://doi.org/10.7551/mitpress/5526.001.0001>
- Lindström, Anna & Sorjonen, Marja-Leena. 2013. Affiliation in conversation. In *The Handbook of Conversation Analysis*, Jack Sidnell & Tanya Stivers (eds), 350–369. Chichester: Wiley-Blackwell.
- Mandelbaum, Jenny. 2013. Storytelling in conversation. In *The Handbook of Conversation Analysis*, Jack Sidnell & Tanya Stivers (eds), 492–507. Chichester: Wiley-Blackwell.
- Mullan, Kerry. 2010. *Expressing Opinions in French and Australian English Discourse* [Pragmatics & Beyond New Series 200]. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.200>
- Pekarek Doehler, Simona. 2016. More than an epistemic hedge: French *je sais pas* ‘I don’t know’ as a resource for the sequential organization of turns and actions. *Journal of Pragmatics* 106: 148–162. <https://doi.org/10.1016/j.pragma.2016.06.014>
- Pomerantz, Anita & Heritage, John. (2013). Preference. In *The Handbook of Conversation Analysis*, Jack Sidnell & Tanya Stivers (eds), 210–228. Chichester: Wiley-Blackwell.
- Sacks, Harvey. 1992. *Lectures on Conversation*. Oxford: Blackwell.
- Sacks, Harvey, Schegloff, Emmanuel A. & Jefferson, Gail. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4): 696–735. <https://doi.org/10.1353/lan.1974.0010>
- Schegloff, Emmanuel A. & Sacks, Harvey. 1973. Opening up closings. *Semiotica* 8: 289–327. <https://doi.org/10.1515/semi.1973.8.4.289>
- Schneider, Stefan. 2007. *Reduced Parenthetical Clauses as Mitigators. A Corpus Study of French, Italian and Spanish* [Studies in Corpus Linguistics 27]. Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.27>
- Schneider, Stefan, Glikman, Julia & Avanzi, Mathieu. 2015. *Parenthetical Verbs*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110376142>
- Stivers, Tanya. 2008. Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction* 41(1): 31–57. <https://doi.org/10.1080/08351810701691123>
- Thompson, Sandra A. & Mulac, Anthony. 1991. The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of Pragmatics* 15: 237–251. [https://doi.org/10.1016/0378-2166\(91\)90012-M](https://doi.org/10.1016/0378-2166(91)90012-M)
- Traugott, Elizabeth Closs. 2008. The grammaticalization of *NP of NP* patterns. In *Constructions and Language Change*, Alexander Bergs & Gabriele Diewald (eds), 23–45. Berlin: Mouton de Gruyter.
- Van Bogaert, Julie. 2011. *I think* and other complement-taking mental predicates: A case of and for constructional grammaticalization. *Linguistics* 49(2): 295–332.

Transcription conventions

(Based on Jefferson 2004; ICOR 2013)⁵

[]	overlapping talk
=	latching talk (alternatively, a pair of equal signs is used when a speaker's turn is broken up in the transcript)
(.)	micropause (0.2 second or less)
(0.0)	time intervals over 0.2 second
stress	stress, via pitch and/or amplitude
:	prolongation of a sound (the longer the colon row, the longer the prolongation)
↑	rising pitch
↓	falling pitch
.	TCU-final falling intonation
,	TCU-final slightly rising intonation
?	TCU-final strongly rising intonation
LOUD	loud volume (relative to the surrounding talk)
◦soft◦	soft volume (relative to the surrounding talk)
-	truncation
>fast<	fast talk (relative to the surrounding talk)
<slow>	slow talk (relative to the surrounding talk)
h	outbreath (the more <i>h</i> , the longer the outbreath)
.h	inbreath (the more <i>h</i> , the longer the inbreath)
(h)	plosiveness (can be associated with laughter, crying, breathlessness, etc.)
£	'suppressed laughter'
()	inaudible talk
(would)	unclear talk or speaker
((mouthings))	transcriber's descriptions
'	standard elision
'	non-standard elision (some sounds are not pronounced)
.tsk	mouth opening sound

5. <http://icar.univ-lyon2.fr/projets/corinte/documents/2013_Conv_ICOR_250313.pdf>

Discourse markers and brain lateralization

Evidence for dual language processing from neurological disorders

Alexander Haselow
University of Münster

This chapter contributes to recent lines of research proposing right-hemisphere dominance of discourse-related language tasks using neurolinguistic data on the incidence of discourse markers in the speech of unilaterally brain-damaged speakers (left- and right-hemispheric damage) compared to data produced by control (unimpaired) speakers. From a more general language-theoretic perspective these data will serve as the basis for the exploration of processing differences between two domains of language structure, one encompassing grammar and semantics, the other one the organization of discourse and interaction, which has important implications for linguistic modeling.

Keywords: discourse markers, pragmatics of communication, discourse structure, neurological impairment

1. Introduction

This paper contributes to ongoing research on the processing of discourse markers (DMs) in the brain, which provides important clues to language organization and the categorization of DMs in linguistic theory. DMs represent one of the showcase examples of the intersection between grammar and discourse-pragmatics and raise questions on where to draw the boundary between grammar and pragmatics, given that DMs are conventionalized linguistic forms with a language-structuring function exhibiting a certain degree of pragmaticity (Diewald 2013) and constraints on their syntactic position and internal ordering (Cuenca & Crible 2019; Haselow 2019), which lends them the character of grammatical items, but operate outside the clause on the discourse level and are thus not subjected to the constraints of clause grammar. Similar to other forms that are clearly outside sentence grammar, such as interjections or vocatives, DMs operate in a domain of language that may

be called “extra-clausal” (Dik 1997), serving functions on a “higher”, “meta” or “macro” level of language structure (e.g. van Dijk 1980; Maschler 2009; Deloufeu 2017; Heine 2019).

This chapter seeks to contribute to recent lines of research proposing a dualistic organization of language and linguistic cognition (e.g. Prat et al. 2007; Pawley 2009; Van Lancker Sidtis 2009; Heine, Kuteva & Kaltenböck 2014; Haselow 2017; Haselow & Kaltenböck 2020), distinguishing between two domains of language structure, a microlevel (=clause-internal) and a macrolevel (=discourse-related). Neurolinguistic data on the incidence of DMs in the speech of unilaterally brain-damaged speakers (left- or right-hemispheric damage, LHD and RHD respectively) compared to data produced by control (unimpaired) speakers will be used to argue for an overlap of structural differences with neurocognitive processing differences between the two levels of language structure in the brain, which carries important implications for linguistic modeling.

Studies of spontaneous speech produced by neurologically impaired speakers have become important cues in exploring the organization of language in the past decades, covering topics such as patterns of impairment or preservation of abilities in the key domains of language competence, such as syntax, semantics, and pragmatics. Most of them explore whether deficits deriving from hemispheric damage in one area of language competence carry over to deficits in another. The present study follows this line of research, seeking to deepen our understanding of the ways in which DM processing is associated with specific hemispheric activity and thus integrated in a set of other linguistic abilities associated with the respective brain structures.

One of the most well-established findings in experimental neurolinguistic research is the basic hemispheric distinction between grammatical, i.e. syntactic-semantic, and pragmatic, discourse-related processing, both of which are associated with different brain structures and appear to follow different principles (e.g. Code 1996, 2005; Beeman & Chiarello 1998). While the left hemisphere (LH) plays a major role for grammatical processing, which involves syntactic structures, semantic relations, and the “machinery” (Heine 2019) dealing with sentence structures based on morphosyntactic and semantic relationships involving the well-known Broca and Wernicke areas (especially Brodman areas 44 and 45) (e.g. Friederici 2004; Friederici et al. 2006), pragmatic processing, i.e. the processing of discourse organization and interaction-related structural phenomena, is mainly associated with the right hemisphere (RH) (e.g. Myers 2001; Lehman Blake 2006; Brady, Armstrong & Mackenzie 2006; Sherratt & Bryan 2012). However, most of these studies are based on relatively broad sets of language-related competences and a wide range of linguistic forms subsumed under umbrella terms such as “pragmatic

abilities” (Code 2005), “discourse production” (Sherratt & Bryan 2012) or simply “discourse representation” (Prat, Long & Baynes 2007). All of these aspects refer to abilities such as preserving the linear order of events, paying attention to the discourse needs of the addressee, respecting turn-taking or maintaining coherence between discourse segments. This chapter narrows the focus on one class of “discourse-related” elements, namely DMs, and thus on a specific class of elements that serve the structuration of ongoing spoken discourse and interaction management (e.g. Pons 2018; Maschler 2009; Crible 2017) and the creation of macrostructures in language (Haselow 2019).

If the processing of discourse-related elements such as DMs is largely independent of the processing of grammatical relations this would provide further evidence for the assumption that linguistic cognition is not organized in a monolithic way, but exhibiting a dualistic organization, with two levels of language structure, a microlevel on which more local, clause-internal relations between single constituents are processed, and a macrolevel on which more global relations outside single clauses, e.g. between segments of discourse, are processed. Existing findings suggest that DMs are associated with RH processing as their use is not related to sentence-internal structures, but refers to discourse processing and interaction management, and they are part of formulaic language, which mainly involves RH processing (Van Lancker Sidtis 2004, 2009; Van Lancker Sidtis & Sidtis 2018). DMs are therefore likely to be underrepresented in the speech of patients with RH impairment. Examination of the incidence of DMs in spontaneous speech produced by neurologically impaired speakers thus provides valuable evidence for our understanding of language organization and processing and puts the relatively broad classification of RH disorders and hemispheric functions to a test by focusing on one particular group of discourse-related elements. This, in turn, allows for a more fine-grained analysis of the correlation between language organization and language processing, based on the neural architecture of the brain.

This chapter is structured as follows. Section 2 provides a brief outline of the *microlevel-macrolevel* model of language structure, discusses recent findings on the hemispheric differences in the processing of linguistic structures, and provides a brief characterization of the language-related abilities of LH-damaged aphasic speakers. Section 3 presents a brief characterization and a classification schema of DMs, Section 4 discusses the design of the study and presents the empirical findings deriving from the study of DM usage in the speech of the three speaker groups (LHD, RHD, control group). The results of the study are discussed and interpreted in Section 5, the conclusions deriving from the study are discussed in Section 6.

2. Microstructures, macrostructures and dualistic processing

2.1 Micro- and macrostructures in language

Language does not only provide principles for combining words into structured syntactic units in which these words become “constituents” and are, as such, formally linked to other constituents within the same structural unit (e.g. phrases, clauses, sentences), but also linguistic means and principles of use for the creation of structures beyond such units (e.g. interjections, vocatives, backchannels, or comment clauses). The structuring principles “outside” the clause or sentence on the interactional or discourse level, and their expression, are typically set apart from (written-biased) mainstream notions of ‘grammar’, which traditionally concern the structure of the (well-formed) sentence as the largest unit to which rules of grammar apply. In order to integrate these “higher-level” aspects of language structure (“higher” in the sense that they operate beyond or “above” the sentence level) more recent functional approaches to grammar operate with a more or less explicit dualistic conception of language structure by distinguishing between a domain referring to clause or sentence grammar, which basically encompasses clausal constituents and the syntactic machinery required for their linearization into clausal/sentential configurations in a propositional format, and one comprising linguistic material that is not part of this machinery, operating “outside” clausal structures, and that is best interpreted “in the context of a grammatical model that takes levels of linguistic organization higher than the clause into account” (Dik [1997: 407] for “extra-clausal constituents”). Table 1 provides some examples of the basic distinctions along these lines proposed in relevant academic grammars and in more recent functional approaches.

Table 1. Examples for implicitly dualistic approaches to grammar

	“Core”/“Basic-level” grammar	“Peripheral”/“Higher-level” grammar
Blanche-Benveniste et al. (1990)	Microsyntaxe	Macrosyntaxe
Dik (1997)	Clausal constituents	Extra-clausal constituents
Biber et al. (1999)	Clausal units	Non-clausal units
Huddleston & Pullum (2002)	Clause structure	Supplements
Debaisieux (2007)	Dépendence grammaticale	Dépendence macrosyntaxique
Kaltenböck et al. (2011)	Sentence Grammar	Thetical Grammar
Haselow (2017)	Microgrammar	Macrogrammar
Heine, Kuteva & Long (2020)	Type 1 processing	Type 2 processing

With the rise of such “expanded” conceptualizations of grammar many expressions in language that have traditionally been excluded from grammar and analyzed under pragmatic frameworks, such as discourse markers (e.g. Schiffrin 1987; Schourup 1999; Aijmer 2002), interjections (e.g. Ameka 1992), or general extenders (e.g. Overstreet 2014) have been (re)interpreted as members of functional categories that serve the structuration of language on a more global or “macro” level, e.g. as *extra-clausal constituents* in Dik’s (1997) *Functional Grammar*, as *supplements* (Huddleston & Pullum 2002), as *theticals* (Kaltenböck et al. 2011; Heine et al. 2015), or as elements of *macrogrammar* (Haselow 2017).

While many of these approaches are primarily used for a functional classification of linguistic expressions, the *microgrammar-macrogrammar* model of language structure and language cognition is particularly suitable as a framework for studies at the grammar-discourse interface as it lends itself for a systematic description of various phenomena relating to the structuration of language outside isolated sentential structures, such as the linearization of segments of speech of whatever syntactic status into coherent units of talk (Haselow 2017: Ch. 5) or the linear order of expressions “outside the clause” (Haselow 2019). Moreover, it is compatible to and integrates neurocognitive and psycholinguistic findings on dualistic processing, providing a grammatical framework for important processing-based distinctions such as that between *novel* and *formulaic speech* (Van Lancker Sidtis 2004, 2009), grammatical and “pragmatic” processing (Wray 2002; Code 2005), or that between propositional processing and discourse processing (e.g. Kintsch 1988; Gernsbacher 1990; Graesser et al. 1994; Greene et al. 1992; McKoon & Ratcliff 1992; Long et al. 2005; Prat, Long & Baynes 2007). The model thus allows for an integrated description of language structures and linguistic cognition, encompassing linguistic as well as psycholinguistic and neurolinguistic findings.

The basic idea underlying two-component models of language structure (grammar) and linguistic cognition in general is that successful communication appears to require structuration of language in two domains: structures associated with syntax and semantics in the structuralist tradition, encompassing morphosyntactic and semantic dependency relationships and propositionality, and relations between structurally autonomous segments of talk on a “higher” structural level, i.e. “above” the clause, e.g. between segments of an utterance, between turns, or between larger discourse units. In other words, the human mind intuitively forms mental representations of microstructures, which are more local, based on the internal structuration of isolated structural units (such as clauses), but also derives representations of global structures of discourse, or “macro” structures, which serve the mental integration of such units in the immediate situation of discourse, i.e. in an individual communicative and sequential context. Macrostructures, in this

sense, contribute to the development of a coherent mental model of ongoing interaction and discourse, allowing speakers not only to produce isolated units based on formal dependency relations, but to engage in organized interaction with another participant and to produce (spoken or written) *text*, that is a piece of meaningful, structured language that serves a particular purpose in an individual discursive context. Both components – micro- and higher-level macrostructuration – are equally required in a communicative act and can be assumed to interact, in various ways, in jointly shaping verbal interaction.

The dualistic approach based on a distinction between micro- and macrostructures in linguistic discourse thus tackles the problem that linear order and “grammaticality” are not generally established through hierarchically organized syntactic structures involving binary morphosyntactic and semantic dependency relationships, as it is usually assumed in sentence-based models of syntactic representation. Moreover, it accounts for the problem that many syntactic configurations, especially in spoken language, include more than “clausal” constituents, basically expressions with procedural meanings in the sense of Blakemore (2002), such as discourse markers (e.g. *well, so, yeah*) or comment clauses (e.g. *I mean, I think*), which serve as interpretive cues beyond lexical-conceptual meanings and help maintaining a coherent representation of what a speaker is doing. An example is (1) from the spoken component of the British section of the *International Corpus of English* (ICE-GB). A and B talk about one of A’s favorite books.

(1)

- 39 A: [I] think it’s a really good book,
 41 B: [uhm]
 42 and do you read the book over and over again?
 43 → A: WELL (.) after a certain length of time [yeah] I would imagine.
 44 B: [yes]
 43 A: because it was really (.) one that always sticks in my mind.
 45 B: °right°.

(ICE-GB S1A-016)

In line 43, speaker A first deals with a restriction of the propositional content of B’s information-seeking request before directly answering it, stating that he reads the book again *after a certain length of time* has passed. This prepositional phrase is followed by an affirmation marker (*yeah*), which serves as a direct response to B’s request at a point at which the modification of the content of the question has been completed. *Yeah* does not establish internal coherence within a single structural unit on the microlevel, but links the utterance-so-far to the prior speaker’s talk and thus to the sequential context in which it is produced, marking it as a direct response to B’s request. Moreover, it indicates a boundary in the utterance-in-progress as it represents the transition from specifying information that is relevant for responding to B’s request to a follow-up part that somewhat mitigates the confirmation and

weakens its epistemic value (*I would imagine*). Other devices to establish structures on the macrolevel are e.g. initial *and* (line 42), which establishes a textual relation between adjacent units of discourse (signalling topical continuity), *well* (line 43), which anticipates a response and thus links the upcoming turn to what the prior speaker said, or *I think* (line 39), which modifies the way in which the upcoming utterance is to be interpreted by the addressee (as an opinion) and thus establishes cognitive alignment.

While we have relatively detailed knowledge as regards the structuring principles on the microlevel of language structure, given the traditional focus of linguistic research on propositional content and its packaging in various types of syntactic and morphological structures, structural relations on the macrolevel offer much potential for further research. Many aspects of such “macro” relations have been studied under various frameworks, e.g. in Halliday’s (1985) and Halliday & Matthiesen’s (2004) systemic *Functional Grammar*, where links are established between sentence-level, propositional elements and categories and metafunctions of language (e.g. interpersonal functions, which are reflected in grammatical choices such as mood or modality), or in *Functional Discourse Grammar* (e.g. Hengeveld & Mackenzie 2008; Keizer 2015), which is organized in a “top-down” manner, starting with the speaker’s communicative intention and from there working its way down to articulation, and which takes the “Discourse Act” rather than the sentence or the clause (i.e. a morphosyntactic unit) as the basic unit of analysis. They are also discussed in various approaches to the so-called “peripheries” of an utterance (see e.g. the contributions in Beeching & Detges 2014) and in works on discourse segmentation (e.g. Hannay & Kroon 2005; Steen 2005; Degand & Simon 2009), which account for the fact that syntactic units are probably not the most optimal units of discourse analysis and that many relations in language hold between discourse units rather than clausal units or clausal constituents. Other lines of current research follow a different course and focus on the possible correlations between neurophysiological processes relating to brain lateralization and language processing and cognitive processing modes for different language-related tasks in order to arrive at a better understanding of why language is structured the way it is and how linguistic resources available for structuration are represented in the mind (e.g. Ullman 2004, 2015; Prat et al. 2007; Lehman Blake 2010; Pallier, Devauchelle & Dehaene, 2011; Boye & Bastiaanse 2018, among many others). Identifying such correlations is a challenging task, given the different research traditions, methodologies, tools and research interests, which hampers comparability and makes it hard to reconcile the different perspectives. With all care and considering that a comparison of the findings produced in the different disciplines is a difficult task it can be concluded that there is increasing evidence for language-related dual process models in all these fields (see Heine [2019] and Haselow & Kaltenböck [2020] for an

overview). This is not surprising, given that many structural-linguistic phenomena and ways of constructing linguistic discourse “resonate” (neuro)cognitive aspects, as it were, and invite for more integrated views on the relation between (neuro) cognitive processes and language structure.

2.2 Language processing and hemispheric differences

Over the last decades neurolinguistic research based on functional imaging techniques and patient observation has shown that a broader set of communicative-linguistic abilities tends to be hemisphere-specific in the sense that they are more strongly associated with neural activity in one of the hemispheres than in the other. As mentioned in Section 1, processing activities related to sentence structure and semantics mainly involve areas of the LH, above all the area known as Broca’s area (Brodmann areas BA 44 and 45) and the frontal operculum (FOP). Both are associated with the processing of hierarchically organized constituents linked by dependency relationships and the processing of embedded structures (e.g. Friederici 2004; Friederici et al. 2003; Sakai et al. 2005; Pallier, Devauchelle & Dehaene 2011). Patients with neurological impairment in these areas tend to suffer from a very similar set of deficits, most notably (i) syntactic impairment preventing speakers from producing novel syntactic, propositional units based on morphosyntactic and semantic dependency relationships and including more than 2–3 words, (ii) impaired lexical access affecting fluency, and (iii) nonlexical vocalizations. However, there is compelling evidence for preserved abilities as regards communicative competence in processing discourse- and conversation-structuring aspects, such as sensitivity to the structure of discourse, the preservation of conversational order (e.g. Marini 2012; Prat, Long & Baynes 2007), and the competence to make appropriate use of certain categories of utterances, such as formulaic or “automatic” speech (e.g. expletives, interjections, discourse markers, greetings, proverbs, idioms, quotes, sentence stems such as *I can’t*, *I want*, *you gotta*, and all kinds of formulas such as recited speech or nursery rhymes, e.g. Van Lancker Sidtis & Postman 2006; Van Lancker Sidtis 2012). Thus, a large proportion of the speech produced by aphasics are fixed expressions (Code 1997; Van Lancker Sidtis & Rallou 2004), which are produced in a quasi-automatic way and represent frequently and conventionally utilized portions of speech in a given language. Unlike novel speech, which is based on compositionality and requires analytic processing, formulaic portions of language tend to be processed holistically (Wray 2002). An example for the high incidence of formulaic speech in the absence of more complex syntactic abilities is (2), which is taken from the *Aphasia Bank* database. The speaker is a 52-year old male aphasic speaker diagnosed for Broca’s aphasia, who uses formulaic units such

as (*oh*) *wait a minute, and then/but then* or *so now* at various points in emergent speech (INV=interviewer, PAR=patient). The presence of such forms has been well documented in clinical descriptions of speakers with impaired LH activity (Van Lancker Sidtis 2001; Code 1997, 2005; Wray 2002) and is one of the most characteristic properties of aphasic speech.

(2)

101 INV: thinking back can you tell me a story about something important
 102 that happened to you in your life?
 103 it could be happy or sad or from any time.
 104 from when you were a kid or more recently.
 105 PAR: um (...) okay.
 106 first long ago [/] um um um <points up> um (.) long ago
 107 <writes> um um
 108 <looks at phone>
 109 <points at paper> okay?
 110 INV: mhm.
 111 PAR: (.) um s::: wait a minute.
 112 Seattle [/] Seattle but will turn no [//] but then go home.
 113 <gesture>
 114 <but one> [//] i- but first (.) Seattle.
 115 <finger:draw_circle> one year.
 116 INV: mhm
 117 PAR: and then go home.
 118 so [//] but [/] but first um (.) me tee [= to]
 119 Seattle.
 120 and [/] and down.
 121 and Brian flying.
 122 INV: mhm.
 123 PAR: <and then> [/] and then relax.
 124 and so first um um
 125 oh wait a minute.
 126 um S:eattle.
 127 and then,
 128 wait a minute.
 129 and the:n flying no [//] and then drive [/] drive
 130 first.
 131 <PAR looks at phone>
 132 INV: mhm.
 133 PAR: um oh wait a minute.
 134 um um (.) wait a minute.
 135 INV: mhm.
 136 PAR: <writes> Seattle.
 137 oh:
 138 well ‡ anyway <points: at paper> flor- f- um l- um s: Seattle
 139 to here. <points at paper>
 140 INV: mhm .
 141 PAR: and oh l- Las Vegas [//] lɔsergis [= Las Vegas]
 143 INV: mm.
 144 mhm.

145 PAR: and then fly home different.
 146 so † um first <finger:one> um um um rocks.
 148 INV: mhm.
 149 PAR: and then a w- um crater [/] crater.
 150 and then go home.
 151 and then after a while (.) work again.
 152 so † no:w um <points:self> one two three four but <finger:count>
 153 one two three <finger:three> <points:self> <head:yes> it's me.
 155 INV: mhm (.) mhm.
 156 PAR: so † no:w working again.
 157 INV: great.
 158 thanks.

(Aphasia Bank: BU07a)

Note that the speech produced by the LHD patient exhibits all the properties discussed above: syntactic impairment (e.g. *but first me to Seattle and down*, lines 118–120, or *now working again*, line 156), short syntactic units rarely spanning more than three words, and impaired lexical access (see e.g. line 138). However, the speaker shows largely preserved skills as regards the ability to organize interaction (e.g. *okay?*) and the use of discourse-connecting devices, which tend to be used appropriately to link discourse segments (e.g. *and then*, *so now*, *but first*) into a relatively coherent whole on the macrolevel (even though the units linked are deficient on the microlevel). For instance, *wait a minute* signals continuation in a context in which the speaker prepares further talk, *and then* indicates a new sequence or event and thus continuity in the narrative.

Note that the deficits reported for speakers with LHD take very different degrees, depending on a variety of factors such as post-onset time, intensity of the damage, or personal factors (see below). Nevertheless, the pervasiveness of formulaic expressions in the speech of LHD patients is a consistent observation that led to the hypothesis that hemispheric representation is mode-specific, the RH using a largely holistic processing mode that allows for the processing of fixed expressions and overused chunks of language and that is associated with abilities deriving from habit learning in general (Squire 2004).

I will now turn to deficits resulting from the impairment of RH activity. Neural substrates in the RH were identified for processing pragmatic aspects of language (e.g. comprehension of idioms, indirect and metaphorical meanings), formulaic speech, and aspects relating to the macrostructure of speech, e.g. discourse organization and interactional aspects of talk such as turn-taking, as well as the respective expressions serving these functions (Lehman Blake 2010). Patients with RH impairment show a strong tendency to use fewer discourse-connecting elements and turn-taking cues, fewer formulaic units (Code 1996; Van Lancker Sidtis 2004; Van Lancker Sidtis & Postman 2006), impaired judgment of the addressee's

communicative needs and problems with taking the perspective of the addressee (Myers 1999), such as awareness of shared knowledge and conversational procedures (Bartels-Tobin & Hinckley 2005; Brady, Armstrong & Mackenzie 2006), turn-taking difficulties (Kennedy et al. 1994), as well as problems in topic maintenance (Caplan & Dapretto 2001) and an impaired ability to structure narratives, e.g. in terms of logical or temporal sequentiality of episodes and indicating relations between structural segments. Marini (2012: 73–75) found that “the participants with RHD produced narratives with normal levels of cohesion and local coherence but with more errors of global coherence and lower levels of informativeness”, that is, they produced significantly more tangential utterances than the control group. This is congruent with a number of studies which report that patients with RHD have difficulties integrating ideas across sentences into a coherent mental model and identifying main ideas and themes (Prat et al. 2007: 284).

Most of the deficits can be analyzed as referring to the macrostructure of discourse: clausal and discourse connections are often impaired and RHD speakers, whose speech includes cohesive errors and fewer linking devices. Some of these devices tend to be overused, such as *and*, which is often used as a mere placeholder for a more specific linking element and thus serves as an all-purpose connective device (e.g. Sherratt & Penn 1990; Davis et al. 1997; Marini et al. 2005). However, these patients have been reported to show little impairment of syntactic skills, producing e.g. a similar degree of clausal embedding than unimpaired speakers (Brady et al. 2006; Marini et al. 2005), even though reduced clause length and impaired lexical processing, e.g. problems with lexical retrieval, have also been observed (Sherratt & Bryan 2012: 227; Barnes et al. 2019).

The major findings in neurolinguistic research, summarized in Table 2, support the hypothesis that the RH serves as neural substrate for discourse- and interaction-related components of language use.

The language-related neuroanatomical distinctions between the LH and the RH just outlined exhibit significant correlations relating to the kind of structural dualism proposed in the micro-/macrostructure model of language in that structuration of language on the microlevel implicates mainly the LH while building a macrostructure of discourse requires substantial involvement of the RH (see also Heine et al. 2015; Heine 2019). In Haselow (2017) it has been shown that many of the functions identified for *microgrammar* and *macrogrammar* correlate with the functions that were identified in the neurolinguistic literature, although the correspondences are, of course, far from being absolute. An overview of the principle correlations is provided in Table 3. Note, again, that the lateralization indicated in the table is not a matter of all-or-nothing, but referring to a hemispheric advantage.

Table 2. Main deficits following from RH damage

Discourse organization		Interactional management	
Deficits	Key references	Deficits	Key references
(i) Discourse contains unnecessary details that lack immediate relevance	Marini et al. (2005), Lehman Blake (2006), Tompkins (2008), Sherratt & Bryan (2012)	Difficulties in respecting turn-taking in conversation	Myers (1994), Chantraine, Joannette & Ska (1998) Hird & Kirsner (2003)
(ii) Problems with topic maintenance and topical coherence, discourse includes frequent divergence from the main topic or point	Caplan & Dapretto (2001)	Difficulties in respecting contextual aspects (e.g. shared knowledge) and in interpreting another speaker's communicative intentions (e.g. literal meanings vs. irony)	Bottini et al. (1994), McDonald (1999), Blake (2009), Champagne -Lavau & Joannette (2009)
(iii) Difficulties with the sequential ordering of events in narratives, impaired ability to formulate structured narratives (e.g. missing episodes, messy ordering of events, repetitiveness), lack of a coherent macrostructure of discourse	Sherratt & Penn (1990), Purdy (2002), Marini et al. (2005)	Difficulties in using formulas of social exchange (e.g. greetings, leave takings)	Van Lancker Sidtis & Postman (2006), Van Lancker Sidtis (2009: 451)
(iv) Disrupted fluency and clarity of discourse; inappropriate use of formulaic expressions and 'nonspecific' elements (e.g. <i>and so on</i>), which signal discourse continuity, but no continuity of meaning, and are often used to fill pauses, repetitions	Sherratt & Penn (1990), Sherratt & Bryan (2012)	Difficulties in comprehending and producing emotional features in speech	Borod et al. (2002), Devinsky (2000), Friederici & Alter (2004), Mitchell & Crow (2005)
(v) Underuse of cohesive ties (connectives) or substitution of ties by <i>and</i> , difficulty in determining the relationship between propositions	Sherratt & Bryan (2012)		

Table 3. Principal correlations between neurolinguistic functions and grammatical functions

Neurolinguistic functions		(Micro-, Macro-)Grammatical functions	
Function and expressions	Hemisphere	Function and expressions	Domain
(i) Syntactic processing: parsing of hierarchical relationships, embedded structures and their propositional content	LH	Creating syntactic structures based on morpho-syntactic and semantic dependency relations and propositionality	microstructure
(ii) Discourse organization: integrating parts of a text into a coherent whole, building up a coherent mental model of discourse	RH	Structuring language beyond local dependency relations, integrating autonomous syntactic segments into a unified whole	macrostructure
(iii) Social context of linguistic communication: taking account of the co-participant and his/her intentions, appreciation of the co-participant's communicative needs, turn-taking	RH	Speaker-addressee relationship: including the addressee into the discourse, referring to shared knowledge/shared understanding, organizing conversational interaction	macrostructure
(iv) Use of formulaic speech (e.g. interjections, discourse markers, expletives, proverbs)	RH	Formulaic expressions outside formal dependency relations, e.g. interjections, discourse markers, response tokens, comment clauses, general extenders	macrostructure

The RH is clearly dominant in the control over most aspects of language structure beyond the sentence and over social and emotional behavior, including social aspects of language use, e.g. the use of expressions serving the organization of speaker-addressee interaction. Formulaic speech, or at least those formulaic units that are syntactically independent in that they are not involved in (sentence-grammatical) dependency relations (see Kaltenböck 2020), is also associated with RH activity, which suggests that the RH operates on the basis of a holistic, noncompositional processing mode (Pawley 2009). The fact that most expressions with a macrostructuring function are noncompositional, unintegrated into formal, hierarchical relationships, and highly conventionalized reflects a further striking correlation between RH activity and *macrostructure*.

As with all dichotomies, the findings and generalizations have to be taken with great care: speech produced by brain-damaged persons is rarely characterized in

terms of ‘all or nothing’ dichotomies, and types of speech are never. For instance, there are individual differences based, for instance, on the cause of the damage, type of disorder, time post-onset, the nature of the lesion, recovery time, or personal characteristics (e.g. age, education, premorbid behavior) (Mackenzie & Brady 2008). Moreover, typologies based on distinctions between the right and left hemisphere are rough dichotomies reflecting statistical tendencies and a particular focus of assessment rather than general all-or-nothing distinctions. A further problem is that findings on RHD speakers tend to be more inconsistent than for LHD speakers, the first varying vastly in the degree of (in)sensitivity to the communicative affordances implemented through dialogic interaction. Finally, it should be considered that there is relatively robust evidence for bilateral activation and more diffuse representation for a number of language-related tasks, e.g. for naming and counting (e.g. Cowell et al. 2000) or for proper nouns (e.g. Brownell & Joanette 1993; Myers 1999).

The correlations between linguistic and neurolinguistic dichotomies are also related to a number of methodological problems. For instance, in the absence of a methodologically sound way of mapping linguistic categories to neural processing, they can only remain vaguely defined. The functional differentiation of the RH is more diffuse than that of the LH, and the same holds for the precise nature of measurable impairments of cognitive subprocesses and communicative difficulties in subfunctions of language associated with RH damage (Tompkins 1995; Myers 1999; Mackenzie & Brady 2008). Nevertheless, in view of the large body of neurolinguistic research on linguistic processing the assumption that the structure of (spoken) language can be reduced to a monolithic system of mental processing and linguistic discourse, which has long dominated linguistic thinking, appears more and more questionable.

2.3 Aphasic speakers and discourse structure

At this point the reader may wonder why a speaker suffering from LHD, whose speech tends to be characterized by severe agrammatism and who is mostly unable to produce coherent syntactic structures, should use devices for structuring language on the macrolevel, given that there is apparently little to be structured. First of all one has to bear in mind that aphasic speakers are not “speechless” or “wordless”, i.e. they do not suffer from a complete breakdown or system failure that would render any macrolevel structuration unnecessary, as shown in (2). While cases close to wordlessness have occasionally been observed (e.g. Howard & Orchard-Lisle 1984, who reported on a patient who had a repertoire of only six words/phrases, e.g. *yes, no, I understand*), they appear to be extremely rare: as Butterworth (1994: 55) states, “[i]t is far more usual for patients to manage some form of sentence production

despite an impairment, even a severe impairment.” Therefore, “[a] traditional classification of aphasia distinguishes the clinical impression of fluent speech from the impression of hesitant speech” rather than ‘no speech’. Even hesitant aphasic speakers may produce as much as 13 words per minute (Berndt 1987). It should also be noted that even highly hesitant aphasics may have well-preserved comprehension skills.

The linguistic competence and the severity of linguistic disorganization vary vastly among aphasic speakers as the degree of agrammatism shows striking individual differences and heavily depends on numerous factors, as mentioned above. Thus, linguistic abilities may range from relatively rich discourse with different degrees and manifestations of agrammatism to different combinations of criterial features, e.g. the omission of grammatical words such as articles, auxiliaries, or pronouns, underuse or absence of grammatical markers for some categories such as case or number, overuse of certain forms or grammatical categories such as third person or present tense, anomalous word order, or paraphasias, that is, aphasic substitution errors or wrong, unintended uses of syllables, words, or inflections during efforts to speak. In other words, there is no general, exhaustive set of deficits. It is therefore important to keep the cohort of LHD patients to be observed in studies on linguistic performance as homogeneous as possible, e.g. with respect to specific types of neural damage and age. For instance, the lesions responsible for *paragrammatic* speech are found in the posterior language regions, especially Wernicke’s area; *agrammatic* speech symptoms (nonfluent speech with syntactically simple utterances and frequent grammatical errors), in turn, are usually a consequence of damages to Broca’s area in the frontal lobe. Approximately 25% of stroke survivors with aphasia exhibit *agrammatism* (Pedersen, Vinter & Olsen, 2004).

In the present study, the focus will fall on male patients between the age of 31 and 62 with Broca’s aphasia. An impression of the kind of speech output produced by aphasics from this cohort has been given in (2) above, which has shown that continuous speech is heavily impaired, including typical features such as short units (rarely longer than three words), simple syntactic structure, and low incidence of function words. Despite these deficits, the speakers tend to be sensitive to the preservation of discourse organization and show efforts in integrating ideas and pieces of discourse into a coherent macrostructure.

To conclude, many linguistic features are preserved both in Broca’s and Wernicke’s aphasia (for a highly valuable cross-linguistic overview of aphasia symptoms see Volume 14, Issues 2–4 of the *Journal of Neurolinguistics*, 2001). Note that aphasic errors are not always systematic, but often context-dependent. For example, Helasuvo, Klippi and Laakso (2001) found that Finnish Broca’s aphasics produced case marking and agreement patterns in some contexts correctly, whereas in other contexts they were produced erroneously. These and other studies show that it is difficult and

often impossible to give a definitive characterization of the grammatical error profile of aphasic speakers. Moreover, sentence structure is sometimes badly distorted both in Broca's and Wernicke's aphasic cases because of difficulties in lexical access to content words. Note, however, that there appears to be agreement on the resistance of high-frequency words to errors in normal and aphasic speakers (Howes 1964; Rochford & Williams 1984; Centeno & Obler 2001), which is possibly a reflection of the automatic, routinized, and overlearned nature of these words due to their early acquisition (Schnitzer 1989) and their high degree of lexical strength (Bybee 1995; Stemberger & MacWhinney 1986), next to frequency.

All this suggests that macrolevel abilities (the structuration of discourse and interaction management) are not a superfluous "extra" that is supposedly irrelevant to speakers who, given the syntactic impairments, produce little speech otherwise. Aphasic LHD speakers depend as much, perhaps even more, on macrostructuration than unimpaired or RHD speakers since they do not produce their speech – however deficitous it is – in a discursive and interactive vacuum and since single fragments of speech need to be integrated into coherent units of discourse beyond sentence grammar.

3. Discourse markers

The class of lexemes analyzed in the present study, DMs, is characterized by a set of formal and functional features. Formally, DMs exhibit syntactic optionality (they can be omitted without affecting grammaticality because they are outside morphosyntactic and semantic dependency relationships holding between clausal constituents, representing formally autonomous, morphologically invariant units), lack of semantic content, and non-truth conditionality (they have no effect on the truth value of the unit they accompany) (Schourup 1999). DMs convey procedural rather than lexical-conceptual meaning (Blakemore 2002), providing a processing cue on how the unit they accompany is to be understood or interpreted within a given communicative context. Moreover, they have no syntactic linking function, which excludes clausal conjunctions (though it includes lexemes that have discourse-structuring functions next to their uses as clausal conjunctions, i.e. which introduce syntactically independent discourse units, such as turn-initial *and* or *but*), and they exhibit a high degree of conventionalization or grammaticalization.

The function of DMs has been conceptualized in different ways. Under a narrow definition, their function is discourse segmentation and the creation of connectivity (Schiffrin 1987; Fraser 1999, 2015). The scope of a DM is thus over two adjacent textual units, each expressing a separate message. In more recent functional, corpus-based studies the nature of the "connective" function and the

types of units involved has come to be defined in broader terms, not only applying to neighboring discourse units, but also to relations between verbalized and non-verbalized, implied messages (Hansen 2006). Moreover, the set of functions includes discourse organization in a wider sense, e.g. interactive functions such as backchanneling, turn-taking, or saving face (González 2005; Crible 2017). Based on these considerations, the term ‘discourse marker’ as it is used here encompasses a broader range of expressions often referred to under specific categorical labels in the literature, such as interjections (e.g. *oh*), linking adverbials/final particles (e.g. final *then* or *though*), general extenders (e.g. *and stuff*), parentheticals/comment clauses (e.g. *I think*), or tag questions (e.g. *isn’t it?*).

There is wide agreement that DMs do not form a class of unifunctional expressions, but exhibit internal subcategorization that reflects functional differences between different types of DMs. A large number of taxonomies has been proposed in the literature (e.g. Pons 2008; Maschler 2009; Crible 2017) which, in spite of different categorical labels, share a broad categorization into basically three functional groups: (i) DMs that primarily serve interaction management, (ii) DMs that primarily serve the establishment of textual relationships and thus the structuration of discourse in terms of rhetorical (e.g. contrast, inference) or action-based relations (e.g. continuity) between units of discourse, and (iii) DMs that serve as explicit cues for utterance interpretation and that serve the coordination of discourse processing between the speaker’s and the addressee’s mind involved in a conversational encounter, as shown in Table 4 (for a more detailed discussion see Haselow 2019).

Table 4. Functional domains of discourse markers

	Domain	General function	Examples
(i)	INTERACTION	organizing turn-taking and interaction (e.g. setting up the conditions for successful uptake of an upcoming message; responding)	<i>listen, ...isn’t it?, yeah</i>
(ii)	DISCOURSE STRUCTURE	indicating the type of (local, global) relationship between discourse units or between a discourse unit and implied meanings or inferences	<i>and, anyway, but, so</i>
(iii)	COGNITION	providing a cue for utterance-interpretation	<i>I think, you know</i>

The category INTERACTION (i) includes all functions related to the management of speaker-addressee interaction and the structuration of the periodic change of speaker roles in interaction, i.e. the transition from one speaker’s turn to another speaker’s utterance. This includes, above all, backchannels or response markers (e.g. *yeah, yes, okay*) when they are followed by more talk, as these structure the transition from prior talk to upcoming talk and often prepare the ground for upcoming

talk, e.g. by acknowledging prior talk or indicating dis/agreement or hesitation before continuing with a new message. DISCOURSE STRUCTURE (ii) involves all functions that serve the structuration of emergent discourse and includes DMs that mark the transition from one unit of discourse to another (local relations), or between a larger set of units (global relations, as with *anyway*), e.g. between topical aspects (topic shift, topic resumption) (see also González 2005; Redeker 2006; Crible 2017). In (3), for instance, initial *yeah* in line 40 indicates acknowledgment of A's turn and thus occurs in the transition zone from one speaker to another, whereas *but* signals the transition to a disaligning move, introducing a counterargument.

(3)

39 A: you could commission prints of yourself.

40 B: **yeah but** I can't afford that kind of thing.

(*International Corpus of English – Great Britain, S1A-015*)

The third category COGNITION (iii) encompasses DM functions relating to the alignment of the co-participants' cognitive states and thus serves structuration in terms of jointly constructed mental representations of emergent discourse. DMs in this domain do so by providing a cue as to how a message is to be interpreted and integrated into the common ground (the set of shared knowledge, beliefs and background assumptions, Stalnaker 2002), e.g. as expressing an opinion rather than a fact (*I think*), or as indicating the speaker's belief that the addressee is able to grasp "the sequential relevance of the statement" (Schourup 1985: 105) and to recognize the implications of an assertion or an expression (*you know*).

(4)

62 A: I was hoping it'd be a week so I could get off the ward for a bit but,

63 B: oh already.

64 A: **well you know** it's nice to have a change you know.

(*International Corpus of English – Great Britain, S1A-099*)

In (4), *well* represents a response to B's expression of surprise, foreshadowing an explanation, *you know* serves as an interpretive cue indicating that B is assumed to be able to recognize the implications deriving from the upcoming utterance since it refers to an idea that everyone is familiar with.

One may discuss whether a more fine-grained classification is possible and necessary (see the discussion in Crible & Degand 2019). For the present purpose, the three clusters appeared sufficient to yield substantial findings as regards the differential use of DMs by RHD and LHD speakers. It should also be noted that a single DM may serve a broader set of functions in different contexts (e.g. Fischer, 2000; Crible, 2017; Crible & Cuenca, 2017), but mostly DMs can be related to core domains in which they prototypically operate.

4. Data: Discourse marker use and hemisphere-specific disorders

The following discussion is based on the assumption that if the use of DMs as devices for creating macrostructures in language is more strongly associated with RH processing their incidence should be more or less heavily reduced with speakers suffering from RHD than with speakers suffering from LHD. Section 4.1 provides a brief outline of the study, Section 4.2 explains the method, Section 4.3 provides a comment on the target objects (DMs), and Section 4.4 presents the results. The data are then discussed in Section 5.

4.1 Aims of the study and database

The general goal of this study is to evaluate the hypothesis about hemispheric differences and to contribute empirical evidence for or against differential processing of micro- and macrostructural aspects of language. Moreover, the study will also provide a more precise characterization of the incidence of DMs in the speech of neurologically impaired speakers: the strong association of discourse-organizational aspects with RH processing should be reflected in a higher incidence of DMs in the speech produced by aphasic speakers suffering from LHD, who operate on the basis of a more or less intact RH, and in a lower incidence of DMs in the speech produced by RHD speakers with an intact LH. The focus will fall on turn-initial DMs, given that turn-beginnings are sensitive moments in emergent discourse in that the change of speaker roles and possible changes in the direction of talk increase the risk of a disintegration of talk. The reduction of the risk, and thus the preservation of the macrostructure of discourse, is considered to be a strong factor motivating the use of DMs. In noninitial positions, DMs may either relate the utterance in which they occur to the discourse context or relate single segments of an emergent turn into a coherent whole and thus serve different functions on different levels.

As mentioned in Section 2, the association between formulaic speech, discourse- and interaction-related aspects of language and RH processing is not a total one, but corresponding to tendencies. Thus, patients with RH disorders or damages can be expected to produce a considerable, though lower amount of linguistic forms and structures related to discourse structure and interaction management as well as fewer instances of formulaic speech than speakers with LH damage or control group speakers. For instance, in Van Lancker Sidtis & Postman's (2006: 418) data the proportion of formulaic expressions in the speech produced by RHD speakers (which includes, among others, DMs) amounts to 17%, compared to 30% with LHD speakers and 25% with unimpaired speakers. Given the variability of linguistic performance with speakers suffering from the same hemispheric damage (LHD

or RHD) reported in the literature a sharp distinction between LHD and RHD speakers is certainly not expectable so that even moderate differences would be an important indicator for processing differences.

4.2 Method

Instances of spontaneous speech produced by speakers with unilateral brain damage were examined and the findings compared to those resulting from unimpaired control speakers (CS). The data on neurologically impaired speakers are based on speech samples deriving from interviews with speakers either classified for LH or RH disorder, which were collected in the *Talkbank* project (MacWhinney et al. 2011) and are accessible through the *Clinical Bank* as part of the *Talkbank*. The speech samples provided by the Clinical Bank are based on elicited free speech based on a standardized set of open questions and picture descriptions, as shown in Table 5.

Table 5. Discourse types and elicitation tasks in the speech samples

	Discourse type	Elicitation task	Example for elicitation instruction
I	Free speech samples	a. Stroke story and coping b. Important event	Do you remember when you had your stroke? Can you tell me about it?
II	Picture descriptions	A. Broken Window B. Refused Umbrella C. Cat Rescue D. Flood	Now I'm gonna show you these pictures. Can you tell me the story with a beginning, a middle and an end?
III	Story narrative	Cinderella	Have you ever heard the story of Cinderella? Tell me as much of the story of Cinderella as you can.
IV	Procedural discourse	Peanut Butter and Jelly Sandwich or other simple sandwich	Tell me how you would make a peanut butter and jelly sandwich.

For the aphasic speakers all utterances were included in the analysis, given the comparatively low number of words produced by these speakers and given that the picture task had a strong dialogic character since the speakers required much more prompts by the interviewers so that turn beginnings are rather frequent. The speech samples from RHD and control speakers were restricted to the parts representing open interviews, given that the picture tasks resulted in longer monologic sequences that did not provide enough contexts for the use of turn-initial DMs in dialogic interaction.

Speech samples from 15 right-handed male speakers in three subgroups were used in this study: five patients with aphasia following LHD, all of them diagnosed

for Broca's aphasia due to a cerebral vascular accident, five patients with RHD due to a cerebral vascular accident, and five control subjects. The LHD patients ranged in age from 42 to 64 (mean = 52.9, $SD = 6.4$) and had an average education of 17 years. The speakers were drawn from a group of patients with a mean post-onset time of 22.6 months. All subjects were evaluated with the WAB-Q Severity score, which ranged from 39.7–67.0 (mean = 56.6, $SD = 9.3$). The age of speakers with RHD ranged from 31 to 62 (mean = 53.2, $SD = 10.2$) with 18 mean years of education. The mean post-onset time was 31.4 months; no information on severity evaluation was available. All speakers were sufficiently able to participate in language testing and verbal interaction and sufficiently fluent to be able to produce an appropriate amount of linguistic micro- and macrostructures. The control speakers' age ranged from 45.06 to 60.09 (mean = 52.72, $SD = 5.1$).

Detailed information on the subjects, the number of turns and the amount of words produced by each speaker is given in Table 6. Note that the number of turns refers only to substantive turns, that is, to those turns that express one or more propositional elements, which excludes mere backchanneling turns. The amount of turns produced by the LHD speakers averaged 18 turns (range = 17–20, $SD = 1.3$), each of the speakers produced a mean of 516 words (range 274–706, $SD = 164.5$). The RHD speakers produced 40 turns on average (range = 38–40, $SD = 0.75$), with a mean amount of 941 words (range 846–1,116, $SD = 91.5$). The average number of turns produced by the control group was 20, with a mean amount of 1,146 words (range: 895–1,347, $SD = 171.2$). The amount of words excludes all filled pauses (*uh, uhm, er*), self-interrupted (e.g. *win-* in *through a win- window*) or unintelligible words (marked as *x, xx* or *xxx* in the transcript, depending on word length) and repetitions.

Table 6. Subject information

Speaker ID	Age	Gender	Diagnosis	WAB-Q severity score	Number of words	Number of turns
Aphasia Bank: Adler13a	52.04	M	Broca	55.8	565	18
Aphasia Bank: BU07a	52.04	M	Broca	51.5	706	17
Aphasia Bank: BU08a	64.06	M	Broca	39.7	616	20
Aphasia Bank: CMU02	42.11	M	Broca	59.6	274	17
Aphasia Bank: Elman03a	55.02	M	Broca	66.2	628	20
Aphasia Bank: Elman11a	52.01	M	Broca	67.0	309	19
Aphasia Total	$M = 52.88$				3,098	111
RHD Bank: Minga05a	56.00	M	RHD	NA	883	40
RHD Bank: Minga11a	55.04	M	RHD	NA	1,116	38
RHD Bank: Minga14a	31.01	M	RHD	NA	887	40
RHD Bank: Minga15a	57.10	M	RHD	NA	846	40

(continued)

Table 6. (continued)

Speaker ID	Age	Gender	Diagnosis	WAB-Q severity score	Number of words	Number of turns
RHD Bank: Minga23a	58.00	M	RHD	NA	912	40
RHD Bank: Nazareth01	62.00	M	RHD	NA	1,000	40
RHD Total	<i>M</i> = 53.19				5,644	238
Control: Capilouto11a	53.05	M	–	–	1,313	20
Control: Capilouto19a	60.09	M	–	–	895	21
Control: MSUC07a	48.01	M	–	–	1,225	20
Control: Wright09a	45.06	M	–	–	950	20
Control: Wright19a	57.06	M	–	–	1,144	19
Control: Wright42a	53.04	M	–	–	1,347	20
Control Total	<i>M</i> = 52.72				6,874	121

The aphasic patients had good comprehension skills, and despite differences in the structural design of their speech, seemed similar in their general linguistic competences and their deficits. As the data indicate, they produced the smallest amount of words compared to the other two groups. All LHD and RHD speakers showed high sensitivity to the interactional context, i.e. they were clearly oriented to dialogic interaction and the presence of another speaker.

4.3 Discourse markers in the speech data

The classification of turn-initial DMs produced by the subjects was based on the formal and functional criteria presented in Section 3. DMs were categorized as either structuring speech in the domain of INTERACTION (e.g. backchanneling, such as *oh*, *yeah*, *well*), DISCOURSE SEGMENTATION (e.g. textual linking, such as *but* or *so*), or COGNITIVE ALIGNMENT (e.g. reference to a shared viewpoint or common ground, as with *you know*). If speakers produced a combination of DMs (e.g. *yeah but*) each DM was counted. Note that the dividing line between the use of a DM for structuration in one of these three domains, and the use as a mere filler and thus as a device for stalling for time (see Aijmer 2002) is very thin, particularly with DMs such as *you know* or *well*. However, this problem holds for any type of speaker and is thus not expected to distort the data.

4.4 Results

Results documenting the incidence of turn-initial DMs in the speech produced by LHD, RHD and the control speakers are shown in Table 7. The RHD group had the smallest proportion of DMs at turn beginnings in their speech ($M = 45.8\%$),

while the LHD and the control group showed a similar proportion ($M = 73.9\%$ and 76%). A two-way ANOVA was performed to assess the differences between the two groups (LHD, RHD) in terms of the number of turns introduced by a DM and revealed a significant main effect of group, $F(1, 10) = 21.08$, $p < .005$. A post hoc Bonferroni test again indicated that the LHD group had a significantly higher proportion of turn-initial DMs than the RHD group, LHD-RHD ($p < .001$). Thus, as regards the two types of hemispheric disorder (LHD or RHD), speakers with LHD showed a stronger tendency to begin their turns with DMs, and thus with a macrostructuring device. No significant difference was found for the mean scores of the LHD and the control group ($p = .86$). As regards the variety of DM types used by the three groups, the mean figures suggest that the control group showed the highest degree of variation ($M = 7.8$). This difference, while suggestive, was not significant.

Table 7. Number and percentage of turn-initial DMs in the three speaker groups

	Number of turns	Turn-initial DMs		N and % of turns introduced by DM(s)	
		Tokens	Types	N	%
Aphasia Bank: Adler13a	18	14	6	10	61
Aphasia Bank: BU07a	17	19	6	16	94
Aphasia Bank: BU08a	20	17	8	14	70
Aphasia Bank: CMU02	17	15	5	13	76
Aphasia Bank: Elman03a	20	24	5	16	80
Aphasia Bank: Elman11a	19	19	7	13	68
LHD Total	111	108	$M = 6$	82	$M = 75$
		$M = 18$			$SD = 11.5$
RHD Bank: Minga05a	40	17	8	16	40
RHD Bank: Minga11a	38	20	6	18	47
RHD Bank: Minga14a	40	24	8	23	58
RHD Bank: Minga15a	40	17	7	15	38
RHD Bank: Minga23a	40	18	7	15	38
RHD Bank: Nazareth01	40	26	8	22	55
RHD Total	238	122	$M = 7.5$	109	$M = 46$
		$M = 20.3$			$SD = 8.8$
Control: Capilouto11a	20	13	7	12	60
Control: Capilouto19a	21	19	6	17	81
Control: MSUC07a	20	16	10	15	75
Control: Wright09a	20	14	6	14	70
Control: Wright19a	19	17	10	14	74
Control: Wright42a	20	19	8	17	89
Control Total	120	98	$M = 7.8$	88	$M = 75$
		$M = 16.3$			$SD = 9.8$

Figure 1 provides a more specific picture, showing the ratio of turns with initial DMs and those without produced by each speaker in percentages. The figure indicates that the proportion of turns that were not explicitly introduced by DMs was highest in the speech of RHD speakers, and that the two other groups used proportionally more turn-initial DMs in order to integrate their turns explicitly into the communicative context, with a mean value of 75% for both groups.

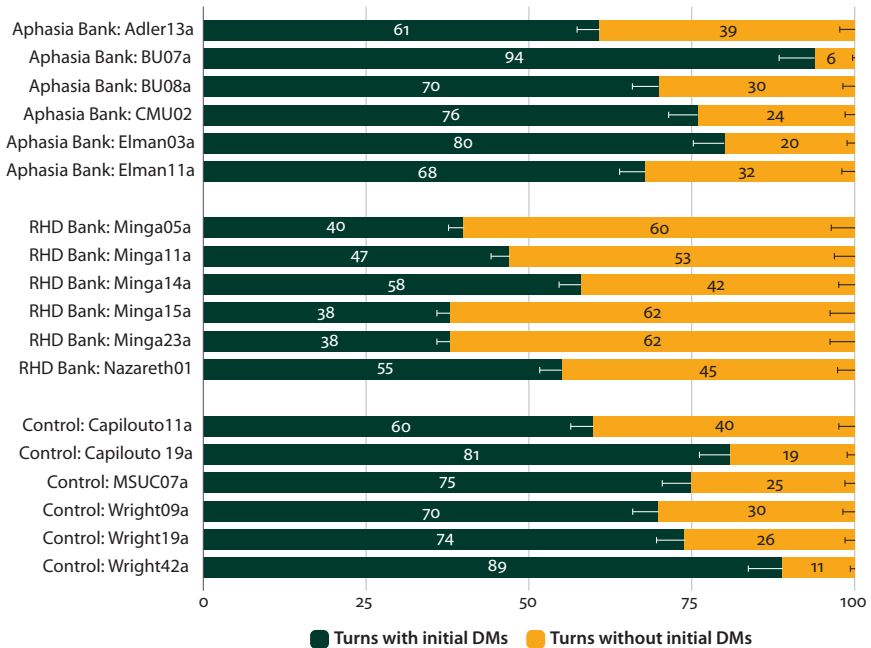


Figure 1. Percentage of turn-initial DMs in all turns

The observation that aphasic speakers (LHD) make more frequent use of DMs than RHD speakers does, of course, not necessarily mean that the first are better at integrating a turn into the broader communicative context. Rather, it is possible that many of the DM uses serve as devices for compensating the lack of grammatical skills that often prevents these speakers from using structural means to make an upcoming turn fit the context. In this sense, DMs may serve as a compensatory device for deficits in tailoring an upcoming turn to the discourse context by using structural means such as elliptical responses or syntactic devices such as object fronting (e.g. A: Do you like peanut butter or do you prefer jelly? – B: Jelly I like), which require mental (syntactic) operations that form one of the major deficits of LHD speakers. Yet, the fact that DMs are used frequently (and adequately) shows that the LHD speakers and the control speakers are well aware of the importance

of explicit macrostructuration by means of DMs after turn-transitions by linking a turn to the contextual conditions, e.g. by marking it as a response or as expressing a contrast to what a prior speaker said.

To examine differences in the types of DMs used by the speakers each DM was assigned to one of the three functional categories discussed above. The incidence of the different types of turn-initial DMs is shown in Table 8. LHD and RHD speakers used proportionately more DMs organizing speech on the interactional level (79% and 80% of all DM uses) than discourse markers structuring speech on the textual level (21% and 19%). This suggests that the bulk of DMs used by both speaker groups were response markers/backchannels (e.g. *oh, yeah, well*) followed by more talk. DMs serving cognitive alignment (*you know, I think*) were almost never used. Speakers of the control group showed a higher degree of variation as regards the functional types of DMs, with DMs serving interaction management and discourse structure being almost equal in proportion (42% and 44%), and with a higher proportion of DMs serving macrostructuration in the COGNITION domain (14%).

Table 8. Frequency of three functional types of DMs by subject group

	Interactional DMs		Discourse-structuring DMs		Cognition-aligning DMs		Total
	Tokens	Types	Tokens	Types	Tokens	Types	
LHD	85 (79%)	$M = 4.3$ $SD = 0.94$	23 (21%)	$M = 2.3$ $SD = 0.94$	0	0	108
RHD	97 (80%)	$M = 4.2$ $SD = 0.68$	23 (19%)	$M = 2.6$ $SD = 0.94$	2 (1%)	$M = 0.2$ $SD = 0.37$	122
CS	41 (42%)	$M = 2.7$ $SD = 0.47$	43 (44%)	$M = 3.6$ $SD = 1.37$	14 (14%)	$M = 1.5$ $SD = 0.5$	98

These findings suggest that one of the major communicative tasks of LHD and RHD speakers considered relevant at turn-beginnings was to explicitly mark a response to the prior speaker's turn before continuing with a message. This appears to be somewhat different for the control group, where responding and indicating the kind of link between two adjacent turns appeared to be equally important, given that the proportion of DMs serving these two major tasks was almost equal.

Figure 2 shows the relative contribution of the three functional types of DMs to all turns produced by the three speaker groups, including those which were not introduced by any DM. LHD and control speakers produced the lowest percentage of turns without initial DMs: only roughly a fourth of all turns (24% and 26%) had no explicit initial cue for macrostructural integration. With RHD speakers more than a half of all turns were produced without initial DMs (54%). ANOVA

two-way measures on the proportion of turns with and without initial DMs revealed a significant group difference $F(2, 15) = 3.85, p < .001$, with a significant difference as regards the absence of turn-initial DMs between LHD-RHD ($p > .001$) and CG-RHD ($p < .001$).

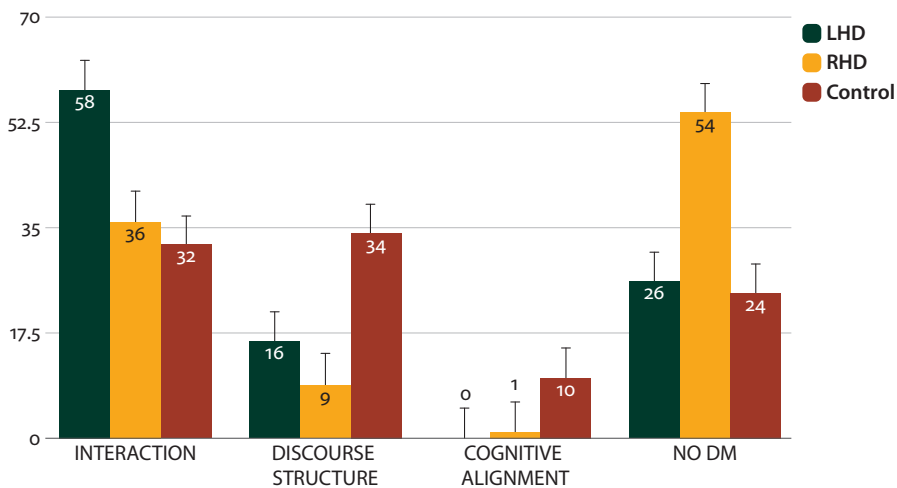


Figure 2. Relative contribution of the three functional types of DMs to all turns produced (in per cent)

As mentioned above, control speakers exhibited the greatest variation in the use of DMs as the different types of DMs are somewhat more evenly distributed, compared to the LHD and RHD speakers. For LHD and RHD speakers interaction management was the most important communicative task at turn-beginnings, given the predominance of DMs serving structuration in the domain of INTERACTION; this tendency was stronger for LHD than RHD speakers, the first of which apparently feel a greater need for showing immediate reactions to a prior speaker's turn.

5. Discussion

The major finding of this study is that RHD speakers used fewer turn-initial DMs than LHD and control speakers in spontaneous speech, which supports findings that an unimpaired RH contributes to the processing of discourse- and interaction-related aspects of language and is strongly associated with structuration on the macrolevel. The observation that aphasic speakers make more frequent use of DMs than RHD speakers is congruent with Kaltenböck's (2020) findings on the uses of the category of "single words" belonging to "Thetical Grammar", such as discourse markers, interjections or vocatives, which were produced by far more

frequently by LHD speakers than by RHD speakers. It is also in line with the large body of evidence for the preservation and the frequent use (often overuse) of “automatic speech” (McElduff & Drummond 1991; Van Lancker Sidtis 2009) with LHD speakers, which includes – next to DMs – overlearned, highly frequent units of speech such as expletives, general extenders or sentence stems (e.g. *I can't, I think*), whose use has often been interpreted as facilitating conversation for aphasic speakers (Oelschlaeger & Damico 1998). The reduced use of DMs by RHD speakers, in turn, coincides with the large body of literature on RH damage and deficits in mental discourse representation and interaction management, as is reflected e.g. in a lack or underrepresentation of discourse connectives, the production of disconnected discourse, and impaired ability to take account of the co-participant and his/her intentions and to appreciate the co-participant's communicative needs (see Section 2). It also supports Wray's (2002) distributed lexical model, in which routine, formulaic, interactional expressions (as well as collocations and memorized texts) are associated with the RH, which operates with a holistic processing mode.

On a language-theoretic level, the findings of this study lend support to the hypothesis that the LH and RH each serve for separate components of language structure: while macrogrammatical abilities appear to be largely preserved after LH damage, RH impairment affects the speakers' ability to structure language on the macrolevel. This observation is fully compatible to other proposals for dual process models of language. Van Lancker Sidtis (2009, 2012), for instance, proposes two modes of linguistic processing and competence, one encompassing the storage and processing of formulaic or fixed expressions, the other one being responsible for the generation of novel utterances. The overlap between microstructures and novel utterances and macrostructures and formulaic speech is not complete, given that macrostructuration is not merely achieved with the use of formulaic units. However, given that structuration of language on the macrolevel includes predominantly formulaic units (e.g. *by the way, I think, you know, and stuff*), most of which serve as devices required for establishing structures outside the microlevel and which tend to be processed holistically and stored as fixed units, the proposals for a dualistic conception of language structure and linguistic processing go into the same direction. A dualistic model offers the advantage of including DMs and other devices serving the structuration of language outside sentence-internal relationships in a broader conception of grammar rather than transferring them to the pragmatic domain, where they form part of a highly diffuse set of “pragmatic markers” with elusive functions, and thus excluding them from a structural description of language.

However, the findings and their interpretation are to be taken with care. First, DMs are, of course, not the only way to establish textual coherence and continuity in interaction, and are by no means required to preserve structural order on the macrolevel. Since speakers with RHD have largely preserved grammatical skills

they are able to use structural devices to establish coherence across turns, such as elliptical responses, or to provide explicit responses (e.g. to yes/no-questions) rather than using response tokens, as illustrated in (5), line 380.

(5)

376 STU: so † how long have you lived on Garner?
 377 PAR: since two thousand two.
 378 STU: nice.
 379 so † you like the area?
 380 → PAR: I do <head-nod:yes>, much better than when I lived in Durham.
 (AphasiaBank: Minga 08)

Yet, the findings show that LHD speakers rely more on the use of DMs in ongoing discourse whereas the RHD group is little consistent in the use of these macro-structuring devices.

It should also be noted that the present study considered only turn-initial DMs and ignored intra-turn DMs, which occur in the speech of both speaker groups, but which are functionally more heterogeneous as they do not only relate the utterance in which they occur to the discourse context and integrate it in a concrete interactional sequence, but may also relate single segments of an emergent turn into a coherent whole.

Secondly, individual variation within the speaker groups with neurological impairment is considerable, especially with RHD speakers, who exhibit a high degree of heterogeneity as regards the precise nature of language-related deficits. In a certain way heterogeneity is, of course, to be expected in any speaker population and cannot be fully eliminated. However, in view of the large discrepancy in the amount of preserved language abilities with neurologically impaired speakers a larger sample would, of course, be desirable, comprising different gender groups, degrees of severity, lesion areas, and a closer consideration of neurological details and post-onset times in order to determine the contribution of each hemisphere to the structuration of language in a more precise way. It is, for instance, still unclear to which extent preserved language-related abilities are due to preserved neural assemblies in the respective damaged hemisphere, or are compensated by the involvement of the other, unimpaired hemisphere. An important question is therefore to which degree macrostructuration involving e.g. the use of different types of DMs by RHD patients is based on residual activities of the RH or on compensatory activities involving the LH. Moreover, hemispheric activities are a matter of degree rather than categorical, that is, many cognitive abilities involve a certain degree of bilateral processing, which is natural since neural networks are not operating in isolated, clearly delimitable areas. This holds especially for language processing, where different operations need to be carried out simultaneously (e.g. phonological,

semantic, syntactic and discourse-related processing or the processing of emotive aspects). Indicators for bilateral processing mechanisms have, for instance, been observed for the processing of idioms, proverbs and proper nouns (Van Lancker 1990; Myers 1999). The precise contribution of each hemisphere and the ways in which processing is mediated by the hemispheres is therefore an important question that needs to be pursued in future research. Moreover, it has repeatedly been argued that other structures, above all basal ganglia and the limbic system, must be considered in the study of formal speech and “pragmatic aspects of language” (Code 2005: 325, Sidtis & Van Sidtis 2018).

Yet, all evidence suggests that the effects of LHD and RHD on macrostructural abilities differ in these sense that these abilities are, to different degrees, impaired after RH damage. While the RH plays a crucial role in processing macrostructural aspects of language, microstructural abilities require at least some residual activity of the LH so that the general associations between different components of language structure and hemispheric activity serve as a meaningful orientation for further research in this field.

6. Conclusion

This study compared the use of turn-initial DMs as devices for the macrostructuration of spoken discourse produced by unilaterally brain-damaged speakers as well as by neurologically unimpaired speakers. The results show significant differences in the use of DMs between speakers with LHD, which do not differ significantly from unimpaired control speakers, and those with RHD. The empirical observation that macrostructuration – as far as the use of DMs is concerned – is the domain in which linguistic skills appear to be largely preserved with LHD speakers, whereas a higher degree of impairment can be observed with RHD speakers, is congruent with the large body of evidence for hemispheric differences in language processing. The findings point to a division of labor between sentence-structuration, or *microgrammar*, which is associated with the LH, and discourse-structuration, or *macrogrammar*, which strongly involves RH activity. In this sense, they have important implications for models of language processing and the processing of DMs in particular, suggesting a major distinction between structures involving morpho-syntactic and semantic relationships within clausal units, and structures involving relationships between utterances and between utterances and the interactional and discursive context in which these are produced. The distinction is certainly far from being clear-cut as both kinds of structures closely interact and given that successful communication requires that, at some point, they converge into one structural

unit in a single act of speaking. In this respect, there is further research potential as regards the different ways in which micro- and macrostructuration complement and affect each other.

Transcription conventions

[]	overlap and simultaneous talk
< >	comment by the transcriber
‡	intonation break
[/]	retracing without correction (e.g. simple repetition)
[//]	retracing with correction (e.g. simple word or grammar change)
=	latching
(.)	micropause
(2.0)	measured pause
;, ::	segmental lengthening according to duration
rea(hh)lly	laugh particles within talk
ABsolutely	strong, primary stress via loudness
<u>really</u>	stress via pitch or amplitude
wo:rd	lengthening
w-	cut off
°word°	produced softer than surrounding talk
.	falling intonation (terminal pitch)
,	continuing intonation
?	rising intonation
¿	a rise stronger than mid-level but weaker than high-terminal pitch

References

- Aijmer, Karin. 2002. *English Discourse Particles. Evidence from a Corpus* [Studies in Corpus Linguistics 10]. Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.10>
- Ameka, Felix. 1992. Interjections: The universal yet neglected part of speech. *Journal of Pragmatics* 18: 101–118. [https://doi.org/10.1016/0378-2166\(92\)90048-G](https://doi.org/10.1016/0378-2166(92)90048-G)
- Barnes, Scott, Toocaram, Sophie, Nickels, Lyndsey, Beeke, Suzanne, Best, Wendy & Bloch, Steven. 2019. Everyday conversation after right hemisphere damage: A methodological demonstration and some preliminary findings. *Journal of Neurolinguistics* 52: 1–16. <https://doi.org/10.1016/j.jneuroling.2019.100850>
- Bartels-Tobin, Lori R. & Hinckley, Jacqueline. 2005. Cognition and discourse production in right hemisphere disorder. *Journal of Neurolinguistics* 18: 461–477. <https://doi.org/10.1016/j.jneuroling.2005.04.001>
- Beeching, Kate & Detges Ulrich (eds). 2014. *Discourse Functions at the Right and Left Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Leiden: Brill. <https://doi.org/10.1163/9789004274822>

- Beeman, Mark & Chiarello, Christine. 1998. Complementary right- and left-hemisphere language comprehension. *Current Directions in Psychological Science* 7(1): 1–8. <https://doi.org/10.1111/1467-8721.ep11521805>
- Berndt, Rita S. 1987. Symptom co-occurrence and dissociation in the interpretation of agrammatism. In *The Cognitive Neuropsychology of Language*, Max Coltheart, Giuseppe Sartori & Remo Job (eds), 221–233. Hillsdale NJ: Lawrence Erlbaum Associates.
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Blake, Margaret Lehman. 2009. Inferencing processes after right hemisphere brain damage: Effects of contextual bias. *Journal of Speech, Language and Hearing Research* 52(2): 373–384. [https://doi.org/10.1044/1092-4388\(2009/07-0172\)](https://doi.org/10.1044/1092-4388(2009/07-0172))
- Blakemore, Diane. 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511486456>
- Blanche-Benveniste, Claire, Bilger, Mireille, Rouget, Christine & Van den Eynde, Karel. 1990. *Le Français Parlé: Études Grammaticales*. Paris: Éditions du CNRS.
- Borod, Joan C., Bloom, Ronald, Brickman, Adam, Nakhutina, Luba & Curko, Elizabeth. 2002. Emotional processing deficits in individuals with unilateral brain damage. *Applied Neuropsychology* 9(1): 23–36. https://doi.org/10.1207/S15324826AN0901_4
- Bottini, Gabriella, Corcoran, Rhiannon, Sterzi, Roberto, Paulesu, Eraldo, Schenone, Pietro, Scarpa, Pina, Frackowiak, Richard & Frith, Chris D. 1994. The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain* 117: 1241–1253. <https://doi.org/10.1093/brain/117.6.1241>
- Boye, Kasper & Bastiaanse, Roelien. 2018. Grammatical versus lexical words in theory and aphasia: Integrating linguistics and neurolinguistics. *Glossa: A Journal of General Linguistics* 3(1): 29. <https://doi.org/10.5334/gjgl.436>
- Brady, Marian, Armstrong, Linda & Mackenzie, Catherine. 2006. An examination over time of language and discourse production abilities following right hemisphere brain damage. *Journal of Neurolinguistics* 19(4): 291–310. <https://doi.org/10.1016/j.jneuroling.2005.12.001>
- Brownell, Hiram H. & Joanette, Yves (eds). 1993. *Narrative Discourse in Neurological Impaired and Normal Aging Adults*. San Diego CA: Singular.
- Butterworth, Brian. 1994. Disorders of sentence production. *Philosophical Transactions of the Royal Society of London B* 346: 55–61. <https://doi.org/10.1098/rstb.1994.0128>
- Bybee, Joan L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–455. <https://doi.org/10.1080/01690969508407111>
- Caplan, Rochelle & Dapretto, Mirella. 2001. Making sense during conversation: An fMRI study. *Neuroreport* 12(16): 3625–3632. <https://doi.org/10.1097/00001756-200111160-00050>
- Centeno, José & Obler, Loraine. 2001. Agrammatic verb errors in Spanish speakers and their normal discourse correlates. *Journal of Neurolinguistics* 14(2): 349–363. [https://doi.org/10.1016/S0911-6044\(01\)00023-9](https://doi.org/10.1016/S0911-6044(01)00023-9)
- Champagne-Lavau, Maud & Joanette, Yves. 2009. Pragmatics, theory of mind and executive functions after a right-hemisphere lesion: Different patterns of deficits. *Journal of Neurolinguistics* 22: 413–426. <https://doi.org/10.1016/j.jneuroling.2009.02.002>
- Chantraine, Yves, Joanette, Yves & Ska, Bernadette. 1998. Conversational abilities in patients with right hemisphere damage. In *Pragmatics in Neurogenic Communication Disorders*, Michel Paradis (ed.), 21–32. Oxford: Pergamon Press.

- Code, Chris. 1996. Speech from the isolated right hemisphere? Left hemispherectomy cases E. G. and N. F. In: *Classic Cases in Neuropsychology*, Vol. 1, Chris Code, Claus-W. Wallesch, Yves Joanette, André Roch Lecours (eds), 319–336. Hove: Psychology Press.
- Code, Chris. 1997. Can the right hemisphere speak? *Brain and Language* 57: 38–59.
<https://doi.org/10.1006/brln.1997.1833>
- Code, Chris. 2005. First in, last out? The evolution of aphasic lexical speech automatisms to agrammatism and the evolution of human communication. *Interaction Studies* 6: 311–334.
<https://doi.org/10.1075/is.6.2.08cod>
- Cowell, Simon F., Egan, Gary, Code, Chris, Harasty, Jenny & Watson, John. 2000. The functional neuroanatomy of simple calculation and number repetition: A parametric PET activation study. *Neuroimage* 12: 565–573. <https://doi.org/10.1006/nimg.2000.0640>
- Crible, Ludivine. 2017. Discourse markers and (dis)fluencies in English and French: Variation and combination in the DisFrEn corpus. *International Journal of Corpus Linguistics* 22(2): 242–269. <https://doi.org/10.1075/ijcl.22.2.04cri>
- Crible, Ludivine & Cuenca, María-Josep. 2017. Discourse markers in speech: Characteristics and challenges for corpus annotation. *Dialogue and Discourse* 8(2): 149–166.
<https://doi.org/10.5087/dad.2017.207>
- Crible, Ludivine & Degand, Liesbeth. 2019. Reliability vs. granularity in discourse annotation: What is the trade-off? *Corpus Linguistics and Linguistic Theory* 15(1): 71–99.
<https://doi.org/10.1515/cllt-2016-0046>
- Cuenca, María Josep & Crible, Ludivine. 2019. Co-occurrence of discourse markers in English: From juxtaposition to composition. *Journal of Pragmatics* 140: 171–184.
<https://doi.org/10.1016/j.pragma.2018.12.001>
- Davis, Albyn, O'Neill-Pirozzi, Therese & Coon, Maribeth. 1997. Referential cohesion and logical coherence of narration after right hemisphere stroke. *Brain and Language* 56: 183–210.
<https://doi.org/10.1006/brln.1997.1741>
- Debaeisieux, Jeanne-Marie. 2007. La distinction entre dépendance grammaticale et dépendance macrosyntaxique comme moyen de résoudre les paradoxes de la subordination. *Faits de Langue* 28: 119–132.
- Degand, Liesbeth & Simon, Anne-Catherine. 2009. On identifying basic discourse units in speech: Theoretical and empirical issues. *Discours* 4. <https://doi.org/10.4000/discours.5852>
- Deulofeu, José. 2017. La macrosyntaxe comme moyen de tracer la limite entre organisation grammaticale et organisation du discours. *Modèles Linguistiques* 74: 135–166.
<https://doi.org/10.4000/ml.2040>
- Devinsky, Orrin. 2000. Right cerebral hemisphere dominance for a sense of corporeal and emotional self. *Epilepsy and Behavior* 1(1): 60–73. <https://doi.org/10.1006/ebeh.2000.0025>
- Diewald, Gabriele. 2013. 'Same same but different' – Modal particles, discourse markers and the art (and purpose) of categorization. In *Discourse Markers and Modal Particles. Categorization and Description* [Pragmatics & Beyond New Series 234], Liesbeth Degand, Paola Pietrandrea & Bert Cornillie (eds), 19–46. Amsterdam: John Benjamins.
<https://doi.org/10.1075/pbns.234.02die>
- Dik, Simon C. 1997. *The Theory of Functional Grammar*, Part 2. Berlin: Mouton de Gruyter.
- Fischer, Kerstin. 2000. *From Cognitive Semantics to Lexical Pragmatics: The Functional Polysemy of Discourse Particles*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110828641>
- Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 31(7): 931–952.
[https://doi.org/10.1016/S0378-2166\(98\)00101-5](https://doi.org/10.1016/S0378-2166(98)00101-5)

- Fraser, Bruce. 2015. The combining of discourse markers – A beginning. *Journal of Pragmatics* 86: 48–53. <https://doi.org/10.1016/j.pragma.2015.06.007>
- Friederici, Angela. 2004. The neural basis of syntactic processes. In *The Cognitive Neurosciences*, Michael S. Gazzaniga (ed.), 789–801. Cambridge MA: The MIT Press.
- Friederici, Angela, Rüschemeyer, Shirley-Ann, Hahne, Anja & Fiebach, Christian J. 2003. The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex* 13: 170–177. <https://doi.org/10.1093/cercor/13.2.170>
- Friederici, Angela & Alter, Kai. 2004. Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language* 89(2): 267–276. [https://doi.org/10.1016/S0093-934X\(03\)00351-1](https://doi.org/10.1016/S0093-934X(03)00351-1)
- Friederici, Angela, Bahlmann, Jörg, Heim, Stefan, Schubotz, Ricarda & Anwander, Alfred. 2006. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences* 103(7): 2458–2463. <https://doi.org/10.1073/pnas.0509389103>
- Gernsbacher, Morton. 1990. *Language Comprehension as Structure Building*. Hillsdale, NJ: Erlbaum.
- González, Momtserrat. 2005. Pragmatic markers and discourse coherence relations in English and Catalan oral narrative. *Discourse Studies* 77(1): 53–86. <https://doi.org/10.1177/1461445605048767>
- Graesser, Arthur C., Singer, Murray & Trabasso, Tom. 1994. Constructing inferences during narrative text comprehension. *Psychological Review* 101, 371–395.
- Greene, Steven B., McKoon, Gail & Ratcliff, Roger. 1992. Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory and Cognition* 18, 266–283.
- Halliday, Michael A. K. 1985. *An Introduction to Functional Grammar*. London: Arnold.
- Halliday, Michael A. K. & Matthiessen, Christian M. I. M. 2004. *An Introduction to Functional Grammar*. London: Hodder Arnold.
- Hannay, Mike & Kroon, Caroline. 2005. Acts and the relationship between discourse and grammar. *Functions of Language* 12: 87–124. <https://doi.org/10.1075/fol.12.1.05han>
- Hansen, Maj-Britt Mosegaard. 2006. A dynamic polysemy approach to the lexical semantics of discourse markers (with an exemplary analysis of French *toujours*). In *Approaches to Discourse Particles*, Kerstin Fischer (ed.), 21–41. Amsterdam: Elsevier.
- Haselow, Alexander. 2017. *Spontaneous Spoken English. An Integrated Approach to the Emergent Grammar of Speech*. Cambridge: CUP. <https://doi.org/10.1017/9781108265089>
- Haselow, Alexander. 2019. Discourse marker sequences: Insights into the serial order of communicative tasks in real-time turn production. *Journal of Pragmatics* 146: 1–18. <https://doi.org/10.1016/j.pragma.2019.04.003>
- Haselow, Alexander & Kaltenböck, Gunther (eds). 2020. *Grammar and Cognition: Dualistic Models of Language Structure and Language Processing* [Human Cognitive Processing 70]. Amsterdam: John Benjamins. <https://doi.org/10.1075/hcp.70>
- Heine, Bernd. 2019. Some observations on the dualistic nature of discourse processing. *Folia Linguistica* 53(2): 411–442. <https://doi.org/10.1515/flin-2019-2016>
- Heine, Bernd, Kuteva, Tania & Kaltenböck, Gunther. 2014. Discourse grammar, the dual process model, and brain lateralization: Some correlations. *Language & Cognition* 6: 146–180. <https://doi.org/10.1017/langcog.2013.3>
- Heine, Bernd, Kuteva, Tania, Kaltenböck, Gunther & Long, Haiping. 2015. On some correlation between grammar and brain lateralization. *Oxford Handbooks Online*. Oxford: OUP. <https://doi.org/10.1093/oxfordhb/9780199935345.013.16>

- Heine, Bernd, Kuteva, Tania & Long, Haiping. 2020. Dual process frameworks on reasoning and linguistic discourse. In Haselow & Kaltenböck (eds), 59–89.
- Helasuvo, Marja-Liisa, Klippi, Anu & Laakso, Minna. 2001. Grammatical structuring in Broca's and Wernicke's aphasia in Finnish. *Journal of Neurolinguistics* 14(2): 231–254. [https://doi.org/10.1016/S0911-6044\(01\)00016-1](https://doi.org/10.1016/S0911-6044(01)00016-1)
- Hird, Kathryn & Kirsner, Kim. 2003. The effect of right cerebral hemisphere damage on collaborative planning in conversation: An analysis of intentional structure. *Clinical Linguistics and Phonetics* 17(4–5): 309–315. <https://doi.org/10.1080/0269920031000080037>
- Howard, David & Orchard-Lisle, Virginia. 1984. On the origin of semantic errors in naming: Evidence from the case of a global aphasic. *Cognitive Neuropsychology* 1(2): 163–190. <https://doi.org/10.1080/02643298408252021>
- Howes, D. 1964. Application of the word frequency concept to aphasia. In *Disorders of Language*, Anthony V. S. DeReuck & Maeve O'Connor (eds), 47–75. London: Churchill.
- Huddleston, Rodney & Pullum, Geoffrey. 2002. *Language Description: The Cambridge Grammar of the English Language*. Cambridge: CUP. <https://doi.org/10.1017/9781316423530>
- Kaltenböck, Gunther. 2020. Formulaic language and Discourse Grammar: Evidence from speech disorder. In *Grammar and Cognition: Dualistic Models of Language Structure and Language Processing* [Human Cognitive Processing 70], Alexander Haselow & Gunther Kaltenböck (eds), 233–266. Amsterdam: John Benjamins. <https://doi.org/10.1075/hcp.70.07kal>
- Kaltenböck, Gunther, Heine, Bernd & Kuteva, Tania. 2011. On thetical grammar. *Studies in Language* 35: 852–897. <https://doi.org/10.1075/sl.35.4.03kal>
- Keizer, Evelien. 2015. *A Functional Discourse Grammar for English*. Oxford: OUP.
- Kennedy, Mary, Strand, Edythe A., Burton, Wendy & Peterson, Connie. 1994. Analysis of first-encounter conversations of right hemisphere damaged participants. *Clinical Aphasiology* 22: 67–80.
- Kintsch, Walter. 1988. The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review* 95, 163–182.
- Lehman-Blake, Margaret. 2006. Clinical relevance of discourse characteristics after right hemisphere brain damage. *American Journal of Speech-Language Pathology* 15(3): 255–267. [https://doi.org/10.1044/1058-0360\(2006\)024](https://doi.org/10.1044/1058-0360(2006)024)
- Lehman Blake, Margaret. 2010. Communication deficits associated with right hemisphere brain damage. In *The Handbook of Language and Speech Disorders*, Jack S. Damico, Nicole Muller & Martin J. Ball (eds), 556–576. Chichester: Wiley-Blackwell. <https://doi.org/10.1002/9781444318975.ch24>
- Long, Debra, Baynes, Kathleen, & Prat, Chantel. 2005. The propositional structure of discourse in the two cerebral hemispheres. *Brain and Language* 95(3), 383–394. <https://doi.org/10.1016/j.bandl.2005.02.004>
- Mackenzie, Catherine & Marian Brady. 2008. Communication difficulties following right hemisphere stroke: applying evidence to clinical management. *Evidence-Based Communication Assessment and Intervention* 2 (4), 235–247. <https://doi.org/10.1080/17489530802615336>
- Maschler, Yael. 2009. *Metalanguage in Interaction: Hebrew Discourse Markers* [Pragmatics & Beyond New Series 181]. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.181>
- Marini, Andrea, Carlomagno, Sergio, Caltagirone, Carlo & Nocentini, Ugo. 2005. The role played by the RH in the organization of complex textual structures. *Brain and Language* 93: 46–54. <https://doi.org/10.1016/j.bandl.2004.08.002>

- Marini, Andrea. 2012. Characteristics of narrative discourse processing after damage to the right hemisphere. *Seminars in Speech and Language* 33(1): 68–78.
<https://doi.org/10.1055/s-0031-1301164>
- McDonald, Skye. 1999. Exploring the process of inference generation in sarcasm: A review of normal and clinical studies. *Brain and Language* 68(3): 486–506.
<https://doi.org/10.1006/brln.1999.2124>
- MacWhinney, Brian, Fromm, Davida, Forbes, Margret & Holland, Audrey. 2011. AphasiaBank: Methods for studying discourse. *Aphasiology* 25: 1286–1307.
<https://doi.org/10.1080/02687038.2011.589893>
- McEldruff, Kathleen & Drummond, Sakina. 1991. Communication functions of automatic speech in non fluent aphasia. *Aphasiology* 5: 265–278. <https://doi.org/10.1080/02687039108248528>
- McKoon, Gail & Ratcliff, Roger. 1992. Inference during reading. *Psychological Review* 99, 440–466.
- Mitchell, Rachel L. C. & Crow, Tim J. 2005. Right hemisphere language functions and schizophrenia: The forgotten hemisphere? *Brain* 128(5): 963–978. <https://doi.org/10.1093/brain/awh466>
- Myers, Penelope S. 1994. Communication disorders associated with right-hemisphere brain damage. In *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders*, 3d edn, Roberta Chapey (ed.), 514–534. Baltimore MD: Williams & Wilkins.
- Myers, Penelope S. 1999. *Right Hemisphere Damage: Disorders of Communication and Cognition*. San Diego CA: Singular.
- Myers, Penelope S. 2001. Communication disorders associated with right hemispheredamage. In *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders*, 4th edn, Roberta Chapey (ed.), 963–987. Baltimore MD: Lippincott Williams and Wilkins.
- Oelschlaeger, Mary & Damico, Jack S. 1998. Spontaneous verbal repetition: A social strategy in aphasic conversation. *Aphasiology* 12: 971–988. <https://doi.org/10.1080/02687039808249464>
- Overstreet, Maryann. 2014. The role of pragmatic function in the grammaticalization of English general extenders. *Pragmatics* 24(1), 105–129. <https://doi.org/10.1075/prag.24.1.05ove>
- Pallier, Christophe, Devauchelle, Anne-Dominique & Dehaene, Stanislas. 2011. Cortical representation of the constituents structure of sentences. *PNAS* 108(6): 2522–2527.
<https://doi.org/10.1073/pnas.1018711108>
- Pawley, Andrew. 2009. Grammarians’ languages versus humanists’ languages and the place of speech act formulas in models of linguistic competence. In *Formulaic Language, Vol. 1: Distribution and Historical Change* [Typological Studies in Language 82], Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali & Kathleen M. Wheatley (eds), 3–26. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.82.01gra>
- Pedersen, Palle, Vinter, Kirsten & Olsen, Tom S. 2004. Aphasia after stroke: Type, severity and prognosis. The Copenhagen aphasia study. *Cerebrovascular Diseases* 17(1): 35–43.
<https://doi.org/10.1159/000073896>
- Pons Bordería, Salvador. 2018. The combination of discourse markers in spontaneous conversations: Keys to undo a gordian knot. *Revue Romane* 53(1): 121–158.
<https://doi.org/10.1075/rro.00008.pon>
- Prat, Chantel S., Long, Debra L. & Baynes, Kathleen. 2007. The representation of discourse in the two hemispheres: An individual differences investigation. *Brain and Language* 100(3): 283–294. <https://doi.org/10.1016/j.bandl.2006.11.002>

- Purdy, Mary H. 2002. Script knowledge following stroke. *Journal of Medical Speech Language Pathology* 10(3): 173–181.
- Redeker, Gisela. 2006. Discourse markers as attentional cues at discourse transitions. In *Approaches to Discourse Particles*, Kerstin Fischer (ed.), 339–358. Amsterdam: Elsevier.
- Sakai, Kuniyoshi L., Tatsuno, Yoshinori, Suzuki, Kei, Kimura, Harumi & Ichida, Yasuhiro. 2005. Sign and speech: Amodal commonality in left hemisphere dominance for comprehension of sentences. *Brain* 128(6): 1407–1417. <https://doi.org/10.1093/brain/awh465>
- Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511611841>
- Schnitzer, Marc L. 1989. *The Pragmatic Basis of Aphasia*. Hillsdale NJ: Lawrence Erlbaum Associates.
- Schourup, Lawrence. 1985. *Common Discourse Particles in English Conversation*. New York NY: Garland.
- Schourup, Lawrence. 1999. Discourse markers. *Lingua* 107: 227–265.
[https://doi.org/10.1016/S0024-3841\(96\)90026-1](https://doi.org/10.1016/S0024-3841(96)90026-1)
- Sherrat, Sue & Penn, Claire. 1990. Discourse in a right-hemisphere brain-damaged subject. *Aphasiology* 4(6): 539–560. <https://doi.org/10.1080/02687039008248506>
- Sherrat, Sue & Bryan, Karen. 2012. Discourse production after right brain damage: Gaining a comprehensive picture using a multi-level processing model. *Journal of Neurolinguistics* 25: 213–239. <https://doi.org/10.1016/j.jneuroling.2012.01.001>
- Squire, Larry R. 2004. Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory* 82: 171–177. <https://doi.org/10.1016/j.nlm.2004.06.005>
- Stalnaker, Robert. 2002. Common ground. *Linguistic Philosophy* 25(5–6): 701–721.
<https://doi.org/10.1023/A:1020867916902>
- Steen, Gerard. 2005. Basic discourse acts: Towards a psychological theory of discourse segmentation. In *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction* [Cognitive Linguistics Research 32], M. Sandra Peña Cervel & Francisco Ruiz de Mendoza Ibáñez (eds), 283–312. Berlin: Mouton de Gruyter.
- Stemberger, Joseph P. & MacWhinney, Brian. 1986. Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition* 14: 17–26. <https://doi.org/10.3758/BF03209225>
- Tompkins, Connie A. 1995. *Right Hemisphere Communication Disorders: Theory and Management*. San Diego CA: Singular.
- Tompkins, Connie A. 2008. Theoretical considerations for understanding “Understanding” by adults with right hemisphere brain damage. *Perspectives on Neurophysiology and Neurogenetic Speech and Language Disorders* 18(2): 45–54. <https://doi.org/10.1044/nnsld18.2.45>
- Ullman, Michael T. 2004. Contributions of memory circuits to language: The declarative/procedural model. *Cognition* 92(1–2): 231–270. <https://doi.org/10.1016/j.cognition.2003.10.008>
- Ullman, Michael T. 2015. The declarative/procedural model: A neurobiological model of language learning, knowledge, and use. In *Neurobiology of Language*, Gregory Hickok & Steven L. Small (eds), 953–968. Amsterdam: Elsevier.
- van Dijk, Teun. 1980. *Macrostructures. An Interdisciplinary Study of Global Structures in Discourse, Interaction and Cognition*. Hillsdale, NJ: Lawrence Erlbaum.
- Van Lancker Sidtis, Diana. 2001. Preserved formulaic expressions in a case of transcortical sensory aphasia compared to incidence in normal everyday speech. *Brain and Language* 79(1): 38–41.

- Van Lancker Sidtis, Diana. 2004. When novel sentences spoken or heard for the first time in the history of the universe are not enough: Toward a dual-process model of language. *International Journal of Language and Communication Disorders* 39: 1–44.
<https://doi.org/10.1080/13682820310001601080>
- Van Lancker Sidtis, Diana. 2009. Formulaic and novel language in a ‘dual process’ model of language competence: Evidence from surveys, speech samples, and schemata. In *Formulaic Language, Vol. 2: Acquisition, Loss, Psychological Reality, and Functional Explanations* [Typological Studies in Language 83], Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali & Kathleen M. Wheatley (eds), 445–470. Amsterdam: John Benjamins.
<https://doi.org/10.1075/tsl.83.11van>
- Van Lancker Sidtis, Diana. 2012. Formulaic language and language disorders. *Annual Review of Applied Linguistics* 32: 62–80. <https://doi.org/10.1017/S0267190512000104>
- Van Lancker Sidtis, Diana & Rallon, Gail. 2004. Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language and Communication* 24: 207–240. <https://doi.org/10.1016/j.langcom.2004.02.003>
- Van Lancker Sidtis, Diana & Postman, Whitney A. 2006. Formulaic expressions in spontaneous speech of left- and right-hemisphere-damaged subjects. *Aphasiology* 20(5): 411–426.
<https://doi.org/10.1080/02687030500538148>
- Van Lancker Sidtis, Diana & Sidtis, John. 2018. The affective nature of formulaic language: A right-hemisphere subcortical process. *Frontiers in Neurology* 9: 573.
<https://doi.org/10.3389/fneur.2018.00573>
- Wray, Alison. 2002. *Formulaic Language and the Lexicon*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511519772>

Vietnamese expletive between grammatical subject and subjectivity marker

Nó at the syntax-pragmatics (discourse) interface

Huy Linh Dao

INALCO & CNRS-EHESS, CRLAO/UMR 8563

In this paper, we undertake a systematic investigation of the peculiar behavior of the optional expletive subject pronoun *nó*_{EXPL} in spoken Vietnamese. The first part of our study consists of an in-depth examination of its syntactic and semantic properties in comparison to those of its referential counterpart. From this overview, it appears that the distribution of *nó*_{EXPL} is jointly determined by syntactic and discourse factors. We then critically appraise the recent analysis by Greco *et al.* (2018a, b), according to which *nó*_{EXPL} imposes a specificity requirement on sentences containing it. We suggest to reinterpret this requirement as a corollary of the assessment by the speaker of a given state of affairs against her epistemic background. Pursuing the line of argumentation developed in Dao (2014) and integrating insights from Speas-Tenny's syntactic approach to Evidentiality (Speas & Tenny 2003; Speas 2004; Tenny 2006), we argue that *nó*_{EXPL} is best treated as an ego-evidential marker. We further show that *nó*_{EXPL} does not intrinsically encode evidentiality but acquires its evidential import via syntactic processes.

Keywords: ego-evidentiality, optional expletive subject, subjectivity marker, Vietnamese, syntax-pragmatics (discourse) interface

1. Introduction

The present study is built upon our preliminary work on expletive subject in Vietnamese (Dao 2014). More precisely, it tackles the syntactic, semantic and discursive characterization of the optional expletive subject pronoun *nó* (*nó*_{EXPL}, hereafter) in modern Vietnamese, an issue which, to the best of our knowledge, has received scant attention in the linguistic literature until recently (cf. Nguyen & Hoang 2011; Dao 2013; Greco *et al.* 2018a, b). Some of these rare studies, albeit

insightful, have primarily focused on the internal properties of *nó*_{EXPL}. For instance, Nguyen & Hoang (2011, hereafter N&N), while highlighting its semantic “emptiness” and its non-referential status, argue that this morpheme, whose usage seems to be confined to spoken language and informal styles (Example 2), must be syntactically and semantically distinguished from its homonymous form – the third singular personal pronoun *nó* (*nó*_{REF}, hereafter) (Example 1).

- (1) a. *Nó mới gặp Mai hôm qua.*
 3SG just see Nprop yesterday
 ‘(S)he just saw Mai yesterday.’
- b. *Mai không muốn nói chuyện với nó.*
 Nprop NEG want talk with 3SG
 ‘Mai doesn’t want to talk to him/her.’
- (2) *Mùa hè rồi! Mặc quần soóc cho (nó) mát!*
 Summer CRS wear shorts so that EXPL be-cool/fresh
 ‘It’s summer already! (We/you) should wear shorts, so that (we/you) will feel fresh/it will freshen up your/our body/it will be fresh.’

On the basis of this type of examples, N&N further claim that the notion of (*grammatical*) *subject* should be maintained in the description of Vietnamese, despite the fact that this language has often been considered a *topic-prominent* language, as is proposed in Li & Thompson’s (1976, 1981) typology (see also Cao 2004), or a *radical pro-drop* language (cf. Huang 1984) without agreement marking on the verb and where arguments can be omitted quite freely. Taking another angle of attack, Dao (2013) points out that although compatible with unaccusative (but not unergative¹ or transitive) verbs, *nó*_{EXPL} is nonetheless limited to inchoative and *irrealis*-resultative contexts (involving change-of-state predicates):

- (3) *Mày làm thế, (nó) vỡ cái bình bây giờ!*
 2SG do so EXPL break_{INTRANS} CLF vase now
 ‘If you do that, the (flower) vase will definitely break!’
- (4) a. *Mày làm thế, thằng bé khóc bây giờ!*
 2SG do so CLF little cry now
 ‘If you do that, the boy will definitely cry!’
- b. * *Mày làm thế, nó thằng bé khóc bây giờ!*
 2SG do so EXPL CLF little cry now
 ‘If you do that, the boy will definitely cry!’

1. *Nó* can combine with unergative ones only if those verbs undergo unaccusativization when provided with suitable lexical material (Dao 2015).

- c. ? *Mày làm thế, nó khóc thành bé bây giờ!*²
 2SG do so EXPL cry CLF little now
 ‘If you do that, the boy will definitely cry!’
- (5) *Mày làm thế, (*nó) mẹ (*nó) sẽ mắng mày đấy!*
 2SG do so EXPL mother EXPL FUT scold 2SG PART
 ‘If you do that, Mom will scold you for sure!’

Those observations are indeed very interesting. Nevertheless, Dao does not give any explanation for this contextual restriction. Moreover, neither N&N (2011) nor Dao (2013) attempt to address questions such as:

- i. Why *nó*_{EXPL} is primarily licensed in impromptu speech?
- ii. Why *nó*_{EXPL} cannot freely co-occur with weather-verbs (Example 6a) but needs to be anchored in a larger (discursive and linguistic) context to be licensed (Example 6b)?
- iii. Why *nó*_{EXPL} requires overt realization of some linguistic material (which can be lexical or grammatical elements, and in some cases, just prosodic cues) or why sentences featuring it are felicitous only if they are interpreted as being indexically linked to the context of utterance (cf. the ungrammaticality of (7a) and the ill-formedness of (8) due to the absence of the sequences in parentheses)?
- iv. Why *nó*_{EXPL} is perfectly acceptable in *realis* contexts (as shown in (8))?

- (6) a. *(*Nó)/(trời) mưa.*³
 EXPL/sky rain
 ‘It is raining.’
- b. *(Nó)/(trời) sắp mưa rồi đấy.*
 EXPL/sky be about to rain CRS PART
 ‘It is about to rain.’
- (7) a. *(Nó)⁴ cháy ngôi nhà *(rồi).*
 EXPL burnt down CLF house CRS
 ‘The house is burning/has burnt down.’

2. For some of our consultants, (4c) is totally fine since the DP *thằng bé* ‘the boy’ is not interpreted as a genuine Agent argument of *khóc* ‘cry’. This appears to be confirmed by the fact that replacing *thằng bé* by *bà cụ* ‘the old lady’ results in illformedness (see Dao 2015 for further discussion).

3. Note in passing that since Vietnamese is an isolating language, the distinction between verbs and nouns is not morphologically encoded. The categorial membership of a given lexical item largely depends on the context in which it is inserted. The form *mưa* is thus ambiguous between the nominal reading (*the rain*) and the verbal one (*to rain*).

4. The examples in (7) and (8) are perfectly natural (even without additional material) if *nó* is used as a referential pronoun, in which case it is interpreted as the *external possessor* of the postverbal DP (see Dao 2015).

b. *Nhìn kia!* (Nó) cháy ngôi nhà rồi.

Look there EXPL burnt down CLF house CRS

‘Look over there! The house is burning/has burnt down.’

(8) *(*Tại nó ăn nhiều đồ béo nên nó* *(*mới*) béo bụng!

Because 3SG eat much stuff greasy so EXPL EMPH fatten belly

‘It is exactly because (s)he ate so much greasy food that her/his belly got fat!’

In the following pages, we will try to answer those questions by showing that *nó* is syntactically a (grammatical) subject but, pragmatically, a subjectivity marker. More specifically, we will lay out an analysis in which *nó*_{EXPL} is intimately related to the speaker and arguably connected to the crucial notion of *evidentiality*. This then amounts to saying that *nó* turns out to be more “indexical” than, or at least not as expletive as, it might appear to be at first sight.

This paper is organized as follows. In Section 2 we provide a detailed description of the main syntactic and semantic properties displayed by *nó*_{EXPL} in comparison with its referential counterpart. In Section 3 we critically assess the recent account by Greco *et al.* (2018a, b) which has also attempted to capture the atypical behavior of *nó*_{EXPL} by pointing out that not only does it occupy a position in the high inflectional layer of the clausal skeleton, but it also imposes a specificity requirement on the eventive content conveyed in the clause where it appears. In Section 4 we briefly recall the gist of our proposal put forward in Dao (2014), on the basis of which we will try to outline a novel formal analysis of *nó*_{EXPL}. In our approach, although *nó*_{EXPL} is externally merged (= base-generated) in a position somewhere in the highest portion of the inflectional layer, as claimed by Greco *et al.*, it can nevertheless move higher to the left periphery via Internal Merge. Moreover, it will be argued to be linked to a functional projection which is situated within this area and which encodes evidentiality. Section 5 closes the paper with some concluding remarks.

2. The syntax and semantics of *nó*_{EXPL}

2.1 Referential *nó*

Let us start by comparing the syntactic properties of *nó*_{EXPL} and its referential counterpart *nó*_{REF}. As illustrated in Examples (1a, b), reproduced below as (9a, b), the latter behaves as a true pronoun and exhibits the same distribution as an ordinary DP. Consequently, it can appear in a full range of grammatical positions: subject (9a), complement of a preposition (9b), direct object (9c).

- (9) a. *Nó mới gặp Mai hôm qua.*
3SG just see Nprop yesterday
'(S)he just saw Mai yesterday.'
- b. *Mai không muốn nói chuyện với nó.*
Nprop NEG want talk with 3SG
'Mai doesn't want to talk to him/her.'
- c. *Họ đã tố cáo nó trước hội đồng chống đạo văn.*
3PL PERF denounce 3SG before Council anti-plagiarism
'They have denounced her/him before the plagiarism Council.'

Furthermore, like other referential pronouns, *nó*_{REF} can be modified by a PP (10a); can be coordinated with another DP or regular pronoun (10b); can undergo topicalization (10c) or occur within the scope of focus sensitive particles such as *chính* 'exactly', *ngay cả* 'even', *chỉ* 'only' (10d). By the same token, *nó*_{REF} can be directly modified by the reflexive elements *bản thân* 'self' and *tự* 'REFL' (10e).

- (10) a. *Nó/tôi/thằng bé của ngày xưa / trong kí ức tôi đã chết rồi.*
3SG/1SG/CLF little of the old days in memories 1SG PERF
die CRS
'The him-her/me/boy of the old days/in my memories has died.'
- b. *Cả nó và tôi/đồng nghiệp cũng thấy Trang cư xử xấu.*
Both 3SG and 1SG/colleague also find Nprop behave bad
'Both (s)he and I/his-her colleague(s) also find that Trang behaved badly.'
- c. *Nó thì/là tôi cho rằng không khiêm tốn.*
3SG TOP/TOP 1SG suppose COMP NEG humble
'As for him-her, I suppose that (s)he is not humble.'
- d. *Chính/ngay cả/chỉ nó thừa nhận đã đạo văn.*
Exactly/even/only (s)he admit PERF plagiarize
'Exactly/even/only (s)he admitted having plagiarized.'
- e. *Bản thân/tự nó/mày/tác giả đó không dám nhận lỗi.*
Self/REFL 3SG/2SG/author DEM NEG dare admit mistake
'(S)he/you/this author himself-herself/yourself didn't dare to admit mistakes.'

In addition to being able to serve as an anaphoric pronoun (11a), *nó*_{REF} shows sensitivity to strong (11b) and weak (11c) crossover effects and can be ambiguous between a strict and a sloppy reading in elliptical construction (11d).

- (11) a. *Tôi gặp con⁵ Phan_i hôm qua. Nó_i lại ăn trộm.*
 1SG meet CLF Nprop yesterday 3SG again steal
 ‘I met Phan_i yesterday. She_i had stolen again.’
- b. *Thằng bé nào_i thì nó⁶_{i/j} nghĩ thầy cô phải khen⁶ ______i*
 CLF little which TOP 3SG think teachers must praise
nhiều hơn?
 much more
 ‘Which boy_i does he_{i/j} think that the teachers must praise ______i more?’
- c. *Thằng bé nào_i thì mẹ nó⁶_{i/j} sẽ mắng?*
 CLF little which TOP mother 3SG FUT scold
 ‘Which boy_i will his mother_{i/j} scold _____?’
- d. *Lan_i đốt thư của nó_i; Mai_j cũng thế.*
 Nprop burn letter of 3SG Nprop also so
 ‘Lan_i burnt her_i letters; Mai did so too (= Mai also burnt her_j own letters or Lan’s).’

It is worthwhile to point out that not only must $nó_{REF}$ “agree in gender” with the DP with which it is co-indexed, but it must also match that nominal expression in terms of “register”. Since $nó_{REF}$ is normally employed to refer to individuals whose social status the speaker may want to belittle or considers to be equal to or lower than hers/his, it is compatible with *thằng bé* ‘the boy’ and *con mẹ già* ‘the old hag’

5. As an animate classifier, *con* can also be placed before a proper feminine noun. Similarly, *thằng* is exclusively used with proper nouns that refer to masculine individuals (but unlike *con*, *thằng* is inherently associated with human nouns). This use is however restricted to informal or colloquial styles and generally conveys pejorative connotations. The same effects can be achieved when those classifiers precede a common noun or an adjective:

- (i) *Con/thằng* *điên/ngu/béo/đĩ*
 CLF[+FEM]/CLF[+MASC] crazy/stupid/fat/slut
 ‘The (one[+FEM]/[+MASC] who is) crazy/stupid/fat/slut’

Such a syntactic-semantic behavior of *thằng* and *con* seems to come close to that of evaluative/qualitative constructions in French and English, such as *cet idiot de docteur*/*that idiot of a doctor* (cf. Milner 1978; Ruwet 1982; Rooryck 2001b, among others). We believe that this constitutes an important clue for the proper characterization of $nó_{EXPL}$ for they clearly hint at the central role played by the speaker and the speech situation (see also Banfield 1982, for whom the analysis of the qualitative construction *that idiot of a doctor* is to be related to the speaker’s perspective active at the sentential level).

6. The co-indexation between the interrogative expression *thằng bé nào* and $nó_{REF}$ is, however, not prohibited if the verb *khen* ‘praise’ is used intransitively. Likewise, their co-referentiality is rendered possible by the presence of a second occurrence of *nó* in object position, functioning as a resumptive pronoun.

but deemed unsuitable to enter into an anaphoric relationship with *bà cụ* ‘the old lady’, as exemplified in (12):

- (12) *Thằng bé_i/con mụ già_j/bà cụ_k nói rằng nó_{i/j/*k} sẽ đến.*
 CLF little/CLF old hag/old lady say COMP 3SG FUT come
 ‘The boy/the old hag/the old lady said that (s)he would come.’

As reported and assumed by Greco *et al.* (2018a), *nó*_{REF} has a default singular specification for number (13a) and hence is incompatible with a nominal expression marked [+plural]. *Nó*_{REF} therefore necessitates the plural markers *bọn* or *chúng* to be licensed (see 13b).

- (13) a. *Con Phan_i thì giám đốc đuổi việc nó_i rồi.*
 CLF Nprop TOP director dismiss work 3SG CRS
 ‘As for Phan, the director has fired her.’
 b. *Bọn con_i Phan thì giám đốc đuổi việc *(bọn/chúng) nó_i rồi.*
 CLF.PL CLF Nprop TOP director dismiss work CLF.PL/CLF.PL 3SG CRS
 ‘As for Phan and her fellows, the director has fired them.’

This agreement constraint seems to me, however, not to be an absolute one as it can be suspended in some contexts. In (14), for instance, the co-indexation between *nó*_{REF} and the plural matrix subject appears not to be entirely ruled out under a distributive reading, even though it is clearly less preferred than the logophoric pronoun *minh*. In this respect, (14) can be said to be semantically equivalent to (15).

- (14) [*Mấy/mọi thằng bé tôi gặp*]_i đều tự cho là chúng
 Several/all CLF little 1SG meet equally REFL think COMP CLF.PL
nó_i/nó_i/minh_i thông minh hơn những đứa khác.
 3SG/3SG/LOG intelligent more PL CLF other
 ‘The boys_i I’ve met all claimed/thought that they_i were/each of them was more intelligent than others.’

Distributive reading = ‘Each of the boys I’ve met claimed/thought that he himself was more intelligent than others.’

- (15) [*Mỗi thằng bé tôi gặp*]_i đều tự cho là chúng nó_{*i}/nó_i/minh_i
 Each CLF little 1SG meet equally REFL think COMP CLF.PL 3SG/3SG/LOG
thông minh hơn những đứa khác.
 intelligent more PL CLF other
 ‘Each_i of the boys I’ve met claimed/thought that he_i himself_i was more intelligent than others.’

2.2 Expletive *nó*

In contrast to *nó*_{REF}, *nó*_{EXPL} has been analyzed as an optional expletive subject that is not semantically selected by the verb, and is inserted for purely syntactic reasons (Nguyen & Hoang 2011; Dao 2013). Furthermore, it can only be found in preverbal subject position and lacks referential content. The first attribute is usually mentioned as being correlated with the fact that *nó*_{EXPL} typically occurs within intransitive environments, such as constructions headed by meteorological predicates (Example 16, but see also Example (6b) and unaccusative one⁷ (Example 17–21), whose sole argument is realized post-verbally:

- (16) *Đạo này, (nó) hay mưa, nắng thất thường quá. Tháng trước,*
 Period DEM EXPL often rain sunny erratic PART Month before
(nó) bão suốt một tuần.
 EXPL storm through one week
 ‘Lately, the weather has been very erratic: one day it’s rainy and the next it’s sunny. Last month, it was stormy for a whole week.’
- (17) *Sau trận bão tối qua (nó) lại đổ một đống cây.*
 After CLF storm night past EXPL again fall down one bunch tree
 ‘After last night’s storm, a bunch of trees fell down again.’
- (18) *Mày đổ nước như thế thì (nó) chết hết mấy con cá của tao.*
 2SG pour water like so TOP EXPL die finish several CLF fish of 1SG
 ‘If you pour water like that, my fishes will all die.’ (Dao 2015)
- (19) *Anh hay thức khuya mà sao (nó) không bạc nhiều tóc*
 2SG often stay up midnight but why EXPL NEG whiten up a lot hair
như em!?
 like 1SG
 ‘You often stay up late at night, but why don’t you have a lot of white hair like me?’
 (Lit. why didn’t your hair whiten up a lot like mine?)

7. Greco *et al.* (2018a, b) have discussed separate instances involving weather verbs, existential verbs and unaccusative verbs. However, while existential verbs arguably constitute a subtype of unaccusative predicates (see, *inter alia*, Levin & Rappaport-Hovav 1995; Alexiadou *et al.* 2004), scholars are not unanimous about the status of meteorological verbs. For Ruwet (1988), weather verbs are unaccusative. This view has been recently challenged by Paykin (2010), who argued that only metaphorically used weather verbs (in French) and weather verbs in personal constructions (in English), followed by a directional particle or a PP, show unaccusative syntax and can have their NP expansions qualified as internal arguments. It should be noticed that some authors analyze the subject of weather verbs as a non-referential “quasi-argument” selected by the verb (Chomsky 1981; Rizzi 1990), whereas others treat it as a true referential argument (Bolinger 1973; Bennis 1986; Pesetsky 1995, *inter alia*).

- (20) *Anh mở tủ lạnh xem trong đó (nó) có/còn thức ăn*
 2SG open refrigerator look in there EXPL exist/still exist CLF eat
hay là (nó) hết sạch (thức ăn) rồi.
 or EXPL finish clean CLF eat CRS
 ‘Open the fridge and look if there are some foods/there are some foods left in it or if there is nothing left in it at all.’

As shown in Examples (4–5) above, slightly modified below as (21–22), *nó*_{EXPL} cannot occur with unergative verbs or transitive ones, since their external argument can only be licensed in preverbal subject position. This seems to indicate that *nó*_{EXPL} competes for the same position as the overt pre-verbal subject.

- (21) a. *Mày làm thế, thằng bé hét bây giờ!*
 2SG do so CLF little scream now
 ‘If you do that, the boy will definitely scream!’
 b. **Mày làm thế, hét thằng bé bây giờ!*
 2SG do so scream CLF little now
 ‘If you do that, the boy will definitely scream!’
 c. **Mày làm thế, nó thằng bé hét bây giờ!*
 2SG do so EXPL CLF little scream now
 ‘If you do that, the boy will definitely scream!’
 d. *Mày làm thế, nó_{REF/*EXPL} hét thằng bé bây giờ!*
 2SG do so 3SG/*EXPL scream CLF little now
 OK ‘If you do that, (s)he will definitely scream at the boy!’
 *‘If you do that, the boy will definitely scream!’
 e. *Mày làm thế, thằng bé nó hét bây giờ!*
 2SG do so CLF little 3SG/*EXPL scream now
 ‘If you do that, the boy will definitely scream!’
 OK if *nó* is used anaphorically; * if *nó* is used as an expletive.
- (22) *Mày làm thế, (*nó) mẹ (*nó) sẽ phạt mày đấy!*
 2SG do so EXPL mother EXPL FUT punish 2SG PART
 ‘If you do that, Mom will punish you for sure!’

The second of the aforementioned attributes of *nó*_{EXPL}, namely its lack of referentiality, makes it unable to receive modification by a PP, as evidenced in (23a). *Nó*_{EXPL} can neither be coordinated with a lexical DP or a referential pronoun (see 23b) nor topicalized (Example 23c). And for the same reason, it cannot be associated with focus operators like *chính* ‘exactly’, *ngay cả* ‘even’, *chỉ* ‘only’ (Example 23d). Needless to say, *nó*_{EXPL} is predicted to be incompatible with the reflexive elements *bản thân* ‘self’ and *tự* ‘REFL’. This prediction is born out, as indicated in (23e).

- (23) a. *Gió to quá nên nó (*trên tầng 9) bay mất mấy cái áo*
 Wind big PART SO EXPL ON floor nine fly lose several CLF shirt
mới mua của con Trăng rồi.
 new buy of CLF Nprop CRS
 ‘The wind blew so strongly that several new shirts of Trang on the 9th floor have flown away.’
- b. *Gió to quá nên nó *(và con cú) bay mất mấy cái áo mới*
 Wind big PART SO EXPL and CLF owl fly lose several CLF shirt new
mua của con Trăng rồi.
 buy of CLF Nprop CRS
 ‘The wind blew so strongly that several new shirts of Trang and the owl have flown away.’
- c. *Gió to quá nên nó (*thì/*là) bay mất mấy cái áo mới*
 Wind big PART SO EXPL TOP/TOP fly lose several CLF shirt new
mua của con Trăng rồi.
 buy of CLF Nprop CRS
 ‘The wind blew so strongly that several new shirts of Trang have flown away.’
- d. *Gió to quá nên (*chính/*ngay cả/*chỉ) nó bay mất mấy cái*
 Wind big PART SO exactly/even/only EXPL fly lose several CLF
áo mới mua của con Trăng rồi.
 shirt new buy of CLF Nprop CRS
 ‘The wind blew so strongly that even several new shirts of Trang have flown away.’
- e. *Gió to quá nên (*bản thân/tự) nó bay mất mấy cái áo*
 Wind big PART SO self/REFL EXPL fly lose several CLF shirt
mới mua của con Trăng rồi.
 new buy of CLF Nprop CRS
 ‘The wind blew so strongly that several new shirts of Trang themselves have flown away.’

Consistent with these observations is the inability of $nó_{EXPL}$ to function as an anaphoric pronoun (24a) or to display crossover effects (compare 24b–c). Obviously, the ambiguity between a strict and a sloppy reading in elliptical construction is not observed either.⁸

8. It is extremely difficult, if not impossible, to construct examples to illustrate this kind of effect, as $nó_{EXPL}$ is restricted to pre-verbal subject position, and as a result, is outside of the elided constituent, namely the VP.

- (24) a. *Tôi lại thấy [con Phan]_i ở chợ trời. Nó_x mới rách ba cái*
 1SG again see CLF Nprop at market sky EXPL new torn three CLF
túi rơm hiệu Louis Vuitton của con Ly.
 bag fake brand Nprop of CLF Nprop
 ‘I’ve seen Phan at the flea market. Three fake Louis Vuitton bags of Ly have
 been torn/ripped recently.’
- b. *[Cái nhà nào]_i thì nó_{i/j/*EXPL} mới cháy tuần trước?*
 CLF house which TOP 3SG/*EXPL new burnt down week before
 ‘Which house burnt down last week?’
- c. *Nó_{*i/j/EXPL} mới cháy [cái nhà nào]_I tuần trước?*
 3SG/EXPL new burnt down CLF house which week before
 ‘Which house burnt down last week?’

A co-indexation relationship can be established between the initial DP and *nó* if the latter is interpreted as an anaphoric pronoun but not as an expletive element. A third reading is however possible, whereby *nó* refers to the owner of the house. Interestingly, as we will show below, the failure of a post-verbal argument to raise past *nó_{EXPL}* can be taken to be symptomatic of the fact that *nó_{EXPL}* occupies an A'-position in the left periphery, acting as an intervening element. Indeed, if the initial DP *cái nhà nào* ‘which house’ is analyzed as being fronted via topicalization from its post-verbal position, *nó_{EXPL}* is excluded altogether. This is in sharp contrast with (24c), where the interrogative expression remains in-situ.

The non-referential and non-anaphoric/cataphoric nature of *nó_{EXPL}* also implies that it is insensitive to feature-matching, be it in terms of gender, number, or “register” (see Examples 25). And logically, it cannot be pluralized by means of the plural markers *chúng/bọn*, as illustrated in (26).

- (25) a. *Kiểu này thì nó chết thẳng bé / con bé thôi!*
 Way DEM TOP EXPL die CLF little CFL little PART
 ‘Like this, the boy/the girl will die for sure!’
- b. *Kiểu này thì nó lại sập mấy ngôi nhà.*
 Way DEM TOP EXPL again collapse several CLF house
 ‘Like this, several houses will collapse again.’
- c. *Kiểu này thì nó lại ngã con cụ già / bà cụ thôi!*
 Way DEM TOP EXPL again fall CLF old hag / old lady PART
 ‘Like this, the old hag/the old lady will fall again for sure.’
- (26) *Kiểu này thì (*bọn/*chúng) nó chết bọn con Phan!*
 Way DEM TOP PL/PL EXPL die CLF.PL CLF Nprop
 ‘Like this, Phan and her fellows will die for sure!’

3. Greco *et al.*'s (2018a, b) analysis of *nó*_{EXPL}

3.1 On the position of *nó*_{EXPL}

In their contribution, Greco *et al.* note that Vietnamese follows the general tendency in many languages of resorting to the third-person singular⁹ pronoun form in configurations where an expletive can be used. They argue that *nó*_{EXPL} must be “located in the topmost portion of the inflectional field, to the right of, that is below, the constituents standardly assumed to occur in the left periphery of the clause” (Greco *et al.* 2018a: 41). On the basis of distributional evidence, they reach the conclusion that *nó*_{EXPL} is situated higher than the highest marker within the pre-verbal temporal and aspectual domain of the clause, namely the future marker *sẽ* (op.cit.: 42–43)¹⁰ (see also Phan 2013):

- (27) a. *Nó sẽ mưa bây giờ đấy.*
 NÓ FUT rain now PART
 ‘It will rain now.’ (their 30a, p. 42)
- b. *Sẽ *nó mưa bây giờ đấy.*
 FUT NÓ rain now PART
 ‘It will rain now.’ (their 30b, p. 43)
- (28) a. *Nó sẽ không có cái bút nào.*
 NÓ FUT NEG exist CLF pen any
 ‘There will be no pens.’ (their 31a, p. 43)
- b. *Sẽ *nó không có cái bút nào.*
 FUT NÓ NEG exist CLF pen any
 ‘There will be no pens.’ (Their 31b, p43)
- (29) a. *Nó sẽ ngã thằng bé.*
 NÓ FUT fall CLF little
 ‘A/The boy will fall.’ (Their 32a, p. 43)
- b. *Sẽ *nó ngã thằng bé.*
 FUT NÓ fall CLF little
 ‘A/The boy will fall.’ (Their 32b, p. 43)

9. Which is, in most cases, the masculine or neuter form.

10. Unless otherwise stated, all the examples drawn from these authors, including gloss and translations, are as given in the original texts. Some of their examples, in particular ones with expletive structures, are rejected or perceived as not fully grammatical by my consultants, probably because, for the sake of the demonstration, they were fabricated without being contextualized.

They further discuss two discourse-related elements which contribute to the marking of information structure in the left periphery of the Vietnamese clause, viz. *thì* and *là*. According to the authors, these are associated with topicalized arguments and adjuncts and can be treated as heads of functional projections which host the topicalized constituents in their specifiers. For brevity and convenience of exposition, their two Examples (33) and (34) (p. 43) are merged into one below as (30).

- (30) a. *Thằng Nam thì/là sẽ ăn cái này đấy.*
 CLF Nam TOP/TOP FUT eat CLF this PART
 ‘As for Nam, he will eat this thing.’
- b. *Cái này thì/là thằng Nam sẽ ăn đấy.*
 CLF this TOP/TOP CLF Nam FUT eat PART
 ‘As for this thing, Nam will eat.’
- c. *Lúc khác thì/là thằng Nam sẽ ăn cái này đấy.*
 Time other TOP/TOP CLF Nam FUT eat CLF this PART
 ‘At another time, Nam will eat this thing.’

Greco *et al.* also claim that *thì* and *là* are strictly ordered, in that the former always the latter, and not the other way round. As indicated in their Example (35) (p. 44), reproduced below as (31), *là* must be lower than *thì* in the left periphery. More particularly, *nó_{EXPL}* is shown to occur in a position lower than *thì* and *là* (Example 32, their 36). In a nutshell, “the expletive *nó* occupies a position somewhere in the highest portion of the inflectional layer, immediately dominated by the functional projections of the left periphery” (p. 44). Their tentative conclusion is schematized as in (33) (their 38).

- (31) a. *Cái này thì lúc khác là thằng Nam sẽ ăn đấy.*
 CLF this TOP time other TOP CLF Nam FUT eat PART
 ‘As for this thing, at another time, Nam will eat.’
- b. **Cái này là lúc khác thì thằng Nam sẽ ăn đấy.*
 CLF this TOP time other TOP CLF Nam FUT eat PART
 ‘As for this thing, at another time, Nam will eat.’
- (32) a. *Trên bàn thì/là nó sẽ không có cái bút nào.*
 On table TOP/TOP NÓ FUT NEG exist CLF pen any
 ‘On the table, there will be no pens.’
- b. *Trên bàn *nó thì/là sẽ không có cái bút nào.*
 On table NÓ TOP/TOP FUT NEG exist CLF pen any
 ‘On the table, there will be no pens.’
- c. **Nó trên bàn thì/là sẽ không có cái bút nào.*
 NÓ on table TOP/TOP FUT NEG exist CLF pen any
 ‘On the table, there will be no pens.’

- (33) $Thì_{[Topic]} > là_{[Topic]} > N\acute{O} > s\acute{ẽ}_{[Future]} > đ\grave{a}[_{Perfect}] > đ\grave{a}ng_{[durative]} > VP$

This analysis calls for some comments. First, while we share Greco *et al.*'s judgment on (31a), we do not think that switching places between *thì* and *là* as in (31b) leads to ungrammaticality, for this sentence sounds fairly natural to our consultants. Besides, the assertion that *là* must be lower than *thì* may appear to be too strong, since not only does it rely on a single example, but the authors also limited themselves to cases where the highest topicalized constituent is an argument and not an adjunct. Moreover, even if some may find (31b) unlikely, this kind of oddity would be ascribable to other factors, such as the fact that *là* can also be analyzed as a copula under certain circumstances or that both *thì* and *là* are found in hypothetical constructions where they are used to introduce the apodosis and to link it to the protasis. Developing such considerations here would go beyond the scope of this paper. Our point is that the reverse order $là > thì$ is not totally impossible. Indeed, the following sentence is judged acceptable by my consultants:

- (34) *Cái này là hôm qua thì Lan đưa cho Mai còn hôm nay thì Lan lấy lại.*
 CLF DEM TOP day past TOP Nprop give to Nprop as for day now TOP
 Nprop take back
 'As for this thing, yesterday, Lan gave it to Mai and today, she took it back from her.'

In an analogous way, in Dao (2019a), we have argued for a highly articulated structure of the Vietnamese left periphery, based on the distributional properties of four C-elements (i.e. which belongs to the CP-field in the architecture of the clause), viz. *rằng, thì, là, mà*, whose order is as follows:

- (35) $[_{ForceP} \text{r\grave{a}ng/m\grave{a}} [_{TopP} \text{l\grave{a}} [_{TopP} \text{th\grave{i}} [_{FocusP} \text{m\grave{a}} [...]]]]]$

The details of this account are not relevant here. What is important is that there are instances where *là* can be higher than *thì* in the left periphery, as shown in (36):

- (36) *Mai đã nói với Tr\grave{a}ng rằng là chuyện đấy thì nó không biết.*
 Nprop PERF say to Nprop COMP TOP story DEM TOP 3SG NEG know
 'Mai told Trang that, as for that story, she didn't know it.'

Second, and more crucially, $n\acute{o}_{EXPL}$ seems to be able to move higher to the left periphery. Consider the following examples:

- (37) a. *Anh mà không đến thì/là con Phan sẽ chửi đấy!*
 2SG COMP NEG come TOP/TOP CLF Nprop FUT insult PART
 ‘If you don’t come, Phan will curse you!’
- b. *Cái đấy mà anh không làm thì/là con Phan sẽ chửi đấy!*
 CLF DEM COMP 2SG NEG do TOP/TOP CLF Nprop FUT insult PART
 ‘If you don’t do that, Phan will curse you!’
- c. *Ngày mai mà anh không gọi điện thì/là con Phan sẽ chửi đấy!*
 Tomorrow COMP 2SG NEG phone TOP/TOP CLF Nprop FUT
 insult PART
 ‘If you don’t call tomorrow, Phan will curse you!’

What these sentences seem to show us is that arguments and adjuncts can be fronted, as observed with topicalization by *thì/là*, and land in the specifier of a left-peripheral projection headed by *mà*. Strikingly, *nó_{EXPL}* can undergo fronting in this context:

- (38) a. *Nó mà vỡ cái bình của bố thì mày chết!*
 EXPL COMP break CLF vase of father TOP 2SG die
 ‘You’ll die if Dad’s vase breaks!’
- b. *Mày làm thế nó mà ngã thằng bé thì mày chết!*
 2SG do so EXPL COMP fall CLF little TOP 2SG die
 ‘You do that, and if the boy falls, you’ll die!’

Recall that above, we have shown in Examples (24b, c), reproduced below as (39), that interrogative expressions are not allowed to move from their post-verbal argument position and raise past *nó_{EXPL}*. This fact can be seen as evidence that *nó_{EXPL}* triggers intervention effects and prevents the interrogative expressions from undergoing A'-movement to the left periphery. This thus lends further support to the hypothesis that *nó_{EXPL}* would already be in the left periphery or, at any rate, manifests some A'-properties.

- (39) a. [*Cái nhà nào*]_i thì *nó_{i/j/EXPL}* mới cháy tuần trước?
 CLF house which TOP 3SG/*EXPL new burnt down week before
- b. *Nó_{i/j/EXPL}* mới cháy [*cái nhà nào*]_i tuần trước?
 3SG/EXPL new burnt down CLF house which week before
 ‘Which house burnt down last week?’

One would expect that Wh-movement in (38) will bring about the same outcome. This prediction seems to hold:

- (40) a. **Cái gì mà nó vỡ ____ thì mà chết?*
 CLF what COMP EXPL break TOP 2SG die
 Lit. ‘What is the thing because of which you’ll die if it breaks?’
- b. **Mày làm thế ai mà nó ngã ____ thì mà chết!*
 2SG do so who COMP EXPL fall TOP 2SG die
 ‘You do that, who is the person because of whom you’ll die if (s)he falls?’

While these data do not explain why *nó*_{EXPL}’s displacement to the Topic Phrases headed by *thì/là*, they clearly suggest that it is somehow linked to the left periphery.

3.2 On the interpretation of *nó*_{EXPL}

We now turn to Greco *et al.*’s account of the interpretative properties of *nó*_{EXPL}. As these authors emphasize,

[A]lthough IP internal, *nó* does contribute to the interpretative content of the clause in which it is inserted. Generally speaking, the nature of the interpretive contribution of *nó* differs from that reported for the left-peripheral expletives with discourse function [...] in that the presence of *nó* does not seem to affect the illocutionary force or the polarity of the sentence, and no effect of emphasis or contrast can be related to the insertion of *nó* in a certain structure. [...] Instead, the semantic contribution of *nó* seems to affect the interpretation of the propositional content expressed by the clause where it appears in terms of **specificity**. (our highlighting)
 (Greco *et al.* 2018a: 45)

More concretely, they claim that “inserting *nó* narrows down the contexts in which the sentence is appropriate in terms of **speaker-related epistemic specificity**” (Greco *et al.* 2018b: 77, our highlighting). Consider the existential sentence in (41) below:

- (41) *Không có ma.*
 NEG exist ghost
 a. ‘Ghosts do not exist.’ (Generic interpretation)
 b. ‘There are no ghosts speaking of a certain place/time.’
 (Contextualized interpretation)

Based on this kind of sentences, the authors argue that the presence of *nó* narrows the interpretation of (41) by eliminating the generic reading, thus restricting the domain of validity of the assertion that “there are no ghosts” to a specific context.

- (42) *Nó không có ma.*
 EXPL NEG exist ghost
 a. # ‘Ghosts do not exist.’ (Generic interpretation)
 b. ‘There are no ghosts speaking of a certain place/time.’
 (Contextualized interpretation)

Greco *et al.* go on to show that the same observation can be made with respect to thetic^{11, 12} sentences, such as (43), where “the insertion of *nó* in thetic sentences is only felicitous in those contexts in which the speaker disposes of sufficient background information to be able to report on the specific event” (Greco *et al.* 2018b: 78, our highlighting). In other words, the presence of *nó* implies that “the eventuality expressed in the clause is specifically identifiable in or anchored to a given context” (op.cit., our highlighting):

- (43) (*Nó*) *Cháy cái nhà kho rồi.*
 EXPL burnt CLF house store CRS
 ‘A warehouse burned.’

11. In Greco *et al.*'s account, sentences featuring *nó*_{EXPL} with an unaccusative predicate are referred to as thetic sentences, which, they assume, assert the existence of an event, owing to the fact that these sentences can be used in an out-of-the-blue context or as an answer to the question *What happened?* However, this line of argumentation runs into problems. First, unaccusative predicates can be eventive or stative (see Dao 2015 for a discussion) but the question *what happened?* seems to require an [+eventive] answer. Second, for our consultants, even sentences with *nó*_{EXPL} and an [+eventive] unaccusative verb are deemed not to be appropriate answers to this question:

- (i) A: *Chuyện gì đã xảy ra?*
 Story what PERF happen
 ‘What happened?’
 B: a. **Nó ngã thành bé.*
 EXPL fall CLF little
 ‘The boy fell.’
 b. *Thành bé ngã.*
 CLF little fall
 ‘The boy fell.’

12. It is not clear to us why Greco *et al.* did not treat existential sentences with *nó*_{EXPL} on a par with the *nó*_{EXPL}-sentences featuring unaccusative predicates. As a matter of fact, the post-verbal NP/DP in Vietnamese existential sentences is arguably the internal argument of existential predicates (see Dao 2015), as evidenced by its ability to appear in pre-verbal subject position:

- (ii) a. *Trên đời này (nó) làm gì có ma.*
 On life DEM EXPL do what exist ghost
 ‘There are no ghosts in this world at all.’
 b. *Trên đời này ma làm gì có.*
 On life DEM ghost do what exist
 ‘There are no ghosts in this world at all.’

Yet, one might argue that in the presence of *làm gì* ‘do-what’ (= ‘How is it even possible that...’), the pre-verbal DP *ma* ‘ghost’ has moved to the left periphery in (iib), instead of being in [Spec, IP/TP]. However, this would in turn mean that the same could be said of *nó*_{EXPL} in (iia).

More notably, the authors further stress that:

[T]he insertion of *nó* is not limited to cases where the place and the time of the event coincide with or are directly related to the utterance context. *Nó* can also be used in contexts where the event takes place in a different location from the utterance location, as long as the speaker can identify the specific event reported. [...] The only requirement for the insertion of the expletive *nó* appears to be the possibility of the speaker having a specific event in mind. In this respect then, the contribution of *nó* seems to be that of conveying Epistemic Specificity (Kasher & Gabbay 1976; Hellan 1981; Farkas 2002). (our highlighting)

(Greco *et al.* 2018a: 48)

While useful and illuminating, this semantic portrayal of *nó_{EXPL}* deserves some comments. Overall, we do agree with (i) Greco *et al.*'s statement that *nó_{EXPL}* can convey speaker-oriented meaning and (ii) their hypothesis according to which the import of *nó_{EXPL}* may consist in introducing a [+specific] feature in *nó_{EXPL}*-sentences with unaccusative predicates. That said, we are not totally convinced by what has been advanced so far for existential sentences. In fact, it is not sure that inserting *nó_{EXPL}* in sentences like (42) would really have the effect of ruling out the generic reading. As we can see in (44)¹³ below (see also footnote 13), this interpretation still remains available even though *nó_{EXPL}* is inserted:

- (44) a. *Trên đời này làm gì có ma.*
 On life DEM do what exist ghost
 'There are no ghosts in this world at all.' (Generic interpretation)
- b. *Trên đời này (nó) làm gì có ma.*
 On life DEM EXPL do what exist ghost
 'There are no ghosts in this world at all.' (Generic interpretation)

This would thus mean that there would be no such "narrowing effects" in existential sentences as previously claimed by Greco *et al.* One might then ask where the so-called "specific context" went to in this case and in which way those *nó_{EXPL}*-sentences, be they built with existential predicates or with unaccusative ones, can be accounted for in a unified fashion. Phrased differently, what would be the common denominator of those two types of *nó_{EXPL}*-sentences if narrowing effects no longer constitute the core property of *nó_{EXPL}* and if one maintains that the *Epistemic Specificity* induced by *nó_{EXPL}* is basically correct? This leads us to another crucial, connected point, which Greco *et al.* did not quite spell out, namely the source from which specificity comes from. Following Fodor & Sag (1982), we believe that such a discourse-related concept should be understood as reflecting the state of knowledge

13. It is of worth to mention that existential sentences presuppose a locative, whether explicitly realized or not, since the existence of an entity can be asserted only with regard to a certain place.

of the speaker (only)¹⁴ in the discourse and, more fundamentally, the speaker's intent to refer. Ionin *et al.* (2004) have further added to this definition the concept of *noteworthy property*. They proposed the following informal definition of specificity:

If a Determiner Phrase (DP) is [...] [+specific], then the **speaker intends to refer** to a unique individual in the set denoted by the NP and **considers** this individual to possess some **noteworthy property**. (our highlighting) (op.cit.: 5)

Leaving the details of this definition aside, for it primarily concerns DPs, what is important for our purpose here is that specificity revolves around the speaker and her involvement in or commitment to her own discourse.¹⁵ The difference between (44a) and (44b) can therefore be thought of in terms of markedness, in that the latter reveals a higher degree of involvement¹⁶ of the speaker than the former. Expressed in more accurate terms, specificity should only be a subcomponent of a larger operation, namely the speaker's assessment of a given situation/state of affairs against her background knowledge. Accordingly, what should be retained from the following conclusion of Greco *et al.* is probably the key role of the speaker and the utterance context. In this respect, the specificity requirement seems to be more relevant to *nó_{EXPL}*-sentences with unaccusative predicates than those with existential ones and can be considered as deriving from the speaker's evaluation of a situation based on her epistemic state. That amounts to saying that *nó_{EXPL}* would rather be the trace of this last indexical operation, whereby it signals the presence of the speaker herself in her own utterance. Greco *et al.* themselves do recognize this *speaker-dependency*, as the passages we have highlighted all show that *nó_{EXPL}* is closely intertwined with the notions of *speaker* and *context of utterance*. It is however intriguing that the authors did not really attempt to seek explanations for this correlation and to formalize it in a principled way.

The specificity effect of *nó* and the fact that it **anchors the proposition to the speaker's context** provides additional empirical support that, while non-referential, expletives can encode speaker-oriented meaning. (our highlighting)
(Greco *et al.* 2018b: 81)

While very attractive, Greco *et al.*'s specificity requirement also leaves a number of questions unanswered or open. For instance, when the authors claim that the insertion of *nó_{EXPL}* introduces a [+specific] feature, it is not clear to us:

14. Whereas definiteness reflects the state of knowledge of both speaker and hearer (cf. Fodor & Sag 1982).

15. This line of reasoning is reminiscent of what Creissels (2008) calls the marking of the "assertor's involvement". See also Mithun (1986), Faller (2002).

16. Informally speaking, (44b) is accompanied by a(n) (inter)personal "coloring".

- (45) i. what kind of feature it is (formal, semantic, pragmatic?);
 ii. whether it is interpretable or uninterpretable;
 iii. how and where the linguistic system computes and interprets it;
 iv. in which way it can be part of $nó_{EXPL}$'s featural make-up;
 v. what $nó_{EXPL}$'s others features are;
 vi. whether $nó_{EXPL}$ inherently carries it when inserted in the syntax or whether the specificity feature is merely a side-effect brought about by the insertion of $nó_{EXPL}$;
 vii. how the presence of this feature and its interplay with $nó_{EXPL}$'s structural properties can be implemented.

Some of these issues have been addressed in Dao (2014), whose main lines will be set out in the next section.

4. $Nó_{EXPL}$, subjectivity, evidentiality

In this section, we will summarize the proposal advanced in Dao (2014)¹⁷ before laying out a novel account of $nó_{EXPL}$. Although the path we followed in that preliminary work is very akin in spirit to Greco *et al.*'s (2018a, b), our characterization of $nó_{EXPL}$ has placed at the heart of its linguistic behavior the role of subjectivity and, more basically, the point of view of discourse participants.

4.1 Dao's (2014) analysis of $nó_{EXPL}$

All the distributional facts depicted in the above sections led us to propose in Dao (2014) that, as opposed to the referential $nó$, $nó_{EXPL}$ does not behave like a full-fledged pronoun and is deficient in a way that is reminiscent of logophoric expressions (cf. Sells 1987). They also are highly reflective of the peculiar properties manifested by $nó_{EXPL}$: while it functions like a syntactic subject, its licensing appears to be strongly sensitive to contextual (i.e. pragmatic) considerations. In this regard, it is not entirely true that $nó_{EXPL}$ is semantically vacuous. Rather, its behavior reminds us of that of a variable which needs to be bound by an Operator.

We argued that $nó_{EXPL}$ occupies the syntactic subject position of the clause and is externally merged in [Spec, TP]. It is further assumed to be D(iscourse)-linked.

17. While they did cite one of my previous studies (cf. Dao 2012) where $nó_{EXPL}$ is only mentioned as a diagnostic test of unaccusativity, Greco *et al.* (2018a, b) apparently did not take notice of Dao (2014), which is devoted to an in-depth investigation of the peculiarities of $nó_{EXPL}$. Note also that Dao (2012) is a largely abridged version of Dao (2013).

More precisely, we suggested that $n\acute{o}_{EXPL}$ encodes the Point of View (PoV) of discourse participants, namely the 1st and 2nd persons (or indexicals), the 3rd person being analyzed as “non-person” (*non-personne*), according to Benveniste (1966 [1946]). In this sense, $n\acute{o}_{EXPL}$ is endowed with the features¹⁸ [-R] (“Referentiality”-Feature), [+d] (“Discourse Participant”-Feature) and [p:_] (“person”-Feature), this last one being unvalued. Following Guéron & Haegeman (2012) (but see also Bianchi 2003), we proposed that PoV heads a functional projection located in the left periphery of the clause. PoV and $n\acute{o}_{EXPL}$ are argued to be engaged in an AGREE relation (cf. Chomsky 2001 and subsequent works), whereby the feature [p: 1st /2nd] borne by PoV values that of $n\acute{o}_{EXPL}$. In our account, $n\acute{o}_{EXPL}$ was not treated as the explicit realization of a *stage topics* as defined in Erteschik-Shir (1997) and extended in Lahousse (2003, 2007) (see also Achard 2009 for an analysis of French impersonal *il*).

As we can see, despite reaching the same conclusion about the syntactic position of $n\acute{o}_{EXPL}$, the approach espoused in Dao (2014) does not make any reference to specificity as does Greco *et al.*'s. Instead, it tries to disentangle the puzzle by putting under close scrutiny the mapping between the syntactic and pragmatic properties exhibited by $n\acute{o}_{EXPL}$. The unusual functioning of $n\acute{o}_{EXPL}$ is thus situated at the syntax-pragmatics (discourse) interface where the contribution of discourse participants is taken into account.

4.2 $N\acute{o}_{EXPL}$ as evidential marker

Let us state straight away, before returning to it at greater length, that the novel account we propose in the remainder of this paper diverges from both Greco *et al.*'s and Dao's (2014) in a number of respects. On syntactic grounds, as argued in the previous sections, $n\acute{o}_{EXPL}$ can arguably move to the left periphery instead of staying in the inflectional domain of the clause. More precisely, while externally merged in [Spec, TP], it can optionally raise to the specifier of a functional projection which will be shown below to encode evidentiality. On semantic/pragmatic grounds, our analysis not only incorporates insightful observations from Greco *et al.*'s works that $n\acute{o}_{EXPL}$ is intrinsically linked to the speaker, but also takes from Dao (2014) the idea that $n\acute{o}_{EXPL}$ constitutes an indexical phenomenon due to its ability to encode the point of view of discourse participants. However, as suggested through the above appraisal of Greco *et al.* (2018a, b), our account takes a step further and concentrates on the speaker and her involvement in her own discourse, or, to put it concretely, her assessment of a given situation against her epistemic background.

18. By contrast, the referential $n\acute{o}$ is treated as bearing the [+R], [-d], and [p:3rd] features.

4.2.1 *Ego-evidentiality, egophoricity*

At first glance, the idea that $nó_{EXPL}$ is related to evidentiality¹⁹ may seem at odds with the isolating nature of Vietnamese, for evidentiality is traditionally viewed as a grammaticalized category, on a par with other verbal categories such as tense, aspect and mood/modality (cf. Aikhenvald 2004). We adopt here the broad view according to which evidentiality is not restricted to verbal inflections (nor must it belong to grammatical systems of closed-class items), but can be expressed across languages by various grammatical and lexical means. To fix terminology, we also stick to the canonical definition that evidentiality marks the source of information the speaker has for her statement. What is relevant to our discussion is the fact that one of the commonly assumed properties of evidentiality is speaker deixis, as stated by Brugman and Macaulay (2015: 216):

Evidentials generally encode person deixis, and are usually described as speaker deictic in that they reflect the speaker's point of view with respect to the situation described in the sentence. All deictic elements presuppose an origo, a point of reference in the domain of time, space, or person; canonically, evidentials encode a speaker origo.

This observation has also been confirmed by Rooryck (2001a: 127), who notes that the notion of *source of information* is “closely linked, if not identical, to that *speaker perspective* and *point of view*”. In this respect, the tight connection of $nó_{EXPL}$ to the speaker and her sphere of personal knowledge²⁰ strongly suggest to analyze it as an evidential marker. More specifically, $nó_{EXPL}$ is best treated as an instance of *ego-evidentiality* in that “it [Ego evidentiality] signals immediate knowledge, unmediated by perception or inference, a subcase of which is self- or *de se* knowledge” (Garrett 2001: 154). Crucially, in line with Garrett, we assume that “ego evidentiality is a pragmatic property of assertion, and not an inherent property of any particular lexical or grammatical words” (op. cit.: 105). In other words, sentences containing $nó_{EXPL}$ indicate that the speaker has immediate knowledge of a proposition. $Nó_{EXPL}$ itself does not intrinsically carry any evidential feature but acquires some indexical features *via* the establishment of an Agree relationship with a dedicated functional head in the left periphery.

19. We cannot possibly do justice to this subject here, since an extremely rich body of literature exists on the issue. The interested reader can refer to, among others, Chafe & Nichols (1986), de Haan (1999, 2005), Dendale & Tasmowski (2001), Rooryck (2001a, b), Aikhenvald (2004, 2014, 2018), Speas (2008), Brugman & Macaulay (2015) and references therein.

20. This characterization is reminiscent of Searle's (2004) *ontological subjectivity*, according to which the knowledge exists only from a first-person point of view (cf. Speas 2010)

In this context, it should be of interest to mention that the description just provided is conceptually not so different from what is known in typological literature under the label of *egophoricity*.²¹ According to Widmer and Zúñiga (2017: 419), this last category refers to “a grammatical category that expresses access to knowledge or, more precisely, to the particular information conveyed in a given utterance”. For these authors, egophoric markers indicate privileged access to knowledge, one form of which is due to “epistemic involvement in the role of a “knower” whose precise relation to the event is not specified” (ibid.). The concept of *privileged access* to knowledge can be understood as “a privileged epistemic relationship that holds between a speech-act participant (SAP) and the knowledge conveyed in a proposition” (op.cit.: 420).

Before turning to configurational aspects of *nó_{EXPL}*, we would like to point out that its evidential import seems not surprising, since often evidentials diachronically spring from either perception verbs and verbs of saying or personal pronouns (Botne 1995) and constitute a property of spoken languages (Rooryck 2001a).²²

4.2.2 *Nó_{EXPL} at the syntax-pragmatics interface*

Evidentiality has been argued by scholars to be syntactically encoded. For instance, Cinque (1999), in his cross-linguistic study of adverbial modifiers, proposes a universal hierarchy of functional projections whereby evidential morphemes and adverbs (as well as adverbial PPs) expressing evidentiality respectively instantiate the head and specifier positions of a left-peripheral projection Evid(ential)P.²³ In his cartographic approach, the left periphery of the clause, viz. its topmost layer (or the C-layer), is populated by a sequence of four hierarchically ordered functional projections. Within this area, where the relationships with the context/the discourse are encoded, EvidP is “sandwiched” between evaluative and epistemic projections:

$$(46) \text{ [Mood}_{\text{speech act}} \text{ [Mood}_{\text{evaluative}} \text{ [Mood}_{\text{evidential}} \text{ [Mod}_{\text{epistemic}} \text{ [(T...)]]]]]]$$

21. We leave aside here the issue of how *ego-evidentiality* and *egophoricity* differ from one another for that would merit a separate study. See Widmer & Zúñiga (2017) for a discussion.

22. Interestingly enough, Rooryck suggests to treat French “ethical datives” as “pronominal-based” evidentials.

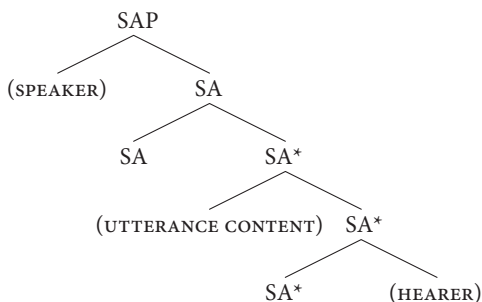
23. But see Blain & Déchaine (2007) for an alternative view. For these authors, evidentials markers are operators that “can enter the sentence at the vP, AspP, IP, or CP levels, with different interpretative consequences.” (Speas 2008: 949).

Combining Cinque's framework with Hale & Keyser's (1993) theory of argument structure, Speas (2004) suggests that each of the four discourse-related projections in (46) licenses in its specifier an implicit argument, to which it simultaneously assigns a corresponding pragmatic/discourse role. The gist of her proposal is summarized in (47):

(47) Position	Pragmatic role
Spec, S(peech) A(ct) P(hrase)]	→ SPEAKER
[Spec, Eval(uative) P(hrase)]	→ EVALUATOR
[Spec, Evid(ential) P(hrase)]	→ WITNESS
[Spec, Epis(temic) P(hrase)]	→ PERCEIVER

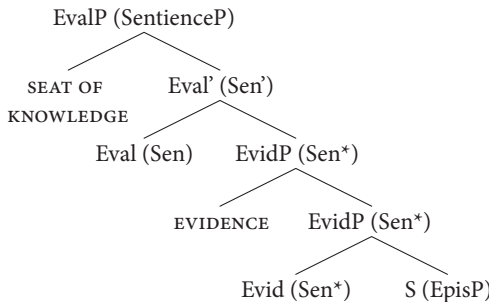
Speas and Tenny (2003) further pursue the idea of a parallelism between the discourse domain (CP) and the thematic domain (vP). Essentially, if thematic roles (θ -roles), such as Agent, Theme, Goal, etc., are to be introduced within the vP and defined in terms of their structural position, then the same mechanism may be applied to the discourse domain: pragmatic/discourse roles (P-roles) are licensed within the CP and structured in the same manner as thematic roles. More specifically, the authors argue that the CP is built up from two domains: The Speech Act Domain and the Sentience/Point of View Domain, the latter being dominated by the former but having scope over the rest of the sentence. As we can observe in (48), the SPEAKER P-role can be conceived of as the Agent of the SAP, the UTTERANCE CONTENT as its Theme, and the HEARER as its Goal. The Sentience Domain, which "encodes judgements and evaluations by a sentient mind on the truth-value of the proposition" (op.cit.: 333), corresponds to the UTTERANCE CONTENT P-role and encompasses Cinque's EvalP, EvidP, and EpisP. In particular, a SEAT OF KNOWLEDGE P-role is introduced in [Spec, SentienceP/EvalP]. This is sketched in (49).

(48) Speech Act Projection in declarative sentences (Speas & Tenny 2003: 320)



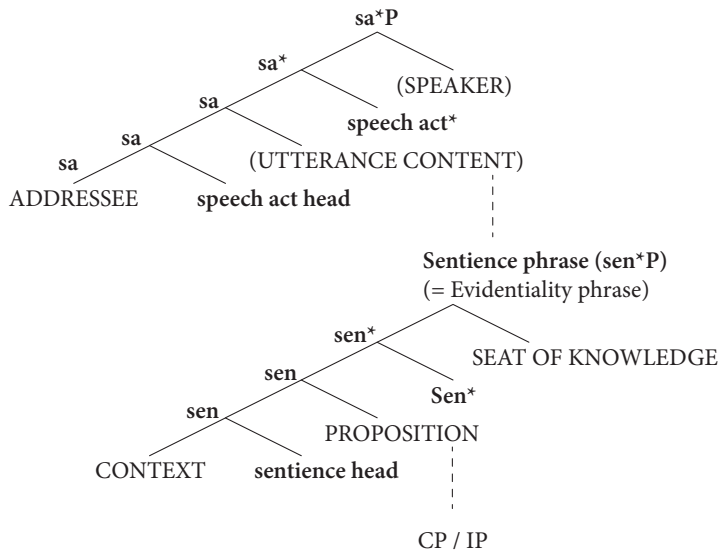
(49) The Sentence Projection

(op.cit.: 334)

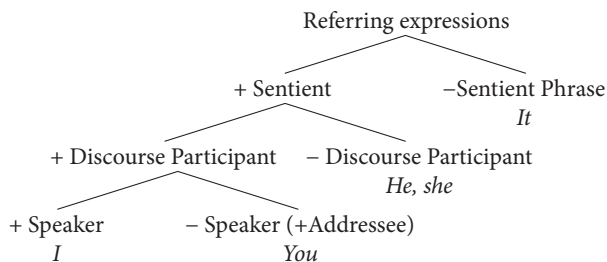


While maintaining the syntactic representation advanced for the Speech Act Phrase, Tenny (2006: 260–261) puts forward another version of that of the Sentence Phrase. The major dissimilarity with Speas & Tenny (2003) probably lies in the fact that she completely dispenses with Cinque’s EvalP and EpisP. For her, Sentence Phrase directly encodes Evidentiality, viz. the evaluation by a SEAT OF KNOWLEDGE of a PROPOSITION with regard to a CONTEXT (see 50). This configuration is further articulated to a feature-based system. Appealing to the universal feature geometry for pronouns laid out by Harley and Ritter (2002), Tenny (op.cit.: 264) adds a [\pm Sentient] feature to it, which gives rise to the modified system in (51). At this point, it goes without saying that Tenny’s feature-based system cannot account for *nó*_{EXPL} for it is clearly not referential. The [\pm R(eferential)] feature is therefore needed to distinguish the two uses of *nó*.

(50) Tenny’s (2006) proposal



(51) Feature geometry for referring expressions (Tenny's (2006) version)



Before showing how our analysis of $nó_{EXPL}$ is implemented using Speas-Tenny's model, let us point out that Vietnamese, like most (South)East Asian languages, has a rich inventory of "sentence-final particles (SFPs) that make subtle distinction in sentence type, stance, evidentiality, and combinations thereof" (Enfield & Comrie 2015: 8). While most of them, if not all, are arguable subjectivity markers and discussing them would provide us with a better understanding of the behavior of $nó_{EXPL}$, three SFPs are of high interest for our purpose, namely *rôi*, *đấy*, *mà*.²⁴ One of their salient features is that they are characteristic of impromptu speech and quite frequently found in $nó_{EXPL}$ -sentences. In Dao (2019b), we have shown that they are strongly indicative of subjectivity and constitute traces of the presence of the speaker in the discourse. For instance, *rôi* can be analyzed as referring to a "currently relevant state" (CRS) in that it indicates that the state of affairs expressed in the sentence containing it "has special current relevance with respect to some particular situation" (Li & Thompson 1981: 240). As shown in (52), *rôi* contrasts with *đã*, which is ambiguous between a past and an aspectual perfect reading. By uttering (52b), the speaker does not focus on the fact that the eventuality "Trang-go-to-Germany" took place in the past. What is relevant for her would be, for instance, that Trang is not present at utterance time, viz. the state resulting from that eventuality. In the same vein, *rôi* can be used in the answer to the question in (53) whereas *đã* is odd. Moreover, the speaker-oriented nature of *rôi* is confirmed by its status as a Positive Polarity Item. In fact, it is anti-licensed in the scope of the negation (cf. Ernst 2009; Dao 2019b) unless it loses its aspectual reading and retains only its CRS value.²⁵ (54) clearly shows the evidential import of *rôi*.

24. Due to space limitations and as it is outside of the scope of this paper, we will not go further into details of how SFPs are structurally represented and hierarchically ordered within the Vietnamese CP-layer.

25. Which means that the negation can only be interpreted as having narrow scope and that *rôi* is located higher in the clausal skeleton.

- (52) a. *Trăng đã đi Đức.*
 Nprop PERF go Germany
 ‘Trang went to Germany/Trang has been to Germany.’
 b. *Trăng đi Đức rồi.*
 Nprop go Germany CRS
 ‘Trang has been to Germany.’
- (53) A: *Xin lỗi Trăng có đấy không ạ?*
 Sorry Nprop exist there NEG PART
 ‘Excuse me, is Trang there?’
 B: a. *#Trăng đã đi ngủ.*
 Nprop PERF go sleep
 ‘Trang went to bed.’
 b. *Trăng đi ngủ rồi.*
 Nprop go sleep CRS
 ‘Trang has gone to bed (she’s sleeping now and she’s not available, you understand?)’
- (54) *(Kiểu này là) Trăng không đến rồi.*
 Way DEM TOP Nprop NEG come CRS
 # ‘(From what I can see/tell,) Trang didn’t/hasn’t come.’
 ‘(From what I can see/tell, I suppose that) Trang won’t come’

Regarding *đấy*, besides its uses as a DP-internal demonstrative and a spatial deictic adverb, it can convey speaker-related meaning²⁶ when occurring sentence-finally. As exemplified in (55), all three uses of *đấy* can be observed in the same sentence. As an SFP, *đấy* indicates that the speaker wants to give more weight to her assertion and/or to emphasize its validity.²⁷ This is in line with the *lococentrism* hypothesis defended by number of scholars for East Asian languages like Vietnamese (Nguyen 1992), Japanese (Tamba 1992), and Chinese (Paris 1992). The important point is that these languages show tendency to shift the deictic center from *ego* to *hic*, thus favoring Place over Person. Therefore, spatial deictics can, in certain contexts, replace indexical pronouns (1st and 2nd persons) to designate a given discourse participant (e.g., *here* means *I*, and *there* means *You*).

26. Note in passing that such behaviors of *rồi* and *đấy* seem to corroborate de Haan’s (1999) observations that Tense/Aspect morphemes and spatial deictics constitute two diachronic sources of evidentials.

27. This « pragmaticalization » process seems to also affect other spatial deictics, such as the proximal *này* and *đây*, as well as the distal *kia* and *kia*.

- (55) *Cái áo đấy thì em mua ở đấy đấy!*
 CLF shirt DEM TOP 1SG buy at there PART
 ‘(Yes, it’s right!) As for that shirt, I bought it there.’

The case of *mà* is more intricate as it is highly polyfunctional (cf. Dao & Do-Hurinville 2013; Dao 2019b). We will mention here only two examples to illustrate its speaker-oriented/evidential meaning. As suggested by the translations in (56), *mà* appears to either signal the speaker’s attitude towards an individual or underscore her claim about a state of affairs. This comes very close to the range of evidential meanings expressed by parentheticals such as *I tell you/I swear/I’m sorry to tell you/I’m afraid* (cf. Rooryck 2001a).

- (56) a. *Trăng mà²⁸ thông mình à?!*
 Nprop MA intelligent PART
 ‘(Oh really? Are you sure?) Is Trang smart? (I’m afraid not)’²⁸
 b. *Tôi rửa bát rồi mà!*
 1SG wash bowl CRS MA
 ‘I’ve already washed the dishes (I tell you/I swear)!’

In the context of this paper, these three discourse particles are taken to instantiate the head of the Sentience/Evidentiality Phrase and license a [+Speaker] SEAT OF KNOWLEDGE. Let us now return to the question of how *nó*_{EXPL} is analyzed in the model sketched above. We basically adopt the syntactic representation proposed in Tenny (2006) for the Speech Act and Sentience Domains. However, as mentioned above, her feature geometry for pronouns should be enriched by an additional feature [\pm R(eferential)] in order to account for expletives. Our analysis thus combines Dao’s (2014) proposal and Tenny’s (2006). Importantly, in contrast to Dao (2014), we assume that *nó*_{REF} and *nó*_{EXPL} are not lexically specified as two distinct items but constitute two uses of the same element. As such, they both stem from a same set of features and are (gradually) specified in the course of syntactic derivation. We propose that *nó* has the basic featural make-up: [\pm R(eferential); \pm Sen(tient); –Sp(eaker); –Ad(dressee); –D(iscourse-participant)], which can actually be reduced to [\pm R(eferential); \pm Sen(tient); –Sp(eaker); –D(iscourse-participant)]. The negative value of the R-feature can be activated only if *nó* is externally merged in subject position, viz. [Spec, TP]. By contrast, its positive value shows up typically (but not only) when *nó* is introduced and licensed within the thematic domain.

28. In this configuration, even though *mà* is not sentence-final on the surface, it arguably occupies the same left-peripheral position (Dao 2019a). The difference between the ultimate surface structure of (56a) and that of (56b) resides in which constituent raises past *mà*: in the former, it is a DP while in the latter, it is at least a TP/IP

Crucially, [-R] entails [-Sen] while [+R] may trigger [\pm Sen]. $N\acute{o}_{REF}$ is [+R; \pm Sen] since it can be used to refer to humans or inanimate objects. Conversely, $n\acute{o}_{EXPL}$ is [-R; -Sen]. The feature [-Sp] may *a priori* seem redundant as it is entailed by [-D]. Nonetheless, as shown above, the connection to the speaker appears to be an inherent property of $n\acute{o}_{EXPL}$. Moreover, the [Sp] feature is also what distinguishes $n\acute{o}_{EXPL}$ from its referential counterpart. The featural make-up of $n\acute{o}_{REF}$ is thus [+R; \pm Sen; -Sp; -D]. What about that of $n\acute{o}_{EXPL}$? We suggest that owing to its discourse-linking nature, $n\acute{o}_{EXPL}$ must have a positive value specification for both [Sp] and [D]. Its set of features therefore contains [+Sp; +D]. In our view, those are strong features that must be licensed (cf. Examples (6)–(8)). How is that obtained? We propose that initially both [Sp] and [D] are negatively specified. $N\acute{o}_{EXPL}$ enters a “matching” process with the Sentience head whose specifier hosts the SEAT OF KNOWLEDGE, which may happen to coincide with the SPEAKER. It is the Sentience head that “overwrites” the features of $n\acute{o}_{EXPL}$ with its own. Note that [-R] cannot be overwritten, even if one may argue that SPEAKER and SEAT OF KNOWLEDGE are by default [+R]. That is because the R-feature on $n\acute{o}_{EXP}$ has been licensed by its structural position. In addition, if that was the case, $n\acute{o}_{EXP}$ would be no different from a 1st personal pronoun. On the other hand, the [+Sen] feature indicates that the source of information is the speaker herself.

$$(57) \quad [{}_{SA^*P} \text{ SPEAKER } [+Sp, +D] \dots [{}_{Sen^*P} \text{ SEAT OF KNOWLEDGE } [+Sen, +Sp, +D] \text{ Sen}^* \boxed{[+Sen, +Sp, +D]}] \dots [{}_{TP} n\acute{o}_{EXPL} [-R, \boxed{+Sen, +Sp, +D}]] \dots$$

$$n\acute{o}_{EXPL} [-R, -SEN, -SP, -D] \rightarrow n\acute{o}_{EXPL} [-R, +SEN, +SP, +D]$$

What will happen if the SEAT OF KNOWLEDGE is not the SPEAKER, but the ADDRESSEE (recall that ADDRESSEE is [-Sp; +D]), as in interrogatives? We would like to suggest that in such a case, the features of $n\acute{o}_{EXP}$ are jointly “overwritten” by the Sentience and Speech Act heads:

$$(58) \quad [{}_{SA^*P} \text{ SPEAKER } [+Sp, +D] \text{ SA}^* \boxed{[+Sp, +D]}] \dots [{}_{Sen^*P} \text{ SEAT OF KNOWLEDGE } [+Sen, -Sp, +D] \text{ Sen}^* \boxed{[+Sen, -Sp, +D]}] \dots [{}_{TP} n\acute{o}_{EXPL} [-R, \boxed{+Sen, +Sp, +D}]]$$

5. Conclusion

In this paper, we have attempted to offer a novel analysis of *nó*_{EXP}, an optional expletive subject pronoun in spoken Vietnamese. On the basis of a thorough overview of its main syntactic and semantic properties by comparison with its referential counterpart, we have suggested that the discourse properties manifested by *nó*_{EXP} should also be taken into account in order to get a better understanding of its peculiar behavior. Upon closer inspection, it appears that while it is externally merged in the subject position of the clause, viz. [Spec, TP/IP], its licensing strongly depends on the context and pragmatic factors. We have provided a critical assessment of Greco *et al.*'s (2018a, b) work, according to which the semantic contribution of *nó*_{EXP} to the interpretation of the clause containing it can be explained in terms of a specificity requirement. We have shown that not only can *nó*_{EXP} move to the left periphery, but it also does not seem to systematically impose such a requirement on sentences where it appears. Rather, the Epistemic Specificity advocated by these authors should be reinterpreted as a corollary of the speaker's involvement in her own discourse, or, more concretely, her assessment of a given state of affairs against her epistemic background. Based on our previous study of *nó*_{EXP} (Dao 2014), we have argued that, due to the speaker-dependency it exhibits, *nó*_{EXP} is best treated as an ego-evidential (or egophoric) marker. Incorporating insights from Speas-Tenny's syntactic approach to Evidentiality (Speas & Tenny 2003, Speas 2004, Tenny 2006), we have put forward an analysis in which both referential and expletive uses of *nó* derive from a same set of features. *Nó*_{EXP} is characterized by a featural make-up comprising [-R(eferential), +Sen(tient), +Sp(eaker), +D(iscourse-Participant)], which sharply contrasts with that of the referential *nó* ([+R; ± Sen; -Sp; -D]). More precisely, *nó*_{EXP} has been argued to acquire its evidential/speaker-related import via a "matching" process with the heads of two functional projections located in the left periphery, namely Sentience/Evidentiality Phrase and Speech Act Phrase. Essentially, while the [-R]-feature on *nó*_{EXP} is assumed to proceed from its structural position, all three other features, whose values are originally negative, appear to be "overwritten" by the Sentience and Speech Act heads.

Acknowledgements

I am very grateful to Dao Quynh Huong, Nguyen Kim Quy and Hoa Van Vuong for their judgments of construed data as well as of data I found on the web. I am also indebted to Albert Alvarez Gonzalez and Dominique Klingler for helpful comments and criticisms, and to Danh Thành Do-Hurinville for stimulating discussions on the topic. It goes without saying that all remaining errors, inaccuracies, omissions or misinterpretations of previously published data are my sole responsibility.

Abbreviations

2SG	second person, singular	FUT	future
3SG	third person, singular	LOG	logophoric
3PL	third person, plural	Nprop	proper noun
CLF	nominal classifier	NEG	negation
COMP	complementizer	PART	particle
CRS	currently relevant state	PERF	perfect
DEM	demonstrative	PL	plural
EMPH	emphatic	REFL	reflexive
EXPL	expletive	TOP	topicalizer

References

- Achard, Michel. 2009. The distribution of French intransitive predicates. *Linguistics* 47(3): 513–558. <https://doi.org/10.1515/LING.2009.018>
- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: OUP.
- Aikhenvald, Alexandra Y. 2014. The grammar of knowledge: A cross-linguistic view of evidentials and the expression of information source. In *The Grammar of Knowledge: A Cross-linguistic Typology*, Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds), 1–51. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780198701316.003.0001>
- Aikhenvald, Alexandra Y. 2018. *The Oxford Handbook of Evidentiality*. Oxford: OUP. <https://doi.org/10.1093/oxfordhb/9780198759515.001.0001>
- Alexiadou, Artemis, Anagnostopoulou, Elena & Everaert, Martin. 2004. *The Unaccusativity Puzzle*. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780199257652.001.0001>
- Banfield, Ann. 1982. *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. London: Routledge & Kegan Paul.
- Bennis, Hans. 1986. *Gaps and Dummies*. Amsterdam: Amsterdam University Press.
- Benveniste, Emile. 1946. Structure des relations de personne dans le verbe. *Bulletin de la Société de Linguistique* XLIII, fasc. 1, n°126. Reprinted in *Problèmes de linguistique générale*, Vol. 1, 225–236. Paris: Gallimard, 1966.
- Bianchi, Valentina. 2003. On finiteness as logophoric anchoring. In *Temps et point de vue/Tense and Point of View*, Jacqueline Guéron & Liliane Tasmovski (eds), 213–246. Nanterre: Université Paris X.
- Blain, Eleanor M. & Déchaine, Rose-Marie. 2007. Evidential types: evidence from Cree dialects. *International Journal of American Linguistics* 73: 257–91. <https://doi.org/10.1086/521728>
- Bolinger, Dwight. 1973. Ambient “it” is meaningful too. *Journal of Linguistics* 9(2): 261–270. <https://doi.org/10.1017/S0022226700003789>
- Botne, Robert. 1995. The pronominal origin of an evidential. *Diachronica* 12: 201–221. <https://doi.org/10.1075/dia.12.2.03bot>
- Brugman, Claudia M. & Macaulay, Monica. 2015. Characterizing evidentiality. *Linguistic Typology* 19(2): 201–237. <https://doi.org/10.1515/lingty-2015-0007>
- Cao, Xuan Hao. 2004. *Tiếng Việt, sơ thảo ngữ pháp chức năng* (Vietnamese, an outline of Functional Grammar). Hà Nội: Nhà xuất bản Giáo dục.

- Chafe, Wallace L. & Nichols, Johanna. 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood NJ: Ablex.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, Michael Kenstowicz (ed.), 1–52. Cambridge MA: The MIT Press.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. Oxford: OUP.
- Creissels, Denis. 2008. Remarks on so-called “conjunct/disjunct” systems. Paper presented at the *3rd Syntax of the World’s Languages Conference*, Berlin, 25–28 September.
- Dao, Huy Linh. 2012. Intransitivité scindée, passif et sujet impersonnel en vietnamien. In *Actes du 31^e Colloque International sur le Lexique et la Grammaire*, Jan Radimský (ed.), 49–55. České Budějovice: Université de Bohême du Sud à České Budějovice.
- Dao, Huy Linh. 2013. Intransitivité scindée, passif adversatif et sujet explétif en vietnamien. *Faits de Langues* 41: 55–78. <https://doi.org/10.1163/19589514-041-01-900000005>
- Dao, Huy Linh. 2014. On expletive subject pronoun in Vietnamese: Syntax or pragmatics, or both? Paper presented at the International Conference *Syntax of World’s Languages 6* (SWL 6), Pavia, Italy, 8–10 September.
- Dao, Huy Linh. 2015. Inaccusativité et diathèses verbales: Le cas du vietnamien. PhD dissertation, University Nouvelle Sorbonne – Paris 3.
- Dao, Huy Linh. 2019a. Dénouer la complexité de la périphérie gauche de l’énoncé vietnamien: Aspects syntaxique et sémantique. In *Complexité des structures et des systèmes linguistiques: Le cas des langues romanes*, Ivaylo Burov & Giuliana Fiorentino (eds), 219–240. Sofia: CU Romanistika.
- Dao, Huy Linh. 2019b. De la subjectivité et de l’intersubjectivité linguistiques en vietnamien. Paper presented at the Seminar *Subjectivité et langage*, 25 October, INALCO.
- Dao, Huy Linh & Do-Hurinville, Danh Thành. 2013. Les structures hypothétiques en vietnamien: *Mà* comme marqueur de non-assertion. *Travaux Linguistiques du CerLiCo* 25: 227–244.
- de Haan, Ferdinand. 1999. Evidentiality and epistemic modality: Setting boundaries. *Southwest Journal of Linguistics* 18: 83–101.
- de Haan, Ferdinand. 2005. Encoding speaker perspective: Evidentials. In *Linguistic Diversity and Language Theories* [Studies in Language Companion Series 72], Zygmunt Frajzyngier, Adam Hodges & David S. Rood (eds.), 379–397. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.72.18haa>
- Dendale, Patrick & Tasmowski, Liliane. 2001. Introduction: Evidentiality and related notions. *Journal of Pragmatics* 33: 339–348. [https://doi.org/10.1016/S0378-2166\(00\)00005-9](https://doi.org/10.1016/S0378-2166(00)00005-9)
- Enfield, Nick J. & Comrie, Bernard. 2015. Mainland Southeast Asian languages: State of the art and new directions, In *Languages of Mainland Southeast Asia: The State of the Art*, Nick J. Enfield & Bernard Comrie (eds), 1–27. Berlin: De Gruyter Mouton.
- Ernst, Thomas. 2009. Speaker-oriented adverbs. *Natural Language and Linguistic Theory* 27: 497–544. <https://doi.org/10.1007/s11049-009-9069-1>
- Erteschik-Shir, Nomi. 1997. *The Dynamics of Focus Structure*. Cambridge: CUP.
- Faller, Martina T. 2002. Semantics and Pragmatics of Evidentials in Cuzco Quechua. PhD dissertation, Stanford University.
- Farkas, Donka F. 2002. Specificity distinction. *Journal of Semantics* 19: 213–243. <https://doi.org/10.1093/jos/19.3.213>
- Fodor, Janet Dean & Sag, Ivan A. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5: 355–398. <https://doi.org/10.1007/BF00351459>

- Garrett, Edward John. 2001. Evidentiality and Assertion in Tibetan. PhD dissertation, University of California.
- Greco, Ciro, Phan, Trang & Haegeman, Liliane. 2018a. On *nó*, an optional expletive in Vietnamese. In *Null Subjects in Generative Grammar*, Federica Cognola & Jan Casalicchio (eds), 31–51. Oxford: OUP.
- Greco, Ciro, Haegeman, Liliane & Phan, Trang. 2018b. Expletives and speaker related meaning. In *Order and Structure in Syntax, II: Subjecthood and Argument Structure*, Michelle Sheehan & Laura R. Bailey (eds), 69–93. Berlin: Language Science Press.
- Guéron, Jacqueline & Haegeman, Liliane. 2012. *Je est un autre*. Subject positions, point of view and the neuter pronoun *tet* in West Flemish. In *Functional Heads*, Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds), 81–91. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780199746736.003.0005>
- Hale, Ken L. & Keyser, Samuel Jay. 1993. *On argument structure and the lexical expression of syntactic relations*. Ms, MIT.
- Harley, Heidi & Ritter, Elizabeth. 2002. Person and number in pronouns: A feature geometric analysis. *Language* 8(3): 482–526. <https://doi.org/10.1353/lan.2002.0158>
- Hellan, Lars. 1981. On semantic scope. In *Ambiguities in Intensional Contexts*, Frank Heny (ed.), 47–81. Dordrecht: Reidel.
- Huang, C. T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531–574.
- Ionin, Tania, Ko, Heejeong & Wexler, Kenneth. 2004. Article semantics in L2 acquisition: The role of specificity. *Language Acquisition* 12(1): 3–69. https://doi.org/10.1207/s15327817la1201_2
- Kasher, Asa & Gabbay, Dov M. 1976. On the semantics and pragmatics of specific and non-specific indefinite expressions. *Theoretical Linguistics* 3: 145–190. <https://doi.org/10.1515/thli.1976.3.1-3.145>
- Lahousse, Karen. 2003. La complexité de la notion de topique et l'inversion du sujet nominal. *Travaux de Linguistique* 47: 111–136. <https://doi.org/10.3917/tl.047.0111>
- Lahousse, Karen. 2007. Implicit stage topics. *Discours* 1. <https://doi.org/10.4000/discours.117>
- Levin, Beth & Rappaport Hovav, Malka. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge MA: The MIT Press.
- Li, Charles N. & Thompson, Sandra A. 1976. Subject and topic: A new typology of language. In *Subject and Topic*, Charles N. Li (ed.), 457–489. New York NY: Academic Press.
- Li, Charles N. & Thompson, Sandra A. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley CA: University of California Press.
- Milner, Jean-Claude. 1978. *De la syntaxe à l'interprétation*. Paris: Seuil.
- Mithun, Marianne. 1986. Evidential diachrony in Northern Iroquoian. In *Evidentiality: The Linguistic Coding of Epistemology*, Wallace Chafe & Johanna Nichols (eds), 89–112. Norwood NJ: Ablex.
- Nguyen, Phu Phong. 1992. La deixis en vietnamien. In *La deixis*, Mary-Annick Morel & Laurent Danon-Boileau (eds), 177–186. Paris: PUF.
- Nguyen, Van Hiep & Hoang, Thi Thu Thuy. 2011. Về chủ ngữ giả trong tiếng Việt” (On dummy subject in Vietnamese). *Ngôn Ngữ* 270: 47–58.
- Paykin, Katia. 2010. Il pleut des idées reçues: NP expansions of weather verbs. *Linguisticae Investigationes* 32(2): 253–266. <https://doi.org/10.1075/li.33.2.08pay>
- Paris, Marie-Claude. 1992. Démonstratifs et personne en chinois standard. In *La deixis*, Mary-Annick Morel & Laurent Danon-Boileau (eds), 165–175. Paris: PUF.
- Pesetsky, David. 1995. *Zero Syntax*. Cambridge MA: The MIT Press.

- Phan, Trang. 2013. Syntax of Vietnamese Aspect. PhD dissertation, University of Sheffield.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge MA: The MIT Press.
- Rooryck, Johan. 2001a. Evidentiality, Part I. *GLOT International* 5: 125–133.
- Rooryck, Johan. 2001b. Evidentiality, Part II. *GLOT International* 5: 161–168.
- Ruwet, Nicolas. 1982. *Grammaire des insultes et autres études*. Paris: Seuil.
- Ruwet, Nicolas. 1988. Les verbes météorologiques et l'hypothèse inaccusative. In *Grammaire et histoire de la grammaire. Hommage à la mémoire de Jean Stéfanini*, Claire Blanche-Benveniste, André Cherval & Maurice Gross (eds), 383–402. Aix-en-Provence: Publications de l'Université de Provence.
- Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18(3): 445–479.
- Searle, John R. 2004. *Mind: A Brief Introduction*. Oxford: OUP.
- Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114: 255–276. [https://doi.org/10.1016/S0024-3841\(03\)00030-5](https://doi.org/10.1016/S0024-3841(03)00030-5)
- Speas, Margaret. 2008. On the syntax and semantics of evidentials. *Language and Linguistics Compass* 2: 940–965. <https://doi.org/10.1111/j.1749-818X.2008.00069.x>
- Speas, Margaret. 2010. Evidentials as generalized functional heads. In *Edges, Heads, and Projections: Interface Properties* [Linguistik Aktuell/Linguistics Today 156], Anna Maria Di Sciullo & Virginia Hill (eds), 127–150. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.156.10spe>
- Speas, Margaret & Tenny, Carol. 2003. Configurational properties of point of view roles. In *Asymmetry in Grammar* [Linguistik Aktuell/Linguistics Today 57], Anna Maria Di Sciullo (ed), 315–344. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.57.15spe>
- Tamba, Irène. 1992. Démonstratifs et personnels en japonais. Deixis et double structuration de l'espace discursif. *La deixis*, Mary-Annick Morel & Laurent Danon-Boileau (eds), 187–195. Paris: PUF.
- Tenny, Carol. 2006. Evidentiality, experiencers, and the syntax of sentience in Japanese. *Journal of East Asian Linguistics* 15: 245–288. <https://doi.org/10.1007/s10831-006-0002-x>
- Widmer, Manuel & Zúñiga, Fernando. 2017. Egophoricity, involvement, and semantic roles in Tibeto-Burman languages. *Open Linguistics* 1: 419–441.

The final particle *like* in Northern English

A particle of reformulation in the context of interenunciative readjustment

Sylvie Hancil

University of Rouen Normandy

The study of the use of quotative *like* has been the object of increasing interest from linguists over the last two decades (see Romaine & Lange 1991; Buchstaller 2002, among others) as it is a recent and global phenomenon. By contrast, final *like* is more established and restricted to varieties in the British Isles. It is possible to further this research by relying on the NECTE corpus (Newcastle Electronic Corpus of Tyneside English), since in Northern English, especially in Newcastle English, the particle is developing a new function of discourse marker; it is associated with semantic values of anaphor, cataphor, intensifier and filler. The final position of the marker can be held responsible for the development of a significant number of interactional forces between speaker and co-speaker, which reveal the discursive strategy of the speaker towards the co-speaker (Pomerantz 1984). The study will be complemented by the analysis of the final particle *like* in the light of Brown & Levinson (1987)'s politeness theory and it will be shown that its use is a combination of negative and positive politeness within the utterance.

Keywords: final particle, anaphor, cataphor, intensifier, filler, interactional forces, politeness theory

1. Introduction

Discourse *like* has been the object of considerable research over the last twenty years in North American Englishes (Fuller 2003; D'Arcy 2005, 2008, 2017; Kastronic 2011; Tagliamonte 2016), alongside in British English (Miller & Weinert 1995; Levey 2006; Bartlett 2013; Truesdale & Meyerhoff 2015), Australasian English (Sharifian & Malcom 2003; Miller 2009), and in Irish English/Hiberno-English (Siemund, Maier & Schweinberger 2009; Nestor et al. 2013; Corrigan 2015; Schweinberger 2015; Diskin 2017; Corrigan & Diskin 2019).

It can be observed that discourse *like* can appear initially, medially and finally in the clause, as in (1), (2) and (3) respectively:

- (1) *Like* if you want to work in Macau.
 (2) Friends in there they're **like** laughing.
 (3) I hadn't a clue **like**. (Corrigan & Diskin, forthc.: 4)

So far much has been written on quotative *like*, which is a recent and global phenomenon. It is a fact that the study of quotative *like* by different linguists (Romaine & Lange 1991; Buchstaller 2002, inter alia) has shown how this marker could contribute to the dynamic rendering of speech and thought. Initially used in *valspeak*, a dialect employed in Californian English, the marker can now be said to be fully integrated into the colloquial conversation of almost English native speakers under forty, all over the world.

But very little has been written on final *like*, which is more established and typically restricted to varieties in the British Isles. The purpose of this paper is to close the gap and shed some more light on a typical feature of Northern English grammar, i.e. the use of adverbials in sentence-final position. Among the various possible adverbials that can be encountered (e.g. *but*, *like*, *mind*), *like* represents a highly interesting marker since corpus-based evidence provided by the analysis of the NECTE corpus (Newcastle Electronic Corpus of Tyneside English) shows that it has semantically and pragmatically evolved.

From a geographic point of view, final *like* can be linked with the British Isles and is said to be typically a northern feature (Hedevind 1967: 237; Miller & Weinert 1995: 368; Truesdale & Meyerhoff 2015: 9–10; Corrigan & Diskin 2019: 1). Among the features characteristic of the North-East of England are discourse markers used in sentence-final position, as underlined by Beal, Burbano & Llamas (2012):

Repetition of the grammatical subject at the end of a sentence, known as right dislocation, is a discourse feature used throughout the North-East. Examples from Tyneside, Wearside and Teeside are:

I'm a Geordie, me, like.

That's awful that !

He's lush him.

She's a liar, her.

(Beal, Burbano & Llamas 2012 : 93)

Consider also 'Ay up, the accent's on good job prospects' (*Daily Mail* 19 June 1996, on the move to telesales to Sheffield and Leeds); and 'Regional riches put one over on standard English, *like*' (*Guardian* 16 March 1998). Irritating though such [exclamations and discourse markers] may be (and 'Howay the lads ! for Geordies !') it is nonetheless interesting to note how they do suggest a spoken idiom, and features

rarely transposed to the printed page. Unlike visual signs such as cloth caps, braces and whippets, linguistic stereotypes also sometimes distinguish very broadly different types of Northerners (e.g. Yorkshire *tykes*, Scousers and Geordies) and their local speech.

The evolution of the final particle is worth being further explored in the NECTE corpus since in Northern English, especially in Geordie English, it has developed into a discourse marker with a great variety of meanings, extending from an anaphoric/cataphoric value to an intensifying value, via a phatic value. The final position of the marker can be held responsible for the development of a number of interactional forces between the participants, which also reveal the speaker's discourse strategy towards the hearer (Pomerantz 1984). The study will be complemented by the analysis of sentence-final *like* in the light of the theory of politeness elaborated by Brown and Levinson (1987).

2. A preliminary study

The final use of *like* (Andersen 2001: 216) would seem to be simply a variant of the discourse marker *like*, on par with the initial or medial occurrences used in discourse. In fact, it is not the case. Final *like* displays different characteristics which set it apart from the occurrences of *like* in other places of the utterance.

Historically, final *like* has different origins. As opposed to occurrences of initial/medial *like* that have scope on the foregoing discourse, final *like* has scope predominantly on the preceding discourse, thereby modifying the preceding segment. Jespersen (1954) tried to explain this by proposing that final *like* is a variant or a later development of the suffix *-like* which was attached originally to adjectives or adverbs. Little by little the suffix grew independent and “may now be added to any substantive and is frequently added to adjectives” (Jespersen 1954: 417), meaning ‘similar to’ or ‘characteristic of’. This form later evolved into the present-day final particle which is associated with colloquial language. Consequently, final *like* grew out of the suffix *-like* and not from the comparative preposition as its American counterpart did. It would be more judicious to identify the forms as two superficially similar forms, a comparative preposition, on the one hand, and a suffix, on the other hand, which fulfil equivalent functions, even though they have followed quite different developmental paths.

Before examining the Geordie English data in more detail, we shall start our analysis by looking over the definition of the marker in a few dictionaries. Among the various dictionaries consulted (Oxford Advanced Learner's Dictionary (OALD), Longman Dictionary of Contemporary English, Cambridge Advanced Learner's Dictionary, Oxford English Dictionary (OED)), only the OED and the

OALD point to the existence of this marker, without explicitly acknowledging its sentence-final position. Both underline that the marker is used in colloquial English but they slightly differ in their semantic interpretation: the OED states it is a way of qualifying a speaker's previous statement, whereas the OALD insists it is a way of providing more explanation on what was previously stated:

Like : dial. and vulgar. Used parenthetically to qualify a preceding statement: = 'as it were', 'so to speak'. Also, colloq. (orig. U.S.), as a meaningless interjection or expletive.

1801 tr. *Gabrielli's Myst. Husb.* III. 252 Of a sudden like.

1815 SCOTT Guy M. vi, The leddy, on ilka Christmas night ...gae twelve siller pennies to ilka puir body about, in honour of the twelve apostles like.

1826 WILSON, John *Noctes Ambrosianæ* (from Blackwood's Mag.) 1855 I.

179 In an ordinar way like.

1838 LYTTON, Edward *Alice, or the mysteries* II. iii, If your honour were more amongst us, there might be more discipline like.

1840–41 DE QUINCEY, Thomas *Style* II. Wks. 1862 X. 224 'Why like, it's gaily nigh like to four mile like'.

1911 A. BENNETT *Hilda Lessways* I. vi. 49 He hasn't passed his examinations like... He has that Mr. Karkeek to cover him like.

1961 *New Statesman* 22 Sept. 382/2 'You're a chauvinist,' Danny said. 'Oh, yeah. Is that bad like?'

1966 *Lancet* 17 Sept. 635/2 As we say pragmatically in Huddersfield, 'C'est la vie, like!' (OED, 1989, second edition online)

like : used in very informal speech, for example when you are thinking what to say next, explaining sth, or giving an example of sth: It was kind of scary, like. (OALD)

In the OED, we note that a number of these citations are from northern and Scottish dialects (Scott, Bennett). This definition also seems to conflate final and medial/initial 'like' in suggesting it has a US origin, whereas final 'like' is British and is said to have been affected by Irish/Scots in-migration historically (Beal & Corrigan 2009: 231–232).

Among the grammarians, Jespersen (1954) notices that final *like* is employed parenthetically by inferiors speaking to superiors "to modify the whole of one's statement, a word or phrase, modestly indicating that one's choice of words was not, perhaps, quite felicitous" (Jespersen 1954: 417). Jespersen's (1954) description reflects other comments on the use of final *like* that mainly describe final *like* as a hedge marker indicating vagueness (Partridge 1984: 264). More recent analyses of final *like* contradict this view of final *like* being viewed with mainly a

hedging function. Miller and Weinert (1995: 388–390), for example, contend that final *like* “does have some retroactive focusing power, but more importantly, it (...) is ‘clearing up misunderstanding’” (Miller & Weinert 1995: 388–389), affirming further that such instances point to relevant arguments. Fleischman and Yaguello (2004: 132) develop a different analysis of final *like*. Even though Fleischman and Yaguello (2004) stand in line with Miller and Weinert (1995) that final *like* has scope over the preceding segment of the utterance, they suggest that it is not employed to offer objections; it disambiguates the scope of the underlined segment so that the co-speaker is guided in the interpretation. The most recent analysis of final *like* is elaborated by Corrigan and Diskin (2019). They work on Irish English and suggest that the marker has scope over discourse-new entities rather than discourse-old entities.

Let us consider now how the examination of the NECTE corpus confirms this semantic analysis.

3. Data

Northern Eastern English is examined in the *Newcastle Electronic Corpus of Tyneside English* (NECTE), an almost 1 million-word corpus, which is based on two pre-existing corpora representing audio-recorded speech: one of them from the late-60s’ *Tyneside Linguistic Survey* (TLS) project, which includes Tyneside speakers talking about their life stories and their attitudes to the local dialect; and the other from the 1994 *Phonological Variation and Change in Contemporary Spoken English* (PVC) project, which comprises dyads of Tyneside friends and relatives talking about a large variety of topics.

4. Theoretical background

Because language and action are temporally linked, the linguistic structures are at the service of the organization of actions, which has led Auer (2005) to speak of the grammar of spoken language as an on-line grammar: the shaping of syntax occurs in real time. The temporal feature of grammar has two consequences: projection and emergence, as underlined by Pekarek Doehler (2011):

Projection refers to the property of one segment of discourse (an action or part of an action, or a grammatical structure) to prefigure possible trajectories of the next (actional or grammatical) segment (Pekarek Doehler 2011: 46)

Even though projection is said to foreshadow a range of possible upcoming trajectories (Pekarek Doehler 2011: 47), it will be shown that in the case of final *like*, the tendency is almost unilateral: the co-speaker is guided towards a specific interpretation of the speaker's utterance.

In addition, the linguistic structures are constantly adapted to the constraints and the needs of actions along the process of their production. The findings are in complete accordance with Hopper's (1987) notion of emergent grammar. Far from a static interpretation of grammar as a set of motionless structures and rules, it is viewed as a dynamic bundle of constructions in constant adaptation, at the service of usage in the process of interactions.

5. Semantic values of sentence-final *like*

There are 276 tokens of the particle *like* in final position in the data. The following semantic values can be distinguished as follows.

5.1 Anaphoric value

Sentence-final *like* has developed an anaphoric interpretation. The speaker comes back to the previously mentioned information and attracts the hearer's attention to it. There are 66 occurrences of this semantic value in the data, namely 23.9%. The semantic value of the marker can be paraphrased as « you know », as displayed in (4) and (5):

- (4) A: *they just don't care whether they're educated or whether they're not you know now*
 B: *yeah that's true*
 A: *that's my opinion like*
 B: *yeah yeah do you think it's very different from when you went to school*
 A: *around here you know then* (tlsg24)
- (5) A: *but otherwise the Geordie I just keep tiv it but i divn't thi I divn't think it's as bad as what what some are using you know I mean when you hear it on the telly on the interviews oh i you know like.hm you laughs at stuff well you i divn't know i was you know what i mean you know*
 B: *yes*
 A: *things like that I well i think that i just think it's awful me like*
 B: *do you you don't like it do you not I mean* (tlsg27)

Let us have a look at Example (4). By using sentence-final *like*, the informant A makes a specific link with some chunk of discourse: he goes back to his previously stated opinion, which can be paraphrased as “They just don’t care whether X or Y” and he also specifies that his comment can only be assigned to him as it reflects his own personal subjective interpretation of the situation (“that’s my opinion”). In Example (5), by using final *like*, the informant A, who talks about the use of Newcastle English, goes back to the previous argument when he heard somebody using Geordie English on T and makes the comment it is awful.

5.2 Cataphoric value

In addition, the speaker can anticipate the information he is going to develop; sentence-final *like* is used with a cataphoric value, as illustrated in (6) and (7). The number of tokens with this semantic value amounts to 53, which makes 19.2% of the total.

- (6) A: *yes do you think you were glad to leave school or not </u><u who="informantTLSG01"> in a way yes I was dying to get to work*
 B: *aye to start to earn some money like*
 A: *and of course I mean the the poor conditions in the home*
 B: *yeah ehm* (tlsg01)
- (7) A: *what about thing things like discipline in schools you know do you think they're B: well i think they're too soft on them me really like*
 B: *do you*
 A: *uhhuh honestly because of the way they're walking around now you know* (tlsg27)

Let us have a closer look at (6). By employing sentence-final *like*, the informant B points to forecoming information, which will explain why he wanted to quit home to earn money, because he was living in poor conditions. In Example (7), by using final *like*, the informant A, who talks about discipline in schools, points to foregoing discourse and makes a comment about discipline.

5.3 Intensifying value

In the case of the intensifying value, the speaker refers back to the information previously developed and underlines it. It can be paraphrased by *really*, as shown in (8) and (9). There are very few tokens of this semantic value: 12, namely 4.3% of the total. This figure is the lowest in the data.

- (8) A: *well take a loan out <pause/> you take a loan out don't you <pause/> our lee's just took a loan out he took a loan out for three grand him <pause/> it's costing him eight*
 B: *how did he*
 A: *he he he <pause/> he's got to pay back like*
 B: *aye eh <pause/> they check <pause/> what how you're working and that as well*
 A: *he took his out <unclear/> man he aye i know* (pvc08)
- (9) A: *i'm now thinking about ehm <pause/> possibly buying my own house*
 B: *yes aye the the rents for these are pretty high aren't they*
 A: *aye that's that's the thing*
 B: *they're going up even further like*
 A: *yes that's the thing and ehm* (tlsg27)

Let us analyse Example (8). The informant A just informs the interviewer that Lee has taken a loan for an important sum of money and then, by using sentence-final *like*, he underlines the fact that the sensible thing to do next is to pay back in order to avoid further problems of money in the future. In Example (9), by using final *like*, the speaker B underlines that the rents are pretty high.

5.4 Phatic value

In the case of the phatic value, the marker is completely desemanticised and is used as a filler with a view to keeping the conversation going. The marker has acquired a purely phatic value, as exemplified in (10) and (11). There are 145 tokens of this semantic value in the corpus, namely more than half of the total of the occurrences of sentence-final *like* (52.5%). This is the most represented semantic value in the data.

- (10) A: *no you say turn cree*
 B: *mm*
 A: *like a hen cree*
 B: *yes*
 A: *what else is there besides a he a hut like*
 B: *well just huts and all sorts pigeon cree <laughter"/>* (tlsg22)
- (11) B: *scotswood road yeah what ehm what job do you do there like*
 A: *millwright*
 B: *yeah are you ehm have you finished your apprenticeship like*
 A: *aye i've been two year now* (tlsg30)

Contrary to previous studies, it turns out that in Northern English in Newcastle, the majority of the semantic values is the phatic value and the intensifying value is rather rare (less than 5%) in the data.

Now that we have gone through the various semantic values of sentence-final *like*, we shall try to explain how its use is pragmatically motivated in the flow of conversation.

6. Discourse strategy and interactional forces

The postposing of constituents is directly related to the speaker's discourse strategy, and can be accounted for in interactional terms.

Indeed, the genre of the text directly influences its communicative goal: informal conversations involve exchanges between interlocutors, so they are hearer-oriented. The syntax, i.e. the linear order of constituents, adapts itself to these interactional forces. Various studies in discourse analysis have long established that the ideal place for the hearer to manifest his point of view about the speaker's sentence is at the end of the speaker's sentence (see Pomerantz 1984, *inter alia*).

This ideal place has specific characteristics. The interaction of semantic, pragmatic, syntactic and prosodic units is defined as a complex unit called Complex Transition Relevance Place, as defined by Ford and Thompson (1996) below:

Intonation and pragmatic completion points select from the syntactic completion points to form what we will call « Complex Transition Relevance Places. The term 'turn unit', then, will refer to a unit which is characterized by ending at a CTRP.

(Ford & Thompson 1996: 154)

According to Ford and Thompson's (1996: 155) data analysis, there is a high proportion of speaker change at CTRPs (71%) but it is possible to find a significant number of occurrences of speaker change in non-CTRPs, in which syntactic completion points do not coincide with intonational completion points (47%) and for which speaker change can only be explained in interactional terms (Ford & Thompson 1996 : 159).

In our data, the presence of sentence-final *like* systematically generates speaker change in the corpus of NECTE. Every single occurrence of speaker change occurs after a non-CTRP.

6.1 Encouragement of the hearer's agreement

6.1.1 Examples of agreement

In the NECTE corpus, there is a high proportion of sentence-final *like* which is followed by markers of agreement such as *aye* (Example (12)), *mm* (Example (13)) and *uh huh* (Example (14)), as illustrated below:

- (12) B: *the eh*
 A: *doddingston ales*
 B: *was it*
 A: *aye doddingston ales was deuchar's beer aye*
 B: *i didn't know that like*
 B: *aye*
 A: *lockside* (pvc02)
- (13) A: *ehm just have what you want stand behind the bar*
 B: *come out completely blitzed*
 A: *<unclear/> i didn't know if i was still in the boat when i got home it was like*
<pause/> definitely haven't got sea legs like
 B: *mm*
 A: *<unclear/> felt like* (pvc01)
- (14) A: *the holiday's over <pause/> you know what i mean like*
 B: *uh huh </u>*
 A: *it's just <unclear/> <pause/> and then you've probably saved that money*
up "oh i want to go on holiday again" (pvc08)

6.1.2 Markers of disagreement

Nevertheless, it is possible to find a few occurrences of sentence-final *like* immediately followed by markers of disagreement such as *no*, as shown in (15):

- (15) B: *aye*
 B: *no the kemp twins*
 A: *they were good like but <pause/> it would have been good just to see them*
 (pvc01)

6.2 Markers of the hearer's attitude

In addition, you can identify occurrences of sentence-final *like* followed by markers of the hearer's attitude. Here are examples in which the hearer evaluates the validity of the previously mentioned information (Example (16)), considers the situation as being disgusting (Example (17)) or even funny (Example (18)):

- (16) A: *she says she was <unclear/> she's dead quiet when she's by herself she was just like*
 B: **I think** *she's great*
 A: *she wasn't even giving the <unclear/> for a change*
 B: *oh she's really canny* (pvc01)
- (17) A: *i'm looking forward to lying in bed for the next couple of days like*
 B: *craig you're making us sick*
 A: *<vocal desc="laughter"/>*
 B: *it's not <unclear/>*
 A: *it's me that's on the sick* (pvc01)
- (18) B: *i would be eh <pause/> just when i got my first bike sixteen <pause/> when i went in and i i'll never ever*
 A: *v forget it we must have looked great if anybody could have gotten a photo there <pause/> it would have made the headlines here like* B: *<vocal desc="laughter"/> but don't give <unclear/> now you know* (pvc02)

7. Politeness theory and sentence-final *like*

It would now be worth showing how the various meanings of sentence-final *like* can be understood in the light of Brown & Levinson's (1987) theory of politeness.

By definition, negative politeness is addressed to the hearer's negative face, primarily by attenuating the imposition by the speaker of a potential Face-Threatening Act (FTA). Because the speaker is keen to orientate the communication in a certain way, he needs to give the hearer subtle evidence of his discourse interpretation. In English, a discourse marker is said to privilege the initial syntactic position to allow the speaker to indicate his point of view about the situation of enunciation (see Kiparsky 1995 and Abraham 1991, *inter alia*). By contrast, various studies such as Pomerantz (1984) have emphasized the importance of the role of the final syntactic position for the manifestation of the hearer's point of view. In order to make the interactants' views coincide, the speaker cannot but use this syntactic position to indirectly guide the hearer's opinion. By moving away from the traditional initial syntactic position, the speaker minimizes his point of view, and, in a way, distantiates from it, which contributes to make the hearer believe that she is the one who elaborates her own point of view and that it is not imposed on her.

Besides, the use of sentence-final *like* can also be said to be a manifestation of positive politeness. This type of politeness is used to attend to the hearer's interests, wants and needs. So the speaker uses sentence-final *like* to claim common ground and cooperation. This marker can often be interchangeable with a familiar address

such as *man*, which shows that the speaker wants to employ ingroup markers in order to include the hearer in his sphere of enunciation. When sentence-final *like* is used with an anaphoric value, the speaker points to the coherence of his discourse: he makes the link between *like* and the previously given information, which is already known by the hearer. The speaker shows his certainty as regards the content of the previous sentence. The speaker's attitude is a confirmation for the hearer that she can trust the speaker's judgment, which satisfies the hearer's positive face. Likewise, when sentence-final *like* is employed with a cataphoric value, the speaker makes a link between *like* and the coming information, which is new to the hearer, thereby fitting well with the hearer's expectations and her positive face. When sentence-final *like* has an intensifying value, it also partakes of the same strategy of cooperation between the two interactants. This linguistic marking of enthusiasm is very often a way for the speaker to encourage a positive reply from the hearer. When sentence-final *like* is completely desemantized and fulfils the role of a filler, it can still be understood as a marker of cooperation in that it shows the hearer when it is her turn to speak, thereby maintaining the rhythm of the conversation between the two interlocutors. All in all, the use of sentence-final *like* in positive politeness can be considered as a Face-Enhancing Act¹ for the positive faces of both the speaker and the hearer, which is in accordance with Beeching's (2002) insightful remarks about the pragmatic nature of hedges in French discourse.

The combination of negative and positive politeness in the use of sentence-final *like* is a good illustration about the way power and solidarity can be manipulated by the speaker to persuade the hearer how to pursue the conversation in a certain direction.

8. Conclusion

To conclude, the synchronic study of sentence-final *like* in the NECTE corpus shows that there has emerged a discourse marker fulfilling various semantic meanings, which contribute to enrich the interactional forces between the interactants.

It would be worth exploring the data further by introducing sociolinguistic criteria to determine how the discourse marker has become integrated into the grammatical paradigm within the different strata of the population in Geordie English. Another area that would need further study is the way *like* can be interchanged with *but* in final position to see how the two variants stand in competition and if one is to be replaced by the other.

1. The expression was introduced by Kerbrat-Orecchioni (1997) in her criticism of Brown & Levinson's (1987) theory of politeness.

All In all, the study of final particles (FPs) has a lot to offer for linguistic description and studies on discourse structure. There are at least six domains in which the study of final particles sheds new light on linguistic theory:

- i. **Theories of Grammar:** FPs are elements of spoken grammar and thus relevant for theories of the structure of spoken language. Since their word class affiliation is indeterminate, they represent a challenge for linguistic categorization.
- ii. **Syntactic Theory:** The syntactic status of FPs is far from clear. The role of the final position (or “right periphery”, or “post-field”) in the construction of a sentence or a structural unit of any type (an “utterance” or “unit of talk”) needs to be rethought in order to encompass FPs.
- iii. **Cognitive Linguistics:** Final particles are prototypical procedural markers, indicating how the linguistic unit they are attached to is to be understood by the addressee. They can index the cognitive status of a message (e.g. surprise/unexpectedness, plausibility) and provide different cues on how to process the message they accompany.
- iv. **Discourse Analysis and Conversation Analysis:** FPs are used to structure jointly produced discourse. They have a cohesive function that needs to be explored and related to the conditions of interactive discourse production. FPs structure conversational turns in that they mark transition-relevant places and thus play a role in the turn-taking system that regulates interpersonal communication.
- v. **Discourse-Pragmatic Variation:** Both the initial and the final position are preferred places for the indication of information that is relevant for the processing of a message. More attention needs to be paid to the question if, and in what way, the initial and the final position differ in their communicative function.
- vi. **Grammaticalization:** FPs are based on a variety of source lexemes which, over time, developed heterosemes in final position, where they acquired metapragmatic functions and lost much of the original semantic content. Similarities in the development of particular types of FPs in different languages (e.g. FPs based on conjunctions) suggest typologically consistent, recurrent diachronic pathways.

References

Corpus

NECTE corpus (Newcastle Electronic Corpus of Tyneside English) (accessed September 2009).

Primary sources

- Cambridge Advanced Learner's Dictionary*, edition online.
Longman Dictionary of Contemporary English, edition online.
Oxford Advanced Learner's Dictionary (OALD), edition online.
Oxford English Dictionary (OED), 1989, second edition online.

Secondary sources

- Abraham, Werner. 1991. Discourse particles in German: How does their illocutive force come about? In *Discourse Particles* [Pragmatics & Beyond New Series 12], Werner Abraham (ed.), 203–252. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.12.o8abr>
- Andersen, Gisle. 2001. *Pragmatic Markers and Sociolinguistic Variation: A Relevance-theoretic Approach to the Language of Adolescents* [Pragmatics & Beyond New Series 84]. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.84>
- Bartlett, Joanne. 2013. 'Oh I just talk normal, like': A corpus-based, longitudinal study of constituent final *like* in Tyneside English. *Newcastle Working Papers in Linguistics* 19(1):1–21.
- Beal, Joan, Burbano Elizondo, Lourdes & Llamas, Carmen. 2012. *Urban North-Eastern English: Tyneside to Teeside*. Edinburgh: EUP.
- Beeching, Kate. 2002. *Gender, Politeness and Pragmatic Particles in French* [Pragmatics & Beyond New Series 104]. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.104>
- Brown, Penelope & Levinson, Stephen C. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511813085>
- Buchstaller, Isabelle. 2002. He goes and I'm like: The new quotatives revisited. In *Internet Proceedings of the University of Edinburgh Postgraduate Conference 2002*. <<http://www.lel.ed.ac.uk/~pgc/archive/2002/proc02/buchstaller02.pdf>> (3 December 2020).
- Corrigan, Karen. 2015. 'I always think of people here, you now, saying like after every sentence': The dynamics of discourse-pragmatic markers in Irish-English. In *Pragmatic markers in Irish English* [Pragmatics & Beyond New Series 258], Carolina Amador-Moreno, Kevin McCafferty & Elaine Vaughan (eds), 37–64. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.258.o2cor>
- Corrigan, Karen & Diskin, Chloé. 2019. 'Northmen, Southmen, comrades all'? The adoption of discourse like by migrants north and south of the Irish border. *Language in Society* 4(5): 745–773. <https://doi.org/10.1017/S0047404519000800>
- D'Arcy, Alexandra. 2005. *Like: Syntax and Development*. PhD dissertation, University of Toronto.
- D'Arcy, Alexandra. 2008. Canadian English as a window to the rise of like in discourse. *Anglistik* 19(2):125–40.

- D'Arcy, Alexandra. 2017. *Discourse-pragmatic Variation in Context: Eight Hundred Years of LIKE* [Studies in Language Companion Series 187]. Amsterdam: John Benjamins.
<https://doi.org/10.1075/slcs.187>
- Diskin, Chloé. 2017. The use of the discourse-pragmatic marker *like* by native and non-native speakers of English in Ireland. *Journal of Pragmatics* 120: 144–157.
<https://doi.org/10.1016/j.pragma.2017.08.004>
- Fleischman, Suzanne & Yaguello, Marina. 2004. Discourse markers across languages? Evidence from English and French. In *Discourse across Languages and Cultures* [Studies in Language Companion Series 68], Carol Lynn Moder & Aida Martinovic-Zic (eds), 129–147. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.68.08fle>
- Ford, Cecilia & Sandra Thompson. 1996. Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In: Elinor Ochs, Emanuel A. Schegloff & Sandra A. Thompson (eds.), *Interaction and Grammar*. Cambridge: Cambridge University Press, 134–184.
- Fuller, Janet M. 2003. Use of the discourse marker *like* in interviews. *Journal of Sociolinguistics* 7(3): 365–377. <https://doi.org/10.1111/1467-9481.00229>
- Hedevind, B. 1967. *The Dialect of Dendale in the West Riding of Yorkshire* [Acta Universitatis Upsaliensis Societas Anglistica Upsaliensi 5]. Uppsala: University of Uppsala.
- Kastronic, Laura. 2011. Discourse *like* in Quebec English. *University of Pennsylvania Working Papers in Linguistics* 17(2): 105–114.
- Kerbrat-Orecchioni, Catherine. 1997. A multilevel approach in the study of talk-in-interaction. *Pragmatics* 7(1): 1–20. <https://doi.org/10.1075/prag.7.1.01ker>
- Jespersen, Otto. 1954. *A Modern English Grammar on Historical Principles, Part 6 : Morphology*. London: Allen & Unwin.
- Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In *Clause Structure and Language Change*, Adrian Battye & Ian Roberts (eds), 140–169. Oxford: OUP.
- Levey, Stephen. 2006. The sociolinguistic distribution of discourse marker *like* in preadolescent speech. *Multilingua* 25: 413–441. <https://doi.org/10.1515/MULTI.2006.022>
- Miller, Jim. 2009. *Like* and other discourse markers. In *Comparative Studies in Australian and New Zealand English: Grammar and Beyond* [Varieties of English Around the World G39], Pam Peters, Peter Collins & Adam Smith (eds), 317–338. Amsterdam: John Benjamins.
<https://doi.org/10.1075/veaw.g39.18mil>
- Miller, Jim & Weinert, Regina. 1995. The function of *like* in dialogue. *Journal of Pragmatics* 23(4): 365–393. [https://doi.org/10.1016/0378-2166\(94\)00044-F](https://doi.org/10.1016/0378-2166(94)00044-F)
- Nestor, Niamh. 2013. The positional distribution of discourse *like*: A case study of young Poles in Ireland. In *Linguistic and Cultural Acquisition in a Migrant Community*, David Singleton, Vera Regan, & Ewelina Debaene (eds), 49–74. Bristol: Multilingual Matters.
<https://doi.org/10.21832/9781847699909-005>
- Partridge, Eric Honeywood. 1984. *A Dictionary of Slang and Unconventional English*. London: Routledge & Kegan Paul.
- Pekarek Doehler, Simona. 2011. Emergent grammar for all practical purposes: The on-line formatting of left and right dislocations in French conversation. In: Peter Auer & Stefan Pfänder (eds.), *Constructions: Emerging and Emergent*. Berlin: De Gruyter Mouton, 45–87.
- Pomerantz, Anita. 1984. Agreeing and disagreeing with assessments: Some features found in preferred/dispreferred turn shapes. In *Structures of Social Action, Studies in Conversation Analysis*, J. Maxwell Atkinson & John Heritage (eds), 57–101. Cambridge: CUP.

- Romaine, Suzanne & Lange, Deborah. 1991. The use of *like* as a marker of reported speech and thought. *American Speech* 66: 227–278. <https://doi.org/10.2307/455799>
- Schweinberger, Martin. 2015. A comparative study of the pragmatic marker *like* in Irish English and in south-eastern varieties of British English. In *Pragmatic markers in Irish English* [Pragmatics & Beyond New Series 258], Carolina Amador-Moreno, Kevin McCafferty & Elaine Vaughan (eds), 114–34. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.258.05sch>
- Sharifian, Farzad & Malcolm, Ian G. 2003. The pragmatic marker *like* in English teen talk: Australian aboriginal usage. *Pragmatics and Cognition* 11(2): 327–44. <https://doi.org/10.1075/pc.11.2.07sha>
- Siemund, Peter, Maier, Georg & Schweinberger, Martin. 2009. Towards a more fine-grained analysis of the areal distributions of non-standard features of English. In *Language Contacts meet English Dialects: Studies in Honour of Markku Filppula*, Esa Penttillä & Heli Paulasto (eds), 19–46. Newcastle upon Tyne: Cambridge Scholars.
- Tagliamonte, Sali A. 2016. *Teen talk*. Cambridge: CUP. <https://doi.org/10.1017/CBO9781139583800>
- Truesdale, Sarah & Meyerhoff, Miriam. 2015. Acquiring some like-ness to others. *Te Reo* 58: 3–28.

On pragma-semantics of expressives

Between words and actions

Suren Zolyan

Immanuel Kant Baltic Federal University, Kaliningrad, Russia /
National Academy of Sciences, Yerevan, Armenia

Expressives can be understood as: (1) a manifestation of speaker's emotional or affective attitudes; (2) a peculiar type of performatives; (3) a linguistic entity expressing or displaying what is the case. The third conception is taken as the basic one; we attempt to develop it, taking into account Bühler's representational theory of language and the theory of performatives (Austin, Searle). Expressives are closely connected with behavioral patterns, situations, and actions. An actualization of verbal behavior (a performative component of expressives) meets its counterpart in the verbalization of action (a descriptive component). We suggest to describe pragma-semantics of expressives as based on prototypical behavioural frames.

Keywords: semantics, pragmatics, performatives, expressive speech acts, expressives, context, social meanings, cognitive frames

1. The theoretical background: Expressives and the theory of meaning

Our approach is based on the general assumption that meaning is a result of a conjunction of linguistic and extra-linguistic systems in the process of social interaction and communication. In particular, this can be described as procedures for correlating linguistic expressions with sets of actual and potential contexts. We suggest describing expressives as latent behavioral patterns with socially determined meanings. The so-called descriptive meaning of expressives can be considered as a result of accommodation of a precedential frame to some actual situation. With this approach, it becomes possible to relate logical and semantic characteristics with cognitive and pragmatic ones. Verbalizations of mental representations, as well as non-verbal and verbal actions, are considered as mutually affecting processes, not limited by operations on signs and texts, but also regulated by patterns of behavior and interpersonal interaction.

The processing of linguistic information presupposes not only linguistic competence but also knowledge of social rules and conventions regarding usage and contextual dependencies of linguistic entities. The semantics of linguistic structures is generated as a dynamic context-dependent derivative of linguistic, social, cognitive, and referential factors. The dynamic approach reflects the capacity of an utterance to produce new meanings in the process of its generation and functioning.

This can be described as a mechanism for correlating multimodal texts and utterances with sets of actual and latent contexts and possible worlds, described by these texts and utterances. With this approach, it becomes possible to relate logical and semantic characteristics to cognitive and pragmatic ones. The verbalization of communicative intentions and semantic interpretation of utterances, as well as non-verbal actions and speech acts are considered as mutually affecting processes, not limited by operations on signs and texts, but also regulated by patterns of behavior and interpersonal interaction. Lexical meaning can be considered as an initial level for the multi-dimensional semantic structure of context-sensitive utterances, or as a starting mechanism for meaning production within a communication.

The concept of “meaning” is fundamental not only in linguistics and semiotics but also in M. Weber’s sociological theory: for him, the subject of sociology is the understanding of “the meaning of behavior”: “Action is social insofar as, by virtue of the subjective meaning attached to it by the acting individual (or individuals), it takes account of the behavior of others and is thereby oriented in its course” (Weber 1964: 88). As one can see, the Weberian “subjective meaning” is attached to such actions as are aimed at others and thus acquire social relevance. Social actions are treated as a kind of communicative operation. This approach allows us to convert actional meanings into their verbal description through linguistic meanings, and vice versa: linguistic meanings of some text can explicate or even create some state of affairs (possible worlds).

Despite the fact that Weber’s theory had no references to linguistics, one can find its counterpart in linguistics too. First of all, Bronislaw Malinowski’s contextual theory can be mentioned. Malinowski challenged the very existence of lexical meaning apart from its social and cultural environment:

Precisely as in the reality of spoken or written languages, a word without linguistic context is a mere figment and stands for nothing by itself, so in the reality of a spoken living tongue, the utterance has no meaning except in the context of situation.
(Malinowski 1923: 307)

For different reasons, L. Hjelmslev also advocated the contextual approach:

The so-called lexical meanings of certain signs are nothing more than artificially isolated contextual meanings, or artificial paraphrases of such. No sign has any meaning if isolated; every sign-meaning arises in a context.

(Hjelmslev 1961[1943]: 41)

Thus, a word has to be considered not as an isolated entity but only within and with the context of its occurrence. Regarding the so-called context-free meaning of a word, it is, as a rule, the meaning in the most typical context (on the recent developments on this subject see: Jaszczolt 2005). Of course, Malinowski and Hjelmslev were referring to different types of context – extra-linguistic and linguistic (textual); however, both of these are relevant and should be taken into account. Malinowski even attempted to reconcile them: this

[...] landed us in another apparent antinomy: words are the elements of speech, but words do not exist. Having once recognized that words have no independent existence in the actual reality of speech, and having thus been drawn towards the concept of context, our next step is clear: we must devote our attention to the intermediate link between word and context, I mean to the linguistic text.

(Malinowski 1935: 23)

This can be understood to mean that *the link between word and context* is to be explicated through some special verbal entities: texts, utterances, indices and markers. This understanding of linguistic meaning may be more closely identified with interjections than with other parts of speech and signals the need to make a theoretical shift: one can consider lexical items with minimal inter-systemic value and maximal affordance to contextualization as a prototypical manifestation of meaning. Such a reverse occurred when referring to expressives (it will be discussed in more detail in the Chapter 3):

In Kaplan's (1999a) words, "It may be that the primary problem in semantics is not what does this or that mean, but rather in what form should we attempt to say what this or that means" (p. 3). He then uses this premise to begin building a semantic theory that is guided by expressives. That is, where Kaplan (1989) sought to assimilate expressives to regular semantics, Kaplan (1999a) seeks to assimilate regular semantics to expressives.

(Potts 2005: 189)

The pragma-semantic description of expressives presupposes that their semantics is to be considered not just a context-dependent entity.¹ It is also and mainly

1. We introduced the term *pragmasemantics* (in: Zolyan 1991: 171) based on the distinction between semantic pragmatics and pragmatic pragmatics, as it was suggested by Max Cresswell: "It is right that such an account should be called 'pragmatics' but perhaps it should be called 'semantic pragmatics'. This is because the way in which the context produces the sense is part of the meaning. Pragmatic pragmatics would then study cases where the meaning itself depended on the context. Unlike semantic pragmatics there would be no rules for getting from the context to the meaning, for if there were we could get from there to the sense and it would just be a case of semantic pragmatics" (Cresswell 1973: 238). Pragma-semantics seemed to us a more convenient term than the phrase semantic pragmatics, and it means semantics of deictic expressions and other lexical units determined by the context of speaking. In English, this term was coined by de Saussure (2007), but concerning the Gricean approach: "..., there is a strong *pragma-semantics*

determined by the contextual regularities. This may clarify the essential characteristic of language activity – how it is possible at the same time to communicate with words in order to do things and express ourselves with words. Even if one uses non-referential verbal items like exclamations, interjections, slurs, etiquette formulas, one needs to choose them in accordance with some behavioral practices and linguistic rules: some social and language conventions and constraints restrict freedom of contextualization. The instrumental and operational view of language presupposes that language is a deed, knowledge and object² and, while linguistic description may be restricted to just one of these aspects, there is a definite necessity to address possible ways in which they may be interconnected. Expressives are a very interesting instance when all these three aspects have to be considered as an integral entity. The semantics of expressives is the result and manifestation of a similar intersection between speech, knowledge, and behaviour, why their study is so essential for understanding the interaction of pragmatics and semantics.

2. Expressives in linguistic theory: Main approaches

The term “expressive” is used in different senses. However, it is possible to identify three main interpretations. Being heterogeneous, in many crucial points, they contradict, but at the same time, in some respect, complement each other (Foolen 2015). We shall try to describe how it is possible to co-relate them and which of the possible descriptions can be considered as prototypical. Three main approaches may be distinguished – (1) expressivity as a characteristic concomitant with any linguistic phenomenon in the process of communication (2) expressives as a special class of speech acts and (3) expressives as a special class of lexical and quasi-lexical units expressing the emotional states and attitudes. After a brief review of these approaches, we address possibilities to find intercorrelations between them.

approach, which is pursued by the followers of Grice and a number of scholars with a referential logical background, and with various degrees of commitment to truth conditionality (e.g., de Saussure 2007). This trend is mainly interested in the construction of meaning by the hearer, through cognitive or formal model” – (Kecskes 2016: 52). We still remain faithful to the first approach and understanding pragma-semantics as an interface between pragmatics and semantics. On the recent developments on the intersections between them and different standpoints see: Recanati 2006; Carston 2008; Saeed 2011; Gauker 2013; Longworth 2017; Depraetere & Salkie 2017; Preyer 2018; Capone 2019.

2. This is a paraphrase of the precursor of structural linguistics, Russian-Polish linguist Baudouin de Courtenay (1904: 536): “If a language should be considered a special kind of knowledge, then it can also be represented, on the one hand, as an action, as a deed, and on the other – as a thing, as an object of the external world”.

2.1 Expressivity and the expressive function of language

Firstly, in the broadest sense, expressives are understood as all phenomena associated with a manifestation of the so-called expressive function of language. According to scholars such as Karl Bühler and Roman Jakobson it expresses speakers' attitudes and emotional states. The expressive function co-exists with the representative (referential) and appellative (phatic) functions. Any speech act represents all these basic functions to some extent; therefore expressives are defined as all linguistic entities where an expression of speaker's attitudes, feelings and emotions prevails over propositional content of an utterance. This vision firstly was put forward by Bühler, when he identified three functions of language: "Though we do not dispute the dominance of the representational function of language, what now follows is suited and intended to delimit it". (Bühler 1990: 35). In addition to the primary representational function, the appellative (*Appell*): 'appeal' to the listener), and the expressive (*Ausdruck*: 'expression of the sender's inner states') were also distinguished by him. As Bühler concentrated on the representational function, there was very little said about the two other functions. However, it is essential that for him, all three functions are inherent characteristics of language and linguistic signs. All of them have to be actualized simultaneously due to the fact that a linguistic sign "...is a *symbol* by virtue of its coordination to objects and states of affairs, a *symptom* (*Anzeichen, indicium*: index) by virtue of its dependence on the sender, whose inner states it expresses, and a *signal* by virtue of its appeal to the hearer, whose inner or outer behaviour it directs as do other communicative signs" (Bühler 1990: 35).

As one can see, this semiotic trichotomy matches directly with the triadic structure of communicative acts: sender – receiver – message. These three functions of the sign act in a complex, and each of them under some circumstances may become the dominant and form a particular area of linguistic research:

These are only phenomena of dominance, in which one of the three fundamental relationships of the language sounds is in the foreground... Each of the three semantic functions of language signs discloses and identifies a specific realm of linguistic phenomena and facts... "Expression in language" and "appeal in language" are partial objects for all of language research, and thus display their own specific structures in comparison with representation in language. (Bühler 1990: 39)

This scheme was developed by Roman Jakobson, who refers to the newly emerging information theory and complements Bühler's approach with the basics of Shannon-Weaver's model of communication (Shannon 1948). He preserved three main Bühlerian functions and introduced three new ones: oriented on code, message and channel of communication. In his approach to expressives R. Jakobson

underlined their emotional background, and he used a synonymic term for it (*emotive*), coined by Anton Marty in 1908:

The so-called emotive or “expressive” function, focused on the addresser, aims at a direct expression of the speaker’s attitude toward what he is speaking about... The purely emotive stratum in language is presented by the interjections. They differ from the means of referential language both by their sound pattern (peculiar sound sequences or even sounds elsewhere unusual) and by their syntactic role (they are not components but equivalents of sentences). If we analyze language from the standpoint of the information it carries, we cannot restrict the notion of information to the cognitive aspect of language. A man, using expressive features to indicate his angry or ironic attitude, conveys ostensible information, and evidently this verbal behavior cannot be likened to such nonsemiotic, nutritive activities as “eating grapefruit” (despite Chatman’s bold simile). (Jakobson 1960: 354)

The approach of Bühler and Jakobson to the all-encompassing pan-linguistic manifestation of the expressive function is now widespread: “Expressivity can be found everywhere in language, in the lexicon, in phonology, morphology and syntax” (Foolen 2015: 479). In its broader sense, this approach can be as a basics for stylistic dimension of language and can also be found in cognitive linguistics:

To some extent, every instance of language use (and every linguistic unit) has conceptual import involving four dimensions: descriptive, expressive/emotive, interactive, and discursive... Expressive/emotive import is internal to the interlocutors, being conveyed but not described. An example would be the expression of pain, e.g., *Ouch* or *Ow!*, differing in the intensity of the pain experienced. These are conventional units of English which express an experience, rather than putting it onstage as a focused object of description. (Langacker 2012: 100)

However, sometimes expressives are understood in the narrow sense, as synonyms of *mimetics* and *ideophones* and identified as “an open lexical class of marked words that depict sensory imagery.” (Akita & Pardeshi 2019: 13).

2.2 Expressives as performatives

According to the abovementioned approach, almost any lexical or phraseological unit under certain conditions can acquire expressive significance. Thus it depends more on the modes and characteristics of its usage. Meanwhile, another approaches to expressives is possible. One can approach to Bühler’s theory from the different perspective, considering language as a social instrument (organon) and speech as a socially significant action. As it have been already mentioned, “the most interesting aspect of Bühler’s schema from a pragmatic perspective is the category of

the “speech action”. Long before Austin, Bühler had shown *how one can do things with word*: “For all concrete speech is in vital union with the rest of a person’s meaningful behaviour; it is among actions and is itself an action” (Bühler 1990: 61). Despite the fact that Bühler used the term *speech act* in a specific way,³ however, it is possible to relate his and the modern understandings with each other. As expressivity is directly related to speakers’ attitudes toward messages, so it was also logical to connect them with characteristics of speech act. Thus, the other option is to consider expressives as a special type of speech act specialized on performing of the expressive function: “a kind of performative concerned roughly with reactions to behaviour and with behaviour towards others and designed to exhibit attitudes and feelings” (Austin 1962: 83). In contrast to the previous approach, when any linguistic item may be endowed with expressivity, this is a particular class of lexical units, consisting mainly of verbs and etiquette expressions, but also including interjections and exclamations. Austin baptized this class “behabitives”. Behabitives can merge with the act itself and sometimes verbal forms of expression may seem to be redundant. However, at the same time, Austin points out a very significant feature that distinguishes behabitives/expressives, from “pure” actions, like – referring to the example of Jakobson and Chatman – eating a grapefruit:

Reactions such as resenting, applauding, and commending do involve espousing and committing ourselves in the way that advice and choice do. But behabitives commit us to like conduct, by implication, and not to that actual conduct.

(Austin 1962: 158)

These phrases are especially liable to pass over into pure performatives where the action which is suited to the word is itself a purely ritual action, the nonverbal action of bowing (‘I salute you’) or the verbal ritual of saying *Huzza* (‘I applaud’) (Austin 1962: 86). This marginal remark can be the key to understanding the correlation between verbal and non-verbal components of language behaviour. For instance, one can imagine a society in which eating grapefruit becomes a manifestation of solidarity. But for this to happen it must acquire a conventional meaning. That is, eating grapefruit must become only a resemblance, or imitation of itself, and can be replaced by some verbal expression (*behabitive*) and interpreted through it.

3. Cf.: “In the Sprachtheorie, ‘speech action’ (Sprechhandlung) means the concrete act of verbal communication (1990: 65–66), in contrast to the ‘speech act’ (Sprechakt): the latter term designates a speaker’s ‘virtual’ meaning intention, a concept that Bühler had taken over from Husserl’s theory of ‘sense-conferring acts’. This problem has to be taken into account when assessing Bühlers use of ‘pragmatic’ concepts, such as speech action, sympractical field and verbal exchange – it would be anachronistic to equate them, for instance, with post-Austin formulations of “speech act theory” or conversational analysis” (Daalder & Musolf 2011: 9–10). See also Musolf (1997).

Austin, as appears from the very name of this class, is inclined to associate it with behavioural patterns, and with them the expressing of feelings and emotions can be associated:

Behabitives include the notion of reaction to other people's behaviour and fortunes and of attitudes and expressions of attitudes to someone else's past conduct or imminent conduct. There are obvious connexions with both stating or describing what our feelings are and expressing, in the sense of venting our feelings, though behabitives are distinct from both of these. (Austin 1962: 159)

Later John Searle suggested a more appropriate name for this class – “expressives”, and this has become common and accepted. For Searle,

The illocutionary point of this class is to express the psychological state specified in the sincerity condition about a state of affairs specified in the propositional content. The paradigms of expressive verbs are ‘thank’, ‘congratulate’, ‘apologize’, ‘condole’, ‘deplore’, and ‘welcome’. (Searle 1976: 12)

The essential novelty of Searle's approach was an explicit reference to the evaluative aspects of expressive utterances: “Behabitives do not seem to me at all well defined (as Austin, I am sure, would have agreed) but it seems to involve notions of what is good or bad for the speaker and hearer” (Searle 1976: 8). These features, evaluated as “good or bad”, are subject to explication through some propositional complement attaching to an expressive: “I congratulate you + on your winning the race”. Singling out the same class of verbs, Searle and Austin proceed from slightly different criteria: Austin referred to the speaker's behavior and his/her reactions to the behavior of others, whereas Searle also referred to psychological states. In both cases, as a kind of performatives, verbs of this class are not subject to truth-conditional semantics, and conditions of sincerity determine their meaning. Obviously, there is a little in common between the approaches of Bühler and Jakobson, on the one hand, and the Austin and Searle, on the other. In the first case, all speech acts are in some degree expressive, and nearly any unit of language allows its expressive usage. In the second, expressives are a rather narrow class of verbs and speech acts. Prototypical units in the first case are interjections that are on the threshold of conventionality, in the second – verbs denoting politeness actions that are strictly regulated by social conventions. At the same time, a common ground also can be found – in both cases, expressives perform verbal actions, inseparable from their linguistic form, and manifesting a speaker's emotions and attitudes.

Presenting Austin's ideas in a systematic way, Searle, nevertheless passed over behavioural aspects of his conception. However, these aspects, which do not fit into the harmonious version of Searle's theory of speech acts, may turn out to be a link that connects the performative theory with the theory of social meaning, in particular, with a possible linguistic interpretation of Weber's understanding of

meaning as the basis of a socially interpreted action. One can recall that the basis of “*understanding sociology*” for Weber is the concept of meaning, oriented to the behaviour of other agents of an actual situation. There is an obvious connection between the Weberian definition of social meaning and Austin’s definition of performatives (op. cit.: 159), which reflects the core features of the speech act as a unity of word and action. Another important point in his approach is an explicit reference to the necessity for semiotic “doubling” – any action can be endowed with social meaning, only if it is represented as some symbolic form (sign, word, sentence, text) – (op. cit.: 86).

It seems to us that these general characteristics have become the basis of David Kaplan’s approach to expressives (Kaplan 1999a), in which descriptive and expressive aspects were brought together. It is natural that during last two decades this approach has stimulated a range of thought-provoking developments (see Potts 2005, 2007; Gutzmann 2013, 2015, 2019) and even became an impetus for re-assessment the very beginnings of the semantic theory of natural language:

In Kaplan’s (1999a) words, “It may be that the primary problem in semantics is not what does this or that mean, but rather in what form should we attempt to say what this or that means” (p. 3). He then uses this premise to begin building a semantic theory that is guided by expressives. That is, where Kaplan (1989) sought to assimilate expressives to regular semantics, Kaplan (1999a) seeks to assimilate regular semantics to expressives. (Potts 2005: 189)

2.3 David Kaplan on expressives and hybrid semantics

The semantic peculiarities of some semi-words, esp., interjections, were mentioned before Kaplan: still St. Augustine was in trouble: how it would be correct to translate the Aramaic derogative term *raka* (Matthew 5:22) and came to conclusion that this word was untranslatable and functioned like a proper name. Similarly to linguists and logicians who came around in XX century, St. Augustine sorted out the class of ‘raka’, ‘alas’, ‘ha’ as *merely express an affection*:

“Hence the view is more probable which I heard from a certain Hebrew whom I had asked about it; for he said that the word [*raka* – S. Z.] does not mean anything, but merely expresses the emotion of an angry mind. Grammarians call those particles of speech which express an affection of an agitated mind interjections; as when it is said by one who is grieved, ‘Alas,’ or by one who is angry, ‘Hah.’ And these words in all languages are proper names, and are not easily translated into another language; and this cause certainly compelled alike the Greek and the Latin translators to put the word itself, inasmuch as they could find no way of translating it” – *On the Sermon on the Mount*, Book I, Chapter 9, § 23;”

(Kraeger 2018: 29; see also: Predeli 2013; Capone 2016)

However, Kaplan was the first who suggested to address these strange words not as marginal stylistic or emotional exceptions, but as a crucial issue of semantics. His integral approach was opposed to the common stands: “Some may have thought wrongly that language could not perform, express, and describe at the same time” – (Kaplan 1999a: 6). Thus, as one can assume, the semantic theory of natural language should explicate how the opposite is possible. As the first step on this direction, the notion of expressive meaning had to be identified. As a proper class of expressives, Kaplan distinguished the lexical items that do not have denotative, or referential semantics (interjections, honorifics, pejoratives), so their meaning can be considered as exclusively expressive. As a prototypical case, Kaplan sorted out *ouch* and *oops* (this subclass was named ironically as a class of *ventilatives* “a poor term for expressives in a narrow sense”, p. 4). (Let’s remind that Jakobson also identified interjections as the most salient entity performing the expressive function). However, Kaplan did not restrict his consideration to just emotive and evaluative frames. The absence of referential force does not prevent the capacity to describe, as well to express correctly or incorrectly some state of affairs:

A descriptive is an expression which describes something which either is or is not the case. Let us call an expression an expressive if it expresses or displays something which either is or is not the case. I say that an expression is descriptively correct if what it describes is the case, and I say that an expression is expressively correct if what it expresses or displays is the case (or, if we take what it expresses or displays to be a state, if the agent indeed is in that state). You need to take account of the context for most of the expressives, because they usually express something about the agent – about the utterer. (Kaplan 1999a: 6)

Thus, instead of capacity to denote, expressives are endowed with the capacity to display and describe simultaneously. The description of these aspects of their semantics became a crucial element for hybrid semantics (Potts 2005, 2007; Gutzmann, 2013, 2015, 2019), since it successfully demonstrated how “the methods of truth-conditional semantics can be extended to cover use-conditional aspects of meaning” (Gutzmann 2015: 6). As a narrow definition, one can use the one suggested as a development of Kaplan’s view: “I refer to as expressives *in the narrow sense*, i.e. expressions that express some emotional and evaluative attitude with a high degree of affect” (Gutzmann 2015: 27) with the provision that we see no reason to insist on “*a high degree of affect*”, it is only one of possible cases: “an expressive can express a state that isn’t attitudinal at all” – (Kaplan 1999a: 14). One can also observe a close relation with patterns of verbal behaviour and its performative components, as well as with semantics of propositional attitudes, esp. in cases

of slurs and pejoratives.⁴ The fact is that through expressives one can lie, feign, play with “words”, use it figuratively, etc.:

One of the notable features of language as a system for conveying information is that you can lie with it. A correct semantics for expressives must provide for the ability to use expressives in this dishonest way (even if one cannot exactly lie with them) by separating the semantic information associated with the expressive, i.e. what is expressed, from the question whether what is expressed is or is not the case. (Kaplan 1999a: 5)

Having in mind this capacity, it seems that to reduce expressives even in the narrow sense to the expression of emotions oversimplifies the situation. Kaplan connected semantics of interjections not with reference, but with the appropriateness or inappropriateness of their actual usage. If an interjection can be used improperly or even figuratively (for example, saying *oops* in a situation of disaster), it implies that this interjection is endowed with something similar to literal meaning. Kaplan introduced an extremely significant complement to the existed semantic theories on expressive function of language: expressives refer to a set of contexts in which the use of this expression is appropriate. So, for exclaiming “*Ouch!*” there will be a set of contexts in which a speaker experiences a state of pain:

What is the semantic information in the word “ouch” on this analysis? The semantic information in the word “ouch” is – more accurately, is represented by – the set of those contexts at, which the word “ouch” is expressively correct (since it contains no descriptive information), namely, the set of those contexts at which the agent is in pain. That set of contexts represents the semantic information contained in the word “ouch”. (Kaplan 1999a: 15–16)

If a speaker does not experience pain, but only pretend to do so, then this exclamation will be insincere, but nevertheless informative. It, however, connects this context with those in which the speaker may experience pain or pretend to display this state of affairs. Thus, the matter it is not the state of pain itself, but the fact that the speaker exhibits this state. We will try to develop this approach by connecting it with the abovementioned Weberian understanding of meaning as directed on other’s attitudes and actions. It seems that Kaplan’s vision of expressives may be developed in this direction in an attempt to clarify the so-called descriptive component of their semantics. In some respects, this can also be considered a challenge to the widespread opinion that expressives are without both the ability to refer and denotational/referential meanings. Paradoxically, the highlighting of the descriptive component also makes more salient some of their performative characteristics.

4. These semantic and pragmatic peculiarities of slurs and pejoratives are discussed in Allan (2016), Anderson & Lepore (2013a, b), Bach (2018), Bianchi (2018), Bolinger (2015), Cella (2016), Croom (2013), Richard (2018).

3. Expressives as meaningful actions and a situational models of behavior

3.1 Expressives: Situations and attitudes

Based on Austin's observations mentioned above, the following clarification can be made: expressives themselves do not express either emotions or any other relationships. They indicate, or rather, display a situation and express the speaker's attitude to this situation. Saying "damn", a speaker shows in a certain way his assessment of the situation occurring, and it is not at all necessary that he really experienced strong emotions. This is a conventional way of expressing some attitude accepted in a given society, and its appropriateness will be evaluated not concerning the speaker's emotional state, but to the extent to which this situation is in line with such an evaluation. From this point of view, it is possible to add some essential details to the ways of interaction between expressive and descriptive meanings. As emphasized by Kaplan, any uttering of an expressive should be correlated with some model or prototypical situation: "I take it that "oops" is an expressive and that it expresses the fact that the agent has just observed a minor mishap" – (Kaplan 1999a: 12). Kaplan dwells in detail on in which situations the use of this pronoun can be considered correct, and in which deviating, but nevertheless understandable (Kaplan 1999a: 12–13). In general, these situations turn out to be dependent on the correct reference of the name mishap in relation to the situation.

Now here we have a case in which you sincerely thought it was a minor mishap. But it wasn't. It wasn't a mishap at all. It was something that was done deliberately and intentionally. You were mistaken in the thought you expressed. Although it is perfectly understandable how you made that mistake, still what you expressed – that you had just observed a minor mishap – was incorrect. (Kaplan 1999a: 14)

Thus, one can continue and imagine something like a negation of the exclamation "oops":

- (1) A: "oops!" –
 B: This is not a mishap. I have done it intentionally.

However, this sounds rather strange. It seems that the negation of 'Oops' could be something like

- (2) "Do not be worry. Nothing has happened".

As about the other case, "*That bastard Kaplan got promoted*", it can be negated not only by negating its descriptive component "It is not true, he was retired", but also:

- (3) Come on, Kaplan is not that bad. And that's no way to speak of a colleague.

This example is borrowed from (Hess 2019: 216); the author concludes:

By using an expressive a speaker is taking license, as it were, to depart from a standard of linguistic decorum; she is now (by default, in non-quotative contexts and barring any perspective shifts, about which see below) committed to this being appropriate. The full meaning of an expressive can only be conveyed as a function of this commitment and the context – as the audience must infer why the speaker takes herself to be warranted in this commitment. (Hess 2019: 213)

This observation demonstrates the relationship between heterogeneous meanings is deeper than the conjunction of the two propositions. Kaplan suggested considering this statement *That damn Kaplan was promoted*

[...] as subject to a twofold – actually expressive and descriptive – characterization: “*That damn Kaplan was promoted*” is going to be expressively correct just in case the speaker has a derogatory attitude toward Kaplan, and descriptively correct just in case Kaplan was promoted. (Kaplan 1999a: 6)⁵

It can be assumed that the descriptive and expressive components in this case are not easy to separate from each other, this rather reminds a blending than logical conjunction. Answering to the question “What did the speaker do?” two options are possible: “He reported the news of Kaplan’s promotion” and “He insulted Kaplan.” Therein is the key difference between these periphrases and what the speaker did: the speaker simultaneously did and said something what the observer can formulate through two different propositions. The main point is that if the speaker had a derogatory attitude toward professor Kaplan and, therefore, negatively evaluated the fact, that Kaplan was promoted, he/she could do it in descriptive form: “This liar (obscurant, plagiarist, racist....) was promoted”⁶; and this way would be more convenient in academy (maybe, it can be equated with Gricean (1975) maxim of *a manner*). Thus, the derogatory epithets *damn*, *fucking*, *dirty* etc. are not

5. An anonymous reviewer made an interesting correction to this statement: “Damn’ is not necessarily negative. Consider this example from the Corpus of Contemporary American English: ‘If I don’t get *that damn Emmy* nomination. There’ll be hell to pay.’ ‘damn’, when used about people, can also encode an attitude of grudging admiration”. This case demonstrates the possible ambivalence and sometimes even enantiosemic nature of evaluatives (slurs, honorifics, etc.), cf. Geurts (2007). However, despite possible change of polarity, the expressive dimension remains untouched. This sentence initiated an ongoing discussion on possible interpretations of slurs and derogative epithets; cf. Williamson (2009); Blakemore (2011, 2015); Hedger (2012, 2013); Camp (2013); Wharton (2016); see also the previous footnote.

6. Imagining a person who relies on dictionaries – how linguistic theories may explain that in this context he/she should choose the meaning “*an obnoxious or despicable person*”; but not *illegitimate child*? cf. <<https://www.collinsdictionary.com/dictionary/english-thesaurus/bastard>>

some function from the meta-meaning “bad”, but are acts of aggressive behavior intentionally violating conversational conventions. The expressivity (speaker’s attitude) may be manifested without expressives, as it was presumed in the theories of Bühler and Jakobson. Therefore, one can explicate descriptive meaning of this utterances as a conjunction of propositions: «The speaker says: “The damn Kaplan was promoted” and thus he/she is insulting Kaplan, as well as demonstrating his/her disappointment with the fact that Kaplan was promoted.⁷ This description of the situation seems to be not sufficient, as it does not capture the main illocutive and perlocutive effects of the utterance: the intention to contempt and thus perform some meaningful social action; its social meaning depends on social practices, as this was presumed in the abovementioned Weberian definition of social action. Despite Kaplan did not concentrated on these aspects of pragma-semantics of expressives, however, he made a very clear reference to this problem at the end of his lecture, maybe, considering it as a prospect for the further developments:

I wish now to return to my distinction between semantic information and social practice, and take up a few additional issues related. A feature that may be semantically encoded in one expression (i.e. may be a part of semantic information) may appear in connection with the use of a different expression only as a result of social practice. For example, one form of address may be demeaning in itself, while another form of address, not demeaning in itself, may be used to demean by the deliberate avoidance of a more respectful form of address that is called for by social practice. We are all familiar with conventionally demeaning forms of address; I suppose that the use of a diminutive to refer to a judge in court would exemplify a case of the second sort. (Kaplan 1999a: 28)

3.2 On descriptive meaning of expressives

As Kaplan has noted, any uttering of expressives (i.e. the word *ouch*) refers to some model situation, or set of contexts, in which this uttering of an expressive is expressively correct, and “that set of contexts represents the semantic information contained in the word “ouch” (Kaplan 1999a: 11). But how one can identify (a) this set of contexts; and (b) the semantic information contained in the word *ouch*? The answer which is given seems to be tautological: “The semantic information in

7. The direct expression of insult “I am insulting you” would be qualified as “an illocutive suicide” (Vendler 1976) despite being polite and non-derogatory. Still, the expressive utterances like “you, bastard” are considered as a conventional way of insulting and humiliating (cf.: Croom 2013; Jeshion 2018). This demonstrates that insults are produced by words, and not through an intention to insult. Cf.: “According to fully non-semantic accounts, slurs don’t insult people, people do; according to semantic accounts (as with guns and killing) there’s a good sense in which the slur itself insults”. (Sosa 2018: 2)

the word “ouch” is – more accurately, is represented by – the set of those contexts at which the word “ouch” is expressively correct (since it contains no descriptive information), namely, the set of those contexts at which the agent is in pain” (Kaplan 1999a: 16). Beside being tautological, this definition does not capture contexts when this expressive is used dishonestly (Kaplan 1999a: 5), when somebody feigns, that he is in pain, or having in mind the possibility of *the empathetic use of “ouch” when we see another person hurt*” (Kaplan 1999a: 12), i.e., contexts, where *the agent is not in pain*. Kaplan is applied the same approach to the semantics of *oops*: “I believe that the linguistic (i.e., syntactic and semantic) relationship between “oops” and “I have just observed a minor mishap” is exactly the same as that between “ouch” and “I am in pain”. (Kaplan 1999a: 15). This led him to the conclusion that “the linguistic differences between the interjection and the sentence seem more syntactical than semantic. The information they convey is the same, but they convey it through different modes of expression” (Kaplan 1999a: 12), so the difference between expressives and assertion also is more stylistic than semantic. Before considering this counter-intuitive conclusion, let us note that the situation with these cases seems to be quite different. (Undoubtedly, Kaplans realized this difference, but consider it as a difference between *subjective and objective* expressives, Kaplan 1999a: 14–15). It is possible to be in pain and describe this situation as “I am in pain,”; but nobody can *observe a minor mishap*; as the word is a generic for the class of non-desirable and non-expected events. One can observe dropping of glass, or falling of a chair and have a feeling that this an occasion to exclaim “oops” properly; and on this way one can identify the event as a context at which the word *oops* is expressively correct to be exclaimed. One can even play with this word to create a macabre joke, etc. having in mind that this usage deviates from these contexts and obtains some additional meaning.

This type of circularity, when event is identified through contexts, and contexts, respectively, through events, seems to be more recursive and auto-referential, than tautological. However, it also reasonably identifies the situation as “a minor mishap,” as Kaplan does. It can be used as a meta-name for the class of cases in where *oops* is expressively correct and presuppose the functions of an observer, or referee, who is in position to evaluate conditions and define appropriateness of usages (we shall refer to it later). This reverses the scheme of semantic relationships between descriptive and expressive meaning: semantic information and descriptive meaning derived from the set of contexts in which the agent exclaims these expressives. Thus, the meaning is primarily derived from usage, but not from the description of true-value conditions, and this approach seems to be more convenient for use-conventional semantic theory. This allows to explicate sources of the descriptive meaning – how it originated and how the appropriateness (felicity) of the uttering of expressive may be estimated.

3.3 Expressives and behavioural patterns

Kaplan's approach can be supplemented with essential notions from the theories of Bühler and Austin. Let us come back to Austin's primary view of expressives (behabitives) as manifestations of some conventional behavioral patterns, situations, and actions (1962: 83). All the usages of any expressives are derived from some typical frame, which represents a pattern (type) for its context-sensitive expressive utterances (tokens). This approach can be extrapolated to the description of expressives as latent behavioral patterns with socially determined meanings (*precedential, or paradigmatic frames*).⁸ This meaning consists in referring to some standard precedential frame or situational paradigm. Accordingly, expressive statements can be described as meaningful actions based on these patterns and gaining their situational significance by referring to this frame (paradigm). Thus, this action (an expressive utterance) must be classified and evaluated in accordance with this paradigm: what this action can mean, at what degree it is appropriate. For this reason, expressives should be considered in a more general perspective than just expressing speakers' emotional states and attitudes. They perform some actual or potential action, which should have some socially accepted lexical or computed contextual meaning. With this approach, it becomes possible to identify the descriptive meaning of expressives, which is closely related to conditions of emerging of this meaning in relation to a given speech act and estimation of its acceptability (or felicity).⁹ Correspondingly, utterances containing expressives can be described as meaningful actions based on these patterns which obtain their situational meaning through a reference to these frames.

From this point of view expressives can be compared with sayings – ready-made expressions which are to be uttered under some conditions when a situation is understood and represented in accordance with some precedential expression

8. This terms are borrowed from Fillmore (1976: 25): "A second notion needed for the concept of framing is that of the prototype or paradigm case... The idea is that in order to perceive something or to attain a concept, what is at least sometimes necessary is to have in memory a repertory of prototypes, the act of perception or conception being that of recognizing in what ways an object can be seen as an instance of one or another of these prototypes. This "situating" process depends not only on the existence of individual prototypes, but also on the character of the whole available repertory of prototypes".

9. Cf.: "We can begin by offering a generalized formulation of the kind of situation which may give rise to an expressive illocutionary act... The factive situation recognized by the speaker can be described as follows: A state of affairs X perceived as factual and judged to have positive or negative value for some person, the patient, brought about by a person, the agent (who may be identical with the patient), and, just in case either the agent or patient role is not filled or both are filled by the same individual, an additional person, the observer". (Norrick 1978: 283).

describing a modeling situation (i.e. *Alea jacta est* – an utterance which is appropriate with regard to some situation when a dramatic and irreversible choice is made). By using this expression, the speaker identifies the situation in which he/she is involved, with the situation in which Caesar once was. This entails some connotations (hope for a successful outcome, like in Caesar’s case, a comparison of his everyday situation with the event of the world history, performing himself as Caesar, etc.). Bühler brings this example as explaining the concept of speech action introduced by him:

Caesar was thus not very inventive, he used a “usual exclamation”; but since then no one who has learned Latin can think of the expression without reference to the River Rubicon and Caesar’s boldness”. The familiar saying has the character of a slogan no matter whether it is a single word or a sentence, a fashionable manner of speaking (an idiom) or a proverb. From here only a slight shift of accent from human destiny to the sayings themselves is required so that we arrive at our goal. Every familiar quotation and everything that is said but is not quite so quotable can be regarded under the aspect of being a human *action*. (Bühler 1990: 60–61)

However, there is a significant difference between the descriptive meaning of sayings and expressives: the contextual features of expressive utterances should coincide with the situation observed and reacted to; one cannot use them to refer to past or future events, but it is possible with regard to timeless and context-free sayings.

Expressives manifest the simplest and immediate forms of merging words and actions. In this case, there is no specific precedent case; its place is taken by descriptive content (or, in Kaplan’s other term, the semantic information), and on this way prototypical contexts of anonymous sayings can be described. Instead of the original precedential context or a situation similar to initial baptism (cf. Donnelan 1972; Kripke 1972) it may be assigned some causal chain of usages and references: from prototypical context to contexts of the occurring speech action/act. At this point, one can modify Kaplan’s definition. The coincidence of two sets of contexts and sets of situations is not the result of one-to-one correspondence between their elements, but a consequence of the fact that the context of the statement necessarily coincides with the context of the situation. They merge. A particular situation in which something happens is the context in which the expressive appropriate in this situation may be expressed correctly. And vice versa, the very fact of expressing an expressive becomes the basis for the situation to be considered as an appropriate context for the uttering of expressive. Thus, the circularity is a property determined by such characteristics as auto-reference and performativity.

Accordingly, expressive utterances of can be described as meaningful actions based on behavioral frames, and they get their situational meaning by referring to them. Expressives should be considered not only as an expression of emotional

states and attitudes but also as a potential action, which should have a certain socially defined and conventionally accepted prototypical meaning. For example, to call someone a scoundrel or a bastard – in some cases, this only expresses a speaker’s attitude. But it can be a social action, in some cases even punishable as a crime (i.e. if it is said about the country’s leader and publicly).

4. “Making sense out of events”

In order to preserve Kaplan’s significant observation on relationship between expressive and descriptive components, but at the same time to avoid of circularity we suggest to make some additional clarification. The vague notion “set of contexts” can be explicated through reference to the prototypical contexts and situations. One can suggest the following description of what is meant by a behavioural precedential frame. The behavioral frame may be understood as it was done by the prominent sociologist Erving Goffman (1974: 10): the “frameworks of understanding available in our society for making sense out of events”,¹⁰ this is congruent with the basic definitions of cognitive and interactional frames given in cognitive semantics.¹¹

As one can see, an expressive is associated not so much with some emotion as with some situation as a whole, where its descriptive content is expressed as a proposition (e. g, “oops” – “*A little mishap e takes place*”, or in explicit way, “*The speaker is declaring: “I have just observed a minor mishap*”). This makes it clear that, unlike lexical units, the linguistic form of expressives should be correlated with structures at the sentential or textual level.

If what is displayed is to either be or not be the case, it must have sentential form. Thus instead of saying that my shriek displays fear, I shall say that it displays the fact (if sincere) that I am in fear... Note also the importance of the availability of the first person in transforming what is displayed into sentential form.

(Kaplan 1999a: 6)

10. Cf.: “I assume that definitions of a situation are built up in accordance with principles of organization which govern events – at least social ones – and our subjective involvement in them; frame is the word I use to refer to such of these basic elements as I am able to identify” (Goffman 1974: 10–11).

11. “A language has both interactional frames and cognitive or conceptual frames... The interactional frames amount to a categorization of the distinguishable contexts of interaction in which speakers of a language can expect to find themselves, together with information about the appropriate linguistic choices relevant to these interactions. One simple example is the greeting frame [...] A part of knowing a language is knowing or recognizing a large number of such frames, and knowing what linguistic choices are relevant for each of them”. (Fillmore 1976: 26)

The expressive utterance necessarily endowed with first-person characteristics, thus it has to be transformed into first-person sentences. Even when it is not a separate exclamation, but, for instance, an epithet or apposition embedded within a sentence, it can be related to some propositional transforms. The precedential frame appears derived from those sentential explications, and it becomes the lexical content of expressive meaning as it is described in dictionaries; as a rule, they describe *expressively correct* contexts:

For example, most dictionaries do not attempt to paraphrase the meaning of *oops*, but rather “define” it by describing the contexts in which it is normally used:

- a. “used typically to express mild apology, surprise, or dismay” – <http://www.merriam-webster.com>;
- b. “an exclamation of surprise or of apology as when someone drops something or makes a mistake” – Collins English Dictionary. (Kraeger 2018: 28)

If following Kaplan’s observation and characterize set of contexts of uttering *oops* as a set of situations of a minor mishap, the frame semantics shows the accuracy of Kaplan’s intuition, mainly when he discussed the possibility of non-literal or non-correct usages. The frame ‘mishap’ is described as inherited from the frame “*catastrophe*”; in the Framenet, it is defined as follows: “The words in this frame involve an *Undesirable_event* which affects the *Patient* negatively. No agent need to be involved” <<https://framenet.icsi.berkeley.edu/>> (25 May 2020). Since the frame does not include *an agent* as an obligatory semantic function, it does not imply intentional action.

As one can see, semantic changes occur within the same mega-frame *catastrophe*. This occurs in all three cases of non-standard use of *oops* that Kaplan cited: the non-literal macabre and polite meanings arise as a neutralization of the feature “*minor*”,¹² and the third case of non-correct relates to missing intentional action with an occasional event. At the same time, as it is presumed by its frame-semantics of *catastrophe* and *mishap*, a patient is to be involved and affected (partially it is taken into account in Kaplan’s example of the polite taking of responsibility and blame-taking – in this case, the speaker/ patient is affected although he/she does not demonstrate it). Maybe, in the case of a minor mishap, this negative affect is not a serious matter and can be ignored.

12. Cf.: “Another nice case that my colleague Seana Schiffrin called to my attention, in which one manipulates the literal meaning of the word *oops* to achieve a certain emotional purpose concerns my saying *oops* when my guest has caused in fact a not so minor mishap, right, I mean in fact, it was the valuable glass heirloom that was knocked off the table and smashed into a thousand pieces. And I say, *oops*, you know, making as if it were a minor mishap. In this case, the reasons are dictated by social practices of blame-taking and blame-placing and so on”. – (Kaplan 1999b: 12).

Even the most approximate description of the frame ‘*Mishap*’ explains the conditions under which the utterance ‘oops’ is expressively correct: that is a case when a patient is affected by undesirable unintentional event but not seriously. Frame accommodation happens through processes of merging an ongoing mishap event with the uttering “oops” event, and the patient of the first frame becomes an agent of the second event. Expressive, if it is expressively correct, is used as a performative function:¹³ it set up a correspondence between a given situation and a situation that may be represented through the precedential frame, as well as mappings the agent of the context-event with the patient of the situation-event. The so-called descriptive meaning of expressives can be considered a result of accommodation of this precedential frame to some actual situation. The semantics of saying and expressives can be considered as a function from some model situation to a situation where these expressions are used. In the case of the “oops,” the general correspondence rule may be represented as a characterizing (or classifying) function:

Uttering of “oops”

The actual situation S (Somebody breaks his glass). ↔ Model situation
(a minor mishap occurs).

This class of situation where the exclamation *oops* is expressively correct; descriptively it coincides with the class of events which can be identified as *a minor mishap*. However, such a description is not complete. There is an essential difference between the units “*a minor mishap*” and “*oops*”. Lexical units are described through the frames of the situations that they denote. Meanwhile, use-conditional entities are defined recursively – by reference to the conditions in which they are used. Therefore, the speaker must be at the same time a participant in the described event.¹⁴

One needs to refer more explicitly to the speaker-centered and context-dependent nature of these semantic and contextual variables: who is uttering “oops”, in which context, and under which circumstances and which situation (world).

The model of context suggested by Potts (2007) and Gutzmann (2015, 2019) may be used in its more simple version:

13. Cf.: “I do believe that the performative function is over and above the semantic expressive function”. (Kaplan 1999a: 12)

14. Compare with roles identified for participants of expressive speech acts: “Another kind of constraint placed upon the successful performance of expressive illocutionary acts requires that the speaker identify the roles of agent (the person responsible for the state), patient, and, in some cases, observer (a person cognizant of the state besides the patient) with himself qua speaker, or with the addressee to whom the illocutionary act will be addressed”. (Norricks 1978: 283)

{“oops”} in (s; t; c; l) == cA (Ac is a speaker/agent of c, Ps – is a patient (or an affected observer) in s; s – situation, t – time, c – context; l – place);

Tc = T_{s+1}, (i.e.: time moment of context of uttering *oops* is immediately follow to moment of mishap);

Lc – Ls; cA = sP (i. e. place and context of uttering *oops* coincides with the observed situation of mishap, excepting reported events).¹⁵

The context-sensitive expressive meaning of the uttering “oops” (pointing out who is the “I” with respect to the given context and establishing some relation with some “S”) is combined with the context-sensitive descriptive meaning of “oops” It is possible to envisage two possible options for the speaker: either the semantic role of patient if the situation S negatively affecting on “I, or the role referee-observer; the second seems to be derivative from the first, as the speaker expresses as if he/she were affected through mishap occurring with others.

This contextualisation may be represent in following way:

Uttering of “oops” in (s; t; c; l)

The actual situation S (Somebody breaks his glass). ↔ Model situation (a minor mishap occurs).

The utterance “oops” is descriptively and expressively correct if: (1) Ac = Ps;

(2). Tc = T_{s+1}; (3) Lc – Ls; (4) cA = sP; (5) S_{t,1} cause *Uttering of “oops” in C (a; t+1; l)*

Such characteristics as the negative affectedness can be considered as built-up component of the model situation and inherited from the frame “mishap” why it should not be duplicated in the semantic description of this expressive. The notion of a precedential frame can be used to reformulate the distinction between objective and subjective expressives (Kaplan 1999a: 12). Their semantics is based on the different procedures of compliance with socially accepted practices of their usage: in the case of subjective expressives (like *ouch*) nobody other than speaker is able to decide on the conformity of utterance with the precedential situation; for objective expressives (like *oops*) all the observers are in an equal position to decide whether some situation can be classified as a minor mishap or not.

15. We exclude an essential participant of the context – this is the judge who determines the appropriateness (or expressive and descriptive correctness) of an utterance (Cf. : Lasersohn 2005, 2007; Gutzmann 2015; 2019). If expressives is as a manifestation of the expressive function of language, the first person, have to perform this semantic role: “We may claim that the judge of a context is automatically fixed to the agent, or to some relevant group or individual intended by the agent, or perhaps something else. In considering this issue, some caution is necessary”. – (Lasersohn 2005: 668). However, if we take into account the mentioned above definition of social meaning given by Weber, that is, focus on addressee, then in the holistic model of context, it will be necessary to include two more participants – the recipient-addressee and the judge-addressee, they may be merged.

This observation may be detailed: expressives are subjective if the speaking agent coincides with the experiencer from the frame-description of the event, and objective in the frame he is either an agent, an addressee, or a patient.

Another differentiation is based on the degree of semantic vagueness/definiteness and contextual dependence. Thus, both the descriptive and expressive/emotive semantics of *oops* and *ouch* are rather definite and are preserved in any context. Even non-standard usages of them (in reported speech, ironical, pretending, figural or even sardonic – *oops* in a case of disaster,) have as a point of departure some well-defined precedential meaning. But many interjections have very poor descriptive component and very flexible contextual semantics. A good example is “*HA*”: this interjection can express any emotion, so its meaning is auto-referential. In uttering this expression, the speaker expresses an emotion which he/she and/or his/her addressee considers as coinciding with the situation of its utterance.¹⁶ As a next step, some specification of meaning can be observed: “*Ha – HA*” presupposes that a situation seems to be funny, that this situation is occurring within the given communicative contexts and that communicants pretend to represent their laughing at it.

The suggested approach to extraction of descriptive meaning from expressives allows us to find essential intersections with the theory of expressive speech acts. The general characteristic suggested by Searle and Vanderveken is applicable to expressives in the Kaplanian sense, and the typology of expressive speech acts may be used as a typological matrix for expressives.¹⁷

16. Cf: ““This is a kind of radical Humpty Dumpty language, in which the word means whatever the user intends it to mean. The reason why it won’t work for normal human beings is because there is very often no independent way of finding out what a person means than through what he says... In general, natural languages will allow as contexts only propositions whose truth could reasonably be expected to be known by those involved in the communication” (Cresswell 1972: 10).

17. Cf: “ Those verbs in English that name expressive illocutionary forces that almost without exception they indicate that there is something good or bad about this state of affairs represented by the propositional content of the expressive... In what follows many of psychological states already carry the belief that object of the state is good or bad, e. g. pleasure and sorrow. Furthermore most of the expressive speech act that have acquired special verbal naming they are essentially hearer-directed. (Searle & Vanderveken 1985 : 211)

5. Conclusion

Through analyses of expressives, we intend to demonstrate that meaning results from a conjunction of linguistic and extra-linguistic systems in the process of social interaction and communication. The highest degree of contextual dependence of expressives provides an opportunity to detect the processes of their conceptualization and contextualization immediately under circumstances of uttering them. Expressives are closely connected with behavioral patterns, situations, and actions; in this sense, they are a proper manifestation of the notion of language as a rule-governed game (Wittgenstein 1958). An actualization of verbal behavior (its performative component) meets its counterpart in the verbalization of action (its descriptive component). Correspondingly, utterances containing expressive can be described as meaningful actions based on these patterns and obtain their situational meaning through a reference to these frames. This approach may reconcile the three abovementioned heterogeneous approaches to expressives. The pragma-semantics of expressives cannot be restricted exclusively to referential or intensional aspects of utterances. Sometimes these are hidden, and they are always interwoven with pragmatic (contextual and extra-textual) variables. The semantics of expressives can be considered a peculiar language game based on social conventions and contextualization. It can be described as a set of rules based on the ability to operate with social and linguistic entities and contexts, in some respect, similar to conventional implicatures (Potts 2007; Lasersohn 2007) and pragmatic presuppositions. The ability to compute appropriate contextual meanings of lexical items is based on shared knowledge of social norms and conventions. The resulting blending of linguistic and socially determined and commonly accepted context-sensitive meanings can be described as a combination of behavioral (Hoffman 1974) and conceptual (Fillmore 1976, FrameNet) frames. In general, such explicit or implicit indexical expressions can be portrayed as a constant function from linguistically and socially determined variables. It is a doubly sensitive function: besides being context-sensitive, it is also circumstances-sensitive. Perhaps this regularly manifested but not unified circumstances-sensitive and a speaker-dependent semantic operational component of utterance can be partially identified with social meanings – as Max Weber defined them. From this point of view, expressives can be considered not as marginal phenomena, but as a core instance of pragma-semantics in operation. Expressives manifest the simplest forms of the relationship between a word and an action. Accordingly, expressions containing expressives can be considered as meaningful actions based on behavioral frames, and they get their situational meaning by referring to these. Expressives are not only as an expression of emotional states and attitudes but also as potential or actual actions, which should have a certain socially defined and conventionally accepted prototypical meaning.

Acknowledgements

I am grateful to two anonymous referees for insightful suggestions, and participants of the workshop International Conference on Current Trends in Linguistics 28–29 March 2019 (University of Rouen) for helpful discussions and comments on my presentation. I owe special thanks to Sylvie Hancil for her efforts to organise this workshop and publish presented papers.

Funding

The research is supported by the Russian Science Foundation, project № 18-18-00442, “Mechanisms of meaning formation and textualization in narrative and performative discourses and social practices.”, Immanuel Kant Baltic Federal University, Kaliningrad, Russia.

References

- Akita, Kimi & Pardeshi, Prashant (eds). 2019. *Ideophones, Mimetics and Expressives* [Iconicity in Language and Literature 16]. Amsterdam: John Benjamins. <https://doi.org/10.1075/ill.16>
- Allan, Keith. 2016. The reporting of slurs. In *Indirect Reports and Pragmatics: Interdisciplinary Studies*, Alessandro Capone, Ferenc Kiefer & Franco Lo Piparo (eds), 211–232. Cham: Springer. https://doi.org/10.1007/978-3-319-21395-8_11
- Anderson, Luvell & Lepore, Ernie. 2013a. Slurring words. *Noûs* 47(1): 25–48. <https://doi.org/10.1111/j.1468-0068.2010.00820.x>
- Anderson, Luvell & Lepore, Ernie. 2013b. What did you call me? Slurs as prohibited words. *Analytic Philosophy* 54(3): 350–363. <https://doi.org/10.1111/phib.12023>
- Austin, John L. 1962. *How to do Things with Words*, James Opie Urmson & Marina Sbisá (eds). Oxford: Clarendon Press.
- Bach, Kent. 2018. Loaded words: On the semantics and pragmatics of slurs. In *Bad Words. Philosophical Perspectives on Slurs*, David Sosa (ed.), 60–76. Oxford: OUP.
- Baudouin de Courtenay, Ivan. 1904. Jazyk i jazyki (Language and languages). In *Enciklopedičeskij slovar'* (Encyclopedic dictionary), Vol. 81, 528–549. Saint-Petersburg: Brockhaus & Efron. (In Russian).
- Bianchi, Claudia. 2018. Perspectives and slurs. In *Beyond Semantics and Pragmatics*, Gerhard Preyer (ed.), 187–200. Oxford: OUP.
- Blakemore, Diane. 2011. On the descriptive ineffability of expressive meaning. *Journal of Pragmatics* 43(14): 3537–3550. <https://doi.org/10.1016/j.pragma.2011.08.003>
- Blakemore, Diane. 2015. Slurs and expletives: A case against a general account of expressive meaning. *Language Sciences* 52: 22–35. <https://doi.org/10.1016/j.langsci.2014.06.018>
- Bolinger, Renée Jorgensen. 2015. The pragmatics of slurs. *Noûs* 51(3): 439–462. <https://doi.org/10.1111/nous.12090>
- Bühler, Karl. 1990. *The Theory of Language: The Representational Function of Language*, translated by Donald Fraser Goodwin. Amsterdam: John Benjamins. <https://doi.org/10.1075/fos.25>
- Camp, Elizabeth. 2013. Slurring perspectives. *Analytic Philosophy* 54(3): 330–349. <https://doi.org/10.1111/phib.12022>

- Capone, Alessandro. 2016. Indirectly reporting and translating slurring utterances. In *Indirect Reports and Pragmatics: Interdisciplinary Studies*, Alessandro Capone, Ferenc Kiefer & Franco Lo Piparo (eds), 233–252. Cham: Springer. https://doi.org/10.1007/978-3-319-21395-8_12
- Capone, Alessandro. 2019. *Pragmatics and Philosophy. Connections and Ramifications*. Cham: Springer. <https://doi.org/10.1007/978-3-030-19146-7>
- Carston, Robyn. 2008. Linguistic communication and the semantics/pragmatics distinction. *Synthese* 165(3): 321–345. <https://doi.org/10.1007/s11229-007-9191-8>
- Cella, Federico. 2016. Slurs: Semantic content, expressive content and social generics. *Phenomenology and Mind* 11: 140–148. https://doi.org/10.13128/Phe_Mi-20114
- Cresswell, Maxwell John. 1972. The world is everything that is the case. *Australasian Journal of Philosophy* 50(1): 1–13. <https://doi.org/10.1080/00048407212341001>
- Cresswell, Maxwell John. 1973. *Logics and Languages*. London: Methuen.
- Croom, Adam. 2013. How to do things with slurs: Studies in the way of derogatory words. *Language & Communication* 33(3): 177–204. <https://doi.org/10.1016/j.langcom.2013.03.008>
- Daalder, Saskia & Musolf, Andrea. 2011. Foundations of pragmatics in functional linguistics. In *Foundations of Pragmatics*, Wolfram Bublitz & Neal Norrick (eds), 229–260. Berlin: De Gruyter. <https://doi.org/10.1515/9783110214260.229>
- Depraetere, Ilse & Salkie, Raphael (eds). 2017. *Semantics and Pragmatics: Drawing a Line*. Cham: Springer. <https://doi.org/10.1007/978-3-319-32247-6>
- Donnellan, Keith S. 1972. Proper names and identifying descriptions. In *Semantics of Natural Language*, Donald Davidson & Gilbert Harman (eds), 356–379. Dordrecht: Reidel. https://doi.org/10.1007/978-94-010-2557-7_10
- Fillmore, Charles J. 1976. Frame semantics and the nature of language. *Annals New York Ac. Sc. Origins and Evolution of Language and Speech* 280(1): 20–32. <https://doi.org/10.1111/j.1749-6632.1976.tb25467.x>
- Foolen, Ad. 2015. Expressives. In *The Routledge Handbook of Semantics*, Nick Riemer (ed.), 473–490. London: Routledge.
- Gauker, Christopher. 2013. Semantics and pragmatics. In *The Routledge Companion to Philosophy of Language*, Gillian Russell & Delia Graff Fara (eds), 18–28. London: Routledge. https://doi.org/10.4324/9780203206966.ch1_2
- Geurts, Bart. 2007. Really fucking brilliant. (Commentary on Christopher Potts, The expressive dimension). *Theoretical Linguistics* 33(2): 209–214. <https://doi.org/10.1515/TL.2007.013>
- Grice, H. Paul. 1975. Logic and conversation. In *Syntax and Semantics, 3: Speech Acts*, Peter Cole & Jerry L. Morgan (eds), 41–58. New York NY: Academic Press.
- Goffman, Erving. 1974. *Frame Analysis: An Essay on the Organization of Experience*. Cambridge MA: Harvard University Press.
- Gutzmann, Daniel. 2013. Expressives and beyond. An introduction to varieties of use conditional meaning. In *Beyond Expressives. Explorations in Use-Conditional Meaning*, Daniel Gutzmann & Hans-Martin Gärtner (eds), 1–58. Leiden: Brill. https://doi.org/10.1163/9789004183988_002
- Gutzmann, Daniel. 2015. *Use-conditional Meaning: Studies in Multidimensional Semantics*. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780198723820.001.0001>
- Gutzmann, Daniel & Henderson, Robert. 2019. Expressive updates, much? *Language* 95(1): 108–155. <https://doi.org/10.1353/lan.2019.0014>
- Hedger, Joseph A. 2012. The semantics of racial slurs: Using Kaplan's framework to provide a theory of the meaning of derogatory epithets. *Linguistic and Philosophical Investigations* 11: 74–84.

- Hedger, Joseph A. 2013. Meaning and racial slurs: Derogatory epithets and the semantics/pragmatics interface. *Language & Communication* 33: 205–213.
<https://doi.org/10.1016/j.langcom.2013.04.004>
- Hess, Leopold. 2019. Expressive meanings and expressive commitments. A case of meaning as use. In *Philosophical Insights into Pragmatics*, Piotr Stalmaszczyk (ed.), 191–221. Berlin: De Gruyter. <https://doi.org/10.1515/9783110628937>
- Hjelmslev, Louis. 1961 [1943]. *Prolegomena to a Theory of Language*. Madison WI: University of Wisconsin Press.
- Jakobson, Roman. 1960. Closing statement: Linguistics and poetics. In *Style in Language*, Thomas A. Sebeok (ed.), 350–377. Cambridge MA: The MIT Press.
- Jaszczolt, Kasia M. 2005. *Default Semantics: Foundations of a Compositional Theory of Acts of Communication*. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780199261987.001.0001>
- Jeshion, Robin. 2018. Slurs, dehumanization, and the expression of contempt. In *Bad Words Philosophical Perspectives on Slur*, David Sosa (ed.), 77–108. Oxford: OUP.
- Kaplan, David. 1989. Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In *Themes from Kaplan*, Joseph Almog, John Perry & Howard Wettstein (eds), 481–563. Oxford: OUP.
- Kaplan, David. 1999a. *The meaning of ouch and oops. Explorations in the Theory of Meaning as Use*. Ms, University of California, Los Angeles.
- Kaplan, David. 1999b. The meaning of ouch and oops. Horison Lecture in Philosophy delivered at UC Berkeley, transcribed by Elizabeth Coppock. <<http://eecoppock.info/Pragmatics-SoSe2012/kaplan.pdf>> (25 May 2020).
- Kripke, Saul. 1972. Naming and necessity. In *Semantics of Natural Language*, Donald Davidson & Gilbert Harman (eds), 253–355. Dordrecht: Reidel.
https://doi.org/10.1007/978-94-010-2557-7_9
- Kecskes, Istvan. 2016. Can intercultural pragmatics bring some new insight into pragmatic theories? In *Interdisciplinary Studies in Pragmatics, Culture and Society*, Alessandro Capone & Jacob Mey (eds). Cham: Springer. https://doi.org/10.1007/978-3-319-12616-6_3
- Kroeger, Paul R. 2018. *Analyzing Meaning: An Introduction to Semantics and Pragmatics*. Berlin: Language Science Press.
- Langacker, Ron. 2012. Interactive cognition: Toward a unified account of structures, processing, and discourse. *International Journal of Cognitive Linguistics* 3: 95–124.
- Lasersohn, Peter. 2005. Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28: 643–86. <https://doi.org/10.1007/s10988-005-0596-x>
- Lasersohn, Peter. 2007. Expressives, perspective and presupposition. *Theoretical Linguistics* 33(2): 223–230. <https://doi.org/10.1515/TL.2007.015>
- Longworth, Guy. 2017. Semantics and pragmatics. In *A Companion to the Philosophy of Language*, Vol. I, Bob Hale, Crispin Wright & Alexander Miller (eds), 107–126. Oxford: Wiley-Blackwell. <https://doi.org/10.1002/9781118972090.ch5>
- Malinowski, Bronislaw. 1923. The problem of meaning in primitive languages. In *The Meaning of Meaning*, Supplement I, Charles Kay Ogden & Ivor A. Richards (eds), 296–336. London: Kegan Paul.
- Malinowski, Bronislaw. 1935. *Coral Gardens and their Magic, Vol. 2: The Language of Magic and Gardening*. London: Allen & Unwin.
- Musolff, Andreas. 1997. Bühler, Karl. In *Handbook of Pragmatics*, Jef Verschueren, Jan-Ola Östman, Jan Blommaert & Chris Bulcaen (eds), 1–15. Amsterdam: John Benjamins.
<https://doi.org/10.1075/hop.3.buh1>

- Norricks, Neal R. 1978. Expressive illocutionary acts. *Journal of Pragmatics* 2: 277–291.
[https://doi.org/10.1016/0378-2166\(78\)90005-X](https://doi.org/10.1016/0378-2166(78)90005-X)
- Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford: OUP.
- Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33(2): 165–197.
<https://doi.org/10.1515/TL.2007.011>
- Predelli, Stefano. 2013. *Meaning without Truth*. Oxford: OUP.
<https://doi.org/10.1093/acprof:oso/9780199695638.003.0007>
- Preyer, Gerhard (ed.). 2018. *Beyond Semantics and Pragmatics*. Oxford: OUP.
<https://doi.org/10.1093/oso/9780198791492.001.0001>
- Recanati, François. 2006. Pragmatics and semantics. In *The Handbook of Pragmatics*, Laurence R. Horn & Gregory Ward (eds), 442–462. Oxford: Blackwell.
- Richard, Mark. 2018. How do slurs mean? In *Bad Words. Philosophical Perspectives on Slurs*, David Sosa (ed.), 155–167. Oxford: OUP.
- Saeed, John. 2011. Pragmatics and semantics. In *Handbooks of Pragmatics*, Wolfram Bublitz, Andreas H. Jucker & Klaus P. Schneider (eds), 461–490. Berlin: Mouton de Gruyter.
- de Saussure, Louis. 2007. Pragmatic issues in discourse analysis. *Critical Approaches to Discourse Analysis across Disciplines* 1(1): 179–195.
- Searle, John R. 1976. A classification of illocutionary acts. *Language in Society* 5(1): 1–23.
<https://doi.org/10.1017/S0047404500006837>
- Searle, John & Vanderveken, Daniel. 1985. *Foundations of Illocutionary Logic*. Cambridge: CUP.
- Shannon, Claude E. A. 1948. Mathematical theory of communication. *Bell System Technical Journal* 27(3): 379–423, 623–656. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Sosa, David. 2018. Introduction. In *Bad Words. Philosophical Perspectives on Slurs*, David Sosa (ed.), 1–5. Oxford: OUP.
- Vendler, Zeno. 1976. Illocutionary suicide. In *Issues in the Philosophy of Language: Proceedings of the 1972 Oberlin Philosophy Colloquium*, Alfred F. MacKay & Daniel D. Merrill (eds), 135–145. New Haven CT: Yale University Press.
- Weber, Max. 1964. *The Theory of Social and Economic Organization*. New York NY: The Free Press.
- Wharton, Tim. 2016. That bloody so-and-so has retired: Expressives revisited. *Lingua* 175–176: 20–35. <https://doi.org/10.1016/j.lingua.2015.08.004>
- Williamson, Timothy. 2009. Reference, inference, and the semantics of pejoratives. In *The Philosophy of David Kaplan*, Joseph Almog & Paolo Leonardi (eds), 137–158. Oxford: OUP.
<https://doi.org/10.1093/acprof:oso/9780195367881.003.0009>
- Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. London: Basil Blackwell.
- Zolyan, Suren. 1991. Semantika i struktura poeticheskoho teksta. (Semantics and structure of a poetic text). Yerevan: Yerevan University Press. (2nd edn, 2014. Moscow: URSS). (In Russian).

PART III

Discourse-related grammatical phenomena

A just amazing marker in French: “Juste”

“Juste excellentissime!” “Juste super heureux!”
“Juste irréal!” “Juste pas possible!”

Danh-Thành Do-Hurinville
Université de Franche-Comté

Juste belongs to those words (*justement, limite, grave, côté, question, niveau...*) that today one cannot do without, both in everyday conversation and in the media! After pointing out that *juste*, from the Latin *justus*, is a transcategorial unit par excellence, as it can be an adjective, a noun, or an adverb, this article deals with its recent adverbial use, identified orally in the early 2000s, officially registered in *Le Petit Robert de la Langue Française* in 2014. I propose that *juste* is a metalinguistic marker which creates double modalisation: it has scope both over an element X which can be an extreme (non-gradable) adjective, a non-gradable noun phrase (verb phrase or prepositional phrase), and over the utterance act. This recent adverbial use of *juste* is analysed in the following configurations: (1) *juste* and extreme adjectives (extreme adjectives formed by the suffix *-issime*, and extreme adjectives formed by the private prefix *-in*); (2) *juste* and non-gradable sequences [*pas* + gradable adjectives]; (3) *juste* and two non-gradable sequences ([*très* / *trop* + gradable adjectives]; [*super* / *hyper* / *ultra* / *supra* / *mega* / *giga* + gradable adjectives]); (4) *juste* and non-gradable [noun / verb / prepositional] phrases.

Keywords: *juste*, extreme adjectives, metalinguistic use, transcategorial marker, magnifying glass effect

1. A general outline of *juste* in French: *Juste* is not really *just*

The most recent adverbial use of *juste* in French as in “*Juste génial!*”, “*Juste excellentissime!*”, “*Juste irréal!*”, “*Juste pas possible!*”, “*Juste super heureux!*”, “*Juste une merveille!*”, “*Juste hors norme!*”, is more and more frequent, both in everyday conversation and in the media. However, considered to be non-standard, this adverbial use is vilified by several blogs and press articles written by supporters of the French

language: some of them went as far as to consider it as a “syntactic pustule”, useless, likely to deteriorate the French language (cf. Garcin 2009).

According to Siouffi *et al.* (2012), this use seems to have emerged in the early 2000s. In principle, it seems difficult to accurately date the appearance of a new use, because in most cases it first imperceptibly slipped into oral language (“fait de parole”), and if it has been popularised by the majority of speakers, it will gradually settle into writing (“fait de langue”), without speakers being aware of its emergence, before it is officially recognised by dictionaries.

Indeed, *Le Petit Robert de la Langue Française* (henceforth PRLF) recorded this new adverbial use of *juste* in its 2014 edition. Online dictionaries as the *Wikidictionnaire* (*C'est juste magnifique! C'est juste pas possible!*) and the *Reverso Dictionary* (*C'est juste incroyable ton histoire*) also recognised this use, but they specify that such use is informal, which *Le Petit Larousse* and the *Hachette* dictionaries have not yet recognised.

In short, this use, corroborates the Saussure’s viewpoint (2005[1916]) on the relationship between “langue” and “parole”: The “fait de parole” always precedes the “fait de langue”.

In the history of all innovation there are always two distinct moments: (1) one in which it arises with individuals; (2) the other one which has become a “fait de langue”, externally identical, but adopted by the community.

(Saussure 2005[1916]: 139)¹

This adverbial use of *juste* with extreme (non-gradable) adjectives such as *génial*, *magnifique*, *merveilleux*, *sublimissime*, was not examined in Bat-Zeev Shyldkrot (2001), Blanche-Benveniste (2001), Leeman (2004), Mellet & Monte (2009), but was studied in Salvan (2014), Do-Hurinville (2018), Do-Hurinville & Dao (2018), Dao, Do-Hurinville & Lafontaine (2020). This use, considered as a calque of *just* in English (*That’s just marvelous!*), is justified by a comparable etymological origin, because *just* and *juste* come from the same Latin source, *justus*. It should be noted that the use of *just* with extreme adjectives in English as in *That’s just marvelous!* was mentioned in Cohen (1969) and Lee (1987, 1991), according to whom *just* means “simply/emphasis”, or is considered as a “focus particle” (cf. König 1991: 122). It took almost half a century before the same use of *juste* in French was studied for the first time in 2014 (cf. Salvan). This means that these two units have not evolved in the same way. It seems that the frequency of use of *just* with extreme adjectives in English is much higher than that of *juste* in French. To get

1. The translation is mine, the original text being: « Dans l’histoire de toute innovation on rencontre toujours deux moments distincts : (1) celui où elle surgit chez les individus ; (2) celui où elle est devenue un fait de langue, identique extérieurement, mais adopté par la collectivité »

an idea, one can look at the English Examples (1), (3), (5), (8), (13), (17), (21), (26) and (27) with *just*, and their translation into French (2), (4), (6), (7), (9) to (12), (14) to (16), (18) to (20), (22) to (24) proposed by *Linguee*, *Online English-French Dictionary*, *Reverso Context*.

- (1) It's **just amazing** that one small request has made such a difference in people's lives (Linguee)
- (2) C'est **incroyable** de penser ce qu'une simple demande peut faire pour améliorer la vie des gens (Linguee)
- (3) It's **just amazing** to see the people who line up to watch these films, because they know or they feel that if they don't watch it now, they'll never get to watch it again [sic] (Linguee)
- (4) Les files d'attente sont **ahurissantes** parce que les gens savent ou pensent que s'ils ne voient pas ces films tout de suite, ils ne les verront jamais (Linguee)

While this use of *just* as a focus particle has existed for a long time in English, this use of *juste* in French is not yet stable in writing, and the translation from *just* in English into *juste* in French is not systematic. Depending on the context, *just amazing*, in (1) and (3) in English, is translated into French, either by *incroyable* in (2) or *ahurissantes* in (4).

- (5) I **just** love this film! (Dictionnaire anglais-français)
- (6) J'adore ce film, **tout simplement** ! (Dictionnaire anglais-français)
- (7) J'adore ce film, **un point c'est tout** ! (Dictionnaire anglais-français)

In the context of (5) in English, the translator proposes two possible French versions: *just* is rendered by the collocational unit *tout simplement* in (6), or by the locution *un point c'est tout* in (7). This collocational unit and this locution, in final position in (6) and (7), and not preposed to the verb phrase as *just* in English, and function as Discourse Markers (DMs). Separated from the rest of the utterance by a comma, they have scope over the entire utterance and are not part of the propositional content of that utterance. They indicate the speaker's stance towards the hearer.

- (8) It's **just amazing** (Reverso Context)
- (9) C'est **tout simplement incroyable** (6) (Reverso Context)
- (10) C'est **juste incroyable** (4) (Reverso Context)
- (11) C'est **Ø incroyable** (4) (Reverso Context)
- (12) C'est **juste stupéfiant** (2) (Reverso Context)

- | | |
|--|-------------------|
| (13) It's just incredible | (Reverso Context) |
| (14) C'est Ø incroyable (8) | (Reverso Context) |
| (15) C'est tout simplement incroyable (4) | (Reverso Context) |
| (16) C'est vraiment incroyable (2) | (Reverso Context) |
| (17) It's just unbelievable | (Reverso Context) |
| (18) C'est tout simplement incroyable (3) | (Reverso Context) |
| (19) C'est Ø pas croyable (2) | (Reverso Context) |
| (20) C'est tout bonnement incroyable (2) | (Reverso Context) |

From the sentences in English with *just* in (8), (13) and (17), *Reverso Context* proposes three possible French translations: (i) the use of adverbs other than *just* such as *vraiment*, *tout simplement*, *tout bonnement* (15 occurrences); (ii) the use of an adjective without an adverb (14 occurrences); (iii) the use of *juste* (6 occurrences). It is clear that the use of *juste* is rather marginal (6 times), compared to the other two possibilities without *juste* (29 times).

- | | |
|--|--|
| (21) Just do it! | (<i>Dictionnaire anglais-français</i>) |
| (22) Vas-y ! | (<i>Dictionnaire anglais-français</i>) |
| (23) Lance-toi ! | (<i>Dictionnaire anglais-français</i>) |
| (24) Fais-le ! | (<i>Linguee et Reverso Context</i>) |
| (25) Juste fais-le | (title of one of Soprano's songs) |
| (26) Just Eat | (advertisement) |
| (27) Just Boris: A Tale of Blond Ambition | (Sonia Purnelle, 2012) |

Let's now look at three other contexts with *just* in English in (21), (26) and (27), which are not translated by *juste* in French. Example (21) in English with *just*, which is an injunctive expression, is rendered in French without *juste* in (22) to (24). Example (25), the title of a Soprano's song, is certainly a calque of (21), but (25) has not yet come into common use. In the context of (26), which is a British online food order and delivery service, and of (27), which is the title of a book by Purnelle, it is absolutely not possible to use *juste* in the French translation.

2. *Juste*, a transcategorial marker

According to the PRLF, *juste* originated in the middle of the twelfth century from the Latin *justus* formed by *jus* “right” and the nominative suffix *-tus*, which can have three meanings: “law, equity”, “legal, conform with the law”, “normal, fitting, regular”. Functioning respectively as adjective, noun, adverb, *juste* is a transcategorial² unit par excellence.

2.1 *Juste* as an adjective

Juste, first identified as an adjective, expresses three meanings in the following chronological order:

Table 1. The main adjectival meanings of *juste*

Three main adjectival meanings of <i>juste</i>		
Meaning 1	Meaning 2	Meaning 3
“Idea of justice” (mid-twelfth century) (Example 28)	“Idea of righteousness” (late thirteenth century) (Example 29)	“Idea of narrowness, insufficiency” (17th century) (Example 30, 31)

- (28) *Cet avocat défend une cause juste.*
(This lawyer is defending a just cause)
- (29) *Trouver le mot juste.*
(Find the right word)
- (30) *Cette jupe est trop juste : tu devrais prendre la taille au-dessus.*
(That skirt is too tight: you should go for the next size up)
- (31) *Tu pourrais me dépanner de 50 euros, je suis un peu juste en ce moment.*
(Can you lend me €50, I’m a bit short at the moment)

From the chronological viewpoint, one notes that the “idea of justice” (first meaning, Example 28) logically precedes the “idea of righteousness” (second meaning, Example 29). As for the “idea of narrowness or insufficiency” (third meaning, Example 30 and 31), it seems to have arisen from the fusion of the second meaning with the co-textual elements, *trop* and *un peu*. I suppose that the interaction between the second meaning and these adverbial elements can give rise to the third adjective meaning.

2. For the presentation of the notion of « transcategorialité », see Do-Hurinville & Dao (2016).

2.2 *Juste* as a noun

- (32) *Le juste et l'injuste.*
[The just and the unjust]
- (33) *Durant la guerre, il s'est comporté comme un juste.*
[During the war, his behaviour was that of a righteous person]
- (34) *Dormir du sommeil du juste.*
[Sleep the sleep of the just]

In each of the first two meanings, *juste* may be nominalised to function as (i) a common noun indicating what is conform to justice, equity (Example 32), a person who observes exactly the duties of his religion, or, by extension, a person who judges or acts fairly (Examples 33, 34); (ii) or a proper noun (*Antoine Juste; Juste Aurèle Meissonnier; Juste de Gand; Just Fontaine*).

2.3 *Juste* as an adverb

The table below summarises five possible adverbial uses of *juste*, which appeared in the 17th century, illustrated by Examples (34) to (47).

Table 2. Five adverbial uses of *juste*

Five adverbial uses of <i>Juste</i>				
(a)	(b)	(c)	(d)	(e)
(XVII ^e siècle) "Accurately, exactness" (Example 35)	(XVII ^e siècle) "Exactly, precisely" (Example 36 & 37)	(XIX ^e siècle) "In a too strict manner, in barely sufficient quantity" (Example 38)	(XX ^e siècle) "Simply, only, uniquely" (Example 39 to 41)	(2014) "Really, frankly" (Anglicism, criticised use, Example 42 & 43)

- (35) *Chanter juste; viser juste; voir juste* (≈ avec justesse).
(Sing in tune; aim accurately; get it right)
- (36) *C'est juste le contraire.*
(It's just the opposite)
- (37) *Le téléphone a sonné juste quand je partais.*
(The telephone rang just when I was leaving)
- (38) *Ce breuvage crayeux et sans fraîcheur, tout juste potable, n'a aucune vertu hygiénique.*
(Amiel, 1866, quoted by Mellet & Monte, 2009)
(This chalky and unfresh beverage, only just drinkable, has no hygienic virtue)

- (39) *Je suis juste venu t’embrasser* (1954, quoted by Mellet & Monte, 2009)
(I came just to kiss you)
- (40) *Je ferme juste les fenêtres* (Leeman, 2004)
(I’m just closing the windows)
- (41) *Juste un ami !* (novel title)
(Just a friend!)
- (42) *C’est juste merveilleux / génial / magnifique / sublime / formidable / fantastique.*
(It’s just marvellous / great / magnificent / sublime / wonderful / fantastic)
- (43) *C’est juste un superbe ami !*
(He’s just a great friend)

The five adverbial uses of *juste* can be split into two groups. In the first group, which includes the first three uses, (a), (b) and (c), *juste* functions as an adverb. In other words, it is an adverb that modalises an element of the utterance to which it refers. According to the PRLF and the Wikidictionnaire, these three uses come directly from their adjectival meaning. In the second group which contains the last two uses, (d) and (e), *juste* seems to be pragmatikalised, with illocutionary value.

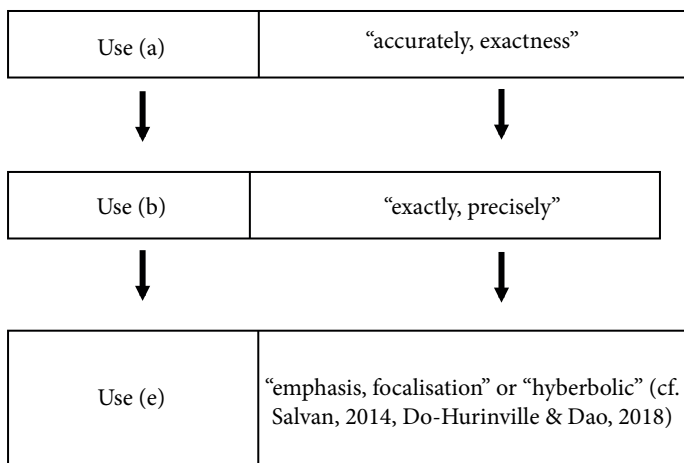
The uses (a) and (b) of *juste*, which remained standard for a long time and probably came from its second adjectival meaning (“idea of rightness”),³ emerge in the seventeenth century. In (a), *juste*, always postponed to the verb phrase (Example 35), can be glossed by “accuracy, exactness, as it should be”. In (b), *juste*, always preposed to the element it modalises, whether it is a noun phrase (36), or a subordinate clause (37), may be glossed by “exactly, precisely”. In (c), which probably derives from its third adjectival meaning (“idea of narrowness, insufficiency”) and appears in the nineteenth century, *juste* functions as an adverb meaning “In a too strict manner, in barely sufficient quantity”, as in (38).

As for the uses (d) and (e), relatively recent and considered as anglicisms, they have been frequently and strongly criticised by several French language supporters. In (d), illustrated by (39) to (41), *juste*, expressing an idea of restriction, can be substituted for “seulement, simplement, uniquement”. According to the PRLF and the Wikidictionnaire, this use of *juste* is an anglicism, which the speaker introduces to minimise the psychological impact on the hearer. In other words, *juste* in (39) to (41) works (i) explicitly as an adverb, modulating the element postponed to it, but (ii) implicitly as a discourse marker (DM), which relates the speaker to the hearer.

3. It’s not the adverb *juste*, but the adverb *justement*, which inherits the first adjectival meaning “idea of justice”, contains the following three meanings: (a) “in accordance with justice”; (b) “with accuracy”, and (c) “to mark the exact concordance of two facts, an idea and a fact” (cf. PRLF)

Its adverb function seems to prevent it from being syntactically placed at the beginning or the end of an utterance, but it must be interpreted pragmatically as a DM.⁴

The use of *juste* in (42) and (43), which represents the use (e), even more recent than that of *juste* in (d), is the main object of this article. This adverbial use of *juste*, called (e), would derive from (b), which comes from (a) according to the following semantic path:



In sum, the semantic evolution of *juste*, as an adjective and as an adverb, can be summarised as follows: justice > righteousness > precision > restriction > emphasis/focus (see Tables 1 and 2).

3. *Juste*, a double modal adverb

This section deals with the last use of *juste* as a double modal adverb. After examining briefly *très*, *vraiment*, *juste* and scalar (gradable) and extreme (non-gradable) adjectives in (3.1), I formulate hypotheses of *juste* and extreme (non-gradable) elements in (3.2).

4. See also Do-Hurinville & Dao (2018).

3.1 Profile of *très*, *vraiment*, *juste* with scalar (gradable) and extreme (non-gradable) adjectives

The examination of Examples (44a) to (46d) allows a better account for the functioning of *très*, *vraiment* and *juste* and the scalar (gradable) adjectives: *chaud* and *heureux*, and the extreme (non-gradable) adjectives: *génial*, *magnifique*, *merveilleux*, *sublimissime*.

- (44) a. *Le café est Ø chaud.*
(The coffee is hot)
- b. *Le café est très chaud.*
(The coffee is very hot)
- c. *Le café est juste chaud.*
(The coffee is barely hot enough)
- d. *Le café est vraiment chaud.*
(The coffee is really hot)
- (45) a. *Paul est Ø heureux.*
(Paul is happy)
- b. *Paul est très heureux.*
(Paul is very happy)
- c. *Paul est juste heureux.*
(Paul is just happy)
- d. *Paul est vraiment heureux.*
(Paul is really happy)
- (46) a. *Ce spectacle musical est Ø génial / magnifique / merveilleux / sublimissime.*
(This musical show is marvellous / great / magnificent / sublime)
- b. *Ce spectacle musical est *très génial (magnifique / merveilleux / sublimissime).*
(This musical show is *very marvellous (great / magnificent / sublime)
- c. *Ce spectacle musical est juste génial (magnifique / merveilleux / sublimissime).*
(This musical show is just marvellous (great / magnificent / sublime)
- d. *Ce spectacle musical est vraiment génial (magnifique / merveilleux / sublimissime).*
(This musical show is really marvellous (great / magnificent / sublime)

Semantically, if the use of *très* (very), an adverb of degree, serves to raise the scalar adjectives *chaud* and *heureux* to the highest degree in (44b) and (45b), its use is, in theory, impossible with *génial*, *magnifique*, *merveilleux*, *sublimissime* as in (46b) because those are extreme (non-gradable) adjectives, which differ from scalar adjectives in that they do not represent a range on a scale, but its ultimate point (cf. Paradis, 2001: 51). In other words, extreme adjectives already contain the idea of

‘very’⁵ in their definitions and thus are qualities “which leave no room for any additions or adornments” (cf. Duffley & Larrivé, 2012: 27).

As for *juste* and *vraiment*, they do not modify the degree of internal intensity of the adjectives, but consist in affirming that the qualities (states) “(be) hot, happy, marvellous” are respectively fair and true, from the speaker’s stance. This is why these adverbs can be combined with scalar adjectives (44c, 44d, 45c, 45d) as well as with extreme adjectives (45c, 46d). However (44c) can be interpreted in two ways: either the speaker estimates that his coffee is “barely hot enough” (not much more is needed to make him completely satisfied), or his coffee is hot as it is necessary, as it should, so he is almost satisfied. In (45c) with *juste*, the speaker notes only a certain satisfaction with Paul. In other words, in (44a) and (45a), the speaker asserts only two states: “to be hot” and “to be happy”, while in (44c) and (45c), these two states with *juste* are barely below those without *juste*.

3.2 Characteristics of *juste* with extreme (non-gradable) X

We have just seen that, from a semantic viewpoint, one cannot add internal intensity to extreme adjectives, which explains why *très* is not compatible with *génial*, *magnifique*, *merveilleux*, *sublimissime* as in (46b). However these adjectives can be modified both by *juste* in (46c) and by *vraiment* in (46d). When *juste* is anteposed to these extreme adjectives, it functions as a double modal adverb, bearing on the utterance and on the utterance act as follows:

- a. Modalisation on the selected extreme adjective (*génial*) of the utterance: *juste* works as a magnifying glass on this adjective (*génial*) to emphasise the intensity inherent in this extreme adjective. It thus creates a magnifying glass effect or a “hyperbolic effect” on it (cf. Salvan, *op.cit.*).
- b. Modalisation on the utterance act: *juste* allows the speaker to point out his stance (his view on this adjective) in relation to this extreme adjective (*génial*) by emphasising the accuracy and the relevance of the selection of this adjective.

By saying *Ce spectacle musical est juste génial (magnifique, merveilleux, sublimissime)* as in (46c), the speaker sends a subliminal message, inciting the hearer to accept his view, which is not the case in (46a) without *juste*. With the choice of these extreme adjectives to formulate a compliment, the speaker prevents, in anticipation, the hearer from reversing this choice, because the speaker indicates that he does not exaggerate, and that there is no other choice more judicious, more relevant than that of the adjective he has selected, because it is the maximum. From this point of view, the use of *juste* in (46c) is interpreted as a metalinguistic use, which allows the speaker to pass on a judgment towards the hearer. One can say that *juste* is an

5. Let’s take *magnifique* as an example which can be glossed by “très beau”.

“emphasizer” (Quirk *et al.* 1985: 447), or a “focus particle” (König, 1991: 121–124) relating to the retained extreme value (represented by a non-gradable adjective) excluding any other lower values.

After examining my corpus,⁶ my hypothesis on the functioning of the recent use of *juste* “Juste X” is as follows: in order for the hyperbolic effect or magnifying glass effect (which comes from the metalinguistic use of *juste* with double modalisation) to occur, the two essential conditions below must be met:

- i. X, which can be an adjective, a noun phrase, a verb phrase, or a prepositional phrase, must be **extreme (non-gradable)** or **interpreted as such**, expressing the highest or lowest degree of a quality, more precisely the opposite extremes of a quality (*freezing, enormous, excellent, marvellous, magnificent, sublime* vs. *boiling, tiny, terrible, hideous, disastrous, dreadful*).
- ii. The co-textual elements (i.e. internal elements) must confirm the **laudatory** or **depreciatory** character of X and of the whole discourse.

If one or both of those conditions are missing, one will be dealing with an ambiguous case where *juste* expresses the restrictive effect. In order to avoid such ambiguity, one must employ the contextual elements (i.e. external elements) in which the utterance occurs.

4. *Juste* and extreme adjectives ⁷

In this section *juste* is examined, first with extreme adjectives (§ 4.1), then with extreme adjectives formed by the suffix *-issime* (§ 4.2), and finally with extreme adjectives formed by the prefix *-in* (§ 4.3).

6. In order to study this new use of *juste* in French (Sections 4 to 7), I have included in my corpus oral (daily conversations, radio and television interviews) and written (written press, Facebook, Twitter, blog, Tripadvisor) examples.

7. Using Google Search, I have compiled a list of extreme adjectives in alphabetical order in which, according to the textual laudatory ou depreciatory elements of the examples, *juste* expresses an effect of non-restrictive meaning, but emphatic or hyperbolic: *Juste abyssal (abominable, admirable, bouleversant, brillantissime, canon, captivant, catastrophique, cauchemardesque, dégueulasse, délirant, démentiel, désastreux, diabolique, éblouissant, économiquissime, effrayant, effroyable, énorme, énormissime, enthousiasmant, épatant, époustouflant, épouvantable, excentrique, excellent, excellentissime, exceptionnel, exquis, extra, extraordinaire, extra, fantasmagorique, fantastique, fantasmagorique, fascinant, féérique, flippant, formidable, génial, génialissime, grandiose, grandissime, gigantesque, gravissime, horrible, hallucinant, impensable, impressionnant, incroyable, infernal, inoubliable, irréel, magique, magistral, magnifique, majestueux, merveilleux, minable, monstrueux, nullissime, odieux, parfait, parfaitissime, phénoménal, prodigieux, rapidissime, ravissant, renversant, sensationnel, simplissime, somptueux, splendide, sublime, sublimissime, superbe, super, terrifiant, terrible, topissime, traumatisant...).*

4.1 *Juste* and extreme adjectives

- (47) « C'est incroyable, c'est **juste sublime** [...] C'est extraordinaire, c'est sublimissime ! » (oral, télé, *France 5*, 2016)
- (48) « Les Jeux olympiques, c'est **juste exceptionnel** » (oral, radio, Teddy Rinaire, Porte-drapeau français, 2016)
- (49) T'es **juste merveilleux** comme dessinateur, franchement, Bravo ! (facebook)
- (50) C'est **juste extra** ! Top ! Génial ! Je kiffe ! rhaaaa gagaga (title, blog, 2015)
- (51) Un point **juste dantesque**. Point of the year? #usopen (title, blog tennis, 2017)
- (52) À vous couper le souffle, **juste magnifique** (title, *Epanews*, 2014)
C'est un feu d'artifice vraiment magnifique.
- (53) **Juste sensationnel** : un endroit magnifique, un accueil de rêve, un massage trop relaxant. À recommander à tous ceux qui veulent s'offrir un havre de paix et détente. C'est sûr on y retournera (title, facebook)

Examples (47) to (53) are oral and written examples. Most of the written examples are headings, (48) to (51), aiming at drawing the reader's attention. *Sublime*, *exceptionnel*, *merveilleux*, *extra*, *dantesque*, *magnifique*, *sensationnel* are all extreme adjectives, which explains why the anteposition of the adverb of degree *très* to those adjectives is not possible. Pragmatically, the use of *juste* is to create a magnifying glass effect on the adjectives, when the speaker sends a message to the hearer to make him understand that this choice of adjectives is judicious and relevant. It should be noted that in these seven examples, the magnifying glass effect produced by *juste* is clearly observed because the following two conditions are reunited: X are extreme adjectives, and the co-textual elements are all laudatory.

- (54) J'ai eu du mal à me plonger dans l'histoire, j'ai quand même voulu finir le livre mais je n'aurais pas dû insister. L'idée de base n'est pas mauvaise mais le livre tire en longueur, la fin est prévisible, les personnages un peu caricaturaux (le psy sur sa belle moto, la commandante garçon manqué et le beau mec qui trompe sa femme...). Et la décision finale de la commandante est **juste abracadabrantesque**... (Critique du roman *Maman a tort* de Michel Bussi 2015)
- (55) Savez-vous à quel pourcentage des dépenses de la France équivaut la paye des élus ? C'est **juste infinitésimal**. La dette de la France, ça se compte en milliards vous savez, pas en centaines de milliers d'euros.

In (54), *abracadabrantesque* means "complètement abracadabrant", and *abracadabrant* means "complètement incroyable". Or, *incroyable* is an extreme adjective (see § 4.3). It is the same for *infinitésimal*, in (55), which means "infiniment petit". In other words, semantically, *abracadabrantesque* and *infinitésimal* are extremely

powerful adjectives. The use of *juste* in these examples is undoubtedly a metalinguistic use, with double modalisation.

- (56) « Mes amis, c’est **diabolique, juste diabolique**. Je me demande si les gens se disent “Ce joli petit film va inciter ma fille de 5 ans à devenir lesbienne, à être ouverte à l’homosexualité ou au moins à la bestialité” [...] Je ne suis pas un conspirateur fou mais je me demande parfois s’il n’y a pas quelque chose de diabolique derrière. Si j’étais le Diable, qu’est-ce que je ferais pour véritablement réduire en miettes le système social dans son intégralité et faire quelque chose de **vraiment, vraiment, vraiment diabolique** pour les enfants de 5, 6, 7 ans de familles catholiques en Amérique? » (*L’Express*, 2014, *La Reine des neiges de Disney*: selon le pasteur Kevin Swanson. Disney)
- (57) « C’était **horrible, juste traumatisant**, je ne sais pas quoi en dire ». Les témoignages affluent, après les deux explosions qui ont endeuillé le marathon de Boston, lundi. Beaucoup de coureurs, qui venaient juste d’en finir avec l’épreuve sportive, se trouvaient toujours sur place et ont pu raconter ce qu’ils avaient vu. « Il y avait des familles autour de nous avec des enfants qui criaient. Nous avons essayé de partir aussi vite que possible », a ainsi témoigné une marathonnienne (20 minutes, 2013)
- (58) « C’est difficile de ne pas être ému, car c’est quelque chose d’exceptionnel. L’intensité est multipliée par je ne sais combien par rapport à ce qu’on vit à terre. La transition est énorme. C’est fou. Il y a quatre ans, j’étais allé sur Foncia (le bateau vainqueur de Michel Desjoyeaux), je pensais donc être préparé, mais en fait pas du tout. Vous réalisez combien vous touchez de gens. C’est **hallucinant, juste extraordinaire** »
(François Gabart, Eurosport.fr, *Quelque chose d’exceptionnel*, 28.01.2013)
- (59) Je ne suis pas **bizarre, juste différent** (facebook)
- (60) Je ne suis pas **fou, juste excentrique** (forum)

At first sight, Examples (56) to (60) are similar, having two adjectives: *juste*, preceded by a comma and placed between the two adjectives, modalises only the second, and not the first one. However, these five examples must be split syntactically and semantically into two groups. The structure of the first group, (56) to (58), is: “It is Adj₁, *juste* Adj₂”, while the structure of the second group, (59) and (60), is: “It is not Adj₁, *juste* Adj₂”. This explains why the interpretation of *juste* differs between the first and the second group.

In (56) to (58), *diabolique*, *horrible*, *traumatisant*, *hallucinant*, *extraordinaire* are extreme adjectives; the use of *juste* in these examples is a metalinguistic use with double modalisation, bearing upon the adjective and upon the utterance act. On the contrary, in (59) and (60), *bizarre*, *différent*, *fou*, *excentrique* are gradable

adjectives. According to the syntactic structure, the speaker first rejects the first properties ($Adj_1 = \textit{bizarre}$ and *fou*) and then emphasises the second properties ($Adj_2 = \textit{différent}$ and *excentrique*) modified by *juste*. In the contexts of those examples, it is probably the use of *juste* with the restrictive effect, glossed by *seulement* or *simplement*, bearing only over the second adjectives.

- (61) **Voyage juste énorme splendide** (title, *Tripadvisor*, 2019)
- (62) **Juste somptueux et merveilleux** (title, *Tripadvisor*, 2018)
 Nous sommes venus entre amis passer 2 jours à St Tropez [...]. Le personnel est très chaleureux, très accueillant et à votre petit soin on se dirait comme « à la maison ». Concernant la partie restauration une vraie partie de plaisir gustative aucune fausse note ! Merci à toute l'équipe de cet splendide établissement.
- (63) **Juste dégueulasse, immangeable** (title, *Tripadvisor*, 2015)
 Archi nul impossible de finir le steak était cramé dure avec un goût de charbon. La salade de la mer c'est bien une salade mais juste de la salade Dans la salade de chèvre il n'y a que du persil, blanc en plus de la salade en sac. On ne prend même pas de dessert tellement c'était horrible.
- (64) **Juste incroyable, irréel, fantasmatique** (title, *Tripadvisor*, 2016)
 Nous décidons au cours de notre séjour de louer un catamaran et de faire du cabotage d'îles en îles. Temps superbe, équipage sympathique, tout est réuni pour que nous passions une bonne journée. [...] Un truc juste incroyable, des petits barracudas et des bébés soles passaient sous nos pieds, cherchant sûrement à récupérer quelques miettes. [...] Nous sommes revenus avec des images plein la tête, nous étions dans un rêve. Si vous passez dans le coin, profitez de chaque instant, ces moments-là sont rares...

Examples (61) to (64) are all headlines in which the author uses a series of two or three extreme adjectives. Adjectives can be juxtaposed as in (61), separated by a comma as in (63) and (64), by the conjunction *and* as in (66). The adjectives are certainly all extreme, but one notices that the adjectives, in (64), go crescendo: “incroyable, irréel, fantasmatique”, or at least up to the most intense or more precise states: “énorme splendide” in (61) or “dégueulasse, immangeable” in (63). In general, an adjective, very often used orally, undergoes a semantic weakening: it is the case of *incroyable* in (64), *énorme* in (61) or *dégueulasse* in (63), in comparison to *splendide*, *magique*, *fantasmatique*. As for *irréel*, which should be interpreted as *fantastique* (cf. PRLF), it seems to be less intense than *fantasmatique*, which justifies its intermediate position in (64). In this example, the reverse order would be more questionable: “?fantasmatique, irréel, incroyable”. In (63), the use of *dégueulasse* expresses a rejection less precise than *immangeable* to describe a meal.

4.2 *Juste* and extreme adjectives with the suffix *-issime*

- (65) **Juste génialissime** (title, *Tripadvisor*, 2016)
Un vol **super génial**, de beaux paysages, pour finir avec un peu de voltige pour un bon bol d'air frais !! Sans oublier un super moniteur !!!
- (66) Un restaurant **juste excellentissime** ! (title, *Tripadvisor*, 2017)
Un restaurant où l'accueil est très pro, sympathique et chaleureux. Les plats sont justes incroyablement bons. Le poulet épicé grillé, accompagné de sa salade verte et de son écrasé de pommes de terres aux olives est juste succulent.
- (67) **Juste énormissime** (title, *Tripadvisor*, 2016)
Marche physique mais qui en vaut largement le coup. Une vue juste incroyable à couper le souffle, je le conseille vivement un bol d'air frais qui fait vraiment du bien !
- (68) **Juste topissime !!!** (title, *Tripadvisor*, 2017)
Suite aux avis du site *Tripadvisor*, nous avons dîné dans ce restaurant hier soir et nous nous sommes régalez. C'était bon, copieux, sympa et surtout avec un rapport qualité / prix inégalable. Restaurant incontournable sur Blanes ! Pour 22,50 euros le menu il y avait 3 entrées à découvrir [...].
- (69) Wouah, comment dire ? Il y a les livres en général et puis il y a les livres de Sarah Dessen. Comme toujours, ce roman m'a totalement transportée et bouleversée. L'histoire est très bien construite et comme à chaque fois que je lis un livre de cet auteur, je n'ai pas pu faire autrement que de le dévorer en une journée. Ce que j'ai tout particulièrement aimé dans ce récit, est le fait que malgré que Macy soit totalement noyée dans son chagrin, elle parvient tout de même à poursuivre sa vie « normalement » après avoir connu l'une des pires douleurs au monde ; la perte d'un être cher. J'ai tout simplement adoré le personnage de Tim qui est **juste parfait**. L'histoire de sa mère m'a également beaucoup émue et j'ai énormément aimé son côté artistique. En bref, une fois de plus Sarah Dessen a touché dans le mille avec ce livre **juste parfaitissime**. Cette histoire est tout simplement vrai, il n'y a pas d'autre mot.
(*Pour toujours... jusqu'à demain*, Sarah Dessen)
- (70) Cette série est **juste nullissime** #TF1 (Twitter, 2012)
- (71) L'ex de Courteney Cox était **juste ridiculissime** dans son costume de Zoro alors qu'il se rendait à une Pre-Halloween party ce weekend
(Halloween 2012, *Les costumes les plus moches des stars*, Staragora, 29.10.2012)

Génial, excellent, énorme, top, parfait are semantically already extreme adjectives with a positive value. However, some speakers, perhaps with some exaggeration, may choose to boost the semantics of these adjectives with the addition of the suffix *-issime*: *génialissime, excellentissime, énormissime, topissime, parfaitissime*, as in (65) to (69). As for *nul* and *ridicule* in (70) and (71), with the suffix *-issime*: *nullissime* and *ridiculissime*, they have become extreme adjectives with a negative value. Therefore, the anteposition of *juste* to those extreme adjectives produces a hyperbolic effect or magnifying glass effect, and its use is metalinguistic with double modalisation.

4.3 *Juste* and extreme adjectives with the prefix *-in*

- (72) **WTA – Serena** : « **L’or olympique, c’est juste incroyable** »
 Double médaillée d’or aux Jeux Olympiques de Londres en 2012 [...], Serena Williams s’est confiée avant le commencement des Jeux de Rio en août prochain. Si l’actuelle numéro 1 mondiale est parvenue à égaler Steffi Graf en nombre de titres du Grand Chelem dans l’ère Open (22), elle avoue que l’or olympique a une teneur toute particulière : « [...] Gagner les Jeux Olympiques, c’est une sensation spéciale. Tu te rends alors compte que tu ne joues pas seulement que pour toi mais pour tout un pays. Je n’aurais jamais pensé y arriver mais je n’y ai jamais renoncé. Gagner l’or olympique, c’est **juste incroyable**. », rapporte TWUSA (TennisActu.net)
- (73) Si la décoration est bien faite, c’est surtout la vue à couper le souffle qui capte le regard. « Se réveiller et voir un panorama de tous les monuments que nous avons visités hier, c’est **juste irréal** », s’émerveille Dawn, la fille de Scott venue elle-même avec sa fille de 13 ans (oral, *Parisien*, 06.07.2016)
- (74) **Pitkowski** : « **C’était juste impensable** »
 Les deux victoires d’Alizé Cornet et de Virginie Razzano à Moscou samedi au premier tour de la Fed Cup ont surpris Sarah Pitkowski. « On pouvait peut-être imaginer que les deux équipes soient à égalité 1–1, mais que la France mène 2–0 face à une équipe composée d’anciennes numéros 1 mondiales, c’était **juste impensable**, explique l’ancienne joueuse de l’équipe de France. [...] » (RMC.Sport, 2011)
- (75) **Juste inoubliable**
 L’AS Monaco n’oubliera jamais ce 21 février. Dans une Disney Arena archi-comble, la Roca Team a battu Chalon (99–74) en finale de la Leaders Cup, décrochant le premier titre en Pro A de son histoire. C’est aussi la première fois qu’un promu s’impose en Leaders Cup ou dans l’ex-Semaine des As. L’exploit est considérable. [...] (Disneyland.Paris Leaders Cup, 2016)

- (76) « C’est **juste impossible** » : très remonté, Nagui remet en place une candidate qui souffle les réponses à son voisin

(*Tout le monde veut prendre sa place* sur France 2, 2018)

Nagui n’a pas vraiment apprécié qu’une candidate de son jeu *Tout le monde veut prendre sa place* souffle la bonne réponse à un concurrent qui pouvait devenir challenger.

- (77) **Juste immangeable**

Nous y sommes allés entre amis, après une journée de ski : quelle erreur. C’était juste ignoble, les plats étaient super épicés, sans goût, ou juste dégueulasses. Y étant allés avec une indienne, il est assez correct d’écrire que c’est un resto indien pourri
(titre, Tripadvisor, 01.03.2016)

Unlike the negative adverb *pas* (see Section 5), the privative prefix *in-* automatically confers a subjective appreciation value to adjectives. According to the laudatory contexts of (72) to (75), *incroyable*, *irréel*, *impensable*, *inoubliable* should therefore be interpreted as “fantastique, fabuleux, extraordinaire, magnifique” (cf. PRLF), which are extreme (non-gradable) adjectives. As for *impossible* and *immangeable* in (76) and (77), they are also non-gradable adjectives. In view of the contexts of (72) to (77), *juste* also produces a hyperbolic effect. Those examples can be glossed by “vraiment” or “tout à fait” (*incroyable*, *irréel*, *impensable*, *inoubliable*, *impossible* and *immangeable*).

5. *Juste* and non-gradable sequences [*pas* + gradable adjectives]

- (78) C’est **juste pas possible** ! (oral)
- (79) **Juste pas bon**, mais pas bon du tout (Tripadvisor, 2017)
Le restaurant n’est pas bon. Mon fils a pris un rumsteck... La viande immangeable. Les haricots verts sont **juste** sans saveurs [...]
- (80) Murray: « Je ne joue **juste pas bien** » (Sportsactu.be, 2017)
Défait à trois reprises sur ses quatre dernières rencontres, Andy Murray est en plein doute. Le n°1 mondial, qui avait fini l’année 2016 en boulet de canon, a recommencé 2017 avec beaucoup plus de difficultés... et il s’en rend compte.
- (81) « 64 policiers nationaux entre 22 et 23 heures je crois, en tout cas au moment de l’attentat, sur la Promenade des Anglais pour sécuriser 40.000 personnes, c’est **juste pas raisonnable**, c’est une grave erreur de la part de la préfecture de Nice », a déclaré M. Lagarde sur RMC. (*Le Parisien*, 2016)

- (82) Zidane : « Ronaldo, c'est **juste pas humain** » (*Sports.fr*, 2015)
 En bon Madrilène d'adoption, Zinédine Zidane n'a que du bien à dire de Cristiano Ronaldo. L'ancien n°10 des Merengue, aujourd'hui entraîneur de l'équipe réserve, a encore fait l'éloge de l'actuelle star du Real, dans une interview accordée au Canal Football Club. « 50 à 55 buts par saison c'est **juste pas humain** », s'extasie « Zizou ».

While [in-adj] in (Section 4.3) is a derivational morphology (the privative prefix merging with the root), the sequence [pas + gradable adjective], without the particle *ne*, in (Section 5) is rather a syntactic construction (cf. Dugas 2014): indeed adverbs can be inserted between *pas* and the adjective or the past participle: “factures pas encore payées”, “travail pas assez rémunéré”, “travail pas bien rémunéré”. One can say that *pas*, which does not merge with the past participle or the adjective, can play the role of negative prefix (cf. Kalik 1971 : 132), unlike *-in*, a real privative prefix.

When one says A is B, one can also say A is [pas B], which signifies A is on the exterior of the attribute B. In other words, one rejects the property B for A. Let's take an example like “un paysage pas réel” meaning that one denies the property “réel” for “paysage”, which can mean it is an artificial landscape. As for “un paysage irréel”, it is a landscape considered to be fabulous, fantastic or magical, as in (73), because the privative prefix *ir-* imparts to the adjective *irréel* a value of positive subjective appreciation.

Furthermore, if certain prefixed adjectives, [in-adj], can be modified by *très* as in “très improductif”, “très irraisonné”, “très inabsorbable” (cf. Dugas, *op. cit.*: 204) or “très incommode”, “très insupportable”, on the contrary, the sequence [pas adj] does not accept to be modified by *très*: [*très [pas adj]]. In other words, [pas adj] is a non-gradable sequence. The use of *juste* with a non-gradable element is therefore metalinguistic with double modalisation.

Noting that in the laudative context of (82), “pas humain” must be interpreted as a superman in the football world, having achieved extraordinary sporting feats arousing limitless admiration from the speaker, unlike “inhumain”, used to point out a cruel and barbaric person, lacking humanity.

The sequence [pas adj], corresponding to X, is categorical negation, emphasising the monological attitude of the speaker, who considers himself as being the only one to have the type of representation that he states (cf. Morel, 1994: 99). The anteposition of *juste* to that sequence focusing on this monological and peremptory attitude of the speaker, who sends a clear message to the hearer, to make him understand that the negative phrases “pas bon”, “pas bien”, “pas raisonnable”, “pas humain” are fair and relevant: the hearer has no choice but to accept those negative phrases.

6. *Juste* and two other non-gradable sequences

In this section I will examine two non-gradable sequences: (i) [*très/trop* + gradable adjectives], and (ii) [*hyper/super/méga/ultra/supra/giga* + gradable adjectives].

6.1 *Juste* and non-gradable sequences [*très/trop* + gradable adjectives]

- (83) Allez-y c’est **juste très beau** !
(titre, avis sur Musée national de Singapour, *Tripadvisor*, 2016)
- (84) Rien à dire c’est **juste très bon** (titre, *Tripadvisor*, 2015)
Très bon et très jolie Restau, avec une formule du midi à un prix responsable pour une telle qualité de produits et services. Bravo.
- (85) Cette fille chante **juste trop bien** ! (Tf1, *The Voice Kids*, 2019)
- (86) Tu es sublime ma chérie ! Tu as l’air **juste trop sympa** ! Tu as tout pour toi quoi (blog)
- (87) T’es pas moche, t’es **juste très difficile** à regarder... (facebook)
- (88) Blagues **juste trop drôles** (facebook)
- (89) Il suffit que tu frappes à la porte, a poursuivi Beverly. Hillary est **juste trop cool, sympa de chez sympa**. Elle ferait n’importe quoi pour moi, elle... »
(*Moi, Charlotte Simmons*, Tom Wolfe, 2015)

Beau, bon, bien, sympa, difficile, drôle, cool are scalar adjectives, but the anteposition of *très* and *trop* to these gradable adjectives constitutes non-gradable sequences as follows: [*très/trop* + gradable adjectives], which corresponds to X (see § 3.3). In other words, *très bon, très beau, trop bien* are semantically equivalent to the following extreme adjectives: “magnifique, merveilleux, parfait”. An utterance like (85) “Cette fille chante trop bien” is comparable to “Cette fille chante magnifiquement”. Consequently, in (83) to (89), the anteposition of *juste* to these non-gradable sequences X can generate the hyperbolic effect. In addition, the co-textual elements of these examples are laudative, which completely confirms the metalinguistic use of *juste* as in § 4.1 with extreme adjectives.

6.2 *Juste* and non-gradable sequences [*hyper/super/méga/ultra/supra/giga* + gradable adjectives]

- (90) Bon, j’avoue, à ce moment-là, je mourais d’envie de raconter ça à quelqu’un.
[...] J’étais **juste super soulagée** de pouvoir confier mon secret à quelqu’un
(*Le Carnet d’Allie 1*, Meg Cabot, Juvenile Fiction, 2012)

- (91) Kerry Washington était **juste hyper mignonne** pour les Oscars 2013...
(Purebreak, 2013)
- (92) Emmanuel Moire, vainqueur avec Fauve Hautot de la 3^e saison de l'émission « Danse avec les Stars » a laissé éclater sa joie à l'issue de la finale samedi soir sur TF1. « On est **juste super méga heureux** » a-t-il lancé sur un tweet. « C'est inouï » a ajouté le chanteur manceau. (Lemainelibre.fr, 02.12.2012)
- (93) Ce dessert est **juste hyper méga calorique** donc une fois que vous l'aurez avalé ou engloutie ne vous pesez PAS !! (sugardises.canalblog.com, 11.03.2015)
- (94) Je suis **juste super méga ultra supra giga heureuse** !!!! Hahahahaha
(twitter, 05.2012)

In the examples above, *super*, *hyper*, *ultra*, *supra*, *mega*, *giga* play the same role as *très* and *trop* in § 6.1, so that the gradable adjectives *soulagée*, *mignonne*, *heureux*, *calorique*, *heureuse*, preceded by those superlative adverbs, form non-gradable adjective phrases as follows: [*super*, *hyper*, *ultra*, *supra*, *mega*, *giga* + gradable adjectives] corresponding to X. The juxtaposition of several superlative adverbs generates the hyperbolic effect, brought to an unbelievable climax as in (94), which is a fabricated example, full of humour.

7. *Juste* and non-gradable [noun / verb / prepositional] phrases

This section illustrates another case: phrases considered to be non-gradable distributed as follows: noun phrases (95) to (99), and (103), verb phrase (100), and prepositional phrases (101) and (102).

- (95) Ce film est **juste splendide**. Mieux que *Gladiator*, c'est **juste le meilleur film du monde**. (AlloCiné, Critique du film *Titus*, 2011)
- (96) **Juste une merveille** ce coucher de soleil (Google)
- (97) **Juste un monstre d'efficacité** !! 🍌🍌 Quel joueur ! (Twitter, 2018)
- (98) Elina Svitolina : « C'est irréal, c'est **juste un moment parfait**. Battre Serena Williams aux Jeux Olympiques, c'est quelque chose de spécial. [...] » (L'Équipe, 2016)
- (99) Roberta Vinci (vainqueur de Serena Williams 2–6, 6–4, 6–4) : « Je ne réalise pas encore mais c'est incroyable. C'est comme un rêve. Je suis très heureuse, même si j'ai un peu de tristesse pour Serena parce qu'elle a une incroyable carrière, des titres dans tous les tournois du Grand Chelem... Mais que dire ? C'est dur d'exprimer mes émotions. Peut-être que demain matin je pourrai dire quelque chose mais là, c'est **juste un moment magique** [...] » (L'Équipe, 2015)

- (100) Après sa troisième défaite d'affilée en finale d'un Grand Chelem, l'Américaine (Serena Williams) a nié toute forme de nervosité et préféré attribuer sa défaite à la grande qualité de tennis développée par Simona Halep. « Qu'est-ce qui vous a le plus surpris aujourd'hui (samedi) ?
Elle (= Simona Halep) a **juste joué un match formidable**. Mais je ne crois pas que ça surprenne quiconque que mes adversaires font parfois des matches formidables contre moi. J'ai essayé différents trucs. Mais rien ne marchait. J'ai surtout fait beaucoup trop d'erreurs. (L'Équipe, 2019)
- (101) Il a chanté **juste à merveille** (Tf1, The Voice Kids, 2019)
- (102) L'Australie comme vous l'avez rêvée ! Des espaces infinis à perte de vue, des rencontres marquantes et des expériences inattendues. Dix-neuf jours fabuleux pour aller à la rencontre de la culture aborigène, découvrir par les voies terrestres, aériennes et maritimes le mythique outback et sa faune spectaculaire, les paysages côtiers uniques et des personnages locaux d'une grande richesse. Une échappée renversante où chaque jour réserve son lot de sensations fortes et de moments magiques. Sans parler des adresses atypiques, luxueuses et exclusives qui composent un écrin majeur tout au long du circuit. Un voyage **juste hors norme** !

In Examples (95) to (102), the element X is represented by noun phrases, verb phrase or prepositional phrases. The noun phrases: “Le meilleur film du monde” in (95), “une merveille” in (96), “un monstre d'efficacité” in (97), “un moment parfait” in (98), “un moment magique” in (99), the verb phrase: “a juste joué un match formidable” in (100), the prepositional phrases: “à merveille” in (101) and “hors norme” in (102), are semantically non-gradable phrases. There are also laudative co-textual elements in all those examples. For all those reasons (see § 3.3): (i) non-gradable phrases, (ii) laudative co-textual elements), this is undeniably a metalinguistic use of *juste* with the hyperbolic effect.

- (103) Hier, contre Simona Halep qui a laissé ses complexes d'infériorité au vestiaire, ce n'était plus une question de tennis ou de physique, **juste une immense championne** (= Serena Williams) qui a vécu un naufrage émotionnel (L'Équipe, 2019)

Examples (100) and (103) deal with the same sporting event, which is 2019 Wimbledon Championships-Women's Singles. The Romanian Simona Halep defeated the American Serena Williams in the final, 6–2, 6–2, to win the Ladies' Singles tennis title. With the laudative context of (100) and the non-gradable verb phrase “a **juste** joué un match formidable”, *juste* must express a hyperbolic effect and not a restrictive effect. In (103), the interpretation of *juste* differs from the one in (100) for the following reason: the noun phrase “une immense championne” in (103) is

undoubtedly non-gradable semantically, but according to the co-textual elements “[...] ce n’était plus une question de tennis ou de physique [...] qui (= Serena Williams) a vécu un naufrage émotionnel”, *juste* must express a restrictive and not hyperbolic effect, unlike (100). In other words, in (103), *juste* must be glossed by “seulement, simplement, rien que”, and not by “vraiment, tout à fait”.

- (104) C’est **juste**₁ l’espoir d’un avenir (Le Parisien, 2019)
 Votre don peut **JUSTE**₂ faire toute la différence dans la vie des réfugiés.
 Le nombre de personnes déplacées de force à travers le monde dépasse aujourd’hui 70 millions. Le HCR, l’Agence des Nations Unies pour les réfugiés, a plus que jamais besoin de votre soutien pour mettre à l’abri et protéger les enfants, les femmes et les hommes déracinés partout dans le monde et leur permettre de se reconstruire.
 Faites un don sur unhcr-undonjuste.fr
 # Undon**juste**₃
 UNHCR, L’Agence des Nations Unies pour les réfugiés

In (104), which is an excerpt from a campaign in a newspaper, launched by UNHCR⁸ in aid of children of refugees, there are three uses of *juste*. If *juste*₃ in “Un don **juste**” is identified as an adjective to qualify the noun phrase “un don”, *juste*₁ and *juste*₂ in “C’est **juste** l’espoir d’un avenir” and “Votre don peut **juste** faire toute la différence...” function as adverbs to focus on the noun phrase “l’espoir d’un avenir” and on the verb phrase “faire toute la différence...”. Certainly this noun phrase and this verb phrase are not non-gradable phrases, but according to the context and the co-text of this example, *juste* must be interpreted as a metalinguistic use. In other words, *juste* must express a hyperbolic effect or a magnifying effect “vraiment, tout à fait”, and not a restrictive effect “seulement, simplement, rien que”, in order to boost donations to the Organisation.

8. Conclusion

Though *juste* in French and *just* in English come from the same Latin source, those two units are far from being perfect equivalents. The frequency of use of *just* in English is much higher than that of *juste* in French, as shown in Examples (1) to (27). If the use of *just* in English with extreme adjectives was examined in 1969 (cf. Cohen), it took almost half a century for this same use of *juste* in French to be studied for the first time in 2014 (cf. Salvan).

8. The *UN Refugee Agency* is a global organisation dedicated to saving lives and protecting the rights of refugees.

The aim of this article was to study this recent adverbial use of *juste*, identified orally in the early 2000s, officially registered in the PRLF in 2014. From the Latin *justus*, *juste* is a transcategorial unit par excellence, as it can be an adjective, a noun, or an adverb. In its recent adverbial use, *juste* functions as a metalinguistic marker which creates double modalisation: it has scope both over the element X (which can be a non-gradable adjective, a non-gradable noun phrase, a non-gradable verb phrase or a non-gradable prepositional phrase), and over the utterance act. In order for the hyperbolic effect or the magnifying glass effect to occur, the following two essential conditions must be met:

- i. The element X, postposed to *juste*, must be **extreme (non-gradable) or interpreted as such**, expressing the highest or lowest degree (the opposite extremes) of a quality.
- ii. The co-textual elements (i.e. internal elements) must confirm the laudatory or depreciatory character of X and of the whole discourse.

If those two conditions are lacking, *juste* expresses the restrictive effect. If one only of those two conditions is missing, the utterance becomes ambiguous in the sense that *juste* can express the restrictive effect (glossed by “seulement, simplement, rien que”), or the hyperbolic effect (glossed by “vraiment, tout à fait”). To remove that ambiguity one must use co-textual elements (i.e. internal elements) to determine the effect produced by *juste*, as shown in Examples (103) and (104).

Acknowledgements

I would like to thank Gaétane Dostie (University of Sherbrooke) and Huy-Linh Dao (Inalco) for their relevant comments and suggestions which allowed me to improve this article. I also thank Eric Corre (University of Sorbonne Nouvelle), who kindly proofread this article.

References

- Bat-Zeev Shyldkrot, Hava. 2001. Grammaticalisation, changements sémantiques et quantification. *Travaux du Cerlco* 14 : 11–34.
- Blanche-Benveniste, Claire. 2001. *Juste et plein*, tout juste et tout plein. In *Langage et Référence. Mélanges offerts à Kerstin Jonasson*, Hans Kronning et al. (eds), 65–74. Uppsala: Uppsala University Press.
- Dao, Huy-Linh, Do-Hurinville, Danh-Thành & Lafontaine, Fanny. 2020. Étude syntaxique et sémantique d’un nouvel emploi de *juste*: *c’est juste inouï* ! In *DIA du français actuel*, Gaétane Dostie, Pascale Hadermann & Florence Lefevre (eds), 205–222. Bern: Peter Lang.

- Do-Hurinville, Danh-Thành. 2018. Juste, de l'adjectif à l'adverbe d'énonciation, une unité juste transcategorielle. In: Danielle Leeman & Céline Vaguer (eds.), *Quand les formes prennent sens. Grammaire, prépositions, constructions, système*. Paris: Lambert Lucas, 235–246.
- Do-Hurinville, Danh-Thành & Dao, Huy-Linh. 2018. Juste un nouvel emploi de *juste*. C'est juste génialissime! *Bulletin de la Société de Linguistique de Paris* 113(1): 103–143.
- Do-Hurinville, Danh-Thành & Dao, Huy-Linh. 2016. La transcategorielle. Une histoire de *limite* sans limite. *Bulletin de la Société de linguistique de Paris* 111(1): 157–211.
- Duffley, Patrick & Larrivée, Pierre. 2012. Collocation, interpretation and explanation: The case of just any. *Lingua* 122: 24–40. <https://doi.org/10.1016/j.lingua.2011.10.008>
- Dugas, Edwige. 2014. [non-Adj] sequences in French: Morphological negation, Syntactic negation, or in-between? *Linguisticae Investigationes* 37(2): 194–208. <https://doi.org/10.1075/li.37.2.01dug>
- Garcin, Jérôme. 2009. C'est juste insupportable. <<http://bibliobs.nouvelobs.com/actualites/20091204.BIB4515/c-039-est-juste-insupportable.html>>
- Kalik, A. 1971. La caractérisation négative. *Le Français Moderne* 39: 128–146.
- König, Ekkehard. 1991. *The Meaning of Focus Particles: A Comparative Perspective*. London: Routledge.
- Leeman, Danielle. 2004. L'emploi de *juste* comme adverbe d'énonciation. *Langue Française* 142: 17–30. <https://doi.org/10.3917/lf.142.0017>
- Mellet, Sylvie & Monte, Michèle. 2009. *Juste/à peine* et la construction de la frontière notionnelle. *Cahiers de Praxématique* 53: 33–56. <https://doi.org/10.4000/praxematique.1085>
- Morel, Mary-Annick. 1994. PAS et NE... PAS en français oral. *Cahiers de Praxématique* 23: 97–116. <https://doi.org/10.4000/praxematique.66>
- Paradis, Carita. 2001. Adjectives and boundedness. *Cognitive Linguistics* 12: 47–64. <https://doi.org/10.1515/cogl.12.1.47>
- Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Salvan, Geneviève. 2014. *Juste la fin du monde*. L'excès juste, ou l'hyperbole exagère-t-elle toujours? *Travaux Neuchâtelois de Linguistique (Tranel)* 61–62: 63–78.
- Siouffi, Gilles, Steuckardt, Agnès & Wionet, Chantal. 2012. Comment enquêter sur les diachronies courtes et contemporaines? *Actes du 3^e CMLF*, 215–226. Lyon: EDP Sciences.

On how the distinction between reciprocal and collective verbs affects (anti-)control

Anna Snarska

State University of Applied Sciences, Włocławek

The phenomenon of Partial Control (henceforth PC), allegedly generating a non-exhaustive referential relation between a matrix argument and a silent subject, has enjoyed the status of a fully-fledged type of control since Landau (2000). The present paper aims to question its existence. Drawing upon the data from both English and Polish, I show how important it is for the theory of PC to properly distinguish between collective predicates and reciprocal verbs. I argue that both types of verbs invariably exhibit Exhaustive Control and the ostensible 'PC-effect' stems from the fact that the reciprocal verbs may co-occur with non-lexical 'with NP' phrase (discontinuous phrase). The data concerning Polish anti-control constructions supports the proposed analysis.

Keywords: partial control, reciprocal verbs, collectivity

1. Introduction

The claim that reciprocal and collective verbs represent two different verbal classes is not particularly easy to discern. On closer inspection, however, it turns out that their contrasting semantic exigencies generate evident changes in the construal of sentences they appear in. Unfortunately, this distinction has gone rather unnoticed in various theories of the phenomenon, where the split between reciprocal and collective predicates should be of great importance since the construction in question (as will be shown below) is most sensitive to the interpretation of number. The phenomenon referred to is the so-called Partial Control (henceforth PC),¹ the exemplification of which is seen in (1):

- (1) John₁ told Mary that he₁ wants [PRO₁₊ to meet in the morning].

1. Wurmbrand (2003) calls it imperfect control.

The name of the construction is self-explanatory – the matrix argument *John* only partially controls the reference of the silent subject (standardly marked as PRO in generative linguistics) in the non-finite clause.² Put differently, *John* is only one of the people that are about to meet and this is reflected in 1_+ notation on PRO. What is important is that this semantic effect can be produced only when the non-finite clause contains a collective predicate which is allegedly represented by, for example, *meet*.³

In order to better grasp the PC meaning, it is always useful to juxtapose this type of control with another, more ‘unexotic flavor’ of control, i.e. Exhaustive Control (henceforth EC):

- (2) John₁ wants [PRO₁ to die].

The relation that exists between *John*, the controller, and PRO, the controllee, leaves no doubt as to the interpretation of the non-finite subject – the reference is exhaustively imposed by *John*; hence, we find the same indices on the two categories.

For the first time brought into linguistic limelight in a seminal work by Landau (2000), PC has steadily gained ground in the last two decades, resulting in a proliferation of new analyses (Hornstein 2003; Jackendoff & Culicover 2003; Barrie & Pittman 2004; Bondaruk 2004; Dubinsky 2007; Dubinsky & Hamano 2007; Landau 2008; Rodrigues 2008; Boeckx, Hornstein & Nunes 2010; Landau 2016, among others).⁴ Inasmuch as they take a different stand on how PC is generated, they all share one view: PC does exist as a grammatical phenomenon and it is triggered by the presence of such collective verbs as *meet*, *congregate* or *assemble*.⁵ The aim of the present paper is to challenge this claim.

The polemic takes on the following structural shape. First, I discuss the difference between collective and reciprocal verbs, drawing upon linguistic evidence

2. I use PRO (Chomsky 1981) for purely expository reasons as most of the linguistic world is accustomed to this terminology. But I might as well use X in place of the non-finite subject.

3. It should also be noted that PC is very much context-sensitive and therefore it constitutes quite a serious challenge for any linguist. Throughout the paper I illustrate PC mainly with desiderative verbs (want, yearn, prefer, etc.) as these allow a relatively easy access to the ‘group’ interpretation. Nonetheless, whenever context is a strong factor at play, judging which sentence is acceptable and which is not will invariably produce conflicting views. Consequently, even describing the phenomenon is by no means an easy task, let alone its explicating.

4. Actually, it was Lawler (1972) who first noticed PC. However, major kudos undoubtedly go to Landau (2000) for marking a quantum leap in the study of this atypical species of control.

5. That *meet* is a collective verb is firmly established in Landau (2000), even though it is not entirely true. This theoretical imprecision may seem to be of minor importance, but it is not, as will be demonstrated in the forthcoming sections.

from both English and Polish. The English data is either consulted with seven native speakers of the language from the US or extracted from the Corpus of Contemporary American English. The Polish linguistic input comes from my own introspective judgments, further tested for acceptability with four other native speakers and consulted with the National Corpus of Polish. I show that *meet* is in fact an instance of the reciprocal verb and only *congregate*, *assemble* and *gather* can be classified as truly collective predicates.⁶ Undoubtedly, this ‘tiny’ change in the perception of the discussed categories has a bearing on the whole theory of PC as it calls into question the theoretical foundations of the module of grammar that was created specifically to handle this type of control.⁷ Should PC prove non-existent, it entails a conceptual redundancy in the linguistic world.⁸ And this is exactly what the present paper implies. The remainder of it features a proposal of how to account for the ‘more-than-one’ interpretation in non-finite clauses with both collective and reciprocal verbs. There is no need to postulate the existence of PC as a new subtype of control since the explication of the ‘₁₊ meaning’ can do without it. I argue that in a sentence with a reciprocal *meet* what we experience is a typical EC relation with a matrix argument fully controlling the reference of the non-finite subject PRO. The ostensible ‘PC-effect’ stems from the fact that *meet* takes a silent ‘with NP’ phrase which delivers an additional participant of the meeting. As regards truly collective predicates, the situation is rather straightforward inasmuch as they fail to license such an interpretation. The rationale behind it is that collective verbs are compatible only with semantically plural subjects and I contend that in the so-called PC constructions the non-finite subject is invariably semantically singular, pace Landau (2000). Finally, the paper concludes by referring to data provided by Polish anti-control constructions which are supposed to lend weight to the proposed solutions.

It is my hope that even if one harbors doubts as to the rightness of the above reasoning, the paper will at least challenge the descriptive correctness of PC.

6. For the purpose of this paper, I confine my attention solely to three examples of collective predicates as these instances are used in Landau (2000) to validate his claim about the existence of PC. Landau also mentions the adverb *together*, which, when added to a verb, transmutes it into a collective predicate. However, since this is not a lexical predicate (designated by just one word), I will leave such cases aside.

7. Interestingly enough, *meet* is the most often exploited example of a collective predicate in all the works on PC.

8. And the cognoscenti in this field know that the redundancy is really massive. To have an idea of it, it is enough to look into Landau’s (2000) meticulously constructed control machinery.

2. Collective predicates as opposed to inherently reciprocal verbs

A typical property of collective verbs is that they are congruent only with semantically plural subjects and are never true of the individual members of the group (Link 1983; Dowty 1987), as witnessed by the examples below:⁹

- (3) a. They all assembled/gathered/congregated in the morning.
 b. *Every person assembled/gathered/congregated in the morning.

As such, they will never allow distributive entailments (Champollion 2015) that warrant the presence of the conjunction:

- (4) a. John and Mary died. → John died and Mary died.
 b. John and the family gathered/congregated/assembled in this old church.
 → *John gathered in this old church and the family gathered in this old church. *John congregated in this old church and the family congregated in this old church. *John assembled in this old church and the family assembled in this old church.¹⁰

(4b) clearly demonstrates that *gather/congregate/assemble* cannot be true of *John* which is semantically singular, but the second part of the entailment, with a semantically plural *family*, becomes licit.¹¹

If collective predicates tolerate only semantic plurality of their subjects, this entails that they can also co-occur with semantically plural but syntactically singular NPs. That this is true indeed is confirmed by (5):

- (5) The family/team/committee assembled/gathered/congregated in the morning.

At first blush, *meet* seems to mimic the behavior of the verbs above in that it takes a semantically plural subject (6a), it is incompatible with a semantically singular subject (6b), which entails its intolerance of the conjunction in the distributive reading (6c), but it may co-exist with a syntactically singular subject, provided it is semantically plural (6d):

9. In fact, what constitutes a collective predicate and what does not is still a matter of debate. For an excellent review of the discussion, see Champollion (2015). The present paper relies on rather standard criteria for identifying collectivity.

10. It has to be mentioned that the sentence John and Mary gathered/congregated/assembled in this old church was unanimously judged by my informants as unacceptable, even though the subject is semantically plural. To them these predicates clearly require a subject with the 'more-than-two' meaning.

11. It has to be pointed out that the sentence John and Mary gathered cannot be analyzed, as it is unacceptable to the speakers of English.

- (6) a. They met in church.
 b. *John met in church.¹²
 c. John and Mary met. → *John met and Mary met.
 d. The family met in church.

The paradigm above indicates a unidirectional thinking – *meet* must be a collective predicate. However, what signals a need for a change of mindset is the fact that the verb is not obligatorily accompanied by a semantically plural subject (although it may take one). In such a case it is forced to take a ‘with NP’ phrase, i.e. a discontinuous phrase (Dimitriadis 2008; Siloni 2012):

- (7) John met with Mary in the morning.

Truly collective predicates can never afford this strategy:

- (8) *John gathered/congregated/assembled with Mary.¹³

These facts can be taken as an indication of a different status of *meet* from that of truly collective verbs. *Meet* is not a collective predicate (although it can disguise its identity when accompanied by a semantically plural verb, as seen in (6a) and (6c)), but an inherently reciprocal verb which invariably involves two participants/arguments playing the same role in the event.¹⁴ That reciprocal verbs express a symmetric relation is seen in (9):¹⁵

12. With Bowers (2008) finding the sentence acceptable, I decided to double-check it. In an extraction carried out on January 30th, 2018 from the Corpus of Contemporary American English (COCA), I did not find any instances of *meet* with a semantically singular subject (out of 300 results). Furthermore, my informants did not accept the sentence.

13. It should be noted, though, that these predicates can occur with ‘with NP’ phrase, however, in that case they necessarily take semantically plural subjects, as seen in (a). This inference is based on data extracted from the COCA on January 30th, 2018. In the study the sequence *gather/gathered/gathers* with occurred 244 times, the sequence *congregate/congregated/congregates* with occurred 10 times and the sequence *assemble/assembled/assembles* with occurred 96 times. With none of these predicates was a semantically singular subject used. Importantly, the ‘with NP’ phrase in such constructions is comitative and not discontinuous. The difference between the two is that the former is an adjunct, whereas the latter is an argument (see Dimitriadis 2004; Siloni 2012). This entails that the comitative phrase can be dropped, as shown by (b) below:

- a. They congregated/assembled/gathered with Mary in this old church.
 b. They congregated/assembled/gathered in this old church.

14. Other reciprocal verbs include *kiss*, *argue*, *marry* or *hug*. However, their analysis goes beyond the scope of the present study which focuses mainly on *meet*, erroneously referred to by the proponents of the existence of PC as a collective verb.

15. But see Kruitwagen et al. (2017) who cast doubt on this assumption.

- (9) a. John met with Mary in the morning.
 b. Mary met with John in the morning.

As might be expected, symmetry entails that if we change the position of the two arguments, truth values of the sentence will be preserved (Dimitriadis 2008).

Summing it up so far, we have just seen that *congregate/assemble/gather*, on one hand, and *meet*, on the other hand, represent two different types of verbs. The former, being collective predicates, disallow the discontinuous phrase, whereas the latter, since it represents a reciprocal verb, requires it.

The analyzed linguistic data pertains to English. Let us now examine Polish empirical facts in the same respect.

The Polish equivalent of the English *meet* is *spotykać się* (with *się* being a reflexive pronoun). The verb, just like in English, at first seems to exemplify a collective predicate, as shown by the following sentences:

- (10) a. *Oni spotkali się po południu.*
 They met REFL after noon
 ‘They met in the afternoon.’
 b. **Janek spotkał się po południu.*
 John met REFL after noon.
 ‘John met in the afternoon.’
 c. *Jan i Maria spotkali się. *Jan spotkał się i Maria*
 John and Mary met REFL → John met REFL and Mary
spotkała się.
 met REFL
 ‘John and Mary met. → *John met and Mary met.’
 d. *Cała drużyna spotkała się w kościele.*
 whole team met REFL in church
 ‘The whole team met in church.’

(10a) shows that the predicate is compatible with semantically plural subjects, in (10b) it cannot accept the subject which is semantically singular, the distributive reading is clearly not an option in (10c) and the final example presents the verb’s tolerance of semantically plural nouns.

However, similarly to its English counterpart, *spotykać się* has to be obligatorily followed by a discontinuous phrase (‘with NP’ phrase), when accompanied by a semantically singular subject. This indicates that it is an inherently reciprocal verb:

- (11) *Janek spotkał się z Marysią.*
 John met REFL with Mary
 ‘John met with Mary.’

Since the two arguments of the verb are in a symmetric relation, they can easily change their position, carrying the same truth values of the sentence:

- (12) *Marysia spotkała się z Jankiem.*
 Mary met REFL with John
 ‘Mary met with John.’

With respect to *gromadzić się* and *zbierać się*, the corresponding Polish expressions of *congregate*, *gather*, *assemble*, the examples below clearly reflect their status as collective predicates:

- (13) a. *Dziesięć dziewczyn zgromadziło się/zebrało się w kościele.*
 Ten girls gathered REFL/assembled REFL in church
 ‘Ten girls gathered/assembled in church.’
- b. **Każda dziewczyna zebrała się/zgromadziła się w kościele.*
 Every girl gathered REFL/assembled REFL in church
 ‘Every girl gathered/assembled/congregated in church.’
- c. *Janek z drużyną zgromadził się/zebrał się w kościele.*
 John with team gathered REFL/assembled REFL in church
 ‘John and the team gathered/assembled in church.’
- d. **Janek zebrał się/zgromadził się w kościele i drużyna zebrała się/zgromadziła się w kościele.*
 John gathered REFL/assembled REFL in church and team gathered REFL/assembled REFL in church
 ‘John gathered/assembled in church and the team gathered/assembled in church.’
- e. *Drużyna/rodzina zebrała się/zgromadziła się po południu.*
 team/family gathered REFL/assembled REFL after noon
 ‘The team/family gathered/assembled in the afternoon.’

Starting from the very top, (13a) features collective predicates with a semantically plural subject, in (13b) the same verbs cannot co-occur with a subject that is semantically singular, (13c–d) show that the distributive interpretation, facilitated by the presence of the conjunction *and* in (13d), is unavailable to the collective predicates¹⁶ and in the final Example (13e) there is a licit co-existence of the semantically plural subject and collective verb.

All in all, there is no doubt that English and Polish are alike in that they both clearly distinguish between truly collective predicates and reciprocal verbs. Additionally, what is important to us, the two languages deem *meet/spotykać się* a reciprocal verb.

16. However, note that the distributive entailment for the collective subject *team* in (13d) is sanctioned. Obviously, it reflects the fact that the subject is semantically plural. The sentence in (13c) was chosen instead of *Janek z Marią zgromadził się w kościele* ‘John and Mary gathered in church’ since the latter is unacceptable in Polish. English is no different in this respect (see (4b)).

Armed with this assumption, we are now in a position to offer a fresh perspective on the (alleged) existence of PC.

3. The ‘PC-might-be-a-ghost’ approach

The proposal that I would like to put forward boils down to the following claim: PC does not exist. Since this can hardly be satisfactory, let us then go deeper into the question of what exactly we are dealing with in non-finite clauses with phonetically unrealized 1_+ subjects. Two scenarios need to be given due consideration.

As far as the structures with reciprocal *meet* are concerned, what one witnesses in (1), repeated here for convenience in (14), is a typical EC, where the matrix argument fully controls the reference of PRO:

- (14) John₁ wants [PRO₁ to meet in the morning].

The ‘more-than-one’ meaning of PRO stems from the fact that the reciprocal co-occurs with a discontinuous ‘*with NP*’ phrase which, in such sentences, undergoes ellipsis:

- (15) John₁ wants [PRO₁ to meet ~~with somebody~~ in the morning].

The deleted phrase constitutes a missing participant of the prospective meeting, apart from *John*, obviously.

The same reasoning can be applied to cases with Polish *spotykać się* ‘meet’:

- (16) *Janek₁ zamierza [PRO₁ spotkać się po południu].*
 John intends to-meet REFL after noon
 ‘John intends to meet in the afternoon.’

The infinitive is also followed by the phonetically deleted ‘*with NP*’ phrase which, nonetheless, keeps its full semantic status, delivering a partaker of the act of the meeting:

- (17) *Janek₁ zamierza [PRO₁ spotkać się z kimś po południu].*
 John intends to-meet REFL with somebody after noon
 ‘John intends to meet in the afternoon.’

With respect to the other scenario referring to collective verbs, the situation is far simpler, as these predicates do not license a subject marked for semantic singularity. What the present paper postulates is that PRO in the so-called PC environments is semantically singular, not plural. In other words, this is an instance of EC, where the matrix argument has a full referential power over PRO. Therefore, collective verbs are not able to legitimize such a subject, as shown by (18):

(18) *John₁ wanted [PRO₁ to assemble/congregate/gather at six].

This incapability is confirmed by instances in Polish:

(19) **Janek₁ chciał [się PRO₁ zgrupować/zebrać w kościele].*
 John wanted REFL to-congregate/to-assemble/to-gather in church
 ‘John wanted to congregate/assemble/gather in church.’

However, if *John* in the examples above is replaced with a semantically plural subject, the sentence immediately becomes acceptable in both languages:

(20) a. The ladies₁ wanted [PRO₁ to assemble/congregate/gather at six].
 b. *Panie₁ chciały [się PRO₁ zgrupować/zebrać w kościele].*
 They wanted REFL to-congregate/to-assemble/to-gather
 in church
 ‘The ladies wanted to congregate/assemble/gather in church.’

Needless to say, the constructions are still an instance of EC and not PC. The ‘more-than-one’ meaning of PRO stems from the fact that this time it is controlled by a semantically plural subject. The relation makes it possible for the non-lexical subject to co-exist with the embedded predicate.¹⁷

In conclusion, the claim that PC is non-existent does seem to stand up to scrutiny when juxtaposed against both English and Polish data.¹⁸

17. The two sentences can be ambiguous, though. The interpretation discussed covers only one of the two accessible scenarios. The other context assumes that the ladies wanted to assemble/gather/congregate with some other (non-lexical) individuals salient in the context. Under this interpretation PRO is still under control of the matrix subject (EC), with other partakers of the gathering/assembly/congregation being licensed by pragmatics.

18. The present paper concentrates on the infinitival complements of verbs but it is interesting to observe that the reasoning adopted here may be equally well applied to infinitival complements of adjectives and nouns, as in (a) and (b), respectively:

- a. John₁ is willing [PRO₁ to meet with somebody in the morning].
 b. John’s₁ willingness [PRO₁ to meet with somebody in this old church]

Unsurprisingly, Polish is no different in this respect:

- c. *Marysia₁ jest skłonna [PRO₁ spotkać się z kimś po południu].*
 Mary is inclined to-meet REFL with somebody after noon
 ‘Mary is inclined to meet in the afternoon.’
 d. *Marysia₁ ma ochotę [PRO₁ spotkać się z kimś po południu].*
 Mary has desire to-meet REFL with somebody after noon.
 ‘Mary feels like meeting in the afternoon.’

Furthermore, as expected, collective verbs are not tolerated in such environments both in English, as in (e-f), and Polish (g-h):

Let us now turn our attention to Polish anti-control constructions so that the proposal made here could gain even more ground.

4. The ‘more-than-one’ meaning of PRO in Polish anti-control constructions

Polish, in general, has two ways of introducing infinitival complement clauses. It may either embed them as bare clauses, as shown by (21):

- (21) *Marysia wolala iść do kina.*
 Mary preferred to-go to cinema
 ‘Mary preferred to go to the cinema.’

or have them occur with the complementizer *żeby* ‘so that’:

- (22) *Marysia chce, żeby odwiedzić Janka.*
 Mary wants so-that to-visit John
 ‘Mary wants for John to be visited.’

The former case will invariably feature PRO as being co-referential with the matrix argument, a typical instance of EC:

- (23) *Marysia₁ wolala [PRO₁ iść do kina].*
 Mary preferred to-go to cinema
 ‘Mary preferred to go to the cinema.’

However, surprisingly, whenever the complementizer appears, it forces referential disjointness of PRO from the matrix argument:

- (24) *Marysia₁ chce, [żeby PRO₂ odwiedzić Janka].*
 Mary wants so-that to-visit John
 ‘Mary wants for others to visit John.’

-
- e. *John₁ is willing [PRO₁ to congregate/assemble/gather in the morning].
 f. *John’s₁ willingness [PRO₁ to congregate/assemble/gather in this old church]
 g. **Marysia₁ jest skłonna [PRO₁ zgromadzić się/zebrać się po południu].*
 Mary is inclined to-gather REFL/to-congregate REFL
 after noon
 ‘Mary is inclined to gather/congregate/assemble in the afternoon.’
 h. **Marysia₁ ma ochotę [PRO₁ zgromadzić się/zebrać się po południu].*
 Mary has desire to-gather REFL/to-congregate REFL after noon.
 ‘Mary feels like gathering/congregating/assembling in the afternoon.’

In case we remove *żeby* ‘so that’ from the sentence, EC is immediately restored.¹⁹

- (25) *Marysia₁ chce, [PRO₁ odwiedzić Janka].*
 Mary wants to-visit John
 ‘Mary wants to visit John.’

In light of these facts, the phenomenon observed in (24) has been dubbed anti-control (Brandt et al. 2016) or obviation (cf. Farkas 1992; Dobrovie-Sorin 2001).²⁰

Let us then check how PRO is interpreted when inserted in *żeby*-complements with reciprocal and collective verbs:²¹

- (26) a. *Janek₁ chce, [żeby PRO₂ spotkać się po południu].*
 John wants so-that to-meet REFL after noon.
 ‘John wants for others to meet in the afternoon.’
 b. *Janek₁ wolał, [żeby PRO₂ zgromadzić się/zebrać się w tym starym kościele].*
 John preferred so-that to-gather REFL/assemble REFL in
 this old church
 ‘John preferred for others to gather/congregate in this old church.’

Interestingly, both examples validate the ‘more-than-one’ interpretation of PRO. As regards (26a), the non-lexical subject is controlled by contextually salient individuals that are not present in the sentence. And here two settings have to be considered. Under the first scenario, PRO may stand for a semantically plural subject (e.g. they, the family, the team, etc.), in which case it need not take a discontinuous phrase since the semantic plurality of the subject is enough. The other scenario assumes that PRO is semantically singular (e.g. maybe it is Mary who is about to meet with somebody); yet, the missing participant of the meeting is delivered by the silent discontinuous phrase:

19. Not all verbs in Polish subcategorize for both non-finite *żeby*-complements and complementizer-free complements. This paper focuses on those that do. For a more exhaustive classification of the Polish control verbs, see Bondaruk (2004: 203).

20. However, despite blocking co-reference between PRO and the matrix subject (subject control), *żeby* ‘so that’ does not affect object control. Yet, the latter is beyond the scope of this paper, so I leave it for future research.

21. At this point it should be noted that *żeby*-complements with *spotykać się* ‘meet’ and *zebrać się/zgromadzić się* ‘to gather/congregate/assemble’ are extremely rare to find in the National Corpus of Polish. The extraction I carried out on 30th January, 2018 produced only 2 results for the former and 0 results for the latter. However, the examples presented here have been tested for acceptability with native speakers of Polish.

- (27) *Janek₁ chce, [żeby PRO₂ spotkać się z kimś po południu].*
 John wants so-that to-meet REFL with somebody after noon.
 ‘John wants for others to meet in the afternoon.’

In either case we are still dealing with EC, much as the control is imposed by the non-lexical entity. In either case there is no need to resort to PC.

As far as anti-control with collective verbs in (26b) is concerned, one has to recall that the same structure but without *żeby* ‘so that’ is unacceptable in (19), repeated here for convenience in (28):

- (28) **Janek₁ chciał [się PRO₁ zgromadzić/zebrać w kościele].*
 John wanted REFL to-congregate/to-assemble/to-gather in church
 ‘John wanted to congregate/assemble/gather in church.’

Surprisingly, the anti-control setting eliminates this effect as the sentence in (26b) is a licit sentence. The reasoning behind it is that when the complementizer forces referential disjointness of PRO from the matrix subject, nothing can prevent the former from taking a semantically plural controller. Undoubtedly, it is still EC, but with a non-lexical controller designated by pragmatics.^{22 23}

Overall, the anti-control constructions with collective verbs and reciprocal *spotykać się* ‘meet’ indicate that the idea entertained in this paper remains a viable option for those that question the existence of PC.

22. An anonymous reviewer asks if this control relation is similar to arbitrary control. In arbitrary control there is no argumental controller irrespective of whether the context is linguistic or pragmatic. In other words, anyone in general can control reference of the silent subject. Yet, in (26b), although the controller is non-lexical, it can easily be identified by the pragmatics of the situation and as such, it will exhaustively (fully) control the reference of PRO. Hence, I treat it as a case of EC.

23. An anonymous reviewer notices that an English ECM construction with want, e.g. I want her to go, may also be taken to illustrate anti-control. What is interesting (and promising), this example supports the analysis presented in the present paper. Embedding meet/congregate/assemble/gather in this linguistic context produces the following results:

- a. I want her to meet in this old church.
- b. *I want her to congregate/assemble/gather in this old church.

In both sentences the object her will always strongly impose its referential power on the non-lexical subject in the embedded clause; hence, the subject cannot be co-referential with the matrix subject. Since the object is semantically singular, the embedded subject must also be marked by the same number. Sentence (a) is acceptable inasmuch as additional participants of the meeting are provided by a discontinuous silent phrase following meet. The other construction, on the other hand, produces a semantic clash, as the collective predicates require a semantically plural subject and her fails to provide the required feature (unfortunately, collective predicates do not take discontinuous phrases).

5. Conclusion

To recapitulate, the empirical evidence presented points towards the conclusion that there are strong grounds for doubting the descriptive adequacy of a phenomenon that has been wrongly, as this paper argues, named PC. The contention that PC accompanies collective verbs, including *meet*, does not seem to hold water as it has been proven that in PC contexts *meet* is in fact a reciprocal verb taking a null discontinuous phrase and truly collective verbs do not license PC. With both types of verbs it is EC that appears. The Polish data firmly supports these findings.²⁴ Accordingly, it seems pointless and theoretically costly to keep the idea of having PC alive. It is only natural then that for a credible analysis of the so-called PC to emerge, we must examine more data from different languages and explore new ways of thinking about the phenomenon. This paper was meant as only one of such explorations.

Acknowledgements

I am most grateful to two anonymous reviewers for their valuable comments and criticisms that have contributed greatly to improving the final version of the present paper.

References

- Barrie, Michael & Pittman, Christine M. 2004. Partial control and the movement towards movement. *Toronto Working Papers in Linguistics* 22: 75–92.
- Boeckx, Cedric, Hornstein, Norbert & Nunes, Jairo. 2010. *Control as Movement*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511761997>
- Brandt, Patrick, Trawiński, Beata & Wollstein, Angelika. 2016. (Anti-)Control in German: Evidence from comparative, corpus- and psycholinguistic studies. In *Co- and Subordination in German and Other Languages*, Ingo Reich & Augustin Speyer (eds), 77–98. Hamburg: Helmut Buske.
- Bondaruk, Anna. 2004. *PRO and Control in English, Irish and Polish: A Minimalist Analysis*. Lublin: KUL.
- Bowers, John. 2008. On reducing obligatory control to movement. *Syntax* 11(2): 125–143. <https://doi.org/10.1111/j.1467-9612.2008.00103.x>

24. I am most grateful to the anonymous reviewer for pointing out that Greek shows medio-passive morphology on reciprocal *meet* and that the analysis presented here holds for Greek as well. However, since Greek has no infinitives and I am not an expert on the language, I leave this issue for further exploration.

- Champollion, Lucas. 2015. Distributivity, collectivity and cumulativity. In *Companion to Semantics*, Lisa Matthewson, Cecile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds). Hoboken NJ: Wiley & Sons. <<https://ling.auf.net/lingbuzz/002133>> (4 december 2020).
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Kluwer.
- Corpus of Contemporary American English*. <<https://www.english-corpora.org/coca>> (30 January 2018).
- Dimitriadis, Alexis. 2004. *Discontinuous reciprocals*. Ms, Utrecht Institute of Linguistics.
- Dimitriadis, Alexis. 2008. Irreducible symmetry in reciprocal constructions. In *Reciprocals and Reflexives: Cross-linguistic and Theoretical Explorations*, Ekkehard König & Volker Gast (eds), 375–410. Berlin: De Gruyter.
- Dobrovie-Sorin, Carmen. 2001. Head-to-head merge in Balkan subjunctives and locality. In *Comparative Syntax of Balkan Languages*, Maria Luisa Rivero & Angela Ralli (eds), 44–73. Oxford: OUP.
- Dowty, Dawid R. 1987. Collective predicates, distributive predicates and “all”. In *Proceedings of the 1986 Eastern States Conference on Linguistics (ESCOL)*, Fred Marshall (ed.), 97–116. Columbus OH: Ohio State University.
- Dubinsky, Stanley. 2007. *On the syntax of exhaustive control and the calculus of events*. Ms, University of South Carolina.
- Dubinsky, Stanley & Hamano, Shoko. 2007. A window into the syntax of control: Event opacity in Japanese and English. *University of Maryland Working Papers in Linguistics* 15: 74–79.
- Farkas, Donka F. 1992. On obviation. In *Lexical Matters*, Ivan Sag & Anna Szabolcsi (eds), 85–109. Stanford CA: CSLI.
- Hornstein, Norbert. 2003. On control. In *Minimalist Syntax*, Randall Hendrick (ed.), 6–81. Oxford: Blackwell. <https://doi.org/10.1002/9780470758342.ch1>
- Jackendoff, Ray & Culicover, Peter W. 2003. The semantic basis of control in English. *Language* 79(3): 517–556. <https://doi.org/10.1353/lan.2003.0166>
- Kruitwagen, Imke, Poortman, Eva B. & Winter, Yoad. 2017. Reciprocal verbs as collective predicate concepts. In *Proceedings of the Forty-Seventh Annual Meeting of the North East Linguistic Society (NELS 47)*, Andrew Lamont & Katerina Tetzloff (eds), 201–210. Amherst MA: GLSA.
- Landau, Idan. 2000. *Elements of Control. Structure and Meaning in Infinitival Constructions*. Dordrecht: Kluwer.
- Landau, Idan. 2008. Two routes of control: Evidence from case transmission in Russian. *Natural Language and Linguistic Theory* 26: 877–924. <https://doi.org/10.1007/s11049-008-9054-0>
- Landau, Idan. 2016. Agreement at PF: An argument from partial control. *Syntax* 19: 79–109. <https://doi.org/10.1111/synt.12118>
- Lawler, John. 1972. *A Problem in Participatory Democracy*. Bloomington IN: Indiana University Linguistics Club.
- National Corpus of Polish*. <<http://www.nkjp.pl>> (30 January 2018).
- Rodrigues, Cilene. 2008. Agreement and flotation in partial and inverse partial control configurations. In *New Horizons in the Analysis of Control and Raising*, William D. Davies & Stanley Dubinsky (eds), 213–230. Dordrecht: Springer.
- Siloni, Tal. 2012. Reciprocal verbs and symmetry. *Natural Language and Linguistic Theory* 30: 261–320. <https://doi.org/10.1007/s11049-011-9144-2>
- Wurmbrand, Susanne. 2003. *Infinitives: Restructuring and Clause Structure*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110908329>

The rise of cause/reason adverbial markers in Yaqui (Uto-Aztecan)

Albert Álvarez González

University of Sonora

This paper deals with three different connectives (*betchi'ibo*, *bwe'ituk*, *po(r)ke*) in Yaqui, which are used to introduce a clause that provides the cause/reason for which the situation denoted in the associated clause is carried out. The goal of this paper is to describe the morphosyntactic diversity present in Yaqui cause/reason adverbial clauses and to explain this diversity considering the evolutionary paths that have originated the three different cause/reason markers in Yaqui. These three markers are associated with three different sources: *betchi'ibo* originates from a postpositional marker that extended its use from noun phrases to clausal constituents; *bwe'ituk* comes from discourse via the recruitment of a discourse marker of discontinuity (the discourse marker *bwe*); and *po(r)ke* arises from language contact with Spanish.

Keywords: cause/reason adverbial clauses, discourse marker, postposition, switch-reference system, syntactic calque

1. Introduction

This study is concerned with three different forms (*betchi'ibo*, *bwe'ituk*, *po(r)ke*) in Yaqui (Uto-Aztecan), which are used to introduce a clause that provides the cause/reason¹ for which the situation denoted in the associated clause is carried out. Interestingly, the three causal adverbial clauses marked by these three different markers exhibit different internal structures, in particular regarding the coding of

1. A difference is often made between cause and reason, according to whether the situation denoted in the main clause is carried out intentionally (the situation denoted in the adverbial clause is thus a reason to act) or not intentionally (the situation denoted in the adverbial clause is thus a cause) (Vester 1983: 45; Luraghi 1989: 295; Givón 2001: 335). However, since Yaqui does not make a morphosyntactic distinction between cause and reason, I have conflated both under the label cause/reason.

the subject of the adverbial clause and the position of the causal marker. The goal of this paper is thus to describe the morphosyntactic diversity present in Yaqui cause/reason adverbial clauses and to explain this diversity considering the evolutionary paths that have given rise to the three different cause/reason markers in Yaqui. It will be proposed that these three markers are associated with three different sources: (i) The marker *betchi'ibo* is originated from a postpositional marker that extended its use from noun phrases to clausal constituents. (ii) The marker *bwe'ituk* is originated from discourse, via the recruitment of a discourse marker (the particle *bwe*). (iii) The marker *po(r)ke* is originated from language contact with Spanish. The paper is thus interested in the genesis of syntactic complexity and the development of clausal integration in the domain of cause/reason adverbial clauses in Yaqui, illustrating three different ways for creating an interclausal connective of cause/reason, and trying to show how the internal structures identified in synchrony are fashioned by the origins of the new adverbial markers and by the diachronic processes derived from these sources. Although this paper is focused on the morphosyntax of causal adverbial clauses in Yaqui, and clearly does not pretend to offer a detailed and plausible semantic/pragmatic account of communicating interclausal causality in Yaqui, it will also be shown that the origin of each causal marker determines in some way the semantic/pragmatic differences between the three markers.

This paper is organized as follows. In Section 2, I provide general information about Yaqui and the data used in this study. Section 3 presents the different cause/reason adverbial clauses that can be used in present-day Yaqui, whereas Section 4 shows the old cause/reason adverbial clauses documented in colonial times. The comparison of these past and present-day constructions reveals that important changes have been introduced in the expression of interclausal causality, in particular the switch-reference system associated with past strategies has been lost and new constructions have been developed with the use of new and more explicit cause/reason adverbial markers. In Section 5, I explore the origin of these new markers as well as the evolutionary paths that have originated the present-day cause/reason adverbial clauses, trying to show how the diachrony of cause/reason adverbial clauses can explain the differences in the internal structures identified in synchrony, specially regarding the coding of the dependent clause subject and the position of the new cause/reason adverbial markers. Lastly, the final remarks 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 (USA) by approximately 500.² This language belongs to the Taracahitan branch of the Sonoran group within the Southern Uto-Aztecan languages. Table 1 presents the different languages belonging to this sub-part of the Uto-Aztecan family.

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 Cahita³ 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 consider them as three varieties of the same language: the Cahita language, as named in the *Arte de la lengua cahita*, first available description of this linguistic group elaborated around 1630 (Álvarez 2018, see Section 4).

2. 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.

3. This exonymic glossonym comes from the word *kaita* meaning ‘nothing’.

Yaqui⁴ is an agglutinative language with a 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	<i>inepo</i> , = <i>ne</i>	<i>nee</i>	<i>in</i> , <i>nim</i>
2SG	<i>empo</i> , =' <i>e</i>	<i>enchi</i>	<i>em</i>
3SG	<i>aapo</i>	<i>aapo'ik</i> , <i>a=</i>	<i>aapo'ik</i> , <i>a=</i>
1PL	<i>itepo</i> , = <i>te</i>	<i>itom</i>	<i>itom</i>
2PL	<i>eme'e</i> , =' <i>em</i>	<i>enchim</i>	<i>em</i> , <i>enchim</i>
3PL	<i>bempo</i> , = <i>mme</i>	<i>aapo'im</i> , <i>am</i>	<i>bem</i> , <i>bempo'im</i>

Examples (1) and (2) show that the basic order in the transitive construction is SOV. They also show that 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), due to the differential object marking existing in Yaqui. 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 17th century by an anonymous Jesuit (probably the priest missionary Tommaso Basilio, Álvarez 2018: 222–224) and edited by Eustaquio Buelna in 1890, and from the letters written by the Yaqui leader Juan Banderas 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).

4. Johnson (1962), Lindenfeld (1973), and Dedrick and Casad (1999) have proposed general descriptions of Yaqui grammar.

3. Cause/reason adverbial clauses in Modern Yaqui

3.1 *betchi'ibo*

One possibility to express a cause/reason adverbial clause in Yaqui is using the form *betchi'ibo*, both for same-subject situations as in (3) and for different-subject situations as in (4).

- (3) *Wikosa-ta ne baajta-k [in ousi jibwa-ka betchi'ibo]*⁵
 belt-ACC 1SG.NOM loosen-PFV 1SG.POSS a_lot eat-PFV BECAUSE
 'I loosened my belt because I ate a lot.'
- (4) *In maala bwaana [in sim-bae betchi'ibo]*
 1SG.POSS mother cry 1SG.POSS go.SG-DES BECAUSE
 'My mother is crying because I am going to leave.'

In these examples, we see that the adverbial connective *betchi'ibo* is located in the adverbial clause-final position, on the right side of the adverbial clause. Considering the coding of the arguments in the dependent clause, there is no switch-reference system associated with the use of *betchi'ibo*, in the sense that its presence is not conditioned by the fact that the subject of the adverbial clause is identical or distinct to the subject of the main clause. In both cases, the connective *betchi'ibo* is present. The subject of the adverbial clause is in a possessive form, implying that the clause introduced by *betchi'ibo* is deranked (Stassen 1985), that is, the structure is different to the structure of the main clause, since the subject is not in the nominative form. As for the verb of the adverbial clause, it is finite (marked as perfective in (3) and as desiderative/near future in (4)).

3.2 *bwe'ituk*

Another possibility for expressing cause/reason adverbial clauses in Yaqui implies the use of *bwe'ituk* as exemplified in (5) and (6).

- (5) *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.'

5. In the different examples provided in this paper, cause/reason adverbial clauses will be put in brackets.

- (6) *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.'

Comparing with cause/reason adverbial clauses marked by *betchi'ibo*, we can see that the adverbial connective *bwe'ituk* is located in the adverbial clause-initial position, on the left side of the adverbial clause. As in the case of *betchi'ibo*, there is no switch-reference system associated with the use of *bwe'ituk*, but in this case, the adverbial clause introduced by *bwe'ituk* is balanced (Stassen 1985), 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 can receive the same Tense-Aspect-Mood marking than with *betchi'ibo* (desiderative/near future in (5), perfective in (6)).

3.3 *po(r)ke*

The third possibility for introducing a cause/reason adverbial clause in Yaqui corresponds to the use of *po(r)ke*, as illustrated in (7) and (8).

- (7) *Ne kaita nooka [porke ne lotti-la]*
 1SG.NOM nothing speak BECAUSE 1SG.NOM get_tired-RES
 'I don't say anything because I am tired.'
- (8) *Aapo sii-ka [poke empo a'abo noite-k]*
 3SG.NOM go.SG-PFV BECAUSE 2SG.NOM here come-PFV
 'He left because you came.'

We can observe that there is no morphosyntactic differences between cause/reason adverbial clauses introduced by *bwe'ituk* and those introduced by *po(r)ke*. In both cases, the causal connective is in the adverbial clause-initial position, the adverbial clause is balanced, the verb in the dependent clause is a finite verb, and no switch-reference system is present.

4. Cause/reason adverbial clauses in Old Cahita

As mentioned above, 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*, from the early 17th century. 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 shall refer to it here as Old Cahita.⁶

4.1 Cause/reason adverbial clauses in Old Cahita

Almost all adverbial clauses in Old Cahita exhibited a switch-reference system. From (9) to (14), the constructions provided in Buelna (1890: 72) for illustrating the cause/reason adverbial clauses from Old Cahita are presented.

Same Subject

- (9) *Emchi ne noctehoa, [emchi eria teca]*
 2SG.ACC 1SG.NOM teach 2SG.ACC love SUB
 ‘Porque te amo, te enseño’⁷
 ‘I teach you because I love you’
- (10) *Emchi ne viiu-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.’
- (11) *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 shall whip you because I shall love you.’

6. 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 clauses in Yaqui.

7. 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 and that mainly come from narrative texts included in Silva et al. (1998), Estrada et al. (2004) or Buitimea (2007).

Different subject

- (12) *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.’
- (13) *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.’
- (14) *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 shall 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 temporal-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 9) and future situations (Example 11)
- *ituca* was also used in present (Example 12) and future situations (Example 14) but in the cases of no 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 13)
- *tuca* was used in co-referential and past situations (Example 10) but it could also be used instead of *tuco*, for no co-referential and past situations (Example 13).

If we compare the cause/reason adverbial clauses illustrated in (9)–(14) from Old Cahita with the cause/reason adverbial clauses of Modern Yaqui exemplified from (3) to (8), we can observe several interesting differences:

- In Old Cahita, there was a system of switch-reference, which does not exist 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, since *betchi’ibo*, *bwe’ituk* and *po(r)ke* are used for both same-subject (SS) and different-subjects (DS) situations.
- This old switch-reference system was accompanied by temporal-aspectual distinctions. Different markers were used 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 no longer exists as the connectives *betchi'ibo*, *bwe'ituk* and *por(r)ke* can be used for all temporal-aspectual situations. In that respect, it can be observed that the neutralization of the switch reference system was apparently incipient in Old Cahita, since *tuca* could be used in past situations for both SS and DS.⁸

- The old cause/reason adverbial clause exhibited a higher degree of syntactic integration in relation to the main clause. The old cause/reason adverbial clauses were deranked (Stassen 1985), like present-day cause/reason adverbial clauses introduced by *betchi'ibo*, however the coding of the arguments of the adverbial clause was different. Contrary to *betchi'ibo* cause/reason adverbial clauses in which the subject is expressed as possessive in cases of subject coreferentiality (3) and subject non-coreferentiality (4), the old cause/reason adverbial clauses used zero anaphora for same subjects, and accusative marking for different subjects. As for cause/reason adverbial clauses introduced by *bwe'ituk* and *po(r)ke*, the subject of the adverbial clause is in the nominative case (Examples (5), (6), and (7), (8), respectively) for both, same subjects and different subjects.
- The old adverbial connectives (*teca*, *tuca*, *tuco*, *ituca*) were in the adverbial clause-final position, like in the case of present-day cause/reason adverbial clauses introduced by *betchi'ibo* and unlike the adverbial connectives *bwe'ituk* and *po(r)ke*, which are in the adverbial clause-initial position.

4.2 The multifunctionality of *teca* in Old Cahita

The final particle *teca* was not only used to mark cause/reason clauses in Old 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 in (15) and (16), purpose clauses in (17) and (18), and temporal posteriority clauses ('before' clause) in (19), always in same-subject situations.

Conditional clause in Old Cahita (Buelna 1890: 66–68)

- (15) *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'

8. 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 Old Cahita exhibited a switch-reference system, which has been completely neutralized in Modern Yaqui, except for temporal adverbial clauses, in which the switch reference is still observable (*-ka(i)* 'when, same subject', *-(k)o* 'when, different subject'), even though not always in a systematic way (Álvarez 2009; Guerrero 2019).

- (16) [*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’

Purpose clause in Old Cahita (Buelna 1890: 82)

- (17) *uaquim ne iepsa-c [misa-ta bit-naque teca]*
 here 1SG.NOM arrive-PFV mass-ACC look-FUT SUB
 ‘He llegado aquí á oír misa’
 ‘I have arrived here to hear the mass’
- (18) *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’
- (19) *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 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 specific 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 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,⁹ 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) in the absence of temporal co-referentiality between both clause situations (that is, there is no future marking in the main clause, as in (17) and (18)) or anterior to the main clause situation (cause/reason interpretation) in the case of temporal co-referentiality between both clause situations, (both the main clause and the adverbial clause exhibit the future suffix as in (11) and (14)).

9. In the case of ‘before’ adverbial clauses, the dependent clause situation is posterior to the main clause situation but this temporal interpretation of posteriority is explicitly expressed by the temporal adverbial marker *quehe* ‘not yet’ located in the initial position within the adverbial clause in (19). The presence of this temporal adverb indicates that the use of *teca* by itself implies the anteriority of the dependent clause situation, which is, in this case, canceled and changed to posteriority by means of *quehe*.

This sequentiality may be then literally interpreted as temporal 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 and 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 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 under-specification 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 (16) 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 with the dependent verb (if we consider the transcription used in the *Arte*)¹⁰ and, more importantly, its final vowel *a* is elided, something that is very pervasive in Old 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

10. Nowadays in Yaqui, the suffix *-tek(o)* is used to mark conditional adverbial clauses and it is clearly a bound morpheme attached to the dependent verb. The same can be said for another temporal adverbial clause marker of Modern Yaqui, the perfect participial suffix *-taka(i)* (or ‘anterior converb’ marker, following the terminology proposed by Haspelmath 1995).

clauses in Old 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 (10) and (13) the particle changes to *tuca*).

Based on these features and on the fact that the last relics of the old switch-reference system are found nowadays in the temporal adverbial marking (*-ka(i)* ‘when, same subject’ vs *-(k)o* ‘when, different subject’ (see footnote 8), in Álvarez (2009) I have proposed that the markers of cause/reason adverbial clauses in Old Yaqui are probably resulting from a bimorphemic combination in which, originally, suffixes *-te* and *-tu* were associated with the temporal marking present in the two clauses being connected (*-te* ‘non-past’ vs. *-tu* ‘past’), and suffixes *-ko* and *-ka* were temporal markers of switch reference (*-ko* ‘when, different subject’ vs. *-ka* ‘when, same subject’), that is, dependent clause markers used for indicating referential coherence between the situation denoted by the main clause, and the dependent clause situation, which is temporally related to the main situation. This marking system of clausal dependency would thus combine a marker of temporal (co)-referentiality between the situations denoted by the clauses being connected (*-te*, *-tu*) and a marker of participant (co)-referentiality between the subjects of these clauses (*-ko*, *-ka*). As we can see above from the examples of old cause/reason adverbial clauses (from (9) to (14)), these distinctions between *-te* and *-tu* on one hand, and between *-ko* and *-ka* on the other hand, were already becoming blurred in the cause/reason adverbial clause of Old Cahita, even though the switch-reference system was still in place.

The combination of those markers was thus used for backgrounding the dependent clause situation to the main clause situation, and for linking both clause situations in terms of consecutiveness, sequentiality. The activation of pragmatic inferences from context was needed for the interpretation of the specific meaning relation (temporality, causality, conditionality, purpose) between the two situations.

In sum, the comparison between the past strategies used in Cahita for cause/reason adverbial clauses and the present-day strategies identified in Yaqui reveals that the switch-reference system has been lost and that new markers have been originated for the more explicit expression of interclausal causality. The internal structures of the old and present-day adverbial clauses are not the same, and, as it will be seen in the next sections, the differences are clearly related to the origins of these constructions.

5. The origin of cause/reason adverbial markers in Yaqui

5.1 The origin of *po(r)ke*

In her study of Yaqui syntax, Lindenfeld (1973: 84) only identifies for introducing causal adverbial clauses the uses of *bwe'ituk* and of *po(r)ke*. She reports that *po(r)ke* is borrowed from Spanish, in which the form *porke* (written as *porque*) is used as causal adverbial clause marker, as illustrated in (20) for same subject and (21) for different subject.

- (20) *Estoy contento [porque estoy de vacaciones]*
 I.am happy because I.am PREP holidays
 'I am happy because I am on vacation.'
- (21) *Estoy contento [porque ya no estás enfermo]*
 I.am happy because already NEG you.are sick
 'I am happy because you are no longer sick.'

Lindenfeld (1973: 84) also pointed out that this Spanish-borrowed syntactic connective is clearly the strategy most commonly used for rendering a causal subordination, contrary to the native Yaqui adverbial marker *bwe'ituk*, which is rarely used according to her. However, it should be highlighted that in fact nowadays the presence of *bwe'ituk* in oral and written narratives seems to be frequent (Buitimea 2007), and *po(r)ke* is virtually non-present in those cases. In conversations and more informal speaking situations, however, the use of *po(r)ke* is real, although *bwe'ituk* is also very frequent in these cases.

This situation is clearly explained by the fact that the influence of language contact is more present in conversations than in narratives. Several studies have suggested that, in intense language contact situations, conversations are a frequent locus for code switching (Pfaff 1982; Brody 1987, 1995; Myers-Scotton 1993), which may serve as a trigger for borrowing (Myers-Scotton 1993; Torres 2002).

As for the morphosyntactic features of the Yaqui cause/reason adverbial clauses introduced by *po(r)ke*, the same features present in the Spanish cause/reason adverbial clauses in (20) and (21) are found: *po(r)ke* is used in the adverbial clause-initial position, the adverbial clause is balanced, the verb in the dependent clause is a finite verb, and no switch-reference system is present. This seems to exemplify a case of syntactic copy, in which the syntactic element that is borrowed, is accompanied by the morphosyntactic features of the source construction, including the position of the adverbial connector, the coding of the arguments of the adverbial clause, and the finiteness of the dependent verb.

5.2 The origin and evolution of *betchi'ibo* in Yaqui

The story of *betchi'ibo* is much longer than that of *po(r)ke*, and shows an interesting syncretism (Álvarez 2015). Indeed, besides its use as a cause/reason adverbial marker, other uses of *betchi'ibo* are present nowadays in Yaqui. It can also be used: (i) as a postposition associated with different semantic roles (cause, purpose and beneficiary/maleficiary) and (ii) as a purposive adverbial marker.

5.2.1 Other uses of *betchi'ibo* in Yaqui

5.2.1.1 Benefactive/malefactive postposition

The most common postpositional use of *betchi'ibo* corresponds to the benefactive/malefactive, in which the oblique complement introduced by the postposition is marked as accusative (suffix *-ta* in (22)), except if this postpositional object is plural (23) since the accusative suffix does not appear with plural nouns, as mentioned in Section 2.

- (22) *Jiak jitoa-ta te Potam-po jinu-bae [Peo-ta betchi'ibo]*
 Yaqui medication-ACC 1PL.NOM Potam-LOC buy-DES Pedro-ACC POSP
 'We are going to buy Yaqui medication in Potam for Pedro'
- (23) *María posoim [yoeme-m betchi'ibo] ya'a-bae*
 María pozole man-PL POSP do-DES
 'María is going to do *pozole*¹¹ for the men.'

With pronominal objects, the forms do not correspond to accusative, since Yaqui has a special pronominal paradigm for postpositional phrases, as illustrated in (24).¹²

- (24) *¿Jitása empo [ae betchi'ibo] yeu pu'a-la?*
 What 2SG.NOM 3SG.OBL POSP for_outward collect-RES
 ¿What have you chosen for him?

The animate noun functioning as the postpositional object can have a malefactive meaning, if the main predicate implies a negative effect, as in (25).

- (25) *[Chikul-im betchi'ibo] ne jogo-ta jinu-k*
 rat-PL POSP 1SG.NOM poison-ACC buy-PFV
 'I bought poison for the rats'

11. Pozole is a traditional Mexican soup.

12. The present-day paradigm for pronominal objects of postpositions is as follows: *ne/nee* '1SG', *ee* '2SG', *aa/ae* '3SG', *ito* '1PL', *emo* '2PL', *ame* '3PL'.

As seen in these examples, the benefactive/malefactive meaning of *betchi'ibo* is associated with active main predicates and animate postpositional objects.

5.2.1.2 Causal postposition

The element *betchi'ibo* can also be used to introduce a causal complement. In this case, the noun phrase that is object of the postposition, denotes the cause/reason why the situation denoted by the predicate is carried out. The coding of the causal postpositional objects is the same marking (accusative suffix *-ta* only with singular noun postpositional objects) than with the benefactive/malefactive postpositional objects, as can be observed in (26) and (27).

(26) [Em kujteer-im *betchi'ibo*] e koòkoi weche-k
 2PL.POSS anger-PL POSP 2SG.NOM sick fall-PFV
 'You got ill because of your angers.'

(27) Maria sioka [ili yoem-ta *betchi'ibo*]
 María be_sad little man-ACC POSP
 'María is sad because of a young man.'

As seen in these examples, the causal meaning of *betchi'ibo* is obvious with in-active predicates, since the lack of agentivity blocks the benefactive/malefactive interpretation.

5.2.1.3 Purposive postposition

Besides its postpositional uses as benefactive/malefactive and cause, *betchi'ibo* are also found nowadays for introducing an oblique complement of purpose. Again, the coding of the postpositional object remains the same, as can be seen in (28) with a singular object (suffix *-ta*) and in (29) with a plural object (no accusative suffix).

(28) Manto saawa koba meje-ku oòre-wa [taiwechia-ta *betchi'ibo*]
 manto¹³ leaf head forehead-LOC put-PASS fever-ACC POSP
 'The *manto* leaf is applied to the forehead for fever.'

(29) Chikul aaki-ta nuuse [naka wantiam *betchi'ibo*]
 mouse pitaya-ACC take.IMP ear pain.PL POSP
 'Take the *pitaya*¹⁴ for earache!'

As seen in these examples, the purposive meaning is associated with intentional active predicates and inanimate postpositional objects.

13. The scientific name of *manto* is *ipomoea carnea* subsp. *fistulosa*, which corresponds to a species of morning glory.

14. The scientific name of this *pitaya* is *mammillaria microcarpa*, which is a very common fish-hook cactus in Sonora (Mexico).

5.2.1.4 Purposive adverbial marker

Besides its use as causal subordinator illustrated in (3) and (4), the element *betchi'ibo* can also be used as an adverbial subordinator associated with the notion of purpose. In (30) and (31), the dependent clause is a purposive adverbial clause, which denotes a situation that is presented as the aim for which the action denoted in the main clause takes place. In (30), the subject of the main clause is the same as the subject of the adverbial clause and this case of co-referentiality between both clauses implies zero anaphora (the subject of the dependent clause is not expressed), whereas the subjects are different in (31), and in this case, the subject of the dependent clause is coded as accusative.

(30) *Tomi-ta ne bu'uria [wakas-im jinu betchi'ibo]*
 money-ACC 1SG.NOM accumulate cow-PL buy SUB
 'I am saving money in order to buy cows'

(31) *Maria in bwa'am-po joyo-ta ya'a-k [nee ko'okoi]*
 María 1SG.POSS food-LOC poison-ACC do-PFV 1SG.ACC sick
wet-ne betchi'ibo]
 fall-FUT SUB
 'María poisoned my meal for me to get ill.'

Regarding the finiteness of the dependent verb, the only verbal marker that it seems possible to attach to it, is the suffix *-ne* 'future', although it is optional and it tends to appear more often in cases of different-subject situations, as can be seen if (30) and (31) are compared.

In Álvarez (2015), I have analyzed the wide syncretism of *betchi'ibo* from both a synchronic and a diachronic point of view. I have proposed that in synchrony, *betchi'ibo* is used to introduce the motive why the situation denoted by the main predicate happens. This motive can be a cause, a reason, a beneficiary, a maleficiary, a purpose, depending on the type of main predicate and the type of postpositional object.¹⁵ In diachrony, I have defended the idea that the case syncretism in Yaqui is better explained by the evolutionary path CAUSE > BENEFICIARY > PURPOSE,¹⁶ and that the adverbial functions are the result of the extension of the postposition from nominal to clausal complements. In the following sections, I focus on the origin of *betchi'ibo* and on the rise of the cause/reason adverbial meaning.

15. Additionally, *betchi'ibo* has also developed a temporal goal meaning when the object of the postposition is a temporal adverb. This use corresponds to a new extension from the purposive use (Álvarez 2015).

16. Luraghi (2001, 2005a, 2005b) has proposed two different paths for explaining the case syncretism between cause, beneficiary and purpose: CAUSE > BENEFICIARY > PURPOSE, and BENEFICIARY > PURPOSE > CAUSE. I refer the reader to Álvarez (2015) for a presentation of all the arguments considered in support of the CAUSE > BENEFICIARY > PURPOSE path for Yaqui.

5.2.2 Postpositional and spatial origins of *betchi'ibo*

Yaqui presents two types of postpositions: free and bound. As Dedrick and Casad (1999: 173, 193) have pointed out, many Yaqui free postpositions are complex and exhibit as first element a form *be-*, such as in (32).

- (32) *bepa* 'over, above'
betuk(u) 'under, beneath'
betana 'from, at one side of'
bewichi 'alongside of, at the same time'
bena 'like'

This form *be-* comes from a third person singular pronoun reconstructed in Proto-Uto-Aztecan as **pi*, which is part of a pronoun copy construction (Langacker 1977). This syntactic type illustrated in (33) has been reconstructed for the syntax of postpositions in PUA.

- (33) COMPLEMENT-ABS 3.PRON-POSP

Some of the *be*-postpositions in Yaqui are no longer segmentable synchronically, even though they are clearly complex. In the case of *betchi'ibo*, it is still possible to identify the elements involved in its formation. Besides the *be-* form, two bound postpositions are involved: the locative postposition *-(e)tchi* 'on, in', and the ablative postposition *-bo* 'from'. Both bound postpositions are exemplified in (34) and (35), respectively.

- (34) *Baawe mayoa-tchi*¹⁷ *kooyo-m si bu'u*
 sea side-LOC shell-PL a_lot too_much
 'At the seaside, there are many shells.'
- (35) *Jaku'u-bo=sa empo wee*
 where-from-INTERR 2SG.NOM come
 'Where do you come from?'

The origin of *betchi'ibo* is thus associated with spatial meanings in which the movement from a source/origin (the form *-bo*) is combined with a static location (the form *-(e)tchi*), resulting in a complex postposition associating directionality and stativity.

17. Nowadays, the reduced form *-(e)t* is more frequently used in order to express this locative meaning, especially if any linguistic element is following the postposition, something that was already frequent in the 17th century (Buelna 1890: 100)

5.2.3 *The uses of betchi'ibo in Old Cahita*

The only uses documented for *betchi'ibo* in the *Arte* (Buelna 1890) are the causal and benefactive postpositional uses, as illustrated in (36) and (37), respectively.

- (36) [*emo / emo-t vetzivo*] *ne vebi-ua-c*
 2SG.OBL/2SG.OBL-LOC POSP 1SG.NOM whip-PASS-PFV
 'Por ti, me azotaron.'
 'For you, I was whipped.'
 Buelna (1890: 107)
- (37) *Dios i-ca aniuu-ta [ito-t vetzivo] a-ieua-c*
 God this-ACC world-ACC 2PL.OBL-LOC POSP 3SG.ACC-create-PFV
 'Dios crió todo este mundo por nosotros ó para nosotros'
 'God created this world for us.'
 Buelna (1890: 107)

According to the information provided in Buelna (1890: 107), this postposition was used to introduce a causal or a benefactive complement, expressed in the Spanish translation by *por* (cause) and *para* (benefactive) and the equivalent form in Latin was *propter*.¹⁸ Interestingly, the object of this postposition was differentially marked when taking a pronominal form. In this case, the postposition could be combined directly to the oblique pronoun, or attached to the locative *-t(zi)*,¹⁹ as can be observed in (36). This differential marking was probably present at the origin of the distinction between cause and benefactive meanings. The cause interpretation could have been triggered by the oblique pronoun, and the benefactive by the oblique pronoun suffixed by the locative *-t(zi)*, if we consider that the locative marker 'on, in' can be interpreted as benefactive via a metaphor based on the idea that the beneficiary is the locus of the action, and the benefit can be identified *on/in* him/her. Additionally, the ablative form *-bo* 'from' in *betchi'ibo* can easily be associated with the causal meaning via another metaphor, which is widely attested in world's languages: "causes are origins" (Lakoff & Johnson 1980). A cause can thus be conceptualized as the source of an event or situation.

The presence of both a locative and an ablative marker in the etymology of *betchi'ibo* could be implying that both cause and benefactive interpretations were

18. The author of the *Arte* states about *vetzivo*: "significa lo mismo que por ó para" ('it means the same than *por* or *para*, "corresponde al *propter*" ('it corresponds to *propter*) (Buelna 1890: 107). It is worth noting that according to Luraghi (2005b), the preposition *propter* in Latin had a locative source ('next to') and evolved first as a causal preposition and, as for Vulgar Latin, developed into a benefactive and purposive preposition. A similar evolution is proposed for *betchi'ibo* in Álvarez (2015).

19. The author of the *Arte* states: "se junta á los ablativos de los pronombres, unas veces inmediatamente, otras mediante el *tzi*, con nombres rige acusativo" (Buelna 1890: 107). The ablative forms mentioned in the *Arte* correspond to oblique pronouns, which were in the period of the *Arte* *ino* '1SG', *emo* '2SG', *aie* '3SG', *ito* '1PL', *emo* '2PL', *ame* '3PL'.

simultaneously active at the very origin. However, some arguments can also be proposed in favor of the CAUSE > BENEFICIARY path.

- The notion of benefit implies the notion of cause, in the sense that when one acts for someone’s benefit, the beneficiary becomes the reason why one acts. The beneficiary can thus be viewed as a reason to act. On the contrary, not all-human causes can be considered as a beneficiary. The notion of cause is thus more basic than the notion of beneficiary, and it would be expected that the latter derives from the former.
- Causes can be animate and inanimate, whereas beneficiaries are prototypically animate, most often human. Examples of *betchi’ibo* in the *Arte* always show a human postpositional object. When the main predicate is active like in (37), both causal and benefactive interpretations are possible (this possibility is indicated by the uses of *por* and *para*²⁰ in the Spanish translation), and when the predicate is inactive like the passivized predicate in (36), only the causal interpretation is possible. This shows that the general meaning of *betchi’ibo* was causal, and that the benefactive meaning is restricted to a certain type of predicate.
- The purposive meaning is commonly derived from the benefactive meaning via an extension of use from animate to inanimate objects (Luraghi 2005a). Recall that the purposive uses of *betchi’ibo* are not mentioned in the *Arte*, in which only human postpositional objects are exemplified. In that regard, the Catechism accompanying the *Arte* in Buelna (1890) contains an interesting use of *betchi’ibo*. Combined with the inanimate indefinite pronoun *jita* ‘something’, it is used as a causal interrogative pronoun (38), suggesting that the causal meaning is prior to the purposive meaning, and probably also prior to the benefactive.

- (38) *Hita vechivo tua iorem-tu-c itom iautzia Jesu-Cristo?*
 Something POSP truly man-VBZ-PFV 1PL.POSS father Jesus Christ
 ¿Por qué se hizo hombre el Hijo de Dios? (Buelna 1890: 242)
 ‘Why did our Father Jesus Christ become man?’

Interestingly, nowadays, *jita betchi’ibo* is used as a purpose interrogative marker ‘for what?’, and the most common cause interrogative marker is *jaisaakai* ‘why?’.

20. According to Corominas (1954), the difference between *por* denoting the cause and *para* denoting the beneficiary and the purpose took place in Modern Spanish, since in Medieval Spanish the preposition *por* could express the three semantic roles. The preposition *para* came out in Spanish after the Middle Age, created out of *por a*. The benefactive/purpose preposition in Spanish is thus the result of the combination of cause and goal, since the form *a* can be associated in Spanish with the spatial domain (directional).

Independently of the simultaneity or directionality at the origin of the cause/beneficiary syncretism, what we know, based on the information contained in the *Arte*, is that *betchi'ibo* was only a postposition, that is, a syntactic element that combines with (pro)nominal, not with clauses.

5.2.4 *The rise to causal adverbial marker*

The evolution from causal postposition to causal adverbial marker is may be a recent development, since the use of *betchi'ibo* as causal adverbial marker is documented neither in Johnson (1962), nor in Lindenfeld (1973). Another option is that this use existed but was not identified because it is not frequent since it seems to be restricted to introducing a negative cause/reason, that is, the motive explaining the affectedness expressed in the main clause. See (3), (4), and (39), (40).

- (39) [In *techoe* *tenku-ka* *betchi'ibo*] *bea ne*
 1SG.POSS badly_presage dream-PFV BECAUSE then 1SG.NOM
kaá-bae-ka juni ne omot tekia-po kibake-k
 NEG-DES-PFV although 1SG.NOM apart_from position-LOC enter-PFV
 'Por el sueño de mal agüero que tuve, no me quedó más que tomar otro cargo.'
 'Because of the ominous dream I had, the only thing left for me was to take another job'

- (40) *Ian e ne-mak wee-ne kajtikaroa-ri-po benasia* [em
 now 2SG.NOM 1SG.ACC-COM go-FUT punish-RES-LOC like 2SG.POSS
neé baitatta'a betchi'ibo]
 1SG.ACC cheat BECAUSE
 'Ahora te irás conmigo como castigo por estarme engañando.'
 'Now you will come with me because you are cheating me.'

If we compare these examples with (26) and (27), it is clear that the same interpretation of negative cause/reason is present in the use of *betchi'ibo* as causal postposition, indicating that the adverbial use comes from the postpositional use via the extension of contexts of use from nominal to clausal complements. This extension results in a hybrid construction exhibiting nominal and clausal features. As a nominal complement, the subject of the adverbial clause is coded as a possessor, which is a clear signal of nominalization (Givón 2001: 25). As a clause, the verb of the adverbial clause is finite and the object is in accusative. The nominalization involved in this construction is a clear indication of its postpositional origin. Indeed, it can be proposed that the causal postposition has been recruited in order to introduce a causal clause, thus imposing the possessive marker to the dependent subjects as an inheritance of the postpositional use in which the object is a noun phrase. As Lehmann (1986) has pointed out, the degrees of clausal interconnectivity correspond to a continuum of desententialization going from the clause prototype to the

noun prototype. The possessive marking thus shows that the new causal interconnectivity marking is accompanied by a process in which the clause adjusts to the noun in some way (Lehmann 1986). Additionally, we can see that the possessive marking, the clause-final position of the adverbial marker, and the semantic preference for a negative cause/reason usage of *betchi'ibo* as causal adverbial marker, are thus all features of its original use as causal postposition.

5.3 The origin of *bwe'ituk* in Yaqui

In Álvarez (2019), I have proposed that the causal adverbial marker *bwe'ituk* 'because' is the result of a recent formation process that combines a conversational marker *bwe* and a linguistic element associated with the strategy used in the past for marking cause/reason clauses (*ituca*, see (12) and (14)).²¹ This formation is functionally motivated by the fact that *bwe* is mainly used as a discourse connective that introduces a new topic. Based on this connecting function, the element *bwe* has been recruited from discourse to syntax, in order to participate in the creation of a new interclausal connective that also correspond to a thematic reorientation device: the cause/reason adverbial marker *bwe'ituk*.

5.3.1 *The bwe'ituk formation*

In the light of the past strategies used for cause/reason clauses in Old Cahita illustrated from (9) to (14), it can be proposed that *bwe'ituk* corresponds to the combination of a monosyllabic element *bwe-* and a form that can be easily related to the old strategies. This formation hypothesis is then as follows:

- (41) Formation hypothesis for the current cause/reason adverbial connective:

$$bwe'ituk < bwe + ituc(a)$$

The proposal in (41) implies two main syntactic rearrangements in the expression of the new cause/reason adverbial clause: (i) the change of the adverbial marker position from dependent clause final position (*ituca*) to dependent clause initial position (*bwe'ituk*), (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, which has already been illustrated in Old Cahita with the Example (16), is still a very

21. In Álvarez (2019) I have shown that, besides the creation of the causal connective *bwe'ituk* 'because', the discourse marker *bwe* has also been recruited in Yaqui for the creation of another interclausal connective of discontinuity: the adversative connective *bweta* 'but'.

frequent phenomenon in Yaqui when suffixed words are in positions others than the final position.

5.3.2 *The discourse marker bwe*

No information about old uses of discourse markers or interjections in Old Cahita is given in the *Arte*. As for today, the bilingual Yaqui-Spanish dictionary published by Estrada et al. (2004) registers several discourse markers. One of them is the particle *bwe*, whose lexical entry is given in (42a). From this information, it appears that *bwe* presents several discourse functions (if we consider the different equivalents given in Spanish), and coexists with a longer form, *abwe*.²² According to the information provided in the lexical entry corresponding to *abwe* in (42b), it is a conversational marker used in initial position.

- (42) a. ***bwe part.*** ¡ah!, ¡a poco!, ¡bueno!, ¡jeste!, ¡pues! *Cf. abwe.*
bwe part. ah!, oh really!, well!, er! um! *Cf. abwe.* Estrada et al. (2004: 73)
- b. ***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)

5.3.2.1 *The origin of bwe*

Regarding the origin of *bwe*, it seems to be plausible and tempting to consider that the discourse marker *bwe* comes from Spanish *pues* (this Spanish discourse marker is often proposed as an equivalent in the translations into Spanish) and *abwe* could be a loan from *ah pues*. Indeed, as mentioned above for the case of *po(r)ke*, 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) proposes 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 here is not only from discourse to syntax but also from syntax to discourse in a first stage: IMPERATIVE VERB > DISCOURSE MARKER > INTERCLAUSAL MARKER. 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 at’), no attestation of this

22. 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.

verb in the *Arte*, and no evidence of this old imperative use. Unfortunately again, Lionnet (1977) does not provide any evidence or argument when proposing this verbal origin. Despite this, although the Spanish origin of *bwe/abwe* is very possible and this external source is tempting, an internal source cannot be totally excluded for now.²³

5.3.2.2 *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) insists on the fact that the genesis of clausal conjunctions is constrained by the discourse context that frames its emergence. Therefore, it is important to examine the different discourse contexts in which the element *bwe* may appear in Yaqui discourse, in order to understand the motivations and the constraints implied in the diachronic rise of the interclausal marker *bwe'ituk*.

In Yaqui texts, the particle *bwe* only appears in reported direct speech, 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*. Its most frequent use is as the initial element of a reply, as in (43).

- (43) - ¿Jaisa into bea yee kòòko-si jooa?
 - how and then someone pain-ADVZ do
 ¿Y cómo es que hace daño?
 And how is it that it hurts?
 - *Bwe*, yee kòòko-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)

This initial position can also imply a kind of surprise for the speaker who has to answer the question, as in (44).

23. Consider also that there are phonologically similar loanwords from Spanish that are kept in Yaqui the initial voiceless bilabial, like *pueta* ‘door’ from Spanish *puerta*.

- (44) *-¡Jita ket nee betana nattemae-k, jaboí?*
 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.NOM 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)

The surprise associated with the use of *bwe* is also clear in (45).

- (45) *¡Bwe jiba yepsa-k!*
 PART always arrive-PFV
 '¡A poco siempre llegó!
 'Oh really, s/he arrived at last!

Bwe functions as a reactive particle that can indicate not only surprise but also some problems on the content level of the question. In (46), *bwe* is used in a case in which the respondent know that he is not giving directly the information the questioner asked for, that there is some sort of insufficiency in his reply.²⁴ In this case, the use of *bwe* indicates that the answer is not optimally coherent with the preceding question because the respondent is not really supplying the requested information.

- (46) *- ¡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)

Other uses involve 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 (47).

24. This use of *bwe* seems to represent a discursive use as a qualifier, according to the terminology used for the discourse marker *well* (Svartvik 1980; Carlson 1984: ch. 5) when, in a context of questions and answers, it is used in the initial position of a reply, signaling a problem between the speaker and the addressee, and thus prefacing a reply that is probably insufficient for the questioner. This discursive use has been much discussed in the literature on *well* (see for instance Lakoff 1973: 458–463; Schiffrin 1987: 102–127).

- (47) - *¿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)

Bwe can also be used as a phatic operator indicating hesitation, as in (48) in which the repetition of *bwe* reinforces the speaker's hesitation. In this case, it is used as a pause filler to bridge interactional silence. It marks the speaker's claim to the floor, being a strategy for keeping the floor while looking for something to say or for the right words in which to say it. The pragmatic function of filling a pause is thus made by an element (*bwe*) that is semantically more or less vacuous.

- (48) - *¿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, working at the interpersonal level.²⁵ However, *bwe* can also work at the textual level,

Example (49) represents a case in which *bwe* seems to function 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 this case as it is in the previous examples, functioning here more appropriately as an interclausal connective linking two independent clauses.²⁶ Another interesting point is the fact that the particle *bwe* is not located in (49)

25. Brinton (1995: 380) identifies two broad functions for pragmatic markers: an interpersonal function and a textual function. In the interpersonal level, the speaker expresses subjective attitudes and evaluations as well as acknowledges and maintains a social exchange with the hearer. In the textual level, the speaker structures utterances as text.

26. Indeed, although the Spanish and English translations of the first clause in (48) refer to a temporal adverbial clause, it is in fact a non-promotional passive construction (Lit. it was put on me the headpiece).

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 comment.

- (49) - *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 je'e! ti ne kaa into jiu-bae*
 - PART no_EMPH 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)

In sum *bwe* appears to be a conversational marker, which exhibits 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. 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. In a general way, the element *bwe* can be considered as a discourse-connecting conjunction ("contextual coordinates" in Schiffrin's (1987) terms) whose main function is to introduce a new comment relevant to the discourse. In this sense, *bwe* functions as a discourse organizer that structures the information and that introduces a new comment distinct from the prior discourse. It always appears in an initial position, being the first part of a reactive intervention and thus prefacing a new propositional content. Its most frequent function is interpersonal, but it can also function in the textual level.

5.3.3 *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 in Yaqui because *bwe* is a discourse-connecting marker, serving as a thematic re-orientation device with a cataphoric orientation. Indeed, the presence of *bwe* in the discourse creates the expectation of a new utterance coming next, since *bwe* is oriented forward. The same cataphoric thematic reorientation is strongly associated with cause/reason adverbial connectives, since causal clauses typically follow the main clause (Diessel 2001: 445–446). Causal connectives are thus introducing causal situations, that is, new relevant information explaining why the main clause has been expressed. So, the presence of the interclausal connective *bwe'ituk* to the left side of the causal clause creates the expectation of a new clause expressing the cause/reason for the main clause situation, thus linking 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 have proposed in Álvarez (2019) that the element *bwe* could move from discourse to syntax, in order to participate in the creation of the cause/reason adverbial marker in Yaqui. In this grammaticalization, the textual use of *bwe* exemplified in (49) might represent the bridging context for the recruitment of *bwe* as part of the causal connective *bwe'ituk*, since in this use *bwe* is already functioning at the interclausal level.

5.3.4 Discursive uses of *bwe'ituk* in Yaqui

The first evidence of the *bwe'ituk* formation comes from the Juan Banderas's letters (Dedrick 1985: 146) written at the beginning of the nineteenth century. Interestingly, in this first use exemplified in (50), the element *bwe'ituk* appears in the initial position and its connecting function is not yet interclausal, but still discursive. Although we do not have strictly speaking a conversation in (50), 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 Banderas who is the addressee).

- (50) *'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 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* is used then to introduce

a new 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 might be better rendered by a translation like: ‘Well, we ask you here...’. In this example, *bwe’ituk* is thus used more as a discourse marker, than as a syntactic marker. It does not connect two different clauses in order to create a complex clause, its connecting function is between two chunks of discourse. In fact, it functions as a frame marker, indicating a change of topic. It thus works at the textual level as a text-structuring device, very much like the textual use of *bwe* in (49).²⁷ Indeed, both uses of *bwe* and *bwe’ituk* correspond to a connecting device that introduces a new topic, both are fixed between the two discourse parts that are connected, and the orientation is, in both cases, cataphorical, that is, the connectives are at the left margin of the new discourse topic, functioning then as a clause-initial reorientation device. However, an important difference is that the presence of the old *ituca* in the form *bwe’ituk* indicates that the new information introduced by *bwe’ituk* corresponds to the cause/reason of the situation expressed in the prior discourse. As the element *ituca* is originally an interclausal marker of causality, it thus conveys to the new form *bwe’ituk*, the causal dependency between the two parts of discourse being connected, while the discourse marker *bwe* provides the focal function and the cataphoric reorientation.

In evolutionary terms, the discursive use of *bwe’ituk* illustrated in (50) is probably representing an intermediate stage in the development of the cause/reason adverbial connective out of the discursive particle *bwe*.²⁸ In this use, *bwe’ituk* exhibits the discourse function of *bwe* and the causal meaning of *ituca*, it appears in the initial/focal position as *bwe* does, and it is used to introduce a clause interpreted as the cause/reason of the prior discourse, as *ituca* did.

Example (50) also shows that the creation of *bwe’ituk* as a causal adverbial marker seems to involve not just a change from discourse to syntax (where *bwe* is concerned), but also a change from syntax to discourse (where the *ituca* particle is concerned). Indeed, in this evolution, *ituca* dissociates from final position of the subordinate clause, being attracted to the initial position by the discourse marker *bwe*, and this change of position not only implies for *ituca* a change from anaphoric to cataphoric orientation, but also a change from syntax to discourse, since *bwe’ituk* in (50) is not an adverbial marker, but a discourse marker. The evolution from this discursive use in (50) to the sentential use as an adverbial marker naturally occurs

27. This function as a frame marker has also been identified for the discourse marker *well* in English (see, for instance, Svartvik 1980; Jucker 1997).

28. It is worth noting that it seems that in Mayo, the other Cahita language, this element has not yet been recruited for clausal integration, since Almada (1999: 31) registers *buë(y)tuk* saying that it only functions as an hesitation marker (‘muletilla’): ‘este ...’ ‘umm...’.

when the chunks of discourse connected by *bwe'ituk* are clauses. Once this inter-clausal usage becomes fully conventionalized *bwe'ituk* turns into a causal adverbial marker, which is only used to introduce the cause/reason for the situation expressed in the previous clause.²⁹

Another important information provided by the Example (50) is that the first causal meaning of *bwe'ituk* had to be action-based,³⁰ since the causal meaning in (50) appears when the writer of the letter presents the motive for writing the letter. Interestingly, concerning the use of the cause/reason connective, Dedrick and Casad (1999: 397) point 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 clause (Diessel 2001: 445–6). This initial position could thus be a reminiscence of the discursive use of *bwe'ituk* illustrated in (50). If we look at some examples proposed by these scholars, the causal meaning of *bwe'ituk* is clearly action-based. For instance, (51) is used after the speaker asks the hearer to go somewhere, and *bwe'ituk* thus introduces the justification for what has been said before.

- (51) *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)

29. 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.

30. Three types of causal meaning associated with causal adverbial markers are usually distinguished (Schiffrin 1987; Sweetser 1990). The first causal relation between two clauses is fact-based, as in (i), whereby a ‘cause and result’ relation holds between two idea units (the content or socio-physical domain of usage). The second type is knowledge-based, as in (ii), in which case the causal relation holds when a speaker uses some piece of information as a warrant for an inference (the epistemic domain of usage). The third one is action-based, as in (iii), and in this case, the causal relation holds when a speaker presents a motive for an action being performed through talk (the speech-act domain of usage).

- i. John is home because he is sick.
- ii. John is home because the lights are burning.
- iii. Is John home? Because the lights are burning.

[Schiffrin 1987: 202]

This use is clearly linked to the use of the discourse marker *bwe* mentioned above, since *bwe'ituk* is used here in a conversation (note the two pronominal forms referring to the speech act participants in (51)), and it is also very similar to the use of *bwe'ituk* in (50), since it is used in the initial position of an utterance, with an intense focal character, introducing an event that is interpreted as the cause/reason of the previous speech-act. The two examples of *bwe'ituk* in (52) also taken from Dedrick & Casad (1999), show the same speech-act sense. In both cases, *bwe'ituk* appears after the performance of a speech act (an advice), and it introduces a clause that expresses the reason for performing this speech act (the reason why the advice has to be followed).

- (52) - *a'a beeba-k juni'i tua a'a me'e-nee*,
 it hit-PFV even truly it kill-FUT
 'If you do it, make sure that you kill it,
 - *bwe'ituk junaa'a baakot kaa muku-k-o enchi na'ateho-nee*
 because that snake not die-PFV-if you.ACC accuse-FUT
 'because if the snake does not die it will accuse you
 ... - *weye-m-ta kat beba*
 go-NMLZ-ACC not.IMP hit
 'Don't kill a moving snake'
 - *bwe'ituk waa'a weye-me ji'i-bwa-bae-ka weye*
 because DEM go-NMLZ thing-eat-DES-SUB go
 'because the one that is on the move is looking for something to eat
 Dedrick & Casad (1999: 406)

In these examples, the use of *bwe'ituk* is clearly not factual, that is, the causal clause does not 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 for an action performed through talk. Thus, the cause is here metadiscursive.³¹ Another example of this action-based causality is given in (53), in which *bwe'ituk* introduces the reason why the previous question has been asked.

- (53) - *Jaisa eme inim kaa mekka jo'a-k*
 how 2PL.NOM here NEG away live-PFV
 - '¿A poco viven por aquí cerca?
 - 'Are you really living nearby?
 - *Bwe'ituk ne jakko juni kee enchim bicha-n*
 because 1SG.NOM when also yet_no 3PL.ACC see-IMPV
 - 'Porque yo nunca las había visto
 - 'Because I had never seen them' *Mi abuelo y yo* (Buitimea 2007: 55)

31. 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.

The original causal meaning of *bwe'ituk* as a causal adverbial marker thus seems to have been action-based, and this is clearly due to the discursive origin of *bwe*.

6. Final remarks

This paper has been focused on the rise of cause/reason adverbial markers in Yaqui. It has been shown that three new explicit markers of interclausal causality have been created from three different sources:

- from a causal postposition, since *betchi'ibo* becomes a causal adverbial marker via a syntactic expansion from the postpositional phrase, the object of *betchi'ibo* changing from noun phrases to clauses.
- from discourse, since *bwe'ituk* becomes a causal adverbial marker via the recruitment of a conversational marker from discourse to syntax.³²
- from language contact, since *po(r)ke* becomes a causal adverbial marker in Yaqui via a morphosyntactic borrowing and a syntactic calque from Spanish.

These new cause/reason adverbial clauses imply that the meaning of the clause-combining device has become totally explicit. If we consider the polyfunctionality of the final particle *teca* in Old Cahita, the evolution in the domain of cause/reason adverbial clauses was from a polyfunctional structure (temporal, conditional, purpose and causal interpretations) to monofunctional structures (only causal interpretation),³³ from economy (hidden complexity) to explicitness (overt complexity). This change illustrates thus 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. In this process, the past strategies used for expressing cause/reason adverbial clauses have been lost, along with the switch-reference system associated with them. The loss of the switch-reference system can also be explained by the origin of the new cause/reason adverbial markers, since the domains in which they were born (postposition, discourse marker, Spanish-*po(r)ke* clause) are not subject to switch reference.

32. It also implies 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.

33. 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).

The genesis of the new cause/reason adverbial markers seems to be relatively recent (the first appearance of *bwe'ituk* is from 1830, *betchi'ibo* has not been documented as cause/reason adverbial marker in Johnson (1962) and Lindenfeld (1973), and *po(r)ke* is a recent loanword from Spanish). Two aspects could have been facilitating this explicitness-driven maturation. The first one is the influence of the contact with a language having an explicit cause/reason adverbial marker, i.e. Spanish. The explicitness existing in Spanish *po(r)ke* clauses could have created the need for explicit cause/reason adverbial markers in Yaqui. The other facilitating factor might be associated with writing, 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 Ramat & Mauri (2011: 657) correctly state, “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 new explicit markers of interclausal causality in Yaqui.³⁴ As written language is also less prone to language contact (code switching) than oral language, this could also have participated in the creation of the explicit cause/reason adverbial markers using Yaqui native resources. 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.

Yaqui is thus a language that affords several markers for the realization of interclausal causal connection. At first sight, they may appear as synonymous or intersubstitutable. It is clear that a detailed pragmatic account of communicating interclausal causality in Yaqui would be needed for the identification of nuances and distinct contextual features attached to each causal marker, trying to uncover meaning distinctions, which can be very subtle. However, the origin of each marker provides interesting clues to the dissimilar distribution of causal markers in Yaqui.

According to its origin in language contact, *po(r)ke* is only found in informal oral language, since language contact is more present in informal oral language than in formal oral language or written language. According to its postpositional use in which the cause meaning is associated with the affectedness expressed in the main clause, *betchi'ibo* seems to be restricted to mark adverbial clauses expressing a negative cause/reason, within the socio-physical domain of causality (content sense, see

34. In the consolidation process, the influence of literacy is fundamental. As Miller (2006) argues, complex constructions are acquired after the age of seven and much later, through the strong influence of literacy.

footnote 30). As for *bwe'ituk*, it is the most frequent cause/reason adverbial clause marker in Yaqui narratives, and it has been shown that its original causal meaning as adverbial marker is associated with action-based causality, something due to the discursive origin of the *bwe*-part of *bwe'ituk*. However, as can be observed in (5) and (6), repeated here for convenience, *bwe'ituk* can also express a cause-and-effect relationship between propositions, that is, *bwe'ituk* can be fact-based.

- (5) *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.'
- (6) *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.'

This possibility seems to indicate that from its original speech act sense, *bwe'ituk* has developed other causal meanings, as the content sense illustrated in (5) and (6). This directionality is opposite to what has been proposed by Sweetser (1990) who has argued that, from their original content meanings, causal connectives have diachronically developed new meanings in the more subjective epistemic and speech-act domains. The meaning of *because* is thus described originally as transcending the content or socio-physical domain of usage to the epistemic domain, and, finally reaching its speech act sense.³⁵ The case of *bwe'ituk* proves that the speech act sense can also be the original causal meaning, if the interclausal marker of causality comes from the discourse level, and shows that new causal meanings can be developed from this original speech-act sense, thus illustrating a case of desubjectification.

Regarding the formal aspects of the cause/reason adverbial clauses presented here, since the constructions are always the results of diachronic processes, the synchronic differences in the internal structures are explained by the diachrony of these constructions. The origin of the adverbial markers thus explains their positions within the dependent clause as well as the differential coding of the dependent clause subject.

35. Examples of such developments in the realm of connectives have been presented by König and Traugott (1988), Traugott (1995) and Traugott and Dasher (2005). They illustrate cases of subjectification, in which "meanings become increasingly based in the speaker's subjective belief state/attitude toward the proposition" (Traugott 1995: 31). As Traugott (1995) has shown, subjectification plays a significant role in the grammaticalization processes on the sentence level.

- In the case of *betchi'ibo*, the postpositional origin leaves traces of the nominal category in the adverbial clause (subject marked as possessor) and explains why the adverbial marker occupies in this case the clause-final position, that is, the most postposed place in the clause.
- The clause-initial position associated with *bwe'ituk* is obviously caused by the recruitment of *bwe* that, in accordance with its function as an introducer of a new discourse topic, always occupies the chain-initial position. The first uses of *bwe'ituk* were probably in this chain-initial position,³⁶ 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*. This conversational marker *bwe* not only implies the clause-initial position but also explains why the adverbial clause marked by *bwe'ituk* is balanced, (that is, identical to a main clause, with a subject in nominative), since the clauses introduced by *bwe* are always main clauses.
- In the case of *po(r)ke*, the borrowing of the Spanish adverbial marker is accompanied by the morphosyntactic features of the source construction, and the adverbial clause is thus balanced, just like Spanish-*po(r)ke* clauses.
- In the case of old cause/reason adverbial clauses, the markers *teca*, *tuca/tuco*, *ituca* indicate the dependent status of the clause. So, the dependent clause subject cannot receive the same treatment than the main clause subject, thus being dropped (zero) if same subject and marked as accusative if different subject.

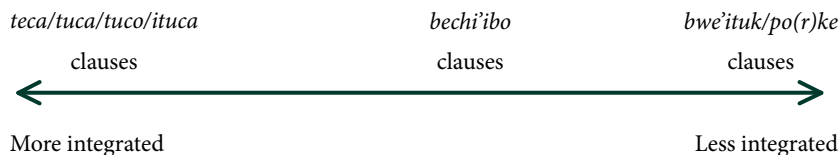
At the end, each construction is what it is, according to where it comes from, and the synchrony of dependent clauses is fashioned by the origins of dependent clause markers and by the diachronic processes derived from these sources.

Finally, the different cause/reason adverbial clauses that have been presented in this paper show different degrees of syntactic integration between the main and the dependent clauses, in particular if we consider the coding of the dependent clause subject. The most integrated adverbial clauses are the ones used in Old Cahita, since the coding of the dependent clause subject implied a switch-reference system in which same subjects were zero-marked and different subjects were

36. Recall the comment made by Dedrick and Casad (§ 5.3.4.) about the initial position of the *bwe'ituk* adverbial clause.

accusative-marked. Compared to the past strategies, *bechi'ibo* cause/reason adverbial clauses are also deranked, but the switch reference is no longer existing and dependent clause subjects are always marked as possessor. As for *po(r)ke* and *bwe'ituk* adverbial clauses, albeit for different diachronic reasons, both are balanced, and the dependent clause subjects are thus marked as nominative, showing the same structures than those of main clauses. The continuum of syntactic integration in Yaqui cause/reason adverbial clauses is schematized in (55):

(55) Syntactic integration continuum in Yaqui reason/cause clauses



The diachronic changes in the adverbial clause subject marking clearly show a change from more to less clause integration, since the evolution was from deranked adverbial clauses to balanced adverbial clauses, thus implying less syntactic dependency, less embedding. If we consider now that explicitness implies a higher degree of semantic dependency (only one possible interpretation of the interclausal meaning, that is, semantic monofunctionality, monosemy), the decrease of syntactic dependency is thus associated with the increase of semantic dependency. More syntactic dependency is thus correlated here with loose semantic dependency (semantic under-specification), and less syntactic dependency with tight semantic dependency (semantic specification).

Abbreviations

1	first person	NEG	negation
2	second person	NMLZ	nominalizer
3	third person	NOM	nominative
ABS	absolute	OBL	oblique
ACC	accusative	PART	discursive particle
ADVZ	adverbializer	PASS	passive
COM	comitative	PFV	perfective
COND	conditional	PL	plural
DET	determiner	POSP	postposition
DEM	demonstrative	POSS	possessive
DES	desiderative	PREP	preposition
DIR	directional	PRON	pronoun

EMPH	emphatic	RES	resultative
IMP	imperative	SG	singular
IMPFV	imperfective	SUB	subordinator
INST	instrumental	TEMP	temporal
INTERR	interrogative	TERM	terminative
LOC	locative	VBZ	verbalizer.

References

- Almada Leyva, Francisco. 1999. *Diccionario yoreme-español, español-yoreme*. Hermosillo: Dirección de Educación Indígena.
- 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 Martínez Baracs (eds), 211–232. Mexico: IIB UNAM/Somehil.
- 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. 2015. Sincretismo entre beneficiario, causa y propósito. Evolución de las funciones de *betchi'ibo* en yaqui. *Amerindia Revue d'Ethnolinguistique Amérindienne* 37(2): 69–96.
- Álvarez González, Albert. 2019. From discourse to syntax : The use of the discourse marker *bwe* in the creation of interclausal connectives in Yaqui. In *Diverse Scenarios of Syntactic Complexity* [Typological Studies in Language 126], Albert Álvarez González, Zarina Estrada-Fernández & Claudine Chamoreau (eds), 217–247. Amsterdam: John Benjamins.
<https://doi.org/10.1075/tsl.126.09gon>
- 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>
- Brinton, Laurel J. 1995. Pragmatic markers in a diachronic perspective. In *Proceedings of the Twenty-first Annual Meeting of the Berkeley Linguistics Society, February 17–20, 1995, General Session and Parasession on Historical Issues in Sociolinguistics / Social Issues in Historical Linguistics*, Jocelyn Ahlers, Leela Bilmes, Joshua S. Guenter, Barbara A. Kaiser & Ju Namkung (eds), 377–388. Berkeley CA: BLS.
- Brody, Mary Jill. 1987. Particles borrowed from Spanish as discourse markers in Mayan language. *Anthropological Linguistics* 29: 507–521.
- Brody, Mary Jill. 1995. Lending the “unborrowable”: Spanish discourse markers in indigenous American languages. In *Spanish in Four Continents: Studies in Language Contact and Bilingualism*, Carmen Silva-Corvalán (ed.), 132–147. Washington DC: Georgetown University Press.

- 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, Lauri. 1984. 'Well' in *Dialogue Games: A Discourse Analysis of the Interjection 'well' in Idealized Conversation* [Pragmatics & Beyond V: 5] Amsterdam: John Benjamins.
<https://doi.org/10.1075/pb.v.5>
- 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>
- Corominas, Joan. 1954. *Diccionario Crítico Etimológico de la Lengua Castellana*. Bern: Francke.
- Croft, William. 1991. *Syntactic Categories and Grammatical Relations*. Chicago IL: The 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* 22: 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, Buitimea Valenzuela, Crescencio, 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.
- Ford, Cecilia E., & Thompson, Sandra A. 1986. Conditionals in discourse: A text-based study from English. In *On Conditionals*, Elizabeth Closs Traugott, Alice G. B. ter Meulen, Judy Snitzer Reilly & Charles A. Ferguson (eds), 353–373. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511753466.019>
- Fraser, Bruce. 1990. An approach to discourse markers. *Journal of Pragmatics* 14: 383–395.
[https://doi.org/10.1016/0378-2166\(90\)90096-V](https://doi.org/10.1016/0378-2166(90)90096-V)
- Givón, Talmy. 2001. *Syntax: An Introduction*, Vol. II. Amsterdam: John Benjamins.
<https://doi.org/10.1075/z.syn2>
- Givón, Talmy. 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*, Albert Álvarez González, Zarina Estrada-Fernández, Claudine Chamoreau (eds), 109–136. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.126.05guc>
- 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.
- Heine, Bernd & Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511613463>
- Johnson, Jean. 1962. *El idioma yaqui*. México: INAH.
- Jucker, Andreas H. 1997. The discourse marker *well* in the history of English. *English Language and Linguistics* 1(1): 91–110. <https://doi.org/10.1017/S136067430000037X>

- König, Ekkehart, & Traugott, Elizabeth Closs. 1988. Pragmatic strengthening and semantic change: The conventionalizing of conversational implicature. In *Understanding the Lexicon: Meaning, Sense and World Knowledge in Lexical Semantics*, Werner Hüllen & Rainer Schulze (eds), 110–124. Tübingen: Niemeyer. <https://doi.org/10.1515/9783111355191.110>
- 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.
- Lakoff, Robin. 1973. Questionable answers and answerable questions. In *Issues in Linguistics: Papers in Honor of Henry and Renee Kahane*, Braj B. Kachru, Robert B. Lees, Yakov Malkiel, Angelina Pietrangeli & Sol Saporta (eds), 453–467. Urbana IL: University of Illinois Press.
- Lakoff, George & Johnson, Mark. 1980. *Metaphors We Live By*. Chicago IL: The Chicago University Press.
- Langacker, Ronald W. 1977. The syntax of postpositions in Uto-Aztecan. *IJAL* 43: 11–26.
- Lehmann, Christian. 1986. On the typology of relative clauses. *Linguistics* 24: 663–680. <https://doi.org/10.1515/ling.1986.24.4.663>
- Lindenfeld, Jacqueline. 1973. *Yaqui Syntax*. Berkeley CA: The University of California Press.
- Lionnet, Andrés. 1977. *Los elementos de la lengua cahita*. México: UNAM.
- Luraghi, Silvia. 1989. Cause and instrument expressions in Classical Greek. *Mnemosyne* VLII, Fasc. 3–4: 294–307.
- Luraghi, Silvia. 2001. Syncretism and the classification of semantic roles. *Sprachtypologie und Universalienforschung* 54: 35–51.
- Luraghi, Silvia. 2005a. Paths of semantic extension: From cause to beneficiary and purpose. In *Historical Linguistics 2003* [Current Issues in Linguistic Theory 257], Michael Fortescue, Eva Skafte Jensen, Jens Erik Mogensen & Lene Schøsler (eds), 141–157. Amsterdam: John Benjamins. <https://doi.org/10.1075/Cilt.257.12lur>
- Luraghi, Silvia. 2005b. Prepositions in Cause expressions. In *Proceedings of the XII International Colloquium on Latin Linguistics* [Papers on Grammar 12/2], Gualtiero Calboli (ed.), 609–619. Bologna: CLUEB.
- 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/B0-08-044854-2/00853-1>
- Miller, Wick R. 1984. The classification of the Uto-Aztecan languages based on lexical evidence. *International Journal of American Linguistics* 50(1): 1–24. <https://doi.org/10.1086/465813>
- Mithun, Marianne. 1988. The grammaticization of coordination. In *Clause combining in Grammar and Discourse* [Typological Studies in Language 18], John Haiman & Sandra A. Thompson (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.

- Ramat, Ana Giacalone & Mauri, Caterina. 2011. The grammaticalization of coordinating inter-clausal connectives. In *The Oxford Handbook of Grammaticalization*, Heiko Narrog & Bernd Heine (eds), 656–667. Oxford: OUP.
- 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.
- Svartvik, Jan. 1980. *Well in conversation*. In *Studies in English Linguistics for Randolph Quirk*, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik (eds), 167–177. London: Longman.
- Sweetser, Eve. 1990. *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: CUP. <https://doi.org/10.1017/CBO9780511620904>
- 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 & Dasher, Richard B. 2005. *Regularity in Semantic Change*. Cambridge: CUP.
- Traugott, Elizabeth Closs. 1988. Pragmatic strengthening and grammaticalization. In *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, Shelley Axmaker, Annie Jaissner & 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>
- Vester, Eline. 1983. *Instrument and manner expressions in Latin*. Assen: Van Cormun.
- 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>

Index

A

- adjective 275–276, 278–279, 282–294, 296–297
- adverbial connector 16, 103, 105, 107, 109, 113–114, 116, 118–119, 325
- anaphora 327, 328
- aphasia 164, 170–172

B

- Basque 57–63, 67–68, 70–75
- brain 153, 157–159, 163, 169, 176, 185

C

- cause/reason adverbial marker 313–314, 325–326, 333, 339, 343–344
- cause/reason adverbial clause 313–314, 317–321, 323–325, 333, 338, 343, 345–347
- Chinese 67, 72, 106, 113
- clause-final 151, 317, 321, 346
- cognition 158–159, 161, 173–174, 181
- cognitive 45, 80, 164, 170, 174, 184, 245–246, 250, 262
- coherence 147, 159, 162, 167–168, 183–184
- collective verb 299–303, 305–307, 309–311
- conversation 128, 130, 146, 152–153, 164, 168, 183, 230, 236–237, 240
- Construction Grammar 46, 132
- cooptation 5, 27–30, 33–49, 151–152

D

- Discourse Grammar 7, 25, 31, 42, 48

- discourse processing 30, 48, 159–161, 164, 166, 170, 173, 182–183, 185
- discourse structure 170, 173–174, 181, 182

E

- English 7, 13, 25, 28, 30, 32, 35–39, 43, 47, 151, 229–233, 235, 237, 239–240, 276–278, 296, 301–305
- emotion 79, 92, 168, 169, 245, 248, 249–250, 252, 253, 255, 256, 260, 262, 266
- expressives 248–252, 254–265
- expressive speech act 266, 336

F

- final particle 229, 241, 321, 333, 343
- formulaic language 159
- formulaic speech 161, 164, 166, 169, 175, 183
- formulaic unit 34, 166
- French 28, 39–43, 46–47, 127, 129–130, 134, 275–278, 296

G

- grammaticalization 4–5, 10–11, 23–24, 26–27, 30, 33–40, 42, 44–49, 57, 60, 65–66, 71–72, 77–78, 82, 92, 95–97, 108, 117, 119, 124, 151–152, 323, 335, 339, 341, 345–346

I

- inference 95, 129, 134, 173, 216
- interjection 80, 232, 255, 259, 266, 336
- intonation 29, 86, 93, 94, 237

J

- Japanese 11, 66, 101, 103–110, 113, 117–119, 124

K

- Korean 11, 77–81, 84–85, 90–91, 96–97

L

- language contact 123–124, 314, 325, 334, 343–344
- language processing 8, 12, 157, 159, 163–164, 184–185
- left periphery 11, 28, 29, 35, 43, 198, 205–211, 215–217, 224

N

- neural 8, 159, 164, 166–167, 170–171, 184, 189
- neurological 8, 158, 159, 164, 176, 184–185
- neurolinguistic 12, 158, 161, 164, 167, 169–170
- Northern English 13, 229–231, 237

P

- parenthetical 25, 33, 47–48, 128, 131, 136–137, 151–153
- partial control 15, 299–301, 311–312
- particle 24, 36, 48, 57–73, 78, 89–90, 103, 106, 107, 113, 131, 199, 220, 222, 229, 231, 234, 241, 253, 276–277, 285, 292, 314, 320–321, 324, 333–337, 340, 343
- Polish 15, 301–309
- politeness 58, 60, 65, 67, 69, 72, 93, 129, 143, 152–153, 239–240, 252

pragmaticalization 23, 46–47,
221
prosody 28, 30, 33, 35, 41,
92–94
prosodic 24, 26–28, 30, 34–35,
41, 77, 92–94, 237
psycholinguistic 161

R

reanalysis 11–12, 29, 33, 105, 10,
110, 113–119, 123
right periphery 28, 33, 36, 43,
48, 241

S

semantics 9, 14, 91–92, 132, 152,
164, 198, 246–248, 253–255,
259, 262–267, 290

sentence-final 13, 220, 222,
230–232, 234–240
speech act 31, 32, 33, 218–219,
222–224, 249, 251, 253, 260,
266, 336, 342, 345
subjectification 92, 97, 345
subjectivity 24, 92, 139, 142,
144, 198, 214, 216, 220, 341
syntax 26, 28, 30, 33, 35, 37, 41,
49, 60, 63, 151, 161, 198, 214,
215, 217, 233, 237, 325, 329,
333–334, 339–340

T

thetical 25, 27, 28, 29, 31–39, 41,
43, 45–49, 153, 160–161, 182
truth condition 25, 37, 248,
252, 254

truth value 136, 172, 218, 304
turn-taking 149, 159, 166–169,
173, 241

V

Vietnamese 13, 195–197, 206–
208, 211, 216, 220–221, 224

Y

Yaqi 16, 313–321, 323–326,
328–329, 333–335, 338–339,
343–345, 347

This book investigates phenomena at the grammar–discourse interface with a strong focus on discourse markers, whose development and concrete uses in a given language tend to be based on a close interplay of grammatical and discourse-related forces. The topics range from the transition of linguistic signs “out of” sentence grammar and “into” the domain of discourse to differences between more grammatical vs. more discourse-pragmatic expressions in terms of structural behavior and cognitive processing, and the different, intricate ways in which the usage conditions and meanings of grammatical constituents or structural units are affected by the discourse context in which they are used. The twelve studies in this book are based on fresh empirical data from languages such as English, Basque, Korean, Japanese and French and involve the study of linguistic expressions and structures such as pragmatic markers and particles, comment clauses, expletives, adverbial connectors, and expressives.



JOHN BENJAMINS PUBLISHING COMPANY