Linguistik Aktuell Linguistics Today 276

Discourse Particles

Edited by
Xabier Artiagoitia
Arantzazu Elordieta
Sergio Monforte

John Benjamins Publishing Company

Discourse Particles

Linguistik Aktuell/Linguistics Today (LA)

ISSN 0166-0829

Linguistik Aktuell/Linguistics Today (LA) provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.

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

Volume 276

Discourse Particles. Syntactic, semantic, pragmatic and historical aspects
Edited by Xabier Artiagoitia, Arantzazu Elordieta and Sergio Monforte

Discourse Particles

Syntactic, semantic, pragmatic and historical aspects

Edited by

Xabier Artiagoitia
Arantzazu Elordieta
Sergio Monforte
University of the Basque Country (UPV/EHU)

John Benjamins Publishing Company
Amsterdam / Philadelphia

DOI 10.1075/la.276

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2022008123 (PRINT) / 2022008124 (E-BOOK)

ISBN 978 90 272 1107 1 (HB)
ISBN 978 90 272 5776 5 (E-BOOK)

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

Table of contents

Part III. The semantic-pragmatics of discourse particles

Introduction

Discourse particles: syntactic, semantic, pragmatic and historical aspects

Xabier Artiagoitia, Arantzazu Elordieta and Sergio Monforte
University of the Basque Country (UPV/EHU)

1. Introduction

Discourse particles, or modal particles as they have also been named in the literature, form a vaguely defined category. Indeed, elements with various syntactic and semantic functions have been grouped under the term *particle*: time, aspect, mode, negation, evidentiality, interrogation, epistemicity and, in general, those related to discourse (Biberauer and Sheehan 2011). In a similar vein, Paul (2014: 77) clearly states that "*particle* is just a cover term resorted to precisely when no satisfying analysis of a particular item can be provided". Nevertheless, attempts to define them are found in the literature; for instance, following Biberauer eta Sheehan (2011), we could characterize discourse particles as follows: elements which have a higher degree of deficiency than other categories in the morphological, phonological, semantic, pragmatic or syntactic field; additionally, they show more than one function in the utterance.

On the other hand, it is common ground that discourse particles are functional constituents which do not modify either the content or the truth-conditions of a proposition and which also serve to express to the addressee that the utterance is shared knowledge, or to convey assumptions or expectations on the part of the speaker or addressee (Thurmair 1989: 2–3; Van Kemenade & Links 2020: 2–3). They are also claimed to form a closed class, something expected for functional elements (Grosz 2016: 336). For the purpose of this monograph, we base ourselves on the following definition of Abraham (2020: 218–219):

– *Modal particles in the narrow sense only surface on the right edge of the left middle field in the setup of the German and Dutch clauses structure.*

https://doi.org/10.1075/la.276.int

- *The fundamental polyfunctionality of modal particles captures not only that they are derived through complex diachronic grammaticalizing processes but that they derive their probing properties as clause (or Common Ground) operators from their lexical (pre-grammaticalizing) source meaning.*
- *Modal particles are grammatical (as opposed to purely lexical) elements.*
- *Modal particles have speaker-deixis status, and, for certain modal particles (most prominently for* **ja**), *also addressee-deixis status, as their primary import.*
- *Modal particles are part of an obligatory system.*

This last property is also defended by other authors (Paul 2014; van Kemenade & Links 2020); that is, although the exclusion of discourse particles does not result in ungrammaticality (as expected considering that they do not modify the content and the truth-conditions of the proposition they are inserted in), their absence brings certain inadequacy with respect to the discourse.

However, here we do not strictly follow the definition of Abraham, since he only considers modal particles that appear in the middle field and, therefore, limits this phenomenon to German and, by extension, to the Germanic languages. Significantly, this volume includes studies on discourse particles of a variety of typologically distinct and/or historically unrelated languages (Basque, English, German, Italian, Laz, Mandarin Chinese, Spanish) and it contains papers which discuss discourse particles located in outer domains rather than in the middle field.

In fact, the interest in that topic has greatly increased in the last years since the domain and study of discourse particles has extended to other languages beyond the Germanic family. If we look at recent works on this subject, we certify that typologically different languages and those belonging to distinct language families have attracted the attention of linguists; for instance: Bangla (Bhadra 2017), Cantonese (Lam 2014), the Cholon language (Alexander-Bakkerus 2011), Italian (Cardinaletti 2011), Japanese (Kuwabara 2013), Tagalog (Tan 2019), Telugu (Balusu 2019), among others. Moreover, the research field of discourse particles has widened and, in addition to its semantics and syntax, other aspects are now being investigated. Proof of such increasing interest can be found in the numerous books dedicated to the topic which have been recently published; just to mention some: *Discourse markers and modal particles: Categorization and description* (Degand, Cornillie & Pietrandrea 2013), *Discourse-oriented syntax* (Bayer, Hinterhölzl & Trotzke 2015), *Final particles* (Hancil, Haselow & Post 2015), *Discourse Particles: Formal Approaches to their Syntax and Semantics* (Bayer & Struckmeier 2017a) and *Information-structural perspectives on discourse particles* (Modicom & Duplâtre 2020). Such works certainly demonstrate that there is currently a live debate on diverse aspects of discourse particles.

The present collection presents an interdisciplinary view on the issue at stake. It tackles aspects previously discussed in the literature such as their diachronic development as well as their syntactic or semantic analysis. Crucially, this book also brings topics less touched upon in the literature but of great interest for linguistic theory, such as the acquisition of discourse particles by children (Chapter 7) or the analysis of elements not usually considered discourse particles but whose historical path and microvariation indicates that they may deserve an analysis in that direction (Chapter 2 & Chapter 3).

2. Issues under discussion in current research on discourse particles

The contributions comprising the present volume are a clear manifestation of current research on discourse particles. More specifically, they have been divided in three main blocks, according to the approach they take: the first part focuses on the historical development of discourse particles; the second one touches on the syntactic field of these particles; and the third one deals with the semantic-pragmatic analysis of particles.

Abraham (1991) is a reference regarding their historical development, as Coniglio states in his contribution to this book. Abraham (1991; 2020: 214) defends the idea that discourse particles derive from distinct categories: adjectives, adverbs, conjunctions, and so on. Interestingly, Coniglio (this volume) argues against this idea and attempts to identify a unified source for all particles, namely adverbs. However, there are far more works which explore the grammaticalization of discourse particles. Hack (2014) looks into the evolution of the particle *pa/po* in Dolomitic Ladin from an adverb to a discourse particle and, from there, to an interrogative marker. Similarly, Bayer (2012) and Pankau (2018) discuss the grammaticalization undergone by the German discourse particle *denn* towards a question particle in some dialects. Other works which discuss the derivation of discourse particles have concluded that they develop from verbs (Haegeman 2014) and even from paratactic clauses (Van Bogaert & Leuschner 2015).

On the other hand, the syntax of discourse particles has mainly received more attention in the last decades. As explained by Bayer and Struckmeier (2017b), most syntactic theories, such as those of the 1970s, lacked the appropriate instruments to carry out a syntactic study of particles; thus, German discourse particles were simply adverbs according to those grammarians, i.e. they were adverbs that did not have all the properties of adverbs. Later, most studies have focused specifically on the syntactic nature or syntactic position of particles (Grosz 2006; Coniglio 2007, 2008; Kuong 2008; Cardinaletti 2011; Kuwabara 2013; Haegeman 2014;

Struckmeier 2014; Del Gobbo et al. 2015; Hinterhölzl & Munaro 2015; Corr 2016; Scherf 2017, among others). However, this does not mean that these issues have been fully clarified and still much remains to be investigated.

Regarding the syntactic position of discourse particles, the neo-performative approach (Heim et al. 2016; Wiltschko 2017; among others) has opened a new way to study particles that do not appear in the middle field as well as to study languages that do not have that area so well-defined. Nevertheless, one must acknowledge that prior to the neo-performative analysis other approaches to the syntactic position of particles occurring both at the left periphery and at the right periphery had been proposed (Izutsu & Izutsu 2013; Haegeman 2014; Del Gobbo et al. 2015).

Finally, we turn to semantic-pragmatics, which is the best studied field in the modal particle category. However, this field continues to play an important role in the analysis of discourse particles, as shown by the number of publications on this subject (113 papers on this matter in the last decade according to the data offered by the Web of Science). In fact, the advances made in semantic-pragmatics in recent years paved the way to offer far more detailed studies on the contribution of these particles. As is well-known, it is not an easy task to pinpoint the precise contribution of discourse particles because they affect different levels or aspects (Abraham 2020) and, hence, a new model was necessary to investigate all these aspects. The model of the table developed recently by Farkas & Bruce (2010) seems to have facilitated this task, as suggested by the number of works applying this model to the analysis of discourse particles, including one in this volume (see Chapter 9).

We will now take a closer look at the main research questions posited above and link them to the issues raised in the chapters of this book. We sketch a short summary of each individual contribution below.

3. Contributions of this book

The first chapter of this volume deals with the grammaticalization process of discourse particles in German and with the pertinent syntactic analysis. Discourse particles have been claimed to derive from different types of donor lexemes (e.g. Abraham 1991); however, Marco Coniglio argues against such a hypothesis. He looks into several German discourse particles for this research, namely: *aber*, *bloß*, *denn*, *doch*, *einfach*, *eigentlich*, *nur*, *ruhig*, and *vielleicht*. He considers that all those elements have adverbial status or that they had it in previous stages of the German language. Based on this idea, he hypothesizes that those lower adverb(ial)s were reanalyzed as higher adverbs by acquiring more functions following the grammaticalization path put forth by Abraham (1991). As a final step in the derivation, their syntactic behavior and internal structure would have been impoverished (cf.

Cardinaletti 2011), thus becoming weak adverbs (or discourse particles) restricted to the middle field.

Following the thread of the historical development of discourse particles, in the second chapter George Walkden and Regine Eckhardt investigate the function of modern English *whether* (*hwæþer* in Old English) in a work written in that period, namely the translation of Boëthius's *De consolatione philosophiae*. *Hwæþer* (old *whether*) could also appear in main sentences at that stage of the language and, although the literature has stated that this strategy or the fronting of the verb were in free distribution and that there was no difference between the two procedures, the authors argue that some uses of *hwæþer* would have been close to some discourse particles since they could have been used to regulate the use-conditional content. Indeed, they deeply analyze the uses of this particle in that text and distinguish five types based on their syntactic environments and semantic-pragmatic contributions. They argue that *hwæþer* could be used to ask about the addressee's beliefs, opinions or own conclusions in some contexts, namely in those where pedagogical questions are licit. This usage would correspond to an intermediate step in the development from a question pronoun to a question complementizer.

To conclude with the historical part of this volume, in the third chapter Manuel Pérez-Saldanya and Jose Ignacio Hualde discuss the discourse particle *es que* in Spanish. While they pay special attention to the use of that Spanish particle, consideration is also given to the development of its counterparts in Catalan and Portuguese, as well as the loan *eske* from Spanish to Basque. They notice that not all usages of *es que* have a similar syntactic structure and, accordingly, the same interpretation. More specifically, they argue that the current particle derives from the sequence V + COMP, which in turn developed in different ways depending on the context of use; in particular, such a sequence can be broken into three groups depending mainly on the referential nature of the covert subject. In fact, they claim that the loss of referentiality of the null subject of *es* increased the usage of *es que* in counter-argumentative contexts and, ultimately, in its function as a discourse particle.

The next part, composed of four chapters, is devoted to the syntax of discourse particles. In the fourth chapter Ömer Demirok and Balkız Öztürk present data from the South Caucasian Laz language, namely of the particles *ya* and *şo*. These particles have been described as *quotative* or *instructional* (Boeder 2002); however, the authors argue that their behavior does not agree with that of other quotatives across languages. They claim that *ya* and *şo* behave as complementizers not only in embedded contexts but also in root clauses based on a number of syntactic properties such as being in complementary distribution with other complementizers and being embedded under certain attitudinal predicates. On the other hand, the authors look into the distribution of those particles under the scope of negation,

mood, causatives and other environments to finally conclude that *ya* and *şo* do not pattern alike. They also tackle the subject of the meaning of those particles. They entertain two hypotheses: (a) that the meaning of these particles remains the same in root and in embedded clauses; and, (b) that they do not show distinct meaning in embedded contexts and that the agreement between some features causes their separate morphological behavior in those contexts. Nevertheless, they put forth an alternative analysis that assigns separate meanings to *ya* and *şo* in both root and embedded contexts.

The fifth chapter analyzes two homophonous particles in Basque, namely *ala*. There are several particles located in the middle field in Basque; nevertheless, both *ala*-s do not behave as those particles located in the middle field; rather, they occupy a peripheral position. Therefore, they should be assigned a different syntactic position. Aitor Lizardi takes this idea and demonstrates that, in addition to the behavior of *ala* as an outer particle, it can also be a tag particle. The author reaches this conclusion by analyzing the phonological, syntactic and pragmatic properties of both *ala*-s and showing that they do not pattern alike among them. Moreover, the author suggests that outer particles are transitive Speech Act heads that take the sentence content as a complement and modify their speech act (Haegeman 2014; Corr 2016). In contrast, tag particles are intransitive Speech Act heads that form their own Speech Act domain. The author therefore differentiates two positions for these particles in the Speech Act domain contra the neo-performative approach (Heim et al. 2016).

Nicola Munaro investigates the syntactic distribution of the interjections *altroché*, *anca massa*, *as capés*, *ciò*, *eterché*, *madona*, *madosca*, (*mo*) *vaca*, (*mo*) *deg*, *però*, *sorbla*, *toh* and also *ah/eh/ih/oh/uh* in several Italian dialects and discusses their occurrence along the left periphery. Based on their syntactic characteristics, the author shows that these Italian interjections do not behave alike as regards their prosodic and syntactic integration. He identifies three patterns: those which are fully independent and, thus, susceptible to occur on their own; those which must be integrated with the associated clause and are intrinsically discourse-linked; and those which can (but need not) be integrated with the associated clause. Considering this characterization and the interpretative contribution of the interjections under study, the author proposes that they lexicalize three syntactic positions: the first group occupies the head of Speech Act2; the second group is base-generated as the head of Evaluative Speaker Phrase; and, finally, the third group occurs either on the head of Speech Act1 or the head of Evaluative Speaker Phrase. He concludes that the distribution along this rigid hierarchy accounts for their strict order and also for the different superficial positions the interjections can occupy.

Waltraud Paul and Shanshan Yan close the syntactic part of the volume by discussing Mandarin Chinese sentence final particles. This sixth chapter analyzes

not only he syntax of particles *le*, *láizhe*$_1$, *ne*$_1$, *bai*$_{mp}$, *ba*$_{Qconfirmation}$, *ma*$_2$, *a*, *láizhe*$_3$, *ei*, *ou*, *ma*$_3$, *zhene*, *ne*$_3$ but also their acquisition. In this case, they also follow a cartographic analysis and distribute these particles in three projections along the Left Periphery. In fact, based on their interpretation and syntactic distribution, they distinguish three groups: LowC or C_1 particles, Force or C_2 particles eta Attitude or C_3 particles. Particles of the same group are found in complementary distribution and it could seem that some of them may appear in two locations; nevertheless, Paul and Yan argue that they display different functions and distribution. On the other hand, the authors show that those C_1 particles play a role in tense and finiteness, contrary to what had previously been said in the literature. Finally, the authors look into the acquisition path of children regarding these particles and they conclude that there is no real evidence for the idea of a step-by-step acquisition, i.e. that children acquire first lower phrases and then higher phrases as postulated by Friedmann et al. (2021). Considering their data, they show that all particles emerge at a similar stage.

Part 3, focusing on the semantic analysis of discourse particles, starts with the article by Kepa Korta and Larraitz Zubeldia on the semantics of the Basque discourse particle *bide*. Unlike the chapter by Lizardi, these authors focus on one of the discourse particles which occur in the middle field. They assume that a *bide*-utterance and its counterpart without the particle assert the same proposition, as it is common in discourse particles; thus, the particle does not modify the truth-conditions of the proposition, but it provides information about the type of evidence the speaker has. Once they introduce the semantic analyzes received by this particle in the literature (van Eys 1879; Euskaltzaindia 1987; de Rijk 2008, among others), Korta and Zubeldia look into two aspects mentioned in the literature about *bide* to clarify its function: the evidential and the doxastic dimensions. After testing the behaviour of *bide* regarding those dimensions and comparing it with the Basque evidential particle *omen*, they conclude that *bide* expresses indirect evidentiality, i.e. the presupposition that the speaker's evidence for the truth of a proposition is indirect and, as for the doxastic dimension, the authors propose that it could belong to pragmatics but with some differences in comparison to the particle *omen*.

The final contribution of this volume reviews the semantic analyzes of some German discourse particles (*ja*, *nicht* and *etwa*) and proposes a new one following Farkas & Bruce's (2010) model. Johannes Schneider pursues to shed light on their contribution to the utterance and to reduce the range of possible meanings they can show. Contrary to the idea formulated in previous works that discourse particles convey the attitude of the speaker towards the proposition, the author argues against this hypothesis and defends that discourse particles modify the speech act of the utterance (cf. Jacobs 1991) by altering a property of the speech act located on the playing board of Farkas & Bruce's (2010) model at the conversation moment.

Acknowledgements

We editors would like to acknowledge that the publication of this book has been possible thanks to the support of the research project *Analysis of some functional particles (modal, negative, complementizer) in Basque grammar* (PGC2018-100686-B-I00) funded by the Spanish Ministry of Science and Innovation (MICINN).

References

Abraham, Werner. 1991. Discourse particles in German: How does their illocutive force come about. In *Discourse particles: descriptive and theoretical investigations on the logical, syntactic, and pragmatic properties of discourse particles in German*, Werner Abraham (ed), 203–252. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.12.08abr

Abraham, Werner. 2020. *Modality in Syntax, Semantics and Pragmatics* [Cambridge Studies in Linguistics]. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781139108676

Alexander-Bakkerus, Astrid. 2011. Some Cholón discourse particles and Quechua homologues. In *Lenguas indígenas de América: morfología y sintaxis*, Andrés Romero-Figueroa (ed), 11–28. Caracas: Mercantil/ Universidad Católica Andrés Bello.

Balusu, Rahul. 2019. Fine tuning the Dravidian left periphery: The three 'complementizers' in Telugu. In *Proceedings of the workshop "Clause typing and the syntax-to-discourse relation in head-final languages"*, Josef Bayer & Yvonne Viesel (eds), 13–38. Konstanz: Fachbereich Linguistik der Universität Konstanz.

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: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199746736.003.0001

Bayer, Josef, Hinterhölzl, Roland & Trotzke, Andreas 2015. *Discourse-oriented syntax*. Amsterdam: John Benjamins. https://doi.org/10.1075/la.226

Bayer, Josef & Struckmeier, Volker. 2017a. *Discourse Particles: Formal Approaches to their Syntax and Semantics*. Berlin: De Gruyter.

Bayer, Josef & Struckmeier, Volker. 2017b. The status quo of research on discourse particles in syntax and semantics. In *Discourse Particles: Formal Approaches to their Syntax and Semantics*, Josef Bayer & Volker Struckmeier (eds), 1–14. Berlin: De Gruyter.

Bhadra, Diti. 2017. Evidentiality and questions: Bangla at the interfaces. PhD dissertation, Rutgers University.

Biberauer, Theresa and Sheehan, Michelle. 2011. Introduction: particles through a modern syntactic lens. *The Linguistic Review* 28(4): 387–410. https://doi.org/10.1515/tlir.2011.011

Boeder, Winfried. 2002. Speech and thought representation in the Kartvelian (South Caucasian) languages. In Tom Güldemann & Manfred von Roncador. (eds), 3–49, *Reported discourse: A meeting ground for different linguistics domains*. https://doi.org/10.1075/tsl.52.04boe

Cardinaletti, Anna. 2011. German and Italian modal particles and clause structure. *The Linguistic Review* 28(4): 493–531. https://doi.org/10.1515/tlir.2011.014

Coniglio, Marco. 2007. German modal particles in the IP-domain. *Rivista di Grammatica Generativa* 32: 3–37.

Coniglio, Marco. 2008. Modal particles in Italian. *Working Papers in Linguistics* 18: 91–129.

Corr, Alice. 2016. Ibero-Romance and the Syntax of the Utterance. PhD dissertation, University of Cambridge.

Degand, Liesbeth, Cornillie, Bert & Pietrandrea, Paola. 2013. *Discourse markers and modal particles: Categorization and description*. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.234

Del Gobbo, Francesca, Munaro, Nicola & Poletto, Cecilia. 2015. On sentential particles: A cross-linguistic study. In *Final particles*, Sylvie Hancil, Alexander Haselow & Margje Post (eds), 359–386. Berlin: De Gruyter. https://doi.org/10.1515/9783110375572-015

Euskaltzaindia. 1987. *Euskal gramatika. Lehen urratsak II*. Bilbo: Euskaltzaindia.

Farkas, Donka. & Bruce, Kim. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27: 81–118. https://doi.org/10.1093/jos/ffp010

Friedmann, Naama, Belletti, Adriana & Rizzi, Luigi. 2021. Growing trees: The acquisition of the left periphery. *Glossa: A Journal of General Linguistics*. 6(1: 131): 1–38. https://doi.org/10.16995/glossa.5877

Grosz, Patrick. 2006. German discourse particles are weak sentence adverbs. In *Proceedings of the 'XXXII Incontro di Grammatica Generativa*, Maria C. Picchi & Alan Pona (eds), 79–91. Florence: Edizioni dell'Orso.

Grosz, Patrick. 2016. Information structure and discourse particles. In *The Oxford handbook of information structure*, Caroline Féry & Shinichiro Ishihara (eds), 336–358. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199642670.013.36

Hack, Franziska M. 2014. The particle *po* in the varieties of Dolomitic Ladin – grammaticalisation from a temporal adverb into an interrogative marker. *Studia Lingüística* 68(1): 49–76. https://doi.org/10.1111/stul.12022

Haegeman, Liliane. 2014. West flemish verb-based discourse markers and the articulation of the speech act layer. *Studia Lingüística* 68(1): 116–139. https://doi.org/10.1111/stul.12023

Hancil, Sylvie, Haselow, Alexander & Post, Margje. 2015. *Final particles*. Berlin: De Gruyter. https://doi.org/10.1515/9783110375572

Heim, Johannes, Keupdjio, Hermann, Lam, Zoe W.M, Osa-Gómez, Adriana, Thoma, Sonja & Wiltschko, Martina. 2016. Intonation and particles as speech act modifiers: A syntactic analysis. *Studies in Chinese Linguistics* 37(2): 109–129. https://doi.org/10.1515/scl-2016-0005

Hinterhölzl, Roland & Munaro, Nicola. 2015. On the interpretation of modal particles in non-assertive speech acts in German and Bellunese. In *Discourse-oriented syntax*, Josef Bayer, Roland Hinterhölzl & Andreas Trotzke (eds), 41–70. Amsterdam: John Benjamins. https://doi.org/10.1075/la.226.03hin

Izutsu, Katsunobu & Mitsuko N. Izutsu. 2013. From discourse markers to modal/final particles. In *Discourse markers and modal particles: Categorization and description*, Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds), 217–236. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.234.09izu

Jacobs, Joachim. 1991. On the semantics of modal particles. In *Discourse particles: descriptive and theoretical investigations on the logical, syntactic, and pragmatic properties of discourse particles in German*, Werner Abraham (ed), 141–162. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.12.06jac

Kuong, Io-Kei J. 2008. Yes/no question particles revisited: The grammatical functions of *mo4*, *me1*, and *maa3*. In *The 20th North American Conference on Chinese Linguistics (NACCL-20)*, Marjorie Chang (ed), 715–733. Ohio: The Ohio State University.

Kuwabara, Kazuki. 2013. Peripheral effects in Japanese questions and the fine structure of CP. *Lingua* 126: 92–119. https://doi.org/10.1016/j.lingua.2012.12.004

Lam, Zoe W. 2014. A complex ForceP for speaker-and addressee-oriented discourse particles in Cantonese. *Studies in Chinese Linguistics* 35(2): 61–80.

Modicom, Pierre-Yves and Duplâtre, Olivier. 2020. *Information-structural perspectives on discourse particles*. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.213

Pankau, Andreas. 2018. The Syntax of '*n* in North East Berlin German. In *the workshop Recent Issues in the Syntax of Questions*. Konstanz: Universität Konstanz, 2018/10/8–10.

Paul, Waltraud. 2014. Why particles are not particular: Sentence-final particles in Chinese as heads of a split CP. *Studia Linguistica* 68 (1): 77–115. https://doi.org/10.1111/stul.12020

Rijk, Rudolf P. G. de. 2008. *Standard Basque: A progressive grammar*. Cambridge MA: MIT Press.

Scherf, Nathalie. 2017. The syntax of Swedish modal particles. In *Discourse Particles: Formal Approaches to their Syntax and Semantics*, Josef Bayer & Volker Struckmeier (eds), 78–99. Berlin: De Gruyter.

Struckmeier, Volker. 2014. Ja doch wohl C? Modal Particles in German as C-related elements. *Studia Lingüística* 68(1): 16–48. https://doi.org/10.1111/stul.12019

Tan, J. 2019. Evidentiality in Tagalog. PhD dissertation, Universidad Autónoma de Madrid.

Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen*. Berlin: De Gruyter. https://doi.org/10.1515/9783111354569

Van Bogaert, Julie & Leuschner, Torsten. 2015. Dutch *('t) schijnt* and German *scheint (') s*: on the grammaticalization of evidential particles. *Studia Lingüística* 69(1): 86–117. https://doi.org/10.1111/stul.12030

Van Eys, Willem J. 1879. *Grammaire comparée des dialectes basques*. Paris: Maisonneuve.

Van Kemenade, Ans & Links, Meta. 2020. Discourse particles in early English: Clause structure, pragmatics and discourse management. *Glossa: a journal of general linguistics* 5(1): 3. 1–23. https://doi.org/10.5334/gjgl.1020

Wiltschko, Martina. 2017. Ergative constellations in the structure of speech acts. In *The Oxford Handbook of Ergativity*, Jessica Coon (ed), 419–446. New York: Oxford University Press.

PART I

Diachronic issues and the development of discourse particles

CHAPTER 1

On the adverbial origin of German modal particles

Marco Coniglio
Georg-August-Universität Göttingen

All German modal particles share important common properties. However, in a diachronic perspective, their origin has often been explained by assuming that they have grammaticalized from different types of lexemes belonging to several word classes: adjectives, adverbs, conjunctions, etc. The present paper intends to dissipate some erroneous assumptions about the grammaticalization paths of modal particles and to offer a novel syntactic approach that explains their origin and development. By following and elaborating on some recent ideas, I will explore the hypothesis that modal particles have an adverbial origin and will provide corresponding evidence. In the syntactic analysis, I will claim that all modal particles originate from specific types of (strong) lower adverb(ial)s that become weak sentential adverbs under reanalysis.

Keywords: adverbial, grammaticalization, modal particle, sentence adverb

Vielleicht sollte man also besser von Abtönungsadverbien sprechen.[1]
(Burkhardt 1994: 148)

1. Introduction

In this paper, I intend to address the issue of the origin of modal particles – henceforth MPs – in German (and possibly in other languages) and – from a syntactic perspective – to answer the question of whether they originate from heterogeneous classes of possible "donor lexemes". While several grammaticalization paths have been proposed for MPs, it is a desideratum from a syntax-theoretical perspective to derive the observed diachronic processes from more general principles. Another

1. "So maybe it would be better to use the term "shading adverbs". (my translation). Note that some authors – among which Burkhardt – use the term *Abtönungspartikeln* 'shading particles', instead of *Modalpartikeln* 'modal particles'.

https://doi.org/10.1075/la.276.01con

important question is how to interpret the diachronic facts for the analysis of the synchronic properties of MPs in German (and in other languages displaying similar particles), such as their restriction to the middle field.

In the following sections of the paper, I will first deal with the definition and the general properties of German MPs (Section 2). In Section 3, the grammaticalization process and the possible donor lexemes of MPs will be presented, first according to the predominant literature and then according to more accurate hypotheses formulated in recent times. Some empirical facts shedding some light on the genuine nature of the donor lexemes of MPs will be presented in Section 4. Section 5 will put forward a novel syntactic analysis of the grammaticalization of MPs, which is compatible with the empirical observations.

2. Definition and general properties of modal particles

MPs are a group of about twenty words in German, which typically display – among others – the following properties (cf. for example Thurmair 1989):[2]

1. Phonology and prosody: MPs are unstressed and prosodically integrated:

   (1) Hast du sie *denn* / **DENN* angerufen?
   have you her PRT PRT called
   'Did you <u>really/finally</u> call her at all?'

2. Semantics and pragmatics: MPs have no propositional meaning (but see Gutzmann 2008, 2009). In (1), for instance, it is difficult to determine the semantic contribution of the particle *denn* (lit. 'then'). This is due to the fact that MPs express the speaker's attitude to the propositional content of the utterance and this mostly depends on the pragmatic context (cf. Schneider, this volume). The particle *denn* typically expresses the speaker's concern or interest (Bayer 2012; also cf. Thurmair 1989: 163ff). It is stated in the literature that MPs operate as modifiers of the illocutionary force of the utterance (Jacobs 1986; Thurmair 1989; Abraham 1991b; Coniglio 2011; Coniglio & Zegrean 2012, but see Autenrieth 2002: 24ff).[3]

---

2. The following items are traditionally classified as MPs: *aber*, *auch*, *bloß*, *denn*, *doch*, *eben*, *eh*, *eigentlich*, *einfach*, *etwa*, *halt*, *ja*, *mal*, *nur*, *ruhig*, *schon*, *sowieso*, *vielleicht*, *wohl* (cf. Thurmair 1989: 21). Other items can be or have been included in this class. For practical reasons, I do not provide a direct translation for each particle in the examples, but I will translate the sense of the particle and will underline the relevant counterpart in the English translation of each example.

3. Following Searle (1976), there are several types of possible illocutionary forces realizing "illocutionary points". This means that MPs do not "modify", in fact they "change" the illocutionary force of an utterance. Nonetheless, I will use the established term "modify" in this paper.

3. Syntax: MPs are syntactically integrated and can only occur in the middle field of the clause, more precisely in the higher part of the I/T-domain (Coniglio 2005, 2011), being typically excluded from the so-called prefield and from the postfield (2). MPs can be stacked, but not coordinated (3), and they cannot be modified (4) (see Thurmair 1989; Müller 2018).

(2) <* *denn* > hast du sie < *denn* > angerufen <* *denn* > ?
PRT have you her PRT called PRT
'Did you <u>really/finally</u> call her?'

(3) Hast du sie *denn* (*und) *überhaupt* angerufen?
have you her PRT and at.all called
'Did you <u>actually</u> (*and) <u>really/finally</u> call her at all?'

(4) Hast du sie (* *sehr*) *denn* angerufen?
have you her very PRT called
'Did you (*very) <u>really/finally</u> call her?'

In this paper, only MPs are investigated, which – as shown – are defined as prosodically and syntactically integrated particles that occur in the middle field (I-particles). Rarely, in German (and in other languages), prosodically and syntactically integrated (phrasal or adverbial) particles can also be found that occur in the prefield (C-particles), such as *ein Glück* 'such a luck', *kein Wunder* 'no wonder', *klar* 'of course', etc. (cf. Frey 2004):[4]

(5) *Klar* hast du eine Meinung!
PRT have you an opinion
'<u>Of course</u> you have an opinion!'

C-particles – as well as all other left-peripheral particles that are syntactically and prosodically non-integrated and thus CP-external – are not the object of the present investigation, which mainly focuses on the origin and grammaticalization process of I-particles in German (and possibly in other languages displaying similar particles).[5]

---

4. As pointed out in Axel (2007: 27ff) and Petrova (2017), a possible paradigm shift in the particle system could be assumed for German (and possibly other Germanic languages). In German, a C-particle system was possibly gradually replaced by an I-particle system. (Sentence-integrated?) C-particles are also found in older Germanic languages, such as Gothic (cf. e.g. Ferraresi 2005), but also in non-Germanic ones. According to Hentschel (1986: 84), Old High German (OHG) particles were more similar to adverbs than to modern MPs. Axel (2007) claims that left-peripheral particles in OHG (and in other Germanic languages) are the remnant of an older system in which particles could mark the clause type and trigger verb fronting (as shown by Ferraresi 2005). More recently, Petrova (2017) argues that OHG C-particles like *inu* and *ia* have similarities to Modern German MPs (in the sense of Coniglio & Zegrean 2012).

5. See, for example, Paul (this volume) or recent works by Haegeman and Hill (2013, 2014), who distinguish different types of left-peripheral particles. Cf. Cardinaletti (2011, 2015) and Dohi

3. Origin and grammaticalization of German modal particles

3.1 Donor lexemes – a short state of the art

The origin and development of the German MPs has been often described as a typical case of grammaticalization, but, here, I will not try to discuss the huge amount of literature written on the grammaticalization process of (individual) MPs (but cf. Hentschel 1986; Abraham 1991a; Burkhardt 1994; Diewald 1997, 1999, 2011; Autenrieth 2002; Molnár 2002; Wegener 2002; Diewald & Ferraresi 2008).

Focusing on the semanto-pragmatic consequences of grammaticalization, it is well-known that MPs develop an increasingly abstract meaning and more pragmatic functions (cf. Abraham 1991a: 373). For Diewald (2011), the notion of "pragmaticalization" – defined as grammaticalization of discourse functions – is crucial in order to explain the origin and development of MPs.

Many works concentrate on the identification of a possible "donor lexeme" (*Spenderlexem*) at the base of the grammaticalization or pragmaticalization process of each MP. Burkhardt's (1994: 138ff) very often quoted work indicates six "typical" starting points for their grammaticalization:[6]

1. conjunctions: *aber*, *doch*, (Low German) *man*, etc.
2. temporal adverbs: *denn*, *eben*, *einmal*, *etwa*, *halt*, *ja*, *jetzt*, *mal*, *noch*, *schon*, etc.
3. sentence adverbs: *eigentlich*, *gefälligst*, *schließlich*, *vielleicht*, etc.
4. "real" adverbs: *einfach*, *fein*, *hübsch*, *irgendwie*, *langsam*, *schlicht*, *zufällig*, etc.
5. predicatives: *dreist* (today uncommon), *ruhig*, etc.
6. degree particles:[7] *bloß*, *nur*, etc.

Many other authors recognize different origins for MPs (cf. Hentschel 1986; Abraham 1991a; Diewald 1997; Molnár 2002, a.o.). More recently, Diewald & Ferraresi (2008) and Diewald (2011) discuss further constraints in the grammaticalization process of MPs, which seem to indicate that MPs do not randomly grammaticalize from one class or another, but follow precise paths.

(2020) on Italian. See also Remberger (2020) for an interesting cross-linguistic classification of different types of discourse and pragmatic markers with a focus on Romance languages.

6. Words that are traditionally and (almost) undisputedly identified as MPs are underlined.

7. I will use "degree particles" to translate the German term "Gradpartikeln" used by Burkhardt (1994) and other authors (see Thurmair 1989, a.o.). However, according to other classifications, *bloß* and *nur* are classified as "Fokuspartikeln" 'focus particles', while the term "Gradpartikeln" is used to only refer to words like *ganz* 'entirely' or *sehr* 'very'.

3.2 Donor lexemes – hypotheses

In this paper, I intend to address the following questions related to the origin and development of MPs:

1. What is the origin of German MPs? Do they really grammaticalize from heterogeneous classes of donor lexemes?
2. What kind of grammaticalization paths do they follow?
3. To what extent can the diachronic facts explain the synchronic syntactic properties of the MPs? E.g., how can their restricted distribution in the middle field be explained based on diachronic data?

As to the synchronic aspects, I will adopt Cardinaletti's (2011) view that MPs are weak sentence adverbs (also Grosz 2005; Coniglio 2005, 2011):[8]

> […] we don't need a new syntactic category 'particle' because modal particles are (deficient) sentential adverbs […]. (Cardinaletti 2011: 494)

Cardinaletti (2007: 100) shows that MPs exhibit typical properties of weak elements, because:[9]

1. they are not accented;
2. they have a limited syntactic distribution;
3. they don't have a propositional content.

From a diachronic perspective, some authors have already argued for an adverbial origin of MPs, as is clear, e.g., from Burkhardt's (1994) quotation at the opening of this paper (and more recently, Cardinaletti 2011; Petrova 2017). However, in his analysis, Burkhardt (1994) eventually considers MPs as *semantische Ableger* ('semantic offshoots') of different word classes.

Elaborating on Cardinaletti's (2011) proposal, I would like to formulate the diachronic assumption that motivates the empirical work in this paper as follows:

---

8. Cf. Abraham (1991a: 332f.): "Categorially, i.e. with respect to distributional properties, MPs come closest to adverbials […]."

9. There is an extensive discussion about the syntactic status of modal particles, in particular about their status as heads or maximal projections (see for example Meibauer 1994; Ormelius-Sandblom 1997a, 1997b, a.o.). For reasons of space, I cannot do justice to this discussion in a few pages. Therefore, I will not present arguments in favour of one thesis or the other, but see Coniglio (2011: 99ff.) for a discussion of some of the proposals made in the literature. In this paper, I will refer for simplicity to Cardinaletti's (2007, 2011) theories, which in a sense represent an intermediate position in the discussion, since she analyzes MPs as phrasal, but as reduced phrases and thus with a head-like behaviour.

(6) *MPs as "weakened" adverbs*
If – synchronically – MPs are weak sentence adverbs, then – diachronically – they must have developed from strong adverbs. (cf. Cardinaletti 2011)

Below, I will argue – also elaborating on Diewald (2011) – that all MPs can only be traced back to adverbs, and to no other word classes. If other word classes are involved as possible donor lexemes, an adverbial intermediate stage must always be assumed (cf. Abraham 1991a; Diewald 1997, 2011; Petrova 2017; but see Autenrieth 2002).

Considering the time of origin of German MPs reported in Table 1, taken from Burkhardt (1994: 140), we observe that most MPs grammaticalized only in the last few centuries (see also Hentschel 1986; Diewald 1997; Autenrieth 2002, gea.o.).[10]

Table 1. Origin of German MPs (adapted from Burkhardt 1994: 140)

OHG	MHG	16th c.	17th c.	18th c.	19th c.	20th c.
denn, doch	da, halt	dreist†, eben, eigentlich, freilich, man, ja, je*, jetzt, nämlich, nicht, nur, wohl	allerdings, aber, auch, einmal, etwa, natürlich, schon, wenigstens	einfach, erst, fein, gleich, hübsch, immerhin, jedenfalls, mal, ohnehin, so, sowieso, überhaupt, übrigens	bloß, eh, gefälligst, noch, nochmal, ruhig, schließlich, schön, vielleicht	irgendwie, dabei, langsam, schlicht, zufällig

As conceded by Burkhardt himself (1994: 139) and argued for in Molnár (2002: 119ff), some dates might not be correct. Nonetheless, the table shows that almost all MPs grammaticalized only in recent time, especially starting from the Early New High German period (ENHG, 1350–1650), the "epoch of incipient subjectivation and epistemification" (Molnár 2002: 23). Abraham (1991a: 336) also links the rise of MPs to the increasing role played by the middle field during the ENHG period.

As to the grammaticalization paths of MPs, there are several interesting proposals in the literature (cf. for instance Hentschel 1986; Burkhardt 1994; Molnár 2002, a.o.). Here, I will espouse one central idea of two main proposals and explore their theoretical consequences for a syntactic approach to the grammaticalization of MPs. The first proposal is the one by Abraham (1991a: 373), who describes the following path for grammaticalization processes in general, and for adverbs and MPs specifically, and thus recognizes an adverbial origin for MPs (cf. Coniglio 2011: 103):

10. OHG = Old High German; MHG = Middle High German; * = no longer a MP; † = outdated. Only elements that are traditionally considered to be MPs are underlined.

(7) localistic > temporal > logical > illocutive/discourse functional
(Abraham 1991a: 373)

Diewald (2011: 377ff) basically shares a similar view of an adverbial origin of the MPs. However, in other works, she leaves open – like Abraham – the possibility that MPs also originate (possibly via further intermediate stages) from degree particles, conjunctions and adjectives (see also Diewald 1997: 73ff., 99).

In this paper, I want to adopt Diewald's hypothesis by formulating it in a stronger way, namely that all MPs are in fact derived – without exclusions – from adverb(ial)s, more exactly from lower adverb(ial)s, i.e. from aspectual or temporal adverbs that occupy the lowest projections in Cinque's (1999) hierarchy of functional projections (which encode information more close to the propositional content of the utterance). What matters is not the original meaning and function of the donor lexeme of a certain MP, but that it be a (low) adverb(ial) at some stage. Thus, I claim that all MPs can grammaticalize only starting from adverb(ial)s, independently of what the origin of this adverb(ial) is. This would lead us to dismiss an approach à la Burkhardt (1994: 138ff), who recognizes several grammaticalization paths for MPs. Furthermore, this analysis has the advantage of explaining the grammaticalization process in terms of "syntactic climbing" in Roberts & Roussou's (2003: 71) terms. According to them, language change is often associated with the reanalysis of a lexical or functional element to a functional element that occupies a higher position in the syntactic structure.

At this point, some clarifications are necessary, which will be discussed in more detail below:

1. MPs have not (directly) derived from adjectives (also cf. Diewald 2011).
2. MPs have derived neither from conjunctions nor from degree particles.
3. Sentence adverbs are rather the result of the grammaticalization process than the donor lexemes of MPs.
4. Temporal adverbs can only become donor lexemes after a "logical" intermediate stage à la Abraham (1991a), in which they lose their temporal meaning through metaphorical processes (cf. Diewald 1997 and the notion of *categorial metaphor*).

4. Which donor lexemes? Some empirical facts

The assumptions underlying this section can be summarized as follows:

(8) *Assumptions*
1. If we assume that MPs are – synchronically – weak(ened) adverbs,
2. and if we prove that all MPs – diachronically – grammaticalized at a time when the corresponding (strong) adverbs were (still) present,

→ then it is legitimate to assume that MPs originated from these (strong) adverbs.

In the following section, I will discuss some candidates for each of the classes proposed by Burkhardt (1994), and summarize the results for the other MPs in the Appendix.[11] I will show that all MPs can be traced back to adverb(ial)s in some way.

For the empirical investigation, I used the following corpora, archives and dictionaries (listed according to the abbreviations used in the following sections):

DTA *Deutsches Textarchiv*. Grundlage für ein Referenzkorpus der neuhochdeutschen Sprache. Herausgegeben von der Berlin-Brandenburgischen Akademie der Wissenschaften, Berlin 2020. URL: http://www.deutschestextarchiv.de/.

DWB *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm*, 16 Bde. in 32 Teilbänden. Leipzig 1854–1961. Quellenverzeichnis Leipzig 1971. Online: http://dwb.uni-trier.de/de/.

FnhdC *Das Bonner Frühneuhochdeutsch-Korpus*, Korpora.org, http://www.korpora.org/FnhdC/.

MWB *Mittelhochdeutsches Wörterbuch online*. Mainzer Akademie der Wissenschaften und der Literatur und Akademie der Wissenschaften zu Göttingen. http://www.mhdwb-online.de/.

ReA *Referenzkorpus Altdeutsch*. Donhauser, Karin; Gippert, Jost; Lühr, Rosemarie (2018). Deutsch Diachron Digital – Version 1.1. Humboldt-Universität zu Berlin. Homepage: http://www.deutschdiachrondigital.de/.

ReM *Referenzkorpus Mittelhochdeutsch (1050–1350)*. Klein, Thomas; Wegera, Klaus-Peter; Dipper, Stefanie; Wich-Reif, Claudia (2016). Version 1.0. https://www.linguistics.ruhr-uni-bo-chum.de/rem/. ISLRN 332–536–136–099–5.

11. As indicated in the Appendix, even a problematic and much-discussed particle such as *ja* can indirectly be traced back to a temporal adverb *ie/je/ye* 'ever' through a possible contamination with the OHG/MHG introductory sentence particle *ia* (Molnár 2002: 101; see also Meibauer 1994: 165, following Paul 1992: 435), as demonstrated by the many ENHG examples in which *ie/je/ye* already had the function of the present German particle (cf. Molnár 2002: 97f).

4.1 "Real" adverbs

That MPs can be derived from genuine adverbs is assumed by hypothesis. Nonetheless, let us consider a typical example of a MP that is considered to have grammaticalized from an adverb by Burkhardt (1994), namely *einfach*.

Einfach

The MP *einfach*, connected to the homophone adverb meaning 'simply, in a simple way', was actually an adjective at its origin (cf. Autenrieth 2002: 176ff.). The adjective is attested very early and still present in German. However, the adverbial use (with the meaning 'just once, not many times') is attested relatively late, with sporadic examples already in the 16th century in the DTA, but increasingly used in the late 18th-early 19th century:

(9) Was er *einfach* ausgegeben/ Kommt ihm über Haubt und Glieder
what he just.once out-given/ comes to.him over head and limbs
Offt mit reichem Wucher wieder. (1704, Abschatz, DTA)
often with rich usury again
'What he has spent only once often returns to him with rich usury over head and limbs.'

The first unambiguous examples of *einfach* with a clear MP interpretation ('just, simply') I could find in the DTA only date back to the 19th century (but Burkhardt 1994 mentions examples from the 18th century):

(10) […] so dass man versucht sein kan anzunehmen, der Künstler habe
so that one tempted be can to-assume the artist has.SUBJ
einfach ein kleines Erzbild copirt von der Art der
PRT a little mineral.sculpture copied of the type the.GEN
clipeati, […]. (1853, Brunn, DTA)
clipeati
'[…] so that one can be tempted to assume that the artist simply copied a small stone sculpture in the style of the clipeati, […]'

Many of the other potential MPs mentioned by Burkhardt (1994) are in fact similar to *einfach* in that they originate from adverbs of manner (most of them being originally adjectives).

4.2 Predicatives

With "predicatives", Burkhardt (1994) refers to adjectives used in predicative functions. Adjectives can be donor lexemes for MPs (also cf. Diewald 2011), but I am arguing that, crucially, they can be donor lexemes only because they can be used

adverbially (similarly to what was shown for *einfach*). In fact, adjectives and adverbs in German cannot always be distinguished morphologically. In the German tradition, this very broad class is referred to as *Adjektivadverbien* 'adjective adverbs'. They are adjectives that can also be used adverbially, in particular as adverbs of manner. Below, I will concentrate on *ruhig*.

Ruhig

The MP *ruhig*, which is connected to the adjective meaning 'quiet', has in fact – as an "adjective adverb" – also an adverbial counterpart, which means 'quietly'. It started as an adjective and developed an adverbial use (only) in the 18th century, with the meaning 'without movement, quietly'. Here is an example of this adverbial usage:

(11) Was du beſtåtiget haſt, laß *ruhig* darinnen ſtecken.
what you fastened have, let still inside stick
'What you fastened, leave it inside (without moving it).' (1724, Fleming, DTA)

Unambiguous MP uses can only be found in the 19th century in the DTA:

(12) Geht *ruhig* nach Hause und kommt wieder in ein Paar Stunden!
Go *PRT* to house and come again in a couple hours
'(Don't worry!/Relax!) Go home and come back in a few hours!'
(1822, Holbergs Lustspiele)

Thus, *ruhig* – similarly to *einfach* – grammaticalized from an adjective via an adverbial stage. Crucially, the particle only appeared when the adverbial use had already grammaticalized.

4.3 Sentence adverbs

As we will see below, sentence adverbs seem to be rather the results of grammaticalization processes than donor lexemes for MPs. We will consider *eigentlich* and *vielleicht* separately, which are often considered as representative examples of MPs originating from sentence adverbs.

Eigentlich

The MP *eigentlich* is often led back to a homophonous sentence adverb (lit. 'actually'), but in fact it cannot be clearly distinguished from it in present-day German. The origin of this MP (and of the adverb) is to be searched for in an adverb of manner meaning 'in a proper way'. This in turn stems from an adjective ('proper'), which acquired an adverbial function and thus became an "adjective adverb" in the 13th century, i.e. earlier than the examples discussed so far (and was still very common later), as illustrated by the following example:

(13) Hie redet der gedültige Hiob aus eigener erfahrung gar *eigentlich*
here speaks the patient Job from own experience really properly […]
'Here the patient Job talks from his own experience in a really good/proper way […]' (1585, Braun, DTA)

The first cases of *eigentlich* used as a MP can be found in the 16th century, i.e. only after the adverbial use had grammaticalized:

(14) Diser Span/ meines einfåltigē verstandts / erhebt sich *aygentlich*
this tension my.GEN simple intellect. GEN elevate himself *PRT*
nicht vber diese Frag (1557, Jacob Andreae, FnhdC)
not over this question
'This tension of my obtuse mind does not really rise above this question'

Vielleicht

It is a common assumption that the MP *vielleicht* (15) grammaticalized from the homophonous sentence adverb with the meaning 'maybe, perhaps':

(15) Er ist *vielleicht* frech!
he is *PRT* impertinent
'He is really impertinent!'

Some authors argue that this happened only in recent times (in the 19th century, according to Burkhardt 1994). However, this late dating would hardly explain the "asseverating" function of the particle as set against to the semantic contribution of the (strong) sentence adverb. In (15), the speaker is not saying that a certain person is possibly impertinent, but that (s)he really is. Much more probable is thus the origin of the MP – exactly like the (strong) sentence adverb – in the MHG adverbial phrase *vil lîhte(r)* ('very easily', figuratively: 'with certainty', etc.), which had a clear asseverative meaning:

(16) […] daz tuot mir nû *vil lîhte* wê. (12th c., Reinmar d. A., MWB)
that does to.me now very easily woe
'[…] this hurts me now very easily.'

Thus, MP and (strong) sentence adverb probably grammaticalized independently from one another starting from the same strong adverb. In the case of *vielleicht* with the meaning 'maybe, perhaps', the asseverative meaning got weakened, but not in the case of the MP *vielleicht*. This amounts to saying that the MP is probably much older than supposed so far. In fact, one can find some potential examples of an early use of the MP in MHG:

(17) […] "sinnelôser Thêodâ, / dû bist *vil lîhte* âne vernunst,/ ein tumber
foolish Theoda you are *PRT* without reason a stupid
esel, âne kunst. /[…]" (ca. 1220, Rudolf von Ems, MWB)
donkey without knowledge
'Foolish Theoda, you are really deprived of reason, a stupid donkey, without any knowledge.'

4.4 Temporal adverbs

Temporal adverbs can become MPs, but only after a logical-connective reinterpretation in the sense of Abraham (1991a). Donor lexemes for MPs are lower temporal adverbs, which are mostly found in lower or aspectual projections in Cinque's (1999) hierarchy (also cf. Cinque 2001, 2006). In fact, they originally have meanings like 'before' (*eh*, *halt*), 'thereafter' (*denn*), 'already' (*schon*), etc. Thus, temporal donor lexemes for MPs emerged by means of an intermediate step in which a certain temporal relation was reinterpreted as logical (cf. Diewald's 1997 use of the notion of "categorial metaphor"), as the clear example of *denn* shows.

Denn
The first necessary step for the grammaticalization of the MP *denn* was the reinterpretation of the temporal adverb *denn* to a logical-connective adverb *denn* (cf. Molnár 2002; Bayer 2012; Ferraresi 2014; and for the origin in the causal meaning of *denn*, see Wegener 2002). This is exemplified by (18) and Figure 1.

(18) [Question to Rumpelstiltskin:]
Wenn du nicht Rippenbiest, Hammelswade, Schnürbein heißt, wie
if you not Rippenbiest Hammelswade Schnürbein are.called how
heißt du *DENN*?
are.called you *DENN*
'If you are not called Rippenbiest, Hammelswade, Schnürbein, what is your name then?'

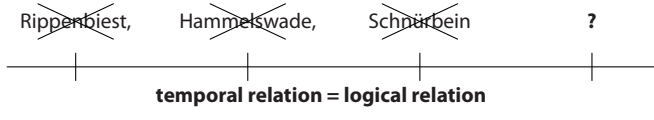


Figure 1. From temporal *denn* to logical *denn*

The MP *denn* originates from the temporal adverb – in OHG, *thanne*/*thenni* 'thereupon, after that, then' (which in turn comes from a locative adverb/demonstrative, cf. Wegener 2002: 384; Diewald 2011: 382; Ferraresi 2014: 81) – only after such a necessary reinterpretation (cf. Diewald 1997: 88; Ferraresi 2014: 80ff; cf. also Cognola & Cruschina 2020 on the grammaticalization of Italian *poi*).

The MP does not originate from the homophonous (logical) conjunction *denn* ('because'), in contrast to what is assumed for instance by Molnár (2002: 54), who proposes a grammaticalization path of the following type:

(19) temporal adverb *dann* > (logical) conjunction *denn* > MP *denn*.
(adapted from Molnár 2002: 54)

Dann/*denn* was hardly used as a conjunction before the 15th century (Ebert et al. 1993: 473). While I think she is right in assuming – in Abraham's (1991a) spirit – the order temporal > logical > illocutionary, she dates the MP use too late. If – as argued for instance by Burkhardt (1994) – the MP already existed in OHG, then it would be anachronistic to claim that the MP originates from the logical conjunction rather than from a logical adverb. In fact, *denn* seems to be one of the oldest MPs (cf. Burkhardt 1994: 140; Diewald 1997: 89; Molnár 2002), possibly showing counterparts in other Germanic languages, such as in Old English (cf. van Kemenade & Links 2020 on OE *þonne*). A more likely interpretation is that – as assumed above – an adverb expressing "logical consequence" originates via a metonymic process from the temporal adverb (see Molnár 2002: 52), but that the causal conjunction and the MP independently originates from this "logical" or connective adverb.

Example (20) could be a very early example for the MP use of *denn*:

(20) Thiu óugun sie imo búntun […] /joh frágetun ginúagi, wer ínan
the eyes they to.him bound and asked enough who him
thane slúagi (9th c., Otfrid IV, 19, 73, in Burkhardt 1994: 143)
PRT beat.SBJCV
'They blindfolded him and asked (often enough) who had beaten him'

4.5 Conjunctions

Despite different opinions in the literature, I intend to argue that conjunctions cannot be donor lexemes for MPs. Below, I will consider the MPs *aber* and *doch* separately.[12]

12. To avoid misunderstandings, I would like to emphasize that by "conjunctions" I am referring to typical uses of these elements in positions that are (only) external to the sentence or the phrase (see *but* in English). It is clear that both *aber* and *doch* have conjunctive uses even in modern German, which are not easily distinguishable from MPs, but these uses are traditionally classified as "conjunctive adverbs". These can also be internal to the sentence (see *however* in English). This internal position (e.g. in the middle field) makes conjunctive adverbs difficult to distinguish from MPs. However, when I say that conjunctions cannot be a source of grammaticalization of modal particles, I intend to refer explicitly to conjunctions and not to conjunctive adverbs.

Aber

It is often assumed that this MP – which has homophones in different word classes and can be literally translated with 'but, however' – is derived from the adversative conjunction. This is stated for example in Burkhardt (1994); Diewald (1997: 80); Molnár (2002); a.o. In (21) below, I tried to summarize Diewald's (1997: 80ff) proposal for the grammaticalization of *aber*, in which she assumes that the MP has its origin in the adversative conjunction (although, with "adversative conjunction", she possibly refers to the conjunctional uses of the adversative adverb):

(21)

Germ. comparative of *ab*	'off, away'
>OHG iterative adverb^{+} *afur*/*abur*	'again, over again, in turn'
> OHG adversative adverb*	'however, though'
> OHG adversative conjunction*	'but, however'
> MHG modal particle*	

$^{+}$ = up to the 18th c.!
* = until today

I would like to argue that the grammaticalization of the MP starting from a conjunction to an adverbial element is unlikely because (1) this would be problematic from a syntactic point of view since it would represent a case of lowering and, in fact, it is not clear which bridging contexts could be responsible for the syntactic reanalysis (conjunctions – operating at sentence level – are confined to the left periphery, while MPs are clause-internal); and because (2) it would represent a case of degrammaticalization (from a more grammatical to a less grammatical element – given its adverbial nature and sometimes still available semantics), which is quite rare, if not impossible (but see for example Norde 2009).

According to Diewald (1997: 81), the first attestations of this MP can be found in the 15th century (thus earlier than proposed in Burkhardt 1994):

(22) herzen myn allerbegirlichistes, liebs Suzelin, wy will ich *aber* so mit
heart my of.all-dearest dear Suzel how will I *PRT* so with
gross froden gwarten der botschafft.
great joys wait.for the message
'My dearest Suzelin, I will wait for your message with such a great joy!'
(1483, Privatbr. d. Ma. 2,76 S., adapted from Diewald 1997: 81)

Diewald (1997: 81) finds very early examples of *aber*, for which she recognizes a connective, but not a MP function. However, we cannot exclude even earlier evidence for this MP, namely in OHG, as possibly exemplified by the sentence in (23). In most of the examples from the corpora, however, it is almost impossible to reliably ascertain the function of each element in the specific context.

(23) […] zi themo hohen hímilriche, In hoho gúallichi theist *avur* thaz
to the high heavenly-kingdom in high glory this-is PRT the
hímilrichi; […] (9th c. , O I, 28, 12ff., ReA)
heavenly-kingdom
'[…] to the high heavenly kingdom. In great glory. This is indeed the heavenly kingdom; […]'

Doch

Originally a demonstrative adverb (Germ. **þau-h*, cf. Kluge 2011: 208), *thoh*/*doh* was already used in OHG in several functions, namely as a conjunction ('but'), as an adversative adverb ('however, nevertheless'), as a concessive adverb ('though, nevertheless') and as a MP, as in the following example (Diewald 1997: 90f; Molnár 2002: 107f; see also Hentschel 1986: 87ff; and Abraham 1991a: 358ff):

(24) […] inti batun inan thaz sie *thoh* tradon sinis qiuuates
and asked him that they PRT thread his.GEN garment.GEN
ruortin […] (9th c. Tatian 82,1, adapted from Diewald 1997: 91)
touch.SBJV
'[…] and asked him that they touch a thread of his garment.'

Its origin from an adversative conjunction is unlikely (*pace* Burkhardt 1994). Not only does the frequency of the adverb in OHG (confirmed by data from ReA) testify to the fact that the adverbial usage was probably at the origin of the grammaticalization of both the conjunction and the MP, but also further considerations already shown for *aber*: namely that, on the one hand, syntactic lowering would be hard to explain and, on the other, that the path from a conjunction to an adverb would constitute a rare case of degrammaticalization.

4.6 Degree particles

As to the question of whether some MPs derived from degree particles, I intend to suggest that degree particles grammaticalized at the same time as MPs, but they are not donor lexemes for MPs. We will consider *nur* and *bloß* below.

Nur

In the literature, the particle *nur* (lit. 'only') is mostly traced back to the homophonous degree particle (Burkhardt 1994; Diewald 1997: 97), which in turn is derived from the OHG locution *ni wâri* ('were it not', MHG *ne wære*). However, at the supposed time of origin of the MP, *nur* could also be used as an adverb. In fact, this use is still attested today, as the following examples show, in which *nur* can be used both in the prefield and in the middle field:

(25) Er ist sehr intelligent.
he is very intelligent
a. *Nur* kann er sich nicht konzentrieren.
PRT can he himself not concentrate
b. Er kann sich *nur* nicht konzentrieren.
He can himself *PRT* not concentrate
'He is very intelligent. He <u>just</u> cannot concentrate.'

In historical examples, the adverb *nur* could even be modified by degree particles (*so* 'so', *ganz* 'fully', etc.), which is a clear indication of its adverbial status:[13]

(26) Dies ist das Verhåltniß alle Zeit. […] So *nur* giebt es in der Zeit
this is the relation of.all time so only exists it in the time
eine Vergangenheit und Zukunft, die sich selbst nicht
a past and future which them selves not
widersprechen; […] (1799, Schlegel, DTA)
contradict
'This is the relation of the whole time. […] <u>But</u>, in the time, there is a past and a future, which do not contradict themselves; […]'

In the DTA, first unambiguous evidence of the MP can be found in the 17th c. (see also Diewald 1997: 97):

(27) ach sehet *nur*, wie hat sie so ein glatte stirn!
oh look *PRT* how has she so a smooth forehead
'Oh, look how smooth her forehead is!' (1669, Simplic. 1, 75, 7, DWB)

Bloß

Bloß ('just') originated from OHG/MHG adjective *blôz* 'naked, unveiled', which then developed an adverbial use (attested until the 19th century). This supports the hypothesis that the MP did not originate from the degree particle (as suggested by Diewald 1997: 87), but rather from the (adjective-)adverb (see Diewald 2011). The following cases exemplify clear adverbial uses, as testified by the fact that they are modified by degree particles:

(28) […] jene aber […] hat die Natur verborgen / daß sie nicht so *bloß*
that however has the nature hidden That she not so merely
sol verschwendet werden. (1699, Abel, DTA)
shall wasted be
'[…] but nature has hidden that, so that it should not be so merely wasted.'

13. There is also an often neglected use of *nur* as a temporal adverb with the meaning 'just (now)' (cf. "ein *nur* gebohren […] töchterlein, […]" 'a newly born little daughter', reported in the DWB), which is attested until the 19th c. and might have also played a role in the grammaticalization of the MP.

(29) Er bat sie, so ganz *bloß* als unbekannter Mahler, sie möchte sich
he asked her so quite simply as unknown painter she might herself
nur völlig frey ihrem Wesen überlassen, […]
only completely free to.her nature leave
'He asked her quite simply as an unknown painter that she only leaves herself completely free to her own nature, […]' (1787, Ardinghello, DTA)

At the same time (possibly in the 19th c.), the MP also grammaticalized (see Diewald 1997: 87):

(30) Meine Briefe gieb *blos* irgend einem herfahrenden Passagier mit.
my letters give PRT any (a) travelling.here passenger with
'Just give my letters to any passenger travelling here.'(1802, Paul, DTA [1960])

4.7 Interim conclusions

Based on the data above, I propose the grammaticalization paths represented in Figure 2 for all MPs considered, which are simpler than many proposals so far.

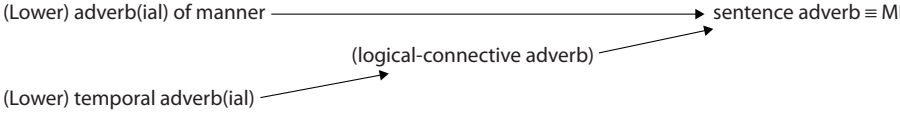


Figure 2. Grammaticalization paths of MPs

The assumptions are in some aspects similar, but – in fact – syntactically stronger than others previously proposed in the literature (see Abraham 1991a; Cardinaletti 2011; Diewald 2011; Petrova 2017). The major claims made here are: (1) only adverb(ial)s and no other lexemes from other word classes are possible (direct) donor lexemes for MPs; and (2) these are defined as lower adverb(ial)s in the sense of modern syntax-theoretical approaches, such as Cinque (1999).

We will see below that, in syntactic-theoretical terms, this stronger assumption can make better predictions than other analyses according to which several donor lexeme types are to be assumed.

5. Syntactic analysis of the grammaticalization of modal particles

From a theoretical perspective, the data discussed above indicate that all MPs have derived from lower adverb(ial)s (cf. Cinque 1999, 2006), which in the course of the grammaticalization process were reanalyzed as higher adverbs (see Roberts & Roussou 2003; van Gelderen 2004; Petrova 2017) and – in contrast to other higher

adverbs, the so-called sentence adverbs – became weak in the sense proposed by Cardinaletti (2007, 2011) (cf. Coniglio 2005, 2011; Grosz 2005). An emerging MP must therefore go through the following steps:

1. The original lexeme is first realized as an adverb(ial) of manner, or a temporal adverb(ial) in the lower I-domain gets a possible adverbial interpretation there (and later gets "logical" interpretation in Abraham's 1991a: 373 sense), thus becoming a possible donor lexeme for a MP.
2. The low adverb(ial) can be fronted to Cinque's (1999) higher I-domain (including projections related to the speech act, mood and modality, etc.) for information- and discourse-structural reasons.
3. The original lexeme undergoes a "semantic split" in the sense of Roberts & Roussou (2003), which leads to the reanalysis of the original adverb(ial) as an illocutionary particle (i.e. as a MP).
4. As a consequence of this reanalysis, the (new) MP is realized in fixed higher syntactic positions in the middle field (see Petrova 2017; cf. the notion of "climbing" in Roberts & Roussou 2003: 71).

The syntactic reanalysis proposed here is graphically shown in Figure 3, in which the syntactic status of the adverb(ial)s is represented on the right and the syntactic operations on the left, respectively (see Axel-Tober & Müller 2017; Müller 2019 for a similar analysis of evidential adverbs).

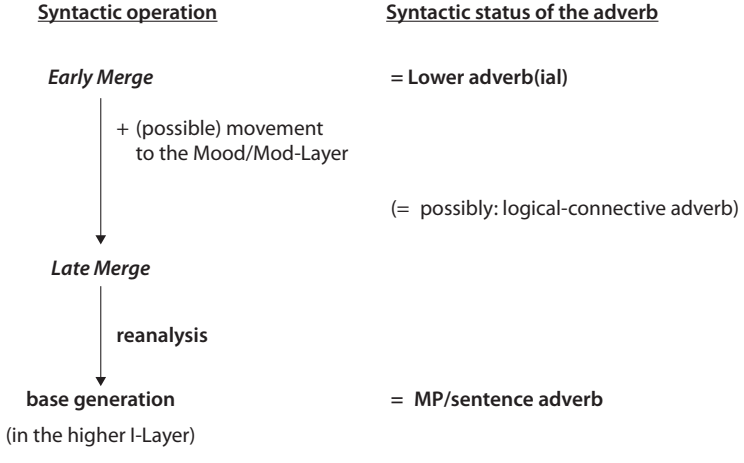


Figure 3. Syntactic reanalysis of low adverb(ial)s to MPs

Let us consider the single steps in the syntactic reanalysis with some examples of the single operations available in each case. As shown, an adverbial use of the donor lexeme must be available at the start, independently of the previous history of this

lexeme. It does not matter whether the adverbial grammaticalizes from an adjective (which is a very common pattern, but not the only one), as in the case of *eigentlich*, or from the adverbial use of a demonstrative, as in the case of *doch*. The adverbial donor lexeme is merged in a (low) base position (*Early Merge*), as illustrated by the following example, in which *eigentlich* follows the negation and therefore must be in a low syntactic position:

(31) Da der König anhielt / sagt er / er könnt dasselb nicht *eigentlich*
when the king stopped said he he could the-same not properly
sagen. (1605, Melander DTA)
say
'When the king stopped, he said that he could not say that exactly.'

As a strong adverb, *eigentlich* was syntactically more flexible and – for information- or discourse-structural reasons – it could move to higher positions in the I-domain (e.g. to a position preceding the negation particle, but still in its semantic scope, (32)) or, under certain circumstances, even to the C-domain (33):

(32) […] ob sie aber alle geritten/ oder etliche zu fusse gangen /
if they however all ridden or some to foot gone
hätte sie so properly not attention had (1659, Bucholtz, DTA)
had.SUBJ she so *eigentlich* nicht acht gehabt.
'[…] but whether they had all ridden or some had gone on foot, she would not have paid such close attention.'

(33) So gar *eigentlich* weiß ich davon nicht zuberichten […].
so completely properly know I there-of not to-report
'I can't tell you about it in a really proper way […].' (1659, Bucholtz, DTA)

When targeting the I-domain, this movement could reach the highest part of Cinque's (1999) projection cascade, where mainly mood and modality projections are found:

(34) [Mod/MoodPs … *eigentlich*$_i$ …] [AspPs … t_i …]]
Higher I-Layer *Lower I-Layer*

This movement – though not obligatory – was possibly at the origin of the subsequent syntactic reanalysis of *eigentlich* and other donor lexemes (see Axel-Tober & Müller 2017; Müller 2019; also cf. van Gelderen 2011 on English, but with a different interpretation of the contexts of the reanalysis). This flexible positioning is also observable with respect to logical-connective adverbs.

This movement paved the way for a semantic and syntactic reanalysis. From a semantic perspective, the potential ambiguity and the fronting of the adverb to

positions dedicated to illocution, mood and modality possibly favored its semantic reinterpretation.

In the following example, the "lexical split" of *eigentlich* with a consequent "lexical to functional reanalysis" – in Roberts & Roussou's (2003) sense – seems to have already taken place. The new MP could be realized in a position preceding the negation and be outside of its scope:

(35) […] aber das hat GOtt *eigentlich* nicht von dem steinern berg/
but that has God *PRT* not of the.DAT stony mountain
oder der gezimmerten stadt gemeinet (1700, Arnold, DTA)
or the.DAT timbered city meant […]
'[…] but God did not really mean that about the stone mountain or the timbered city […]'

I argue that such a lexical split affecting the donor lexeme and leading to the emergence of a MP is associated with syntactic reanalysis and thus with so called "syntactic climbing" in the sense of Roberts & Roussou (2003: 71),[14] whose syntactic approach to grammaticalization is summarized in their central tenet below:

> Since movement is always local and upward, categorial reanalysis is also local and upward. (Roberts & Roussou 2003: 71)

Cartographic theories *à la* Cinque (1999) offer an excellent benchmark for testing this syntactic approach to grammaticalization.

I assume that the moved adverbial element was reanalyzed as an adverb that is base-generated in the higher I-Layer (*Late-Merge*):

(36) [Mod/MoodPs … *eigentlich* …] [AspPs …]]
Higher I-Layer *Lower I-Layer*

This reanalysis can be led back to economy principles. As stated by Roberts & Roussou (2003: 15), "[i]f the trigger is ambiguous, the learner will choose the option that yields the simpler representation". In fact, the simpler representation is the one in which no movement has occurred.

14. Note that, according to Lehmann (2002: 128ff), grammaticalization usually goes hand in hand with the "shrinking" of structural scope ("condensation"), but – as he admits – the widening of semantic scope is also "frequent in the case of tense, aspect and mood operators." (Lehmann 2002: 129). Indeed, Molnár (2002: 19) emphasizes that the MPs do not undergo narrowing of the semantic scope. Their syntactic scope is even extended as a result of the grammaticalization process.

As the result of a grammaticalization process and independently of their donor lexemes, MPs all ended up occupying the leftmost part of the middle field (cf. Abraham 1991a: 372f). In fact, Coniglio (2005 and ff.) shows for present-day German that, while each MP behaves in a different way with respect to the classes of higher adverbs they can precede or follow, MPs tend to occupy very high positions among Cinque's (1999) functional projections. The lowest position they can occupy is the one between habitual and higher repetitive adverbs (data from Coniglio 2005, 2011: 109):

(37) *Distribution of MPs with respect to Cinque's (1999) adverbial classes:*

a. *ja* > $\text{Asp}_{\text{habitual}}$ / $\text{Asp}_{\text{habitual}}$ > *ja*

i. Der Attentäter wird *ja normalerweise* von der Polizei festgehalten.

ii. Der Attentäter wird *normalerweise ja* von der Polizei festgehalten.
'The terrorist is usually arrested by the police.'

b. *ja* > $\text{Asp}_{\text{repetitive(I)}}$ / * $\text{Asp}_{\text{repetitive(I)}}$ > *ja*

i. Der Attentäter ist *ja nochmals* von der Polizei festgehalten worden.

ii. *Der Attentäter ist *nochmals ja* von der Polizei festgehalten worden.
'The terrorist has again been arrested by police.'

Coniglio (2011: 115) argues that MPs – like adverbs – are merged in specifier positions of dedicated MoodPs in the upper part of Cinque's (1999) syntactic cascade of projections. MPs thus tend to occupy positions that are typical for higher adverbs and, from a theoretical perspective, this seems to support the idea of their (low) adverbial origin and also of their adverbial nature.

We noted above that, in contrast to other sentence adverbs that occupy the higher I-domain, MPs are best characterized as being weak, i.e. weak sentence adverbs (Coniglio 2005, 2011: 77ff; Grosz 2005; Cardinaletti 2007). Based on the comparison of the properties of different types of pronouns in languages like Italian and German, Cardinaletti & Starke (1999) propose that pronouns show three different levels of "structural deficiency":

(38) a. Maria gli ha dato ~~gli~~ un libro ~~gli~~. *clitic*
b. Maria ha dato loro un libro ~~loro~~. *weak*
c. Maria ha dato un libro a loro. *strong*
Mary to.him has given to.them a book to-them
'Maria gave him / them a book.' (adapted from Cardinaletti 2011: 502)

In Italian, the pronouns corresponding to English "to them" can be realized as a full pronoun modified by a preposition (as in "a loro"), as a weak pronoun ("loro") or as a clitic pronoun ("gli", which is mainly restricted to the spoken language, since, in written Italian, it means "to him"). Full pronouns are stressed and have a more

flexible position, while weak pronouns (like German personal pronouns) can have a word stress, but they don't get sentence stress. Furthermore they are confined to a high position in the I-domain. Clitic pronouns are neither stressed nor do they display syntactic flexibility, since, in Italian, they are cliticized to the finite verb.

Building on Cardinaletti & Starke's (1999) proposal, Cardinaletti (2007, 2011) extends her analysis to adverbs and proposes an adverbial tripartition in analogy to pronouns (see also Grosz 2005; Coniglio 2005, 2011 for a similar proposal):[15]

(39)	clitic MPs	< MPs	< (strong) adverbs
	(= clitic adverbs)	(= weak adverbs)	
	'n	*denn*	*dann*/*DENN*/(reg.) *denn*

As to the internal syntax, Cardinaletti (2007: 79) proposes the following structure for MPs (also cf. Cardinaletti 2011: 510), which are analyzed as adverbs lacking a full-fledged structure, i.e. as weak adverbs in Figure 4:

Figure 4. Structural deficiency of German MPs (Cardinaletti 2007: 79)

Thus, if – synchronically – MPs are to be analyzed as weak adverbs, then – diachronically – they must derive from strong adverbs that have "weakened". MPs have thus become non-complex sentence adverbs (Cardinaletti 2007). In contrast to strong sentence adverbs,

1. MPs cannot be modified, so that *so nur*, *gar eigentlich*, etc. are no longer grammatical in present-day German.
2. MPs cannot be moved (cf. Abraham 1991a). Strong adverbs are integrated in Cinque's hierarchy, but remain syntactically flexible, e.g. by being able to move to the C-domain. MPs, on the other hand, have lost their syntactic flexibility and occupy fixed positions in "specialized" mood projections in the I-domain (Coniglio 2011: 108ff).
3. MPs have developed different licensing mechanisms and a stronger connection to the left periphery of the sentence, acquiring new illocution-modifying functions (Jacobs 1986; Thurmair 1989, 1993; Abraham 1991b; Zimmermann 2004; Coniglio 2011, 2014; Bayer 2012; Coniglio & Zegrean 2012).

15. As to the clitic form *'n*, it should be noted that this is the only clitic MP present in (sub-) standard German (Wegener 2002; Grosz 2005; Bayer 2012).

6. Conclusions

In this paper, I argued against the hypothesis that MPs have originated from different types of donor lexemes and I showed that all MPs have diachronically derived from lower adverb(ial)s. In the course of the grammaticalization process, the adverbial donor lexemes – by way of acquiring further meanings and functions – have been syntactically reanalyzed as higher adverbs, namely sentence adverbs. The final step of the grammaticalization process was argued to consist in a modification of their internal structure and of their syntactic flexibility. MPs have thus become weak adverbs restricted to the middle field. More generally, it was shown that a generative approach to syntactic change can offer a straightforward and elegant explanation of the grammaticalization process of MPs.

Acknowledgments

I would like to thank the organizers and the audience of the International Workshop on Discourse Particles (University of the Basque Country – UPV/EHU, October 8th 2020), the anonymous reviewers and other colleagues that discussed with me different parts of this paper, in particular: Katrin Axel-Tober, Regine Eckardt, Kalle Müller, Katharina Paul, Stavros Skopeteas, Markus Steinbach.

References

Abraham, Werner. 1991a. The grammaticization of the German modal particles. In *Approaches to Grammaticalization* vol. 2 [Typological Studies in Linguistics 19:2], Elizabeth C. Traugott & Bernd Heine (eds), 331–380. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.19.2.17abr

Abraham, Werner. 1991b. Discourse particles in German: How does their illocutive force come about?. In *Discourse Particles*, Werner Abraham (ed), 203–252. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.12.08abr

Autenrieth, Tanja. 2002. *Heterosemie und Grammatikalisierung bei Modalpartikeln: Eine synchrone und diachrone Studie anhand von »eben«, »halt«, »e(cher)t«, »einfach«, »schlicht« und »glatt«* [Linguistische Arbeiten 450]. Berlin: de Gruyter. https://doi.org/10.1515/9783110960907

Axel, Katrin. 2007. *Studies on Old High German syntax. Left sentence periphery, verb placement and verb-second* [Linguistik Aktuell/Linguistics Today 112]. Amsterdam: John Benjamins. https://doi.org/10.1075/la.112

Axel-Tober, Katrin & Müller, Kalle. 2017. Evidential adverbs in German. Diachronic development and present-day meaning. *Journal of Historical Linguistics* 7(1–2): 9–47. https://doi.org/10.1075/jhl.7.1-2.02axe

Bayer, Josef. 2012. From modal particle to interrogative marker: a study of German *denn*. In *Functional heads* [The cartography of syntactic structures 7], Laura Brugè et al. (eds), 13–28. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199746736.003.0001

Burkhardt, Armin. 1994. Abtönungspartikeln im Deutschen. Bedeutung und Genese. *Zeitschrift für germanistische Linguistik* 22: 129–151. https://doi.org/10.1515/zfgl.1994.22.2.129

Cardinaletti, Anna. 2007. Für eine syntaktische Analyse von Modalpartikeln. In *Gesprochene Sprache – Partikeln*, Eva-Maria Thüne & Franca Ortu (eds), 89–101. Frankfurt: Lang.

Cardinaletti, Anna. 2011. German and Italian modal particles and clause structure. *The Linguistic Review* 28(4): 493–531. https://doi.org/10.1515/tlir.2011.014

Cardinaletti, Anna. 2015. Italian verb-based discourse particles in a comparative perspective. In *Discourse-oriented Syntax*, Josef Bayer, Roland Hinterhölzl & Andreas Trotzke (eds), 71–91. Amsterdam: John Benjamins. https://doi.org/10.1075/la.226.04car

Cardinaletti, Anna & Starke, Michal. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In *Clitics in the Languages of Europe*, Henk van Riemsdijk (ed), 145–233. Berlin: de Gruyter. https://doi.org/10.1515/9783110804010.145

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective* [Oxford Studies in Comparative Syntax]. Oxford: Oxford University Press.

Cinque, Guglielmo. 2001. 'Restructuring' and the order of aspectual and root modal heads. In *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi*, Guglielmo Cinque & Giampaolo Salvi (eds), 137–155. Amsterdam: Elsevier. https://doi.org/10.1163/9780585473949_009

Cinque, Guglielmo. 2006. *Restructuring and Functional Heads. The Cartography of Syntactic Structures*, Volume 4 [Oxford Studies in Comparative Syntax]. Oxford: Oxford University Press.

Cognola, Federica & Cruschina, Silvio. 2020. From time to discourse: a syntactic analysis of Italian *poi*. Handout for Seminari di Linguistica Ca' Foscari – VeReWebLing, 17th November 2020.

Coniglio, Marco. 2005. Deutsche Modalpartikeln: eine syntaktische Analyse. Diplomarbeit, Università Ca' Foscari Venezia.

Coniglio, Marco. 2011. *Die Syntax der deutschen Modalpartikeln: ihre Distribution und Lizenzierung in Haupt- und Nebensätzen*. Berlin: Akademie. https://doi.org/10.1524/9783050053578

Coniglio, Marco. 2014. The fine structure of Force. On the interaction of modal particles with illocutionary force and clause type. In *On Peripheries. Exploring Clause Initial and Clause Final Positions* [Hituzi Linguistics in English 23], Anna Cardinaletti, Guglielmo Cinque & Yoshio Endo (eds), 103–140. Tokyo: Hituzi Syobo.

Coniglio, Marco & Zegrean, Iulia. 2012. Splitting up Force. Evidence from discourse particles. In *Main Clause Phenomena. New Horizons* [Linguistik Aktuell/Linguistics Today 190], Lobke Aelbrecht, Liliane Haegeman & Rachel Nye (eds), 229–255. Amsterdam: John Benjamins. https://doi.org/10.1075/la.190.10con

Diewald, Gabriele. 1997. *Grammatikalisierung. Eine Einführung in Sein und Werden grammatischer Formen* [Germanistische Arbeitshefte 36]. Tübingen: Niemeyer. https://doi.org/10.1515/9783110946673

Diewald, Gabriele. 1999. Die dialogische Bedeutungskomponente von Modalpartikeln. In *Dialogue Analysis and the Mass Media. Proceedings of the International Conference, Erlangen, April 2–3, 1998* [Beiträge zur Dialogforschung 20], Bernd Naumann (ed), 187–199. Tübingen: Niemeyer. https://doi.org/10.1515/9783110943115.187

Diewald, Gabriele. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49(2): 365–390. https://doi.org/10.1515/ling.2011.011

Diewald, Gabriele & Ferraresi, Gisella. 2008. Semantic, syntactic and constructional restrictions in the diachronic rise of modal particles in German: A corpus-based study on the formation of a grammaticalization channel. In *Theoretical and Empirical Issues in Grammaticalization* [Typological Studies in Language 77], Elena Seoane & María José López-Couso (eds), 77–110. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.77.06die

Dohi, Atsushi. 2020. CP-internal Discourse Particles and the Split ForceP Hypothesis. *Lingua* 233: 102757. https://doi.org/10.1016/j.lingua.2019.102757

Ebert, Robert P., Reichmann, Oskar, Solms Hans-Joachim and Wegera, Klaus-Peter 1993. *Frühneuhochdeutsche Grammatik*. Tübingen: Niemeyer. https://doi.org/10.1515/9783110920130

Ferraresi, Gisella. 2005. *Word order and phrase structure in Gothic. A comparative study*. Leuven: Peeters.

Ferraresi, Gisella. 2014. *Grammatikalisierung* [Kurze Einführung in die Germanistische Linguistik 15]. Heidelberg: Winter.

Frey, Werner. 2004. The grammar-pragmatics interface and the German prefield. *Sprache und Pragmatik* 52: 1–39.

Grosz, Patrick. 2005. "Dn" in Viennese German. The Syntax of a Clitic Version of the Discourse Particle "denn". MA thesis, Universität Wien.

Gutzmann, Daniel. 2008. On the Interaction between Modal Particles and Sentence Mood in German. Magisterarb. Johannes Gutenberg-Universität Mainz.

Gutzmann, Daniel. 2009. Hybrid semantics for modal particles. *Sprache und Datenverarbeitung* 33(1–2): 45–59.

Haegeman, Liliane & Hill, Virginia. 2013. The syntactization of discourse. In *Syntax and its limits* [Oxford Studies in Theoretical Linguistics 48], Raffaella Folli, Christina Sevdali & Robert Truswell (eds), 370–390. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199683239.003.0018

Haegeman, Liliane & Hill, Virginia. 2014. Vocatives and speech act projections: a case study in West Flemish. In *On peripheries: Exploring Clause Initial and Clause Final Positions*, Anna Cardinaletti, Guglielmo Cinque & Yoshio Endo (eds), 209–236. Tokyo: Hituzi Shobo.

Hentschel, Elke. 1986. *Funktion und Geschichte deutscher Partikeln:'Ja','doch','halt' und 'eben'*. Tübingen: Niemeyer. https://doi.org/10.1515/9783111371221

Jacobs, Joachim. 1986. Abtönungsmittel als Illokutionstypmodifikatoren. *Groninger Arbeiten zur Germanistischen Linguistik* 27: 100–111.

Kluge, Friedrich. 2011. *Etymologisches Wörterbuch der deutschen Sprache. 25. Auflage*. Berlin: De Gruyter.

Lehmann, Christian. 2002. *Thoughts on grammaticalization. 2nd revised edition* [ASSidUE 9]. Erfurt: Seminar für Sprachwissenschaft der Universität.

Meibauer, Jörg. 1994. *Modaler Kontrast und konzeptuelle Verschiebung: Studien zur Syntax und Semantik deutscher Modalpartikeln* [Linguistische Arbeiten 314]. Tübingen: Max Niemeyer. https://doi.org/10.1515/9783111339795

Molnár, Anna. 2002. *Die Grammatikalisierung deutscher Modalpartikeln. Fallstudien* [Metalinguistica 12]. Frankfurt: Lang.

Müller, Kalle. 2019. Satzadverbien, Evidentialität und Non-At-issueness. PhD dissertation, Eberhard Karls Universität Tübingen.

Müller, Sonja. 2018. *Distribution und Interpretation von Modalpartikel-Kombinationen*. Berlin: Language Science Press.

Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199207923.001.0001

Ormelius-Sandblom, Elisabet. 1997a. The modal particle *schon*: Its syntax, semantics and pragmatics. In *Modality in Germanic Languages. Historical and Comparative Perspectives* [Trends in Linguistics. Studies and Monographs 99], Toril Swan & Olaf J. Westwik (eds), 75–131. Berlin: de Gruyter. https://doi.org/10.1515/9783110889932.75

Ormelius-Sandblom, Elisabet. 1997b. *Die Modalpartikeln "ja", "doch" und "schon". Zu ihrer Syntax, Semantik und Pragmatik*. Lund: Almqvist & Wiksell International.

Paul, Hermann. 1992. *Deutsches Wörterbuch*. 9. Auflage. Tübingen: Niemeyer. https://doi.org/10.1515/9783111546629

Petrova, Svetlana. 2017. Left-Peripheral Modal Particles in Older Germanic? In *Syntax and Semantics of Modal Particles*, Volker Struckmeier & Josef Bayer (eds), 304–331. Berlin: de Gruyter.

Remberger, Eva-Maria. 2020. Discourse and Pragmatic Markers in the Romance Languages. In *The Oxford Encyclopedia of Romance Linguistics*, Michele Loporcaro & Francesco Gardani (eds). Oxford: Oxford University Press. Advance online publication. [http://linguistics.oxfordre.com/].

Roberts, Ian & Roussou, Anna. 2003. *Syntactic change. A minimalist approach to grammaticalization* [Cambridge Studies in Linguistics 100]. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486326

Schifano, Norma & Cognola, Federica. 2020. From macro to nano: A parametric hierarchy approach to the diatopic and diachronic variation of Italian *ben*. In *Syntactic architecture and its consequences I: Syntax inside the grammar* [Open Generative Syntax 9], András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds), 233–250. Berlin: Language Science Press.

Searle, John. 1976. A classification of illocutionary acts. *Language in Society* 5(1), 1–23. https://doi.org/10.1017/S0047404500006837

Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen* [Linguistische Arbeiten 223]. Tübingen: Niemeyer. https://doi.org/10.1515/9783111354569

Thurmair, Maria. 1993. Äußerungsform oder Äußerungsfunktion? Zu den Bedingungen für das Auftreten von Modalpartikeln. *Deutsche Sprache* 21: 22–43.

van Gelderen, Elly. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins. https://doi.org/10.1075/la.71

van Gelderen, Elly. 2011. *The Linguistic Cycle: Language Change and the Language Faculty*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199756056.001.0001

van Kemenade, Ans & Links, Meta. 2020. Discourse particles in early English: Clause structure, pragmatics and discourse management. *Glossa: A Journal of General Linguistics* 5(1): 3. https://doi.org/10.5334/gjgl.1020

Wegener, Heide. 2002. The Evolution of the German Modal Particle "denn". In *New Reflections on Grammaticalization* [Typological Studies in Language 49], Ilse Wischer & Gabriele Diewald (eds), 379–393. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.24weg

Zimmermann, Malte. 2004. Discourse particles in the left periphery. *Proceedings of the Dislocated Elements Workshop, Zentrum für Allgemeine Sprachwissenschaft Berlin, November 2003. ZAS Papers in Linguistics* 35(2): 543–566. https://doi.org/10.21248/zaspil.35.2004.241

Appendix – other MPs

auch (lit. 'also') < OHG *ouh* (< IE *au-*), adversative-iterative ('again'), then augmentative meaning ('in addition'); MP since 17th-18th c. (Diewald 1997: 83f).

eben (lit. 'flat') < OHG *eban(i)* 'flat'. Adjective-adverb > temporal adverb > focus particle. MP (in the 16th c. according to Burkhardt 1994) originating either from a manner adverb, a temporal adverb or a focus particle (cf. DWB; Hentschel 1986; Abraham 1991a: 365; Burkhardt 1994; Diewald 1997: 92f; Autenrieth 2002: 114ff; Molnár 2002; Diewald & Ferraresi 2008). The latter path is chronologically possible, but against the hypothesis presented here.

eh < OHG *ē(r)* 'earlier, before'. Temporal adverb > logical-connective adverb > MP (19th c., Burkhardt 1994).

etwa (lit. 'approximately') < OHG *edde(s)(h)wār* 'somewhere'. Localistic > temporal > logical > MP (17th c., Burkhardt 1994).

halt < OHG *halt(o)* 'more, rather' (Germ. **haldiz* 'more, rather', comparative, cf. Got. *haldis*, Kluge 2011: 389). Already in OHG temporal adverb 'fast, immediate, earlier' and adverb of manner 'rather, (more) preferably' (Diewald 1997: 93). The MP developed from it in MHG. See also Hentschel (1986); Abraham (1991a: 366f) and Autenrieth (2002: 142ff).

ja (lit. 'yes') < OHG *ia* (as C-particle one of the oldest particles, cf. Hentschel 1986; Abraham 1991a: 367ff; Diewald 1997: 95f). However, evidence that the I-particle grammaticalized from the temporal adverb *je* 'ever' (or from a contamination of this adverb with the C-particle) in the 16th c. (see Meibauer 1994: 165 following Paul 1992: 435; Molnár 2002: 101; Petrova 2017).

mal Temporal adverb ('once'), MP in the 18th c. (Burkhardt 1994).

schon (lit. 'already') < OHG *scōno* 'in a beautiful way' (< adj. *scōni*), adverb of manner > in MHG temporal adverb 'earlier than expected' > degree particle (Molnár 2002: 77f). In the literature, it is an open question whether the MP originated from the temporal adverb or from the degree particle. The temporal origin is defended by Meibauer (1994) and Ormelius-Sandblom (1997a, 1997b: 65ff) (also cf. Molnár 2002 and Diewald & Ferraresi 2008). MP since the 16th c. (Molnár 2002: 81).

sowieso ('anyway') adverbial of manner > MP (18th c., Burkhardt 1994).

wohl (lit. 'well') < OHG *wola* 'well' (Germ. **welō-* < IE **u̯el* 'wanted, (as) desired', DWDS). Adverb of manner. MP since the 16th c. Cf. Molnár (2002) on the possible grammaticalization steps (cf. Schifano & Cognola 2020 on Italian *ben*).

CHAPTER 2

A particle-like use of *hwæþer*

Wisdom's questions in Boethius

Regine Eckardt and George Walkden
University of Konstanz

The paper investigates unembedded *hwæþer*-questions in Old English (OE). We argue that they represent an intermediate stage in the development of *hwæþer* 'which of the two' to Modern English *whether*. Syntactically, we find a range of quasi-subordinating uses of *hwæþer* in questions that all have in common that the speaker expresses a pedagogical question. Pedagogical questions are questions the speaker knows the answer to, but is urging the addressee to consider while drawing their own conclusions. In the OE Boethius, *hwæþer* can convey this use-conditional pragmatic flavour for polar questions. It thus comes close in function to other use-conditional particles.

Keywords: pedagogical question, domain of choice, *wh*-pronoun, Old English, reanalysis

1. Introduction

This chapter deals with the syntax, semantics, and pragmatics of Old English questions introduced by the particle-like element *hwæþer*, as in (1).

(1) Hwæðer nu gimma wlite eowre eagan to him getio heora to
Whether now jewels looks your eyes to them attract them.GEN to
wundrianne (OE Boethius 13: 40–41; Godden & Irvine 2009: 266)
wonder.INF
'Does the beauty of jewels attract your eyes, to wonder at them?'

Unembedded *hwæþer*-questions such as (1) have figured prominently in the literature on historical English syntax since Traugott (1972) and Allen (1980); see Section 3 for discussion. It is generally acknowledged that they serve as an alternative to verb-fronting (V1) clauses as a syntactic strategy for forming direct polar questions. What has not been addressed in this literature, however, is the circumstances conditioning the use of *hwæþer*-questions. In particular, the implicit assumption

https://doi.org/10.1075/la.276.02eck

in the literature has so far been that the two strategies are in free variation, with no semantic or pragmatic difference (or at least nothing is said on the issue). This is the lacuna that the present paper aims to fill: did *hwæþer*-questions mean the same as V1 questions, and how does this relate to their syntactic properties?

The pragmatic-semantic-syntactic proposal presented in this paper builds on recent advances in the study of non-canonical questions at the interfaces. Our approach has several features that set it apart from previous research. First, we look at all *hwæþer*-questions, not just unembedded ones. Secondly, we attempt to reach a descriptively adequate characterization of their common pragmatic function across question types, and of the semantic and pragmatic contribution of *hwæþer* itself. The analysis of unembedded *hwæþer*-questions that we present is thus backed by the full range of uses, as opposed to earlier syntactic proposals (see Section 3) that have not always taken the overall picture into account.

The paper is structured as follows. Section 2 presents the data sources and evidence base that we draw upon, the range of uses of *hwæþer* in basic descriptive terms, and some observations on its distribution and pragmatics. In Section 3 we discuss and evaluate previous proposals and research on Old English *hwæþer*-questions. Section 4 presents our own analysis in detail. Section 5 summarizes and outlines further questions.

2. Data sources and evidence base

2.1 The Old English *Boethius*

In this paper we draw data primarily from a single text, the Old English translation of Boëthius's *De consolatione philosophiae* (Consolation of Philosophy), henceforth referred to as the Old English Boethius. Our initial search of the York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE; Taylor et al. 2003) revealed that this single text provides well over half of the total number of unembedded *hwæþer*-questions in the corpus.

The text takes the form of a dialogue between Boethius himself and a female figure who is an anthropomorphic personification of Philosophy. The West Saxon Old English translation is preserved in two manuscripts: MS C, a mixed prose/verse manuscript from the mid-tenth century, and MS B, a prose manuscript dating to the late eleventh or early twelfth century. Traditionally it was thought that King Ælfred himself translated Boethius from Latin, but this is now disputed (Godden 2007; Bately 2009, 2015). Godden & Irvine (2009, I: 146) propose that the text was produced between 890 and 930 by "an unknown writer of substantial learning, not necessarily connected with King Alfred or his court".

For our purposes, what is important is that the Old English Boethius is a very free translation, perhaps better termed a recomposition. The five books of the original, each comprising 11–24 sections, have been reshaped into a single text, composed of 42 sections in MS B. The first person narrator sometimes seems to be Boethius, and at other times an everyman figure, and is also referred to as *Mod* 'mind'; the Philosophy figure, his interlocutor, is gendered using both masculine (*Wisdom* 'wisdom') and feminine (*Gesceadwisnes* 'reason/discernment') nouns, sometimes both at once. The Old English version diverges from the Latin in a number of ways, some trivial, some significant: for instance, the Old English Boethius makes much more use of first- and second-person forms than the Latin original. Godden & Irvine (2009: 50) state that "the author did not intend anything like a literal or even a free translation". For this reason we can safely assume that the syntactic patterns we find in the Old English Boethius reflect autochthonous norms and are not merely artefacts of translation. For a detailed overview of the text's form and substance we refer the reader to Guenther Discenza (2015). The Old English Boethius has been edited several times, by Fox (1864), by Sedgefield (1999), whose version is included in the YCOE, and most recently by Godden & Irvine (2009; see their overview of previous editions in I, 215–221). Our examples are presented in the form in which they occur in this latter critical edition, including translations. Where examples are taken from other Old English texts, they follow the YCOE (Taylor et al. 2003), and the references given are YCOE token IDs.

2.2 Range of syntactic uses

This section surveys five types of *hwæþer*-questions that are attested in Boethius. Type 1 shows *hwæþer* as a *wh*-pronoun in the sense 'which (of the two)'.

(2) hwæþerne woldest þu deman wites wyrþran þe ðone
whether.ACC would you deem punishment.GEN worthier either the
þe ðone unscyldgan witnode, ðe ðone þe þæt wite
that the innocent.ACC punished or the that this punishment.ACC
þolade? (OE Boethius 38: 220)
suffered
'Which (of the two) would you judge worthier of punishment, the one who punished the innocent, or the one who suffered this punishment?'

The case morpheme *ne* in *hwæðerne* underscores its nominal status, and the example is obviously not an instance of the unembedded *hwæþer* clauses described above. The question pronoun *hwæþer* is cognate with Gothic *ƕaþar*, both descended from Proto-Germanic **h^{w}aþeraz* (Nielsen 1998: 78–79; Ringe 2006: 290),

and the question pronoun sense is the only attested sense in Gothic (Parra-Guinaldo 2013: 155–161; Walkden 2014: 146–147). The choice in (2) – between the good man and the evil man – is left implicit, but can also be given explicitly as in the following example (not from Boethius). Again the two alternatives are introduced by *þe … þe* 'either … or'.

(3) þa þæt folc gesamnod wæs þa cwæð Pilatus, hwæþer wylle ge
there the people collected was then said Pilatus, whether want you
þæt ic eowa gyfe þe Barrabban ðe þone hælynd ðe is Christ gehaten?
that I you give or Barabbas or the saviour that is Crist called?
'When the people was assembled, Pilate said: Which one do you want that I should give you, Barabbas or the saviour who is called Christ?'
(cowsgosp, Mt_[WSCp]:27.17.2019)

Type 2 are questions about the addressee's beliefs, opinions or own conclusions, as in (4).

(4) hwæðer þu woldest cweðan þæt he wære unwyrðe anwealdes and
whether you wanted say that he be.SBJV unworthy power.GEN and
weorðscipes (OE Boethius 27: 40–41)
honor.GEN
'would you say that he was unworthy of power and honour?'

These examples have in common a matrix clause with a verb of saying, thinking or belief and an embedded clause that contributes the proposition *p* in question. Type 2 examples also occur as embedded clauses (see Example (10) below) and we argue in Section 4 that such embedded Type 2 examples constitute the bridging examples that allowed for reanalysis and eventually led to modern *whether*.

Type 3 are unembedded *hwæþer*-questions, illustrated by (1) in Section 1 above. The sentence conveys a polar question. Its syntactic structure resembles the one of embedded clauses, in that the verb *getio* occurs late in the clause and is in the subjunctive. While the syntax of Type 2 and 3 differ, the speaker's intentions are the same in either type of question, as we argue in Section 2.4 below.

Examples of type 4 show *hwæðer* as a question complementizer. The embedded clause *hwæðer p* provides a clausal argument for the matrix verb, for instance the verb *acsian* 'ask' in the following examples (see also (8)).

(5) ðry weras … axodon … hwæðer se halga Petrus þær wununge
three men asked whether the holy Peter there dwelling
hæfde (coaelive,+ALS[Peter's_Chair]:109.2346)
had.SBJV
'Three men asked whether Saint Peter lived there'

(6) Sege me nu hwæðer þu mid rihte mæge seofian þina
say me.DAT now whether you with right may lament your
unsælþa (…) (OE Boethius 10: 16–17)
misfortunes …
'Tell me now whether you can rightly lament your misfortunes…'

Following earlier authors, we assume that these examples essentially show the syntax and semantics of embedded *whether* clauses in Modern English (ModE). However, there is a group of such examples that are no longer licit in ModE. They are illustrated in (7) and we class them as a separate Type 5.

(7) Wenst þu hwæðer he mæge yfel don? (OE Boethius 35: 150)
think you whether he may.SBJV evil do?
'Can he [= God] do evil, do you think?'

Syntactically, Example (7) could qualify as a case of an embedded polar question with complementizer *hwæþer*. Semantically, however, the example violates a semantic universal. Karttunen (1977) was the first to point out that the verb *believe* and near-synonyms prohibit embedded questions. The same prohibition was found in more and unrelated languages, and indeed verified in all languages where the construction has been tested (Theiler et al. 2017; Uegaki 2016, 2019). Semanticians therefore hypothesize that the prohibition most likely rests on an incompatibility between the epistemic nature of *believe*-verbs and the interpretation of embedded questions. Turning to the interpretation of data in Boethius, we must thus either propose that the writer and his contemporaries spoke a variety that violates semantic universals, or alternatively that the underlying structure of the – seemingly unproblematic – Example (7) remains yet to be revealed. Our analysis in Section 4 takes the latter course.

Let us finally quantify the types of uses. In Boethius, there is a total of uses of

Type 1 "hwæþer of the two, X or Y?"	$n = 2$
Type 2 "hwæþer you believe that q?"	$n = 27$
Type 3 "hwæþer q?"	$n = 19$
Type 4 embedded questions "I wanted to ask you hwæþer p."	$n = 11$
Type 5 embedded questions "Do you believe hwæþer p?"	$n = 8$

Adding up type 4 and 5, we have 19 examples where *hwæþer* shows in the syntactic position of a complementizer in an embedded clause, as opposed to 48 unembedded *hwæþer* clauses, and two uses of the nominal *wh*-pronoun *hwæþer*.

2.3 More observations on unembedded *hwæþer*-questions

This section presents some further formal facts about Types 2 and 3, the unembedded *hwæþer*-questions, especially ways in which they differ from other unembedded questions found in Old English.

First, while *wh*-questions display verb-second syntax in all the early Germanic languages (Eyþórsson 1995; Walkden 2014: 114–121), unembedded *hwæþer*-questions consistently have the verb in a late position (Traugott 1972: 73; Allen 1980); this holds for all of the examples in Boethius. This fact has been crucial in determining the syntactic analysis of this type of question, to which we will return in Section 3. Van Gelderen (2009: 140 n. 4) challenges the generalization on the basis of two apparent counterexamples with V2 constituent order, one of which is given in (8).

(8) Hwæðer wæs iohannes fulluht þe of heofonum þe of mannum
whether was John.GEN baptism or of heaven.DAT or of man.DAT
'Which was John's baptism: of heaven, or of man?'
(cowsgosp, Mt_[WSCp]:21.25.1438)

However, both examples can be read as instances of Type 1 ('which of the two'), with two non-propositional disjuncts given as alternatives (see Walkden 2014: 148 for discussion).

Secondly, Type 2 and 3 unembedded *hwæþer*-questions more often than not contain the words *nu* (23x in Boethius) or *þonne* (3x in Boethius). These are originally temporal adverbs meaning 'now' and 'then' respectively. Van Kemenade & Links (2020) argue that these elements in Old English have grammaticalized into discourse particles comparable to those found in Dutch and German. These particles are much more common in unembedded *hwæþer*-questions than they are in *wh*-questions in general.

2.4 The pragmatics of unembedded *hwæþer*-questions

This section takes a closer look at the meaning of unembedded *hwæþer*-questions. Previous authors generally propose that they convey polar questions and are in fact synonymous to verb-initial polar questions.[1] However, general principles of

1. There are some exceptions to this. Traugott (1972: 73) characterizes unembedded *hwæþer*-questions as favoured in cases of doubt or incredulity (cf. Fischer et al. 2000: 84), and Mitchell (1985, I: 682) suggests that many such questions are rhetorical. Cf. also Godden & Irvine (2009, I: 196) on the prominent role of questions in general, and rhetorical questions in particular, in the Old English Boethius. None of these authors investigate the meaning of *hwæþer*-questions in detail, however.

synonymy avoidance stand against this assumption (Levinson 2000). We therefore expected to detect a specific pragmatic flavour for unembedded *hwæþer*-questions.

This expectation is met by the data. Unembedded *hwæþer*-questions in Boethius are indeed used as 'Socratic questions' in pedagogic discourse. In the dialogue between Wisdom (W) and Boethius (B), only W ever asks unembedded *hwæþer*-questions (in 48 instances). Our initial example occurs in a debate as to whether the possession of valuables (such as gold or jewels) can make a man happy. W argues that richness is unsuited as a means to achieve universal happiness, as it is necessarily restricted to few, excluding many: "Are the riches of this middle earth worthy of a man when no one can fully have them? Nor can they enrich any man, unless they bring another to poverty." W then provokes B:

(9) Hwæðer nu gimma wlite eowre eagan to him getio heora to
whether now jewels looks your eyes to them attract them.GEN to
wundrianne (OE Boethius 13: 40–41; = (1) above)
wonder.INF
'Does the beauty of jewels attract your eyes, to wonder at them?'

W immediately answers the question herself ("I know that they do so"). This example is typical: W knows the answer in *all* uses of unembedded whether questions, sometimes answering herself, sometimes simply moving on, and sometimes offering B the opportunity to answer. B generally agrees, but opposes the insinuated answers in two cases. The polarity of the answer can be positive or negative.

Wisdom's intention in posing an unembedded *hwæþer* question is to invite the addressee, Boethius, to think about the question by himself. We find this intention in many philosophical dialogues, specifically in classical 'Socratic' questions as a means to provoke independent thinking. The intention is explicated in Type 2 questions like (3) 'would you say that he is unworthy of power and honour?' Questions in verb-final syntax, in combination with particles, are known to convey special speaker intensions in other languages, as for instance in *ob…wohl, was…wohl* questions in German (Eckardt & Beltrama 2019, Eckardt 2020, Truckenbrodt 2006). Unembedded *hwæþer*-questions pattern with these in that they ask for the addressee's opinion about Q. We propose that this speaker attitude played a central role in the diachronic development of unembedded *hwæþer*-questions, and should be captured as use-conditional content (Gutzmann 2015).

Socratic questions in Boethius are limited to unembedded *hwæþer*-questions. Embedded *hwæþer* questions in Boethius (Type 4) can convey information seeking questions, as in the following utterance of B.

(10) Ac ic wolde þe acsian *hwæðer* we ænigne frydom habban, oððe
and I wanted you ask whether we any freedom have or
ænigne anweald hwæt we don hwæt we ne don
any power what we do what we not do
'But I would like to ask you whether we have any freedom, or any power as to what we do or not do … .' (OE Boethius 40: 101–103)

This indirect question act by B is information seeking, not pedagogical or biased.

3. Old English *hwæþer*: The state of the art

3.1 The standard syntactic story

Previous literature on Old English *hwæþer* has largely focused on its syntax in uses of Types 2–3, and in particular the fact that the verb does not occur in second position, as already mentioned in Section 2.3. The conventional analysis, going back in its essence to Allen (1980: 791), is that there are two *hwæþer*s in Old English.[2] One is a NP (or DP) proform meaning 'which (of the two)', and is used in questions of Type 1. The other is a complementizer, and is used in questions of Types 2–5. In current terms, the proform is in Spec,CP and the complementizer is in C^0. Type 1 questions can then be verb-second with subject-verb inversion when the fronted proform is not itself a subject, as the verb is free to move to C^0. In the other types, by contrast, since *hwæþer* occupies C^0, the verb cannot move there, and must remain in a lower position.[3] This fits well with the classic intuition, attributed to den Besten (1989), that asymmetric V2 in languages like German is driven by the complementary distribution of the complementizer and the finite verb, which both in a sense compete for the same position.

Some form of this analysis has been adopted by most subsequent authors writing on Old English *hwæþer*, e.g. Kiparsky (1995: 142), van Gelderen (2009), Parra-Guinaldo (2013), Walkden (2014: 144–155). Van Gelderen (2009) also presents a diachronic scenario for the emergence of complementizer *hwæþer* and its subsequent history in Old and Middle English (cf. also Ukaji 1997, Parra-Guinaldo 2013). The basic narrative is that the pronoun that moves to Spec,CP is reanalyzed

2. Cf. its treatment in Bosworth & Toller's dictionary (1898: s.v. *hwæþer*), which lists it as both a conjunction and a pronoun.

3. Exactly what position the verb occupies in 'verb-late' *hwæþer*-questions (either embedded or unembedded) is a matter of some debate. 'Verb-late' here means simply that the verb surfaces in a position that is lower in the clause than C^0 (and hence later in linear order, since the Old English CP is uniformly head-initial).

in some contexts as an operator first Merged there, an instance of lexical split driven by the Late Merge Principle ("Merge as late as possible"; van Gelderen 2004), and then as a complementizer in C^0, driven by the Head Preference Principle ("Be a head, rather than a phrase"; van Gelderen 2004). The complementizer stage is the one that predominates in Old English, and evidence for this is furnished by the lack of verb movement to the C-domain (van Gelderen 2009: 142).

3.2 Synchronic problems with the standard syntactic story

The consensus analysis sketched in the previous subsection is not unreasonable, but it leaves several questions unanswered both synchronically and diachronically. We will begin by outlining the problems with the synchronic analysis.

First, Old English did not ordinarily permit independent clauses introduced by complementizer elements (e.g. with the unambiguous C^0 elements *þæt* or *þe*).[4] Finding exactly such a phenomenon with *hwæþer*-questions and not elsewhere is at the very least mysterious. Secondly, the proposal provides no explanation for the pragmatic facts adduced in Section 2.4.

Finally, and perhaps most significantly, the idea that the complementizer is in complementary distribution with the finite verb and hence blocks V-to-C movement is not by itself sufficient to derive the clause type asymmetries found in Old English. This is so for two main reasons: (i) there is ample evidence that more than one position for verb movement is needed in Old English main clauses, and a good case can be made that these are distinct head positions in the C-domain (see Walkden 2014, 2017, 2021 and references cited there). Thus, simply saying that the verb and the complementizer compete for a single C^0 position may seem reasonable for questions, but is insufficient to account for clause type asymmetries in Old English more broadly.[5] (ii) Old English embedded clauses are verb-late even when there is demonstrably no overt complementizer present (Walkden & Booth 2020: Section 3). This is the case, for instance, in regular embedded *wh*-questions, where the *wh*-item is uncontroversially in a specifier position in the embedded C-domain. This is also the standard analysis of Old English relative clauses headed by a demonstrative pronoun of the *se* paradigm.

4. In this sense, Types 2 and 3 *hwæþer*-questions can be viewed as instances of 'insubordination' (Evans 2007; cf. Traugott 2017 for a perspective from the history of English).

5. See Roberts (1996) and Salvesen & Walkden (2017) for discussion of the syntactic positions of complementizers in OE.

We can conclude that the standard analysis of *hwæþer* as a C^0 head in *hwæþer*-questions of Types 2–3 does not derive their synchronic properties.[6]

3.3 Diachronic problems with the standard syntactic story

Turning to diachronic concerns, the first issue is whether all stages in the change of *hwæþer* from a moved pronoun to a first Merged operator in Spec,CP to a head in C^0 (as in van Gelderen's 2009 proposal) are attested. The pronominal stage corresponds to our Type 1, and the head stage corresponds to our types 2–3 (assuming that verb-late correlates with *hwæþer* being a head, as in embedded *wh*-questions). The operator stage, however, does not seem to be attested in Old English. We might expect this to manifest itself as a V2 question with all the other formal properties of Types 2–3, but these do not seem to be attested.

Another issue concerns the context for the change. Assuming, as standard, that grammaticalization involves reanalysis (Campbell 2001: 141; Hopper & Traugott 2003: 59), we might ask what the bridging contexts were that enabled reanalysis. The discussion in van Gelderen (2009) suggests that the change took place in embedded contexts. However, it seems implausible to us that a Type 1 question could be reanalyzed as a question of Types 2–3. These types are simply too different formally (case marking, verbal mood, verb position) and semantically for a potential bridging context to arise. The alternative found in the literature is that Types 2–3 have their origin instead as indirect/embedded questions (Mitchell 1985, I: 681, citing earlier work; Fischer et al. 2000: 54; Walkden 2014: 150), which is what our analysis in Section 4 will also propose.

Finally, even putting all of the above concerns aside, none of the previously proposed scenarios really motivate why *hwæþer*, as a *wh*-pronoun, should all of a sudden turn into a complementizer for *polar* questions. Functionally and logically these two things are different beasts, and saying that *hwæþer* becomes a C^0 element alone does not explain why it should also take on this particular function. In principle, of course, a synchronic analysis of *hwæþer* in 800 needn't provide this, but of two accounts, one that does is superior to one that does not.

---

6. An intriguing alternative to the standard analysis is provided by Berizzi (2010: 129–131), who suggests that there is a silent IS IT THAT between *hwæþer* and the following clause. This is closer in spirit to the account we develop in Section 4, and is somewhat more descriptively adequate than the standard analysis, though there is also a substantial element of stipulation. For present-day English, it is disputed whether *whether* is in C^0 or Spec,CP; see e.g. Nakajima (1996) for the C^0 analysis and Den Dikken (2006) for the Spec,CP analysis. Resolving this debate is beyond the scope of this paper.

Having outlined a number of potential issues for existing accounts of the synchrony and diachrony of *hwæþer*-questions in English, we now put forward our own proposal.

4. The stages and uses of *hwæþer*-questions

This section proposes stages in the diachronic development of *hwæþer*, including the use as pedagogical questions. Our ordering of uses is restricted by the following assumptions.

- In stage 1, *hwæþer* is a question pronoun with the meaning 'which of the two'. This use is the earliest attested one.
- In the final stage, *hwæþer* is a question complementizer for polar questions. This use is attested latest and persists today.
- Reanalysis must have taken place, as the logical type of *hwæþer* is different in the first and last stage.
- Grammars that support different types of *hwæþer*-questions in adjacent stages must be minimally different. Any reordering of stages would stipulate adjacent grammars that differ more.

We submit that these principles allow us to hypothesize the diachronic order of sentence types even without a data record based on corpora. The reasoning could be likened to reasoning in archaeology where a sparse record of specimens can be tied together by assumptions about universal evolutionary processes.

Section 4.1 treats original *hwæþer* in the question pronoun sense. Section 4.2 proposes possible bridging examples, and 4.3 discusses their reanalysis. Section 4.4 argues that type 2 and type 4 examples can be viewed as actualizations of the resulting grammar if we allow for a limited amount of non-canonical steps in syntax/semantics. Section 4.5 relates them to canonical questions with partial *wh*-movement in OE, thus confirming that pedagogical questions in OE come in many varieties. This supports our final proposal in 4.6: unembedded *hwæþer*-questions have grammaticalized the restriction to pedagogical discourse as their use-conditional content.

4.1 Gothic

We take our start from Gothic *ƕaþar* in the sense 'which of two', which we assume stands in for the unattested Northwest Germanic precursor stage of Old English *hwæþer*. *ƕaþar*/*hwæþer* shares the meaning of *which* over a contextually given

domain D of size 2. We propose that D is instantiated by an assignment function *g* that captures deictic parameters in context. D can be specified by appositive clauses (as in the choice between Barabbas or Christ in Example (3) above) or by context alone. The logical type of elements in domain D is determined by the predicate to which *hwæþer* contributes. In syntax, *hwæþer* is base-generated as an argument and then raised to Spec,CP in both embedded and matrix questions.

Stage 1:

Syntax: *Hwæþer* is argument of the matrix clause verb.
Hwæþer is raised to Spec,CP and leaves a coindexed trace t_i.
Semantics: $[\![\text{ hwæþer }]\!]^{w,g} = D$, Psp: $|D| = 2$
$[\![\text{ hwæþer }]\!]^{w,g} = \{A,B\}$ combines with further parts of the sentence by pointwise composition (Hamblin 1973).
Hwæþer is of flexible type. D can be domains of type e or of type <s,t>, as in sentences like *hwæþer do you believe, S or T?*

The following derivation illustrates the syntax and meaning of a simple example (with *ad* being the addressee of the utterance in context). The interpretation of questions proceeds by combining sets of denotations, resulting in a set of possible answers (Hamblin 1973). The shift from declarative to question meaning is usually triggered by question syntax as opposed to declaratives, but we will argue that speakers at this stage would also opportunistically use this mode to compute meanings for sentences in non-canonical syntax.

(11) Hwæþer do you want, Barabbas or Christ?
1. LF structure: [hwæþer$_1$ do you want t_1]
2. $[\![\text{ you want } t_1]\!]^{w,g} = \{ \text{WANT}_w (\text{ad},t_1) \}$
3. $[\![\text{ hwæþer}_1]\!]^{w,g} = \{\text{Barabbas, Christ}\}$
4. $[\![\text{ hwæþer}_1 \text{ you want } t_1]\!]^{w,g}$
$= [\![\text{ hwæþer}_1]\!]^{w,g} \oplus [\![1]\!]^{w,g} \oplus [\![\text{ you want } t_1]\!]^{w,g}$
$= [\![\text{ hwæþer}_1]\!]^{w,g} \oplus \{ \lambda t_1 .\text{WANT}_w (\text{ad} , t_1) \}$
$= \{\text{Barabbas, Christ}\} \oplus \{ \lambda t_1.\text{WANT}_w (\text{ad}, t_1) \}$
$= \{\lambda w.\text{WANT}_w (\text{ad, Barabbas}), \lambda w.\text{WANT}_w (\text{ad, Christ})\}$

The analysis thus predicts the following denotation: {'You want Barabbas', 'You want Christ'}. In an information seeking question, the speaker requests the addressee to tell which of the two is the case. This stage is still attested in Old English, in the form of questions of Type 1.

4.2 Embedded sentences of type 2

The next crucial step focusses on examples that allow for an analysis as part of the Gothic (and pre-Old English) grammar, as well as an analysis closer to modern *hwæþer*. Given that all later versions of *hwæþer* are restricted to sentences in subordinate clause syntax, we must start from sentences where *hwæþer* occurs in an embedded context. We argue that *hwæþer* should moreover be an argument of a verb of belief or opinion, as we find it in the following example (speaking about the transient nature of wealth).

(12) Sege me nu hwæðer þu æfre gehyrdest þæt he angum þara þe
say me now whether you ever heard.SBJV that it to-any those who
ær us wære eallunga þurhwunode. (OE Boethius 29: 8–9)
earlier us was entirely persisted.
'Tell me now whether you have ever heard that it [= wealth] persisted in full for any of those who were before us.'

If we allow for a small irregularity in the explication of the alternatives, the structure of the embedded clause could rest on Gothic *hʋaþar.*

(13) … *hwæþer* ('which') you heard: That wealth stayed with any of those before us (or that it never stayed).

The assumed structure (13) is situated between sentences with an explicated domain of choice ('that it stayed or that it did not stay') and those with an implicit domain of choice. The choice between *p* and *non-p* as a complement of *you heard* could be easily construed from *p* alone. The structure of the embedded question shows *hwæþer* as a complement of *hear*, with an explicated domain 'that S or that non-S'.

(14) [*hwæðer*$_i$ [$_{IP}$ þu gehyrdest t$_i$]] [þæt S (or þæt non-S)]

Bear in mind that we do not claim that *specifically* (12) was a first bridging example. Assuming the structure in (14), the example can be analyzed as in Section 4.1 and yields the following denotation.

(15) {'you heard that wealth stayed with someone before us',
'you heard that wealth never stayed with anyone before us'}

This question provides the complement of the matrix clause *sege me nu* 'Tell me now'. This imperative requests the addressee to tell which of the two propositions is true – which is tantamount to answering the question in (15).

4.3 Reanalysis

Embedded Type 2 examples can be produced and analyzed by the Gothic (and general early Germanic) grammar, assuming that the explicated domain of *hwæþer* can be partially elided (cf. the proposal in Walkden 2014: 154–155). The hearer has to first construe an elided *non-S* in order to interpret ⟦ *hwæþer* ⟧^{w,g} = {S, non-S}. It would be less costly to assume that *hwæþer* directly combines with a proposition S to form {S, non-S}, as proposed for *whether* in Modern English (Hamblin 1973).[7] With this assumption, *hwæþer* no longer is a cataphor but enters an operator-argument relationship with *þæt S*. At LF, the most plausible position for it would be next to the complementizer *þæt.* The new syntactic structure of (13) is given in (16).

(16) [$_{CP}$ *hwæþer*$_i$ [$_{IP}$ þu gehyrdest [t$_i$ þæt S]]
whether you heard.SBJV that S

We have to leave the details of the complex subordination *hwæþer þæt* open. As the CP in (16) is a complement clause of *sege me nu*, *hwæþer* plays a double role as a syntactic subordination (for matrix clause *sege me nu*) and a question word, combining with *that S* in the lowest clause. Plausibly, *hwæþer* is first Merged in the lower Spec,CP and moves to the higher position; this could be taken to correspond to the operator stage of van Gelderen (2009).

(17) LF structure: [$_{IP}$ þu gehyrdest [*hwæþer*$_i$ *þæt* S]]
1. ⟦ hwæþer ⟧^{w,g} = λpλq(q = p ∨ q = ¬p)
2. ⟦ hwæþer S ⟧^{w,g} = λq(q = ⟦ S ⟧ ∨ q = ¬⟦ S ⟧)
3. ⟦ þu gehyrdest ⟧^{w,g} = { λp.HEARD$_w$(**ad**, p) } combines with (2.) pointwise to yield
4. { λw.HEARD$_w$(**ad**, ⟦ S ⟧), λw.HEARD$_w$(**ad**, ¬⟦ S ⟧) }
={'you heard p', 'you heard non-p'}

Combined with the meaning of *sege me nu*, we predict the literal meaning 'Tell me: Did you hear that S, or did you hear that not-S'. The overall sentence is still a request to the addressee to say which of S, non-S they believe to be true (she has hearsay evidence for, to be precise).

Two factors stand in favour of the proposed reanalysis. Firstly, the denotation of *hwæþer* in (17.1) is the denotation of its modern descendant.[8] Secondly, *hwæþer* changed from cataphor to function, which is in line with the general trend of grammaticalization leading towards functional elements.

7. To be precise, Hamblin's denotation for *Is it the case that* can also serve as denotation for the question complementizer *whether.*

8. We use the easier-to-read notation λp{p, ¬p} in the following.

However, the new structure suffers from new syntactic irregularities. The supposed syntax attributes a double status to *hwæþer* as a syntactic complementizer-like element in the higher clause and a semantic operator in the lower clause. This might be a reason why speakers experimented further with the pattern. The next section argues that both type 2 and type 4 examples can be viewed as new variants where syntactic and semantic functions of *hwæþer* are in better match.

4.4 Varieties of actualization: Type 2 and type 5 examples

The present section turns to type 2 and type 5 examples, building on (17) in the preceding section. Let us begin with type 2, unembedded *hwæþer*-questions about speaker's beliefs. We assume that they arise by simply leaving out the matrix sentence in (17).

(18) hwæðer þu woldest cweðan þæt he wære unwyrðe anwealdes and
whether you wanted say that he be.SBJV unworthy power.GEN and
weorðscipes (OE Boethius 27: 40–41)
honour.GEN
'would you say that he is unworthy of power and honour?'

Hwæðer combines with *þu woldest cweðan* where the finite verb precedes the non-finite verb. This order is typical for subordinate clauses in Germanic languages that exhibit the verb-final/verb-second opposition. Type 2 examples thus pattern with the embedded question in (12) and speakers extended the pattern to non-embedded questions. This might suggest that the embedders ('say me') did not add to the semantics of the utterance, but this remains speculative. Using the semantic operations in (17), (18) can be analyzed as follows.

(19) LF structure:
[$_{IP}$ hwæðer þu woldest cweðan [t_i þæt he wære unwyrðe anwealdes and weorðscipes]]
1. $[\![$ hwæðer $]\!]^{w,g}$ = λp{ p, ¬p} is used in its modern sense.
2. $[\![$ þæt he wære unwyrðe anwealdes and weorðscipes $]\!]^{w,g}$
= λw.UNWORTHY$_w$(**He**, **Power&Honour**) =: **p**
The embedded clause contributes the proposition 'he was unworthy of power and honour', abbreviated as **p** in the following.
3. We interpret *hwæþer* in its underlying position, combining with **p**.
$[\![$*hwæðer*-þæt he wære unwyrðe anwealdes and weorðscipes$]\!]^{w,g}$ = { **p**, ¬**p**}
We compute the question meaning of 'Is he unworthy of power and honour?'
4. The matrix clause contributes the predicate 'you want to say q':
$[\![$ þu woldest cweðan $]\!]^{w,g}$ = { λq.SAY(**ad**, q) }

5. Matrix clause and embedded question compose pointwise:
{ λw.SAY(**ad**, UNWORTHY$_w$(**He**, **Power&Honour**)),
λw.SAY(**ad**,¬UNWORTHY$_w$(**He**, **Power&Honour**))}

The resulting question can be paraphrased as 'Do you say that he is unworthy etc., or do you say that he is not unworthy etc. – which of the two is it?' Remarkably, *hwæþer* in C seems to force pointwise composition at the matrix level (qualifying *hwæþer þu woldest cweðan* as a question) although *hwæþer* still semantically combines at the lower clause level. As a result we derive a pedagogical question meaning.

Although this type occurs with highest frequency in Boethius, it shares the markedness of the examples in 4.3: the relation between *hwæþer* and *þæt* is still unclear. Moreover, we must assume that pointwise composition was licensed beyond the CP level of questions.[9] We still believe that the analysis in (19) is on the right track, not least as it allows us to account for examples that seem to violate a semantic universal. We now turn to type 5 examples, illustrated in (20).

(20) Wenst þu hwæðer he mæge yfel don? (OE Boethius 35: 150)
think you whether he may.SBJV evil do?
'Can he [= God] do evil, do you think?'

The example shows an embedded question as a seeming complement clause of *wenan*. The verb *wenan*, cognate to German *wähnen*, has the same meaning as 'believe' in ModE, including erroneous belief. Believe verbs are generally incompatible with question complements (compare **He believed who came*).

Our preceding stage offers an alternative analysis for this type, assuming that (21) shows *hwæþer* overtly in the position we assumed for LF in (19). It is in the standard position to take the embedded clause as its argument. The predicate *wenan* and the matrix question syntax explicate the intended speech act.

(21) Wenst þu hwæðer he mæge yfel don?
think you whether he may.SBJV evil do?
1. LF: [wenst þu [hwæþer he may do any evil]]
2. ⟦hwæþer S⟧^{w,g}
= {'he may do evil', 'he may not do evil'}

9. We follow Uegaki (2019) in assuming that question-embedding predicates take questions as their semantic argument. They thus 'absorb' alternatives (Shimoyama 2006: 158) and stop pointwise composition. While Shimoyama (2006) offers convincing arguments for alternative expansion beyond clause boundaries in Japanese, the general picture for embedded questions in Germanic languages doesn't support alternative expansion as a standard move. For instance, English **Did Tom believe who wins the race?* cannot be interpreted in the sense 'who did Tom believe to be the winner?', as general alternative expansion would predict.

3. $[\![$wenst þu t_i $]\!]^{w,g}$ = { λp_i .BELIEVE$_w$ (**ad** , p_i) }
4. pointwise composition of (2) and (3):
{ λw.BELIEVE$_w$ (**ad**, 'he may do evil'), λw.BELIEVE$_w$ (**ad**, 'he may not do evil')}

Let us assess the properties of the proposed structure. *Hwæþer* no longer co-exists with complementizer *þæt,* and overt and LF positions match. Yet, the syntax-semantics interface in the matrix clause is still non-canonical: the clause shows subject-verb inversion, standardly triggered by a [Q] feature in the matrix clause CP-domain. Yet, if there is such a feature it must necessarily remain uninterpreted.[10] Instead, pointwise composition jumps in again to derive the (desired) question meaning.

In sum, we find that type 2 and type 5 sentences both serve to convey questions about the opinions and beliefs of the addressee (used as pedagogical questions where the speaker knows the answer already). Yet, both ways of expressing this come along with slight irregularities in syntax and semantics. While we cannot decide whether speakers saw these irregularities, we want to point out that OE had a type of questions where all these irregularities were healed. While they were not included in the record in Section 2, we nevertheless want to relate them to our data.

4.5 Irregularities resolved? *Hwæþer*-questions with partial wh-movement

This section presents questions where a matrix *what*-question 'what do you think' combines with a subordinate *hwæþer* question, as we see in (22).

(22) Hwæt wenst du nu, (…) hwæðer he sie swa ungesælig swa se
what believe you now (…) whether he is.SBJV so unworthy as he
þe nanwuht godes næfþ? (OE Boethius 38: 108–110)
who not.any good.GEN not-has
'What do you think now, (…) ? Would he [who has some element of good in him] be as unfortunate as one who had nothing good?'

(22) can be paraphrased as 'what do you think about the following question: Is *he* who has at least some good in him as unworthy as *he* who has no good at all?' Similar examples in modern Dutch, German and Russian are studied as 'partial *wh*-movement' (Reis 2000, Fanselow 2017), and the pattern in (22) is attested for more types of embedded questions in Old English as well. As OE data do not offer evidence for movement, *wh*-doubling or scope marking complementizers in

10. Cf. Godden & Irvine's (2009, I: 196) suggestion that *wenst þu* in Boethius may have the function of an interrogative tag.

general, we favour a base-generation analysis in which the two *wh*-elements do not form a syntactic chain, following Dayal (1994, 2000), Felser (2001). We assume that the *hwæþer*-question rests on the word's newer sense and serves to specify the search domain of *hwæt* in the matrix clause. According to this view, semantic composition proceeds in the following steps.

(23) Semantic composition of (22)
1. ⟦ he sie swa ungesælig swa se þe nanwuht godes næfþ ⟧^{w,g} = **p**2.
2. ⟦hwæþer S⟧^{w,g} = { **p**, ¬**p**}
3. ⟦wenst þu t_i ⟧^{w,g} = { λp_i .BELIEVE$_w$ (ad , p_i) }
4. ⟦hwæt$_i$⟧^{w,g} resumes ⟦hwæþer S⟧^{w,g}, therefore ⟦hwæt$_i$⟧^{w,g} ={ **p**, ¬**p**}
5. question denotation by standard composition of (3) and (4):
⟦hwæt wenst þu t_i ⟧^{w,g}
= {λw.BELIEVE$_w$ (ad , **p**), λw.BELIEVE$_w$ (ad , ¬**p**)}

This question type avoids several of the irregularities of the preceding examples. Firstly, pointwise semantic composition of question pronoun and verbal predicate remains within the CP. Secondly, *hwæþer* is analyzed as a question complementizer that is located in Spec,CP of the embedded clause, as standard would have it. Thirdly, the question type generalizes to other types of embedded questions in OE as well as in other Germanic languages. This corroborates our claim that sentence (22) avoids idiosyncratic steps in the derivation that were needed to account for question types 2 and 5.

4.6 Type 3: Unembedded *hwæþer*-questions

We turn to our initial example 'Hwæðer nu gimma wlite eowre eagan to him getio, heora to wundrianne', where *hwæþer* occurs in a polar question in verb-final syntax. There is no verb of saying or belief, but the speaker still aims to ask a pedagogical question. We propose that this shows a further step in grammaticalization: speaker intentions have fostered into use-conditional content. The question conveys that the speaker already knows the answer and requests the addressee to form their own opinion.

Let *p* = 'jewels attract your eyes' be the prejacent and S be the corresponding clause *gimma wlite eowre eagan to him getio*. We propose that the example involves a homonymous complementizer *hwæþer*$_{exp}$ that conveys use-conditional content. We moreover assume that (1) has a tacit embedding predicate *sege me* 'tell me …' which accounts for the subordinate clause syntax of the prejacent, and is restricted to the expressive complementizer *hwæþer*$_{exp}$. We notate two-dimensional content as <π • ε> with the propositional content π and use-conditional content ε.

Use-conditional content will be promoted to the top level in semantic composition (Potts 2005, Gutzmann 2015). Semantic composition proceeds as follows.

(24) LF with tacit embedding predicate[11]
[[sege me]$_{\emptyset}$ [$_{CP}$ hwæþer$_{exp}$ C^{o} [$_{TP}$ S]]]

(25) $[\![\text{hwæþer}_{exp}]\!]^{w,g}$
= λp < {p, ¬p} • '*sp* knows answer to {p, ¬p} and *sp* requests *ad* to think about {p, ¬p}' >
turns p into question {p, ¬p}, and conveys: 'the speaker knows the answer to {p,¬p} and requests addressee to give their opinion on {p, ¬p}.'

(26) $[\![$ *hwæþer* jewels attract your eyes $]\!]^{w,g}$
= < {'jewels attract your eyes', 'jewels don't attract your eyes'} • *sp* knows answer to Q and *sp* requests *ad* to think about Q>
Q = {'jewels attract your eyes', 'jewels don't attract your eyes'}

(27) $[\![\textit{sege me}\ t_i\ ]\!]^{w,g} = \lambda p_i\ .\text{TELL!}_w\ (\text{ad}\ ,\ p_i\ )$

(28) Composition with question Q, propagate expressive content.
< λw .TELL!$_w$ (ad , Q) • *sp* knows answer to Q and *sp* wants *ad* to think about Q >

The composition in (28) rests on the composition of verba dicendi with questions (Uegaki 2016, 2019). We gloss the contribution of imperatives as TELL![12] The imperative requests the addressee to provide suitable answers. The expressive content conveys that the speaker knows the answer but wants to hear the addressee's opinion.

Treating (1) as a request for an answer may seem too strong, given that Wisdom rarely waits for B to respond. However, the source text contains numerous explicit imperatives of the form '*Sege me* Q' where Wisdom likewise never stops to wait for B to answer.

(29) Ac gesege me nu, ic ascige þe, þu Boetius, hwi þu
and tell me now I ask you.NOM you.VOC Boethius why you.NOM
swa manigfeald yfel hæfdest (…)? (OE Boethius 27: 27–28)
so manifold evil had.SBJV(…)?
'But tell me now, I ask you Boethius, why you experienced such great evil [and so much hardship in office while you held it]?

11. We have represented *hwæþer* as occupying Spec,CP in (24), but nothing rests on this either syntactically or semantically.

12. For a formal treatment of imperatives see Portner (2007).

In fact, B here never has a chance to respond. The first utterance by B after (29) answers an entirely different question that has emerged in the meantime. We therefore assume that requests in (1), as well as (29) and elsewhere, can be overruled by Wisdom's actual aims, much in the same way as indirect speech acts can generally overrule literal speech acts (Searle 1969).

5. Summary and outlook

The Old English Boethius offers a rich spectrum of embedded and unembedded *hwæþer*-questions that we analyzed as stages in the development of a question pronoun (similarly attested in Gothic) to a question complementizer in Modern English. We argue that unembedded *hwæþer*-questions are pedagogical questions. While the analysis of most examples requires special assumptions about their syntax and semantics, we propose a partial ordering where every new construction deviates minimally from the preceding construction(s). Type 1 "*hwæþer* of the two, X or Y?" rests on the earlier question pronoun. We propose that embedded Type 2 examples "Tell me *hwæþer* you believe that q?" are the most likely bridging structure that can be captured by earlier and later grammatical stages. We submit that no other attested construction in Boethius can be captured by an earlier *and* later grammar with fewer extra assumptions. From these, we suggest, Type 2 examples "*hwæþer* you believe that q?" arose by elision of the matrix clause. We assume that type 5 questions "Do you believe *hwæþer* p?"are the first in our ordering to show *hwæþer* in the position of a question complementizer. Semantic composition in type 5 uses pointwise composition of matrix predicate and complement question (as in Type 2 examples), which renders the syntax-semantics interface slightly irregular. Yet, the type patterns with so-called partial movement, following the resumptive account in Dayal (1994).

Type 3 questions *hwæþer p?* in subordinate clause syntax exhibit an expressive homonym $hwæþer_{exp}$ and a tacit matrix imperative 'tell me …'. $Hwæþer_{exp}$ comes close to modal particles, as it serves to contribute use-conditional content. It did not develop into a proper modal particle, as suggested by Coniglio's hypothesis that modal particles rest on earlier homonymous adverbs (Coniglio, this volume).

Finally, Boethius includes embedded questions of the form "I wanted to ask you *hwæþer* p", listed as Type 4 above. These can be captured by *whether* in its modern syntax and semantics. We leave the derivation to the readers.

Our proposal rests on the evaluation of a single historical text, and should be tested against further data. Yet, we maintain that the OE Boethius can be viewed as a diachronic "sediment" where many, perhaps all stages in the grammar of

hwæþer are conserved.[13] Among the documented uses, some rest on a less ideal syntax-semantics interface than others. They leave the syntactic status of *hwæþer* vague, they rest on LF positions of a complementizer that differs from its surface position, or they use semantic modes of composition at non-standard places in the derivation. From a theoretical point of view, it is tempting to hypothesize that these slight irregularities spurred further variation until *hwæþer* in its modern grammar had emerged. This view could also reconcile the conceptual opposition between gradual and categorial change – small categorial changes leading to irregular lexical entries could be conceived of as gradual in that they already bear the seeds of further changes to come. We will leave this theme for future research.

References

Allen, Cynthia L. 1980. *Whether* in Old English. *Linguistic Inquiry* 11: 789–793.

Bately, Janet M. 2009. Did King Alfred actually translate anything? The integrity of the Alfredian canon revisited. *Medium Ævum* 78: 189–215. https://doi.org/10.2307/43632837

Bately, Janet M. 2015. Alfred as author and translator. In *A companion to Alfred the Great*, Nicole Guenther Discenza & Paul E. Szarmach (eds), 113–142. Leiden: Brill.

Berizzi, Mariachiara. 2010. Interrogatives and relatives in some varieties of English. PhD dissertation, Università degli Studi di Padova.

Bosworth, Joseph and Northcote Toller, T. 1898. *An Anglo-Saxon dictionary, based on the manuscript collections of the late Joseph Bosworth*. Oxford: Clarendon. <https://bosworthtoller.com/> (11 March 2021).

Campbell, Lyle. 2001. What's wrong with grammaticalization? *Language Sciences* 23: 113–161. https://doi.org/10.1016/S0388-0001(00)00019-X

Coniglio, Marco. (this volume). On the adverbial origin of German modal particles. *Discourse Particles*, xx–xx.

Dayal, Veneeta. 1994. Scope marking as indirect *wh*-dependency. *Natural Language Semantics* 2: 137–170. https://doi.org/10.1007/BF01250401

Dayal, Veneeta. 2000. Scope marking: Cross linguistic variation in indirect dependency. In *Wh-scope marking*, Uli Lutz, Gereon Müller & Arnim von Stechow (eds), 157–194. Amsterdam: John Benjamins. https://doi.org/10.1075/la.37.07day

den Besten, Hans. 1989. On the interaction of root transformations and lexical deletive rules. In Hans den Besten (ed), *Studies in West Germanic syntax*, 14–100. Amsterdam: Rodopi. [First published as den Besten, Hans. 1983. On the interaction of root transformations and lexical deletive rules. In *On the formal syntax of the Westgermania*, Werner Abraham (ed), 47–131. Amsterdam: John Benjamins].

Den Dikken, Marcel. 2006. *Either*-float and the syntax of co-*or*-dination. *Natural Language and Linguistic Theory* 24, 689. https://doi.org/10.1007/s11049-005-2503-0

13. On further potential analogies between geology and historical linguistics see Walkden (2019).

Eckardt, Regine & Andrea Beltrama. 2019. Evidentials and Questions. In *Empirical Issues in Syntax and Semantics* 12, Chris Pinon (ed), 121–155. Paris: CNRS.

Eckardt, Regine. 2020. Conjectural questions: the case of German Verb-final 'wohl' questions. *Semantics and Pragmatics* 13. https://doi.org/10.3765/sp.13.9

Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and empirical foundations*, Irina Nikolaeva (ed), 366–431. Oxford: Oxford University Press.

Eyþórsson, Þórhallur. 1995. Verbal syntax in the early Germanic languages. PhD dissertation, Cornell University.

Fanselow, Gisbert. 2017. Partial *wh*-movement. In *The Blackwell companion to syntax*, Martin Everaert & Henk van Riemsdijk (eds), 437–492. Oxford: Wiley-Blackwell.

Felser, Claudia. 2001. *Wh*-expletives and secondary predication: German partial *wh*-movement reconsidered. *Journal of Germanic Linguistics* 13, 5–38. https://doi.org/10.1017/S1470542701003178

Fischer, Olga, van Kemenade, Ans, Koopman, Willem, & van der Wurff, Wim. 2000. *The syntax of early English*. Cambridge: Cambridge University Press.

Fox, Samuel. 1864. *King Alfred's Anglo-Saxon Version of Boethius De consolatione philosophiae, with a literal English translation, notes, and glossary*. London: H. G. Bohn.

Godden, Malcolm. 2007. Did King Alfred write anything? *Medium Ævum* 76: 1–23. https://doi.org/10.2307/43632294

Godden, Malcolm, & Irvine, Susan (eds). 2009. *The Old English Boethius: an edition of the Old English versions of Boethius's De consolatione philosophiae*. 2 vols. Oxford: Oxford University Press.

Guenther Discenza, Nicole. 2015. The Old English Boethius. In *A companion to Alfred the Great*, Nicole Guenther Discenza & Paul E. Szarmach (eds), 200–226. Leiden: Brill. https://doi.org/10.1163/9789004283763_009

Gutzmann, Daniel. 2015. *Use conditional meaning*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198723820.001.0001

Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language* 10(1): 41–53.

Hopper, Paul J. and Traugott, Elizabeth C. 2003. *Grammaticalization*. 2nd ed. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1(1): 3–44. https://doi.org/10.1007/BF00351935

Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In *Clause structure and language change*, Adrian Battye & Ian Roberts (eds), 140–169. Oxford: Oxford University Press.

Levinson, S. C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge MA: MIT Press. https://doi.org/10.7551/mitpress/5526.001.0001

Mitchell, Bruce. 1985. *Old English syntax*. 2 vols. Oxford: Clarendon. https://doi.org/10.1093/acprof:oso/9780198119357.001.0001

Nakajima, Heizo. 1996. Complementizer selection. *The Linguistic Review* 13, 143–164. https://doi.org/10.1515/tlir.1996.13.2.143

Nielsen, Hans Frede. 1998. *The continental backgrounds of English and its insular development until* 1154. Odense: Odense University Press. https://doi.org/10.1075/nss.19

Parra-Guinaldo, Victor. 2013. Reanalysis of Old English *hwæðer* in the left periphery. PhD dissertation, Arizona State University.

Portner, Paul. 2007. Modals and Imperatives. *Natural Language Semantics* 15(4), 351–383. https://doi.org/10.1007/s11050-007-9022-y

Potts, Christopher. 2005. *The Logic of Conventional Implicature*. Oxford: Oxford University Press.
Reis, Marga. 2000. On the parenthetical features of German *Was … w*-constructions and how to account for them. In *Wh-Scope Marking*, Uli Lutz, Gereon Müller & Arnim von Stechow (eds), 359–408. Amsterdam: John Benjamins. https://doi.org/10.1075/la.37.14rei
Ringe, Don. 2006. *From Proto-Indo-European to Proto-Germanic* [A linguistic history of English vol. 1]. Oxford: Oxford University Press.
Roberts, Ian. 1996. Remarks on the Old English C-system and the diachrony of V2. In *Language change and generative grammar*, Ellen Brandner & Gisela Ferraresi (eds), 154–164. Linguistische Berichte, Sonderheft 7. Opladen: Westdeutscher Verlag. https://doi.org/10.1007/978-3-322-90776-9_6
Salvesen, Christine Meklenborg & Walkden, George. 2017. Diagnosing embedded V2 in Old English and Old French. In *Micro-change and macro-change in diachronic syntax*, Eric Mathieu & Robert Truswell (eds), 168–181. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198747840.003.0011
Searle, John. 1969. *Speech Acts*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139173438
Sedgefield, Walter J. (ed). 1899. *King Alfred's Old English Version of Boethius* De Consolatione Philosophiae. Oxford: Clarendon.
Shimoyama, Junko. 2006. Indeterminate Phrase Quantification in Japanese. *Natural Language Semantics* 14(2), 139–173. https://doi.org/10.1007/s11050-006-0001-5
Taylor, Ann, Warner, Anthony, Pintzuk, Susan, & Beths, Frank. 2003. *York-Toronto-Helsinki Parsed Corpus of Old English Prose*. Heslington, York: University of York. <https://www-users.york.ac.uk/~lang22/YCOE/YcoeHome.htm> (25 March 2021).
Theiler, Nadine, Roelofsen, Floris & Aloni, Maria. 2017. What's wrong with believing whether? In *Semantics and Linguistic Theory* (SALT) 27, Maryland. https://doi.org/10.3765/salt.v27i0.4125
Traugott, Elizabeth C. 1972. *A history of English syntax: a transformational approach to the history of English sentence structure*. New York, NY: Holt, Rinehart & Winston.
Traugott, Elizabeth C. 2017. 'Insubordination' in the light of the Uniformitarian Principle. *English Language and Linguistics* 21(2): 289–310. https://doi.org/10.1017/S1360674317000144
Truckenbrodt, Hubert. 2006. On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics* 32(3): 257–306. https://doi.org/10.1515/TL.2006.018
Uegaki, Wataru. 2016. Content nouns and the semantics of question-embedding. *Journal of Semantics* 33(4): 623–660.
Uegaki, Wataru. 2019. The semantics of question-embedding predicates. *Language and Linguistics Compass*. https://doi.org/10.1111/lnc3.12308
Ukaji, Masatomo. 1997. A History of *Whether*. In *Studies in English Linguistics: A Festschrift for Akira Ota on the Occasion of His Eightieth Birthday*, Masatomo Ukaji, Toshio Nakao, Masaru Kajita & Shuji Chiba (eds), 1236–1261. Tokyo: Taishukan.
van Gelderen, Elly. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins. https://doi.org/10.1075/la.71
van Gelderen, Elly. 2009. Renewal in the left periphery: economy and the complementizer layer. *Transactions of the Philological Society* 107, 131–195. https://doi.org/10.1111/j.1467-968X.2009.01216.x
van Kemenade, Ans, & Links, Meta. 2020. Discourse particles in early English: Clause structure, pragmatics and discourse management. *Glossa: A journal of general linguistics* 5(1): 3, 1–23. https://doi.org/10.5334/gjgl.1020

Walkden, George. 2014. *Syntactic reconstruction and Proto-Germanic*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198712299.001.0001

Walkden, George. 2017. Language contact and V3 in Germanic varieties new and old. *Journal of Comparative Germanic Linguistics* 20(1): 49–81. https://doi.org/10.1007/s10828-017-9084-2

Walkden, George. 2019. The many faces of uniformitarianism in linguistics. *Glossa: a journal of general linguistics* 4(1), 52: 1–17. https://doi.org/10.5334/gjgl.888

Walkden, George. 2021. Do the wealthy stay healthy? Rich agreement and verb movement in early English. *Journal of Historical Syntax* 5 (30), 1–28.

Walkden, George & Booth, Hannah. 2020. Reassessing the historical evidence for embedded V2. In *Rethinking verb second*, Rebecca Woods & Sam Wolfe (eds), 536–554. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198844303.003.0022

CHAPTER 3

The discourse particle *es que* in Spanish and in other Iberian languages

Manuel Pérez-Saldanya and José Ignacio Hualde
Universitat de València / University of Illinois at Urbana-Champaign

We examine the historical reanalysis of the Spanish sequence *es que* (lit. 'is that') as a discourse particle with justificatory and even purely emphatic functions. We argue that the diachronic process involved first the appearance of non-coindexed *pro* as subject of the copula *es* 'is' and, at a later stage, deletion of the empty pronoun and syntactic restructuring. The restructuring was triggered by the use of the construction in counter-argumentative and other contexts where the semantic content of the null category was not easily recoverable. We also consider parallel developments in Catalan and Portuguese, as well as the borrowing of the particle *es que* in colloquial Basque.

Keywords: particle, counter-argumentative, justificatory, reduced clefts, grammaticalization

1. Introduction

In all the Ibero-Romance languages the sequence *es que* 'is+that' (in Spanish) ~ *é que* (Portuguese and Galician) ~ *és que* (Catalan) ~ *ye que* (Asturian) has developed as a justificatory or even purely exclamative discourse particle in certain contexts. The particle *eske*, borrowed from Spanish, is also frequent in colloquial Basque. This grammaticalization phenomenon is a different process from the grammaticalization of the question particle *est-ce que* in French, which contains the same elements, although the overall syntactic reanalysis of the structure has been similar (see, e.g., Waltereit 2018 for a recent analysis).

In Spanish, there is at least one other example of a V+COMP sequence having developed into a discourse particle with evidential value, namely, *dizque* 'it seems, apparently' < *dice* 's/he says' + *que* 'that', which is found in a number of Latin American Spanish varieties (see, e.g., Company 2018). The expression *es que,* however, has a much broader geographical diffusion, as it is found both in European and Latin American Spanish, and has much greater frequency.

https://doi.org/10.1075/la.276.03per

The sequence *es que* – and its variants in the other Ibero-Romance languages – is made up of the third person singular present tense of the copula *ser* 'to be', followed by the complementizer *que* 'that'. When used as a discourse particle, it is not readily translatable as 'it is that', but rather may correspond in English to expressions such as 'well, the thing is', or, depending on the conversational context, even 'well, you know', 'sorry' or '(it's) just that'.

This somewhat vague explanatory or justificatory use of *es que* is found in all the Ibero-Romance languages, including both European and American varieties. As just mentioned, the particle *eske* is also found in the colloquial Basque of bilingual Basque/Spanish speakers. By way of illustration, we give in (1) a Spanish example adapted from España (1996: 135, Example (18)), which we translated into Portuguese, Catalan and Colloquial Basque in consultation with speakers of these languages. The English translation is the same in every case as that provided for Spanish:

(1) a. Spanish
A: ¿Quieres tomar un café?
you.want take a coffee
'Do you want to have a cup of coffee?'
B: *Es que* tengo prisa.
is that I.have rush
'*Sorry*, I'm in a hurry.'

b. Portuguese
A: Quer tomar café?
you.want take coffee
B: *É que* eu tô com pressa.
is that I am with rush

c. Catalan
A: Que vols prendre un cafè?
that you.want take a coffee
B: *És que* tinc pressa.
is that I.have rush

d. Colloquial Basque
A: Kafe bat hartu nahi duzu?
coffee one take want AUX
B: *Eske* presa dut.
eske rush I.have

It should be noticed that the examples in (1) would also be grammatical without *es que*; but, without this element, speaker B's reply would not contain an explicit indication that the proposition expressed is to be interpreted as a justification for turning down speaker A's invitation. Speaker A would have to infer the relevance of speaker's B reply purely from contextual cues.

In the Ibero-Romance languages, the discourse particle in the examples above has transparently resulted from the grammaticalization of the two-word sequence *es/é/és* 'it is'+ *que* 'that' in sentences where the copulative verb lacks an overt subject. In Colloquial Basque, on the other hand, the discourse particle *eske* is clearly a borrowing from Spanish, since the constituent elements do not exist in Basque.

Not all occurrences of *es que* are to be analyzed as a discourse particle, however. The particle coexists nowadays with the sequence from which it arose by reanalysis and with other constructions that show different degrees of grammaticalization of the sequence *es que*. Thus, we do not have a particle, but rather a *pro*+V+COMP sequence, in the example in (2), whose Spanish version in (2a) we have taken from Romera (2009: 150) and have translated into Portuguese and Catalan in (2b) and (2c) respectively, in consultation with native speakers. In this example, *es/é/és* is a verb whose null subject *pro* is coindexed with the noun phrase in the preceding clause (*una cosa/uma coisa/una cosa* 'one thing'):

(2) a. Spanish
Una $cosa_i$ es cierta y pro_i *es que* seguiremos adelante.
one thing is certain and is that we.will.continue ahead

b. Portuguese
Uma $coisa_i$ é certa, (e) pro_i *é que* seguiremos em frente.
one thing is certain (and) is that we.will.continue ahead

c. Catalan
Una $cosa_i$ és certa, i pro_i *és que* continuarem endavant.
one thing is certain and is that we.will.continue ahead
'One thing is true, and *it is* that we will keep on going.'

Not all uses of *es que* are thus syntactically equivalent. As just mentioned, in present-day usage, the particle is found side by side with the *pro*+V+COMP sequence that, arguably, gave rise to it by grammaticalization, the two being homophonous (albeit with some prosodic differences pointed out below in Section 2). This is different from the case of *dizque* 'apparently', mentioned above, which is now phonologically distinct and thus easily distinguishable from the sequence *dicen que* 'they say that' or *dice que* 's/he says that', from which it arose by a process of grammaticalization. An important analytical issue is, thus, that of the correct syntactic analysis of *es que* in different contexts.

Several authors have pointed out that the justificatory construction arose from sentences like those in (2), but generally without providing evidence for this historical link. From a historical perspective, Romera (2009) is a pioneering work. Romera explains the creation of the justificatory construction with *es que* in Spanish as resulting from a process whereby the referentiality of the subject of the copular sentence was progressively weakened and lost. In Pérez Saldanya & Hualde (2021)

the historical evolution of this Spanish construction is also studied from the perspective of grammaticalization theory. As far as we know, there are no parallel studies for either Catalan or Portuguese.

In this paper, we will attempt to explain why sequences with the structure '(it) is that' function as discourse particles in some of their uses in the Ibero-Romance languages. As we will show, the historical trajectory that leads to the formation of this particle has striking similarities with the process by which so-called "insubordinate clauses" develop (Evans 2007).[1] In both cases, the starting point stems from complex sentences with elided elements where there is first a conventionalization of the elision and later a reanalysis.

In the remainder of the paper, we first identify different constructions with *es que* (Section 2). Then, we discuss the possible tests that can be used to determine whether *es que* functions as a V+COMP sequence or as a discourse particle (Section 3). After that, we show that *es que* may function as a particle in sentences with an explanatory function (Section 4). Subsequently, we sketch the historical process that has led to the development to the particle in Spanish, distinguishing three diachronically-ordered stages (Section 5). Our account here essentially follows that given in Pérez-Saldanya & Hualde (2021), where additional historical facts are examined. In Section 6, we consider the facts of Portuguese and Catalan, which substantially parallel those of Spanish. A brief summary is presented in Section 7.

2. Constructions with the sequence *es que*

In the contemporary Ibero-Romance languages we find four types of related constructions where the copula is immediately followed by the complementizer *que* 'that'. These four constructions are the following: (a) pseudo-clefts and similar copulative sentences, (b) reduced clefts, (c) inferential reduced clefts, and (d) constructions with the justificatory particle *es que*. As we will show in detail in Section 5, these four types of constructions are diachronically related and allow us to trace a historical trajectory that culminates with the creation of the discourse particle *es que*. Let us consider now the specific properties of the four constructions, with Spanish examples.

a) The construction with the particle *es que* ultimately derives from pseudo-cleft constructions like that in (3a) and comparable copular sentences like the example in (3b):

1. We are grateful to an anonymous reviewer for pointing out this connection to us.

(3) a. Lo que quiero decirte *es que* tienes que ser
PRON that I.want say.to.you is that you.have that be
más cuidadoso.
more careful
'What I want to tell you is that you have to be more careful.'
b. El problema *es que* tienes que ser más cuidadoso.
'The problem is that you have to be more careful.'

Cleft constructions, and, in particular, pseudo-clefts are special types of copulative sentences. Cleft constructions consist of a free or semi-free relative clause, a copular verb and a cleft phrase that bears sentential focus.[2] In pseudo-clefts the relative clause is in initial position and is followed by the copula and the cleft phrase. The example in (3a) is a pseudo-cleft, which contains a semi-free relative clause (*lo que quiero decirte* 'what I want to tell you'), a copulative verb (*es* 'is') and a cleft phrase, which is a noun clause headed by the complementizer *que* 'that' (*que tienes que ser más cuidadoso* 'that you have to be more careful').

Both clefts and pseudo-clefts are identificational copular sentences; more exactly, identificational specificative sentences.[3] They can be defined as identificational because they establish a relationship of identity between both constituents linked by the copula. They are also specificative because the relative clause introduces a variable that is specified by the cleft phrase. For this reason, the sentence in (3a) can be paraphrased as follows: 'There is one thing x I want to tell you, and x = that you have to be more careful'.

Copular sentences like the example in (3b) above are also identificational and specificative. The only difference they have with respect to pseudo-clefts is that the variable saturated by the post-copular constituent is expressed by means of a definite noun phrase: 'there is a problem x and x = you have to be more careful'.

b) In the constructions exemplified in (3), the constituent that introduces the variable functions as the syntactic subject of the clause. Informationally, the subject is thematic in contrast with the cleft phrase, which is the most informatively relevant element. This thematic character of the subject explains why it can be deleted in those contexts where it can be anaphorically recovered, as in the examples in (2) above and in the following example:

---

2. Considering the order of the elements, three types of cleft constructions are often differentiated: the prototypical cleft, or *it-cleft* in English (*It is a book that I want (not a potato)*), the pseudo-cleft (*What I want is a book*) and the inverted pseudo-cleft (*A book is what I want*).

3. We follow the classification of copulative sentences adopted, among others, by Declerck (1988: 1–54) and Fernández Leborans (1999: 2398–2403). For a classification with some points of divergence and different terminology, see Mikkelsen (2019).

(4) Quiero decirte una cosa y *es que* tienes que ser
I.want say.to.you one thing and is that you.have that be
más cuidadoso.
more careful
'I want to tell you one thing, and it is that you have to be more careful.'

As was already mentioned regarding the examples in (2), in this case the subject of the copular sentence is an empty pronoun *pro*, which is co-referential with the preceding noun phrase *una cosa* 'a thing' and can be paraphrased as 'what I want to tell you'. We will refer to these constructions as reduced cleft sentences, following Declerk (1992: 217) and Fernández Leborans (1992: 236, 239).

c) The third type of construction relevant for our study shares with the one exemplified in (4) the fact that these are also constructions that do not have an explicit syntactic subject; but they differ in that the subject cannot be anaphorically recovered. Consider, for instance, the following dialogue:

(5) A: Trataré de ser más cuidadoso la próxima vez.
I.try.FUT of be more careful the next time
'I will try to be more careful next time.'
B: No es que tengas que ser más cuidadoso, *es que* tienes que
not is that you.have that be more careful is that you.have that
ser más precavido.
be more cautious
'It is not that you have to be more careful, *it is that* you have to be more cautious.'

In the example in (5), there is no constituent in Speaker A's utterance that could serve as an antecedent for the empty subject of the copula in Speaker B's utterance. Nevertheless, we may still assume the presence of a null subject whose referent is interpreted from discourse inferences (Declerck 1992; Fernández Leborans 1992; Delahunty 2001). Notice that, in this case, the construction with *es que* has a corrective character and is used to negate a previous interpretation or explanation (*no es que tengas…sino que* 'it is not the case that you have to… but '), which may be followed, as in the example in (5), by the interpretation that the speaker considers to be preferable (*es que tienes…* 'it is the case that you have to…'). Given this corrective value, an interpretation such as 'the issue' can be inferred for the empty subject of the copula.

These are thus constructions that can also be analyzed as reduced clefts, but of an inferential type, in the sense that they cannot be interpreted based only on their propositional content and require the speaker to infer some kind of relation between the sentence and the discourse context (Delahunty 1990: 20; Declerk

1992: 204; Dufter 2008: 1765–1767). According to this process of inference from the context, the empty subject of the reduced cleft may be equivalent to 'the issue', 'what (really) happens', etc. These inferential constructions present discursive functions involved in illocutionary differentiations, but in them *es que* does not function as a discourse particle.

d) In the last type of *es que* construction that we are considering here, the sequence *es que* functions as a fixed complex unit with a specific meaning of its own. In these constructions, it is difficult and sometimes impossible to posit any null subject at all. Rather, the sequence functions as a discourse particle that introduces a declarative sentence where some type of justification or reaction is expressed (Marín & Cuenca 2012; Cuenca 2013; Pérez Saldanya & Hualde 2021).[4]

As a particle, *es que* has a reactive illocutionary character, in the terms of Roulet (2006), since it introduces an intervention whose interpretation is linked to the preceding one, or, sometimes, to the extralinguistic context. In some cases, *es que* introduces a justification of what has been said or the speech act that has been performed, as in (6a). In some other cases, it introduces an excuse as a reaction to a request, a suggestion, a recrimination, etc., that has been expressed by the interlocutor in the preceding turn, as in (6b). It may also be the case that it adds emphasis in contexts where, for instance, an emotional reaction is being expressed, as in (6c). In the most extreme case of emphatic-emotional usage, it can cooccur with an interjection or exclamatory expression, as in (6d). In that last case, the use of *es que* can be described as "absolute" or "quasi-absolute," since it does not properly introduce any sentence.[5]

(6) a. Dímelo, por favor. *Es que* necesito saberlo.
say.to me.it please is that I.need know.it
'Tell me, please. I need to know.'

b. A: ¿Me podrías ayudar?
to.me you.could help
'Could you help me?'
B: *Es que* no tengo tiempo.
is that not I.have time
'*Well*, I don't have time.'

4. Cuenca (2013) argues that *es que* is a modal particle, rather than a discourse particle, because it modifies the illocutionary content of the utterance that it introduces. She adds, however, that, because of its reactive role, it is functionally similar to a discourse marker.

5. As an anonymous reviewer points out, there are clear similarities between this use of *es que* and certain uses of the expression *as if* in English (Brinton 2014), even if the discourse meaning is different.

c. A: ¿Te gustaría que fuéramos a París?
to.you like.COND that we.go.SBJV to Paris
'Would you like it if we went to Paris?'
B: ¡Claro que sí! ¡*Es que* me encantaría!
clear that yes is that to.me love.COND
'Of course! it would be fantastic!'

d. ¡*Es que*, vamos, qué personaje!
is that come.on what character
'Gee, what a character!'

As Marín & Cuenca (2012) point out, the justification expressed by means of *es que* may have an argumentative character if it is co-oriented with the previous discourse and introduces an argument in its favor, as in (6a); but it may also have a counter-argumentative character, if it is anti-oriented and introduces an argument against what has been said before or against what the interlocutor might expect, as in (6b). Additionally, *es que* may function essentially as an emphatic, intensifying particle, as in (6c)–(d). The limits between the different usages are often diffuse and a single construction may introduce at the same time an argument in favor of what the speaker has previously said and against the explicit or implicit opinion of the interlocutor, as in (7), which is akin to (6a) above:

(7) A: Dímelo, por favor.
'Tell me, please.'
B: Pero…
'But…'
A: *Es que* necesito saberlo.
is that I.need know.it
'*Come on*, I (just) need to know it.'

Because of the "fuzziness" just mentioned, the sequence *es que* is more clearly analyzable as a particle in some cases than in others. For instance, it would be easier to analyze (6a) as containing a null subject *pro* interpretable as 'the reason (I am asking you)'. Nevertheless, we believe there are good reasons to consider that in (6a) as well *es que* functions as a discourse particle. First of all, *es que* is a fixed expression. Secondly, although the construction has the same focalizing value that we find in clefts, this focalization is not oriented solely to the content of what is being said, but, rather, to the argumentation (Fuentes 1997: 245). For this reason, if we simply postulated a null subject with the interpretation 'the reason (for this)' we would not be capturing the entire pragmatic value of the construction. Notice that in (6a) the construction with *es que* does not introduce a simple, neutral explanation of why the speaker is making a request. Rather, the speaker is using this expression

to mitigate the pragmatic cost that the request may have, thus softening a potentially face-threatening request. This is in fact a typical function of many usages of the particle *es que* (España 1996: 134; Porroche 1998; Marín & Cuenca 2012: 77).

In order to strengthen the analysis just sketched, in the next section we consider a number of formal tests that buttress the conclusion that *es que* is a particle in justificatory constructions. Then, in Section 4, we analyze corpus examples of such constructions where, although more than one analysis may be possible, the particle analysis appears to be preferable.

3. V+COMP vs. discourse particle

As just mentioned, in certain contexts, *es que* is a discourse particle (which we will write as *es-que* for the sake of clarity, to distinguish it from other uses). In other contexts, however, it is simply part of a syntactic structure where the copulative verb *es* 'is' takes a clausal complement and has either a full subject or a null subject that can be recovered either anaphorically or through discourse inferences. A number of tests can be used to distinguish among these different syntactic structures; especially between constructions with a *pro* subject (coindexed or inferential) and clauses with the particle *es-que*. We will consider five tests of particlehood: (i) the morphophonological integrity of the sequence, (ii) the possibility of having negation or a topic before *es*, (iii) the lack of tense variation on the copula, (iv) the impossibility of having a left-dislocated topic preceding the sequence *es que*, and (v) its borrowability in Basque.

i. First of all, as pointed out by España (1996), sentences headed by the discourse particle *es-que* differ prosodically from those containing a V+COMP sequence *es que*. Whereas the two-word sequence 'is that' allows making a pause between the two words, the particle *es-que* is a single phonological unit that does not allow pausing between its two syllables. In the reduced cleft in (4), which we repeat in (8), the three dots indicate a possible pause between the copula *es* 'is' and the complementizer *que* 'that'. We analyze this construction as a *pro*+V+COMP sequence where *pro* is coindexed with the subject of the first clause, *una cosa* 'one thing':

(8) *pro*+V+COMP (reduced cleft):
Quiero decirte una cosa y *es… que* tienes que ser más cuidadoso.
'I want to tell you one thing, and *it is*… that you have to be more careful.'

A pause between the copula and the complementizer is also possible in the inferential reduced clefts in (5), repeated in (9), in which the null subject can be recovered from discourse inferences.

(9) *pro*+V+COMP (inferential reduced cleft):
A: Trataré de ser más cuidadoso la próxima vez.
'I will try to be more careful next time.'
B: No es… que tengas que ser más cuidadoso, *es… que* tienes que ser más precavido.
'It is not that you have to be more careful, *it is* that you have to be more cautious.'

Finally, a pause can also be inserted between *es* and *que* in the dialogue in (10), where in speaker B's reply we find a conventionalized pseudo-cleft construction headed by *lo que pasa* 'what happens':[6]

(10) Pseudo-cleft structure with *lo que pasa es* 'what happens is'
A: ¿Me ayudas?
to.me you.help
'Can you help me?'
B: Lo que pasa *es… que* esas cosas no se me dan
PRON that happens is…that those things not REFL to.me they.give
muy bien.
very well
'What happens is that I am not very good at those things.'

By contrast, in (11), speaker B could introduce a pause after *que*, but not after *es*.

(11) Discourse particle *es-que*
A: ¿Me ayudas?
'Can you help me?'
B: **Es… que* esas cosas no se me dan muy bien.
√*Es que*… esas cosas no se me dan muy bien.
'*Well,* I am not very good at those things.'

This prosodic test, due to España (1996), neatly distinguishes instances of *es que* where there is a subject position before the copula, either containing a null subject *pro*, as in (8) and (9), or an overt subject, as in (10), from other instances where *es-que* is a single word (11).

In a similar vein, a further test of wordhood is provided by the possibility or impossibility of inserting other elements between the two syllables of /eske/. A sentence like (12), where the adverb *solo* 'only' is inserted between *es* and *que* can only

6. This construction has a justificatory-contradictory value and is highly fixed (Reig 2011; Pérez Saldanya & Hualde, in press). Nevertheless, the fact that the relative clause appears in subject position allows the insertion of a pause between *es* and *que*.

be interpreted as containing the sequence '*pro*+is+that', where *pro* is interpretable as 'the problem', 'what (really) happens':[7]

(12) Sí, es muy buena persona, *es solo que* es muy pesado.
'Yes, he is a very good person; *it is only that* he is very tiresome.'

ii. A second test is provided by the (im)possibility of having negation before *es*. The particle *es-que* cannot be preceded by negation:

(13) *No *es-que* es muy pesado.
not is-that is very tiresome
'It is not that he is tiresome.'

Such negation is, however, perfectly grammatical in an example such as (14a), where the copula *es* has an explicit subject, and also in (14b), where it has a null subject that is recoverable through the discourse:

(14) a. El problema no es que es muy pesado, sino que…
the problem not is that is very tiresome but that
'The problem is not that he is very tiresome, but that…'
b. A: No me gusta quedar con él porque es muy pesado.
not to.me like go out with him as is very tiresome
'I do not like to go out with him because he is very tiresome.'
B: No es que sea muy pesado, sino que es muy aburrido.
'It's not that he is very tiresome, but *what happens is that* (lit.: but that) he is very boring.'

When *es que* functions as a justificatory particle, it relates to the illocutionary force of the utterance and falls outside of the scope of negation. In this sense it differs from other constructions with *es que*, where the scope of negation may include both the copula and the following subordinate clause.

iii. Thirdly, lack of tense variation can also be used as an argument for the status of *es-que* as a particle with discourse value. Variation in person or tense results in ungrammaticality in the intended interpretation, as shown by the difference in grammaticality between speaker B's replies in (15):

7. Intercalated adverbs between *es* and *que* may mostly appear in the class of inferential short clefts that we have called corrective (see Section 2), including those that negate a prior interpretation or explanation (*no es solo que…* 'it is not only that…') and those that introduce an explanation that the speaker considers more appropriate, as in (12). Out of the 66 examples of *es solo/sólo que* without an overt subject that we have found in CREA, 45 are of the former type and 21 of the latter.

(15) A: ¿Y no ibas con ellos?
and not you.went with them
'And weren't you going to go with them?'
B: *Es/*Era que* eran muy pesados.
Is/was that they.were very tiresome
'(The thing) is/*was that they were tiresome.'

This can be compared with the possibility of tense variation in other constructions either with a null subject (16a) or with an explicit subject (16b):[8]

(16) a. Una cosa era cierta y *era que* eran muy pesados.
'One thing was true and *it was that* they were very tiresome.'
b. Lo que pasaba *era que* eran muy pesados.
PRON that happened *was that* they.were very tiresome
'What happened was that they were very tiresome.'

iv. An additional difference has to do with the (im)possibility of having a left-dislocated topic preceding the sequence *es que* when we consider cases (a) to (c) in Section 2 above. In pseudo-clefts, it is not possible to left-dislocate a constituent from the post-copular subordinate clause to the beginning of the whole sentence. Thus, the object *a nosotros* 'to us' in (17a) can be placed after *que* within its clause, as indicated with the brackets in (17b), but not before the relative, as in (17c).

(17) a. Lo que es cierto *es que* nos importa mucho a nosotros.
PRON that is true is that to.us it.matters much to us
'The truth is that it matters a lot to us.'
b. Lo que es cierto *es que* [a nosotros] nos importa mucho.
c. *?[A nosotros] lo que es cierto *es que* nos importa mucho.

The same facts arise when the construction is a reduced cleft with a co-referential *pro*, namely left-dislocation of a constituent embedded in the *es que* structure is unavailable:

(18) a. Una cosa es cierta y *es que* nos importa mucho a nosotros.
one thing is true and is that to.us it.matters much to us
'One thing is true and *it is that* it matters a lot to us.'
b. Una cosa es cierta y *es que* [a nosotros] nos importa mucho.
'One thing is true and it is that, to us, it matters a lot.'

8. In all types of cleft sentences, the copula may either appear in the present tense or agreeing in time specification with the relative clause or, in the case of reduced clefts, with the preceding clause (Pérez Saldanya & Hualde, in press). The fact that this sort of tense agreement is not possible in certain contexts strongly suggests that that these constructions are not true clefts and that the sequence *es que* has been reanalyzed as a particle.

c. *Una cosa es cierta y [a nosotros] *es que* nos importa mucho.
*'One thing is true and, to us, it is that it matters a lot.'

In contrast, a left-dislocated topic may appear either before or after the particle *es-que*. The sentence in (18c) above would in fact be grammatical in a different interpretation in which we have an emphatic use of the particle *es-que*, which requires a different intonational contour:

(19) a. ¡A nosotros *es-que* nos importa mucho!
b. ¡*Es-que* a nosotros nos importa mucho!
'*Well*, to us it matters a lot!'

In sentences with multiple left-dislocated topics, the particle *es-que* can freely combine with them in any order. Thus, if we add the particle *es-que* to the sentence in (20a), all the possible word orders in (20b)–(d) with topicalization of both the direct object and the subject, are allowed. In (20b) the particle *es-que* precedes both topics. In (20c) the topicalized subject precedes *es-que* and the direct object follows the particle, and in (20d) we have the opposite situation. In (20e)–(f) both the subject and the direct object precede the discourse particle.

(20) a. Yo no aguanto a ese tío.
I not I.stand to that guy
b. [*es-que*] $[yo]_i$ $[a\ ese\ tío]_k$ $[e_i$ no le aguanto $e_k]$
is that I to that guy not him I.stand
[both subject and object fronted, following es-que]
c. $[yo]_i$ [*es-que*] $[a\ ese\ tío]_k$ $[e_i$ no le aguanto $e_k]$
[subject fronted, preceding es-que, object following es-que]
d. $[a\ ese\ tío]_k$ [*es-que*] $[yo]_i$ $[e_i$ no le aguanto $e_k]$
[object fronted, preceding es-que, subject following es-que]
e. $[yo]_i$ $[a\ ese\ tío]_k$ [*es-que*] $[e_i$ no le aguanto $e_k]$
[both subject and object fronted, preceding es-que]
f. $[a\ ese\ tío]_k$ $[yo]_i$ [*es-que*] $[e_i$ no le aguanto $e_k]$
[both object and subject fronted, preceding es-que]
'That guy, I just can't stand him.'

This freedom of word order is somewhat surprising given the fact that originally what we have is a sequence of verb + subordinating conjunction.

On the other hand, *es-que* can only occur before, and not after, the verum focus particle *sí que* 'truly', from *sí* 'yes'+ *que* 'that' (cf. (21b) and (21c)). This suggests that *es-que* occupies a position higher than FocP, as the structure we will later propose in (25b) indicates. For clarity and parallelism with *es-que*, we write this particle as *sí-que* in the following example:

(21) a. *Es-que* Juan *sí-que* mola.
is that Juan yes-that rocks
'Dude(s), Juan rocks, indeed!'
b. Juan *es-que sí-que* mola.
c. *Juan *sí-que es-que* mola.
d. **Sí-que es-que* Juan mola.

v. Finally, an interesting test for particlehood is its borrowability. A consequence of *es-que* having become a discourse particle is its availability in code-switching and borrowing. As Ibarra (2008, 2020: 407) points out, *eske* is very frequent as a sentence-initial element in Basque sentences in colloquial conversations among young bilingual Basque-Spanish speakers. In Basque there is no question of tracing the diachronic path of grammaticalization of *eske*. This element has simply been borrowed as a discourse particle from Spanish. The importance of this instance of borrowing is that it provides very clear evidence regarding the syntactic analysis of Spanish *es que* /eske/ in different contexts.[9] In (22) we provide one of the examples in Ibarra (2020), which contains two instances of *eske*:

(22) *Eske* papel guztik pegauta dare ia e… *Eske* ez da libratzen ezta
eske paper all glued are almost eh eske not AUX free.impf even
papel bat.
paper one
'*Eske*, all the pieces of paper are glued to each other… *eske* not even one is free'

Obviously, the phonological sequence /eske/ can be integrated into a Basque sentence only when it functions as a particle; not when it corresponds to a *pro*+V+COMP sequence in Spanish. The insertability of /eske/ in Basque sentences thus provides us with another test of particlehood, as Examples (23)–(24) show. In (23a), we have a particle and in (24a), we do not. For that reason, a translation of (24a) into Basque keeping the element *eske* is completely ungrammatical:

(23) a. *Es-que* para esas cosas no soy muy hábil.
is-that for those things not I.am very skillful
b. *Eske* gauza hori-etan ez naiz batere trebea.
eske thing those-in not I.am at.all skillful
'*The thing is,* I am not skillful at all in those things.'

(24) a. Una cosa$_i$ es cierta y pro$_i$ *es que* seguiremos adelante.
one thing is certain and is that we.continue.FUT ahead

9. Probably the borrowing of this element from Spanish should be considered akin to the common borrowing of English *so* by bilingual speakers of other languages (e.g. *so, a ver qué pasa* 'so, let's see what happens', in US Spanish).

b. *Gauza bat egia da eta *eske* jarraituko dugu aurrera.
thing one true is and eske follow.FUT AUX ahead
'*One thing is true, and it is that we'll keep going on.'

At least, these examples show us that bilingual speakers can analyze the sequence *es que* as a particle in the Spanish sentence in (23), whereas such an analysis is impossible in (24). Only in those contexts where Spanish *es que* is a particle, can it be borrowed by bilingual speakers. The considerable syntactic typological difference that exists between Spanish and Basque would make the borrowing of a *pro*+V+COMP sequence from Spanish into Basque rather difficult (although an intense contact can of course lead to the borrowing of syntactic constructions, see Harris & Campbell 1997: 124). On the other hand, a discourse particle may be easily borrowed from one language to another, regardless of the syntactic structure of the two languages. Together with the pause-insertion test, this provides one of the clearest tests for particlehood.

From the evidence presented in this section, it seems fairly clear that the phonological sequence /eske/ in Spanish may correspond to different syntactic structures. It is a sequence 'is+that' in clauses where the subject is either an explicit noun phrase or a null subject *pro* that can be recovered anaphorically or through discourse inference. In other instances, it has been reanalyzed as a fixed complex unit with justificatory or reactive functions. In these functions, we no longer have a *pro*+V+COMP sequence, but a discourse particle that modifies the illocutionary force of the utterance, introduces a justificatory or reactive speech act and occupies a high position in the left periphery of the sentence.

For this reason, and because of the syntactic position it occupies, which we show in (25), this particle can be analyzed as the morphological realization of declarative illocutionary force within Rizzi & Bocci's (2017) cartographic analysis. In order to show what we assume to be this syntactic position in more detail, in (25a) we have included a topicalized dative object (*a Juan*) preceding *es-que* and another topicalized noun phrase (*las partículas*), which functions as the subject of the verb, following it, in addition to a focus particle (*sí que*) and a modal adverb phrase (*realmente*). The structure is labelled in (25b):

(25) a. [*A Juan* [*es-que* [*las partículas* [*sí-que* [*realmente* [*le gustan*]]]]]
to Juan is-that the particles yes-that really to.him like
'Juan, it is-that, particles, he really likes them'
= 'As for Juan, *the thing is,* he really likes particles.'

b. [$_{\text{TopP}}$Top [$_{\text{ForceP}}$Force [$_{\text{TopP}}$Top [$_{\text{FocP}}$Foc [$_{\text{ModP}}$Mod [$_{\text{FinP}}$…]]]]

According to Rizzi (1997), among others, speaker-oriented adverbs occupy the position of specifier of Force. The particle *es-que*, which according to our proposal would

be the head of this phrase, can appear with this type of adverb (e.g. *Francamente/sinceramente es que*… 'Frankly/sincerely the thing is that…'). We should also notice that *es-que* cannot be in FocP (Focus Phrase), because there is only one such position, and (25a) shows that *es-que* may cooccur with *sí-que*. In addition, it should be obvious that *es-que* cannot be in TopP (Topic Phrase) or any other peripheral position, such as ModP (Modal Phrase), which are occupied by adverbs with sentential scope such as *realmente* 'really'. Finally, *es-que* cannot be in FinP (Finiteness Phrase), since *es-que* is not a subordinating particle in this example.

The question that arises, from a diachronic perspective, is how the particle developed from a reduced cleft. In Section 5 we examine this process in Spanish, distinguishing three different stages in this evolution. Before doing so, we will consider some examples where there might be structural ambiguity, as both a particle interpretation and an interpretation of the structure of the sentence as containing a null subject would appear to be possible.

4. The status of *es que* in justificatory constructions

The boundaries between different *es que* constructions are clear in prototypical examples, but in real usage we also find examples that allow more than one syntactic interpretation. We will now consider several examples presenting such ambiguity and we will provide reasons for concluding that in fact these seemingly ambiguous examples are better analyzed as containing a particle.

In (26) we present a constructed example of the justificatory use of *es que*:

(26) A: Juan habla francés muy bien.
Juan speaks French very well
'Juan speaks French very well.'
B: *Es que* su madre es francesa.
is that his mother is French
'(*The reason*) *is that* his mother is French.' /
'*Well*, her mother is French.'

In this use, there can be some ambiguity in the analysis of *es que* as a particle or as a *pro*+V+COMP sequence. Although there is no explicit antecedent in the previous discourse, in a possible analysis, we could have an empty subject interpretable as 'the reason' or 'what happens'. On the other hand, a particle analysis is possible and probably preferable, since this example would pass the tests of particlehood, including the pause-insertion test (27a),(b), the invariability test in cases of reported speech (27c) and the Colloquial Basque *eske* test (27d).

(27) a. **Es… que* su madre es francesa.
b. ✓*Es que…* su madre es francesa.
c. Me dijo que Juan hablaba francés muy bien y le contesté que ✓*es*/**era que* su madre era francesa.
'She/he told me that Juan spoke French very well and I replied that *it (the reason) is*/**was that* his mother was French.'
d. ✓*Eske* bere ama frantsesa da.
eske his mother French is

Given its nature as a fixed expression, it is preferable to consider *es que* as a complex unit that modifies the illocutionary force of the utterance that it introduces conveying a justification value (see Siemund 2018). This interpretation excludes the possibility of having a null subject.

The use of *es que* to justify a previous statement is also illustrated in the example in (28), taken from a novel by the modern Argentine writer Manuel Puig (1976). In this case, the speaker offers a justification for his own immediately preceding statement. The English translations that we provide in (28b) and in subsequent examples are taken from the English translation made by Thomas Colchie (Puig 1979).

(28) a. –¿Y qué era que soñabas?
– No me acuerdo para nada. *Es que* estoy intoxicado, pero ya se me pasará. (Puig, *Beso*: 126)
b. – What were you dreaming?
– I don't remember at all. It's that my system is still messed up, but it'll go away soon. (Puig/Colchie: 122)

In this example too, our tests for particlehood show that a particle analysis is preferable. Nevertheless, a *pro* analysis cannot be completely excluded as an alternative. Consider now the following constructed example:

(29) A: ¿Estás bien?
you.be well
'Are you all right?'
B: Sí, *es que* tengo un poco de frío.
yes is that I.have a little of cold
'Yes, *it's just that* I'm a little cold.'

In Example (29), unlike in (28), speaker B does not provide a justification for a previous statement. However, it is reasonable to infer that, in the context of the interaction, speaker B is assuming that speaker A is asking for the reason of speaker B's observable behavior. In one possible analysis, we may thus still postulate the presence of an empty subject pronoun *pro* in B's reply interpretable as 'the reason

for my behavior' or 'what happens to me'. We may have a similar interpretation in (30)–(31), from Puig's novel:

(30) a. – Te ruego, no me hagas reír.
– Pero así te olvidás del dolor, sonso.
– *Es que* tengo miedo de hacerme encima. (Puig: 82)
b. – Come on, don't make me laugh.
– But that way you'll forget about the pain, silly.
– It's just that I'm afraid of going in my pants. (Puig/Colchie 118)

(31) a. – Tendrías que haber almorzado algo.
– *Es que* no tenía nada de ganas.
b. – 'You should have had something for lunch.'
– 'It's just I didn't want anything at all.' (Puig/Colchie: 96)

In (30), the sentence with *es que* offers a justification for an earlier statement by the speaker, and thus it is sensible to assume that it could contain an empty subject interpretable as 'the reason why I don't want to laugh'. In (31) the second sentence can be paraphrased as 'I didn't have anything for lunch because I didn't feel like eating anything', where the proposition being justified can be extracted from the previous discourse. On the other hand, both in (30) and (31) we could have a transfer of *eske* in a Colloquial Basque version, and the most natural placement of a prosodic break is also after *es que*.

In its purely argumentative use, *es que* passes all the particlehood tests. However, it is also possible to analyze the examples as having a null subject *pro* that refers to the reason for either a statement or a specific behavior.

The justificatory force that *es que* provides does not need to apply to an explicit statement in the previous discourse. Thus, in the example in (32), which is based on a real-life occurrence, there is an implied question (*why are you looking under the bed?*) that is answered with *es que*. In this example *es que* justifies the act of looking under the bed.

(32) A: [looking under a bed]
[B approaches A but does not say anything]
A: *Es que* no encuentro mis zapatos negros.
is that not I.find my shoes black
'*Es-que* I cannot find my black shoes' = 'I just can't find my black shoes!'

In all these examples, where the sentence headed by *es que* has justificatory value, the syntactic analysis is ambiguous. In one analysis, we may have an empty subject *pro* interpretable as 'the reason (that explains some statement or behavior)'. In an

alternative analysis we have a particle with justificatory or exclamatory value. The application of the tests for particlehood leads us to prefer the particle analysis:

(33) a. *? *Es… que* no encuentro mis zapatos.
is that not I.find my shoes
b. ✓*Es-que…* no encuentro mis zapatos.
c. *?Es solo que no encuentro mis zapatos.
is only that not I.find my shoes
d. ✓ *Eske* ez ditut aurkitzen nire zapatak. (Colloquial Basque)
eske not AUX find my shoes

The analysis of *es que* as a particle is clearer in strictly counter-argumentative constructions used in replies introducing a reason against some previous statement. A couple of examples from the same novel by Puig as in the examples above are given in (34)–(35):

(34) a. – Seguí un poco más.
– *Es que* con el sueño se me olvida la película, ¿qué te parece si la seguimos mañana? (Puig, Beso: 14)
b. – Go on a little more.
– Just that I get sleepy and forget the film. What do you say, we go on with it tomorrow? (Puig/Colchi: 8)

(35) a. – […] nunca nunca me hablaste de tu mamá.
– Sí, cómo no.
– Por Dios, te lo juro, nunca nunca.
– *Es que* no tengo nada que contar. (Puig 83)
b. – […] you never never talk about your mother.
– Of course, I do, what do you mean?
– I swear to God, never, never.
– Well, maybe I have nothing to tell. (Puig/Colchie 119)

In this counter-argumentative usage as well, more than one analysis appears to be possible, but the interpretation of the *pro* subject becomes less straightforward than in argumentative contexts. In (34) the meaning that would have to be assigned to this element would be something like 'My reason for not continuing speaking', and in (35) 'My reason for not talking about my mother'. The tests of particlehood provided in Section 2 lead us to favor the particle analysis.

Finally, this analysis is even clearer in contexts where *es que* is used essentially as an expressive element, as in the constructed example in (36).

(36) a. A: Y qué me dices de Pedro?
and what to.me you.tell of Pedro
'What about Peter?'
b. B: ¡A ese tío *es que* no le aguanto!
to that guy is that not him I.stand
'That guy, *well*, I can't stand him!'
c. B′: *Es que… es que* no puedo con él.
is that is that not I.can with him
'*The thing is, the thing is* that I just can't deal with him.'

We thus conclude that, in its most grammaticalized usages, *es que* is a particle with a number of different discourse-related functions. In these functions, we no longer have a *pro*+V+COMP sequence, but rather a discourse particle that modifies the illocutionary force of the utterance and occupies a high position in the left periphery of the sentence.

5. The historical process of the formation of the particle *es que* in Spanish

The diachronic evolution of *es que* shows interesting similarities with the phenomenon of insubordination (Evans 2007).[10] Like in the process of insubordination, in the evolution of the particle *es-que* certain elements of a compound clause are deleted. In insubordinate clauses, the entire main clause is deleted, so that the original subordinate clause ends up functioning as an independent clause. Spanish examples would be: *¡Que te vayas*! 'Go away!' (literally: 'that you go-subjunctive'); *Si acabo de entrar* 'But I have just arrived' (literally: 'if I have just arrived'). In constructions with the particle *es-que*, instead, only the subject position of the main clause is eliminated, so that the copulative verb ends up being reanalyzed as a particle, together with the following complementizer.

Even though Evans (2007: 384–386) explicitly states that in cases where part of the main clause remains one cannot speak of insubordination, the process of particle formation that we are examining here shows some interesting similarities with the phenomenon of insubordination. In fact, in the development of the construction with the particle *es-que*, we find the same four steps that Evans (2007: 370–375) distinguishes in the process of insubordination. The main difference is the type of deleted element and the consequences that this has for the final results. In Table 1, we reproduce the four stages that Evans (2007: 370) distinguishes in the diachronic process of insubordination, making a direct comparison with the steps

10. We want to thank an anonymous reviewer for pointing out the parallelism between the development of *es que* as a particle and the phenomenon of insubordination.

Table 1. Historical trajectory of insubordinate clauses and clauses with the particle *es-que*

	Subordination	**Ellipsis**	**Conventionalized ellipsis**	**Reanalysis as main clause structure**
	(1)	**(2)**	**(3)**	**(4)**
Insubordinate clauses	Compound sentence (with a subordinate construction)	Ellipsis of the main clause	Restrictions of the interpretation of elided material	Conventionalized main clause use of a formally subordinate clause
Clauses with the particle *es-que*		Ellipsis of the subject of the main clause		Conventionalized main clause use of a formally compound sentence

in the development of the particle *es-que*. Stages 1 and 3 are similar in both cases, but the other two stages are not.

Insubordinate conditional clauses with exclamative force may in fact be used with a very similar pragmatic meaning as the particle *es-que*, either agreeing with or as a rebuttal contradicting a previous statement (Gras 2020; Gras & Elvira-García 2021), as in the following examples:

(37) A: El Valencia ha vuelto a ganar la liga.
the Valencia has returned to win the league
'Valencia has won the league again.'
B: a. ¡*Es que* son los mejores!
is that are the best
'They are the best!'
b. ¡*Si* son los mejores!
if are the best
'*Of course*, they are the best!'

(38) A: Vamos de excursión este sábado, ¿vale?
go.IMP of trip this Saturday ok
'Let's go on an outing this Saturday, ok?'
B: a. *Es que* han dicho que va a llover.
is that have said that goes to rain
'*Well*, they said it's going to rain.'
b. ¡*Si* han dicho que va a llover!
if have said that goes to rain
'*But* they've said it's going to rain!'

In the evolution of *es que* as a particle we find the four constructions and stages that we defined in Section 2: (1) pseudo-cleft sentences and copulative equivalents; (2) reduced clefts; (3) inferential reduced clefts, and (4) sentences with the particle *es-que*. The common thread that explains this diachronic evolution is that the postcopular

constituent is focalized, often contrastively, through all stages. This focalized nature explains the fact that, by contrast, the syntactic subject of the copular clause may be deleted when it is anaphorically recoverable (stage 2). It also accounts for the possibility of using elliptical sentences with inferential meaning to introduce contrastive foci that negate previous interpretations or explanations, expressing more appropriate ones (stage 3). Finally, it also explains how *es que* may become a particle that introduces justificatory or reactive utterances that are also focalized (stage 4).

On the other hand, a common element in all constructions without an overt subject or without any subject at all is their linking with the preceding discourse or the context of the discourse. This linkage is what enables the recovery of an empty subject in reduced clefts, what allows for the interpretation of the subject through inferences in inferential reduced clefts and what explains the reactive illocutionary value of the particle *es-que*.

Finally, we may note that the historical path in Table 1 accounts for the cross-linguistic differences that we mentioned in the introduction, including the fact that whereas in the Ibero-Romance languages the process has arrived at the most grammaticalized stage, in other languages such as English this evolution has not been completed yet.

5.1 First Change: From a pseudo-cleft to a reduced cleft

If the evolution outlined in Table 1 is correct, we would expect to find the explanatory use of *es que* to develop historically once pseudo-clefts became common constructions in Spanish and, especially, when reduced clefts reached a certain frequency in their use. That is what the historical evidence shows. Pseudo-clefts with *es que* are very unusual in Medieval Spanish, but their frequency increases noticeably in the 16th century (Romera 2009: 155–157, Pérez Saldanya & Hualde, in press). See Figure 1, based on Pérez-Saldanya & Hualde (in press), which may be consulted for the data in tabular form and details about the corpus and the analysis:

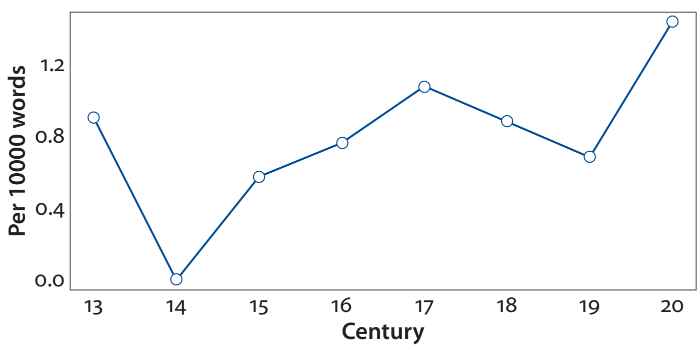


Figure 1. Frequency of pseudo-cleft constructions per 10000 words in the diachronic Spanish corpus analyzed in Pérez-Saldanya & Hualde (in press).

Reduced clefs show a somewhat similar, although not identical, diachronic pattern. These constructions are almost non-existent in medieval texts, their frequency reaching its maximum in the 16th and 17th centuries. See Figure 2, based on Pérez Saldanya & Hualde (2021). Notice that the x-axis starts in the 15th century, since we have not found earlier examples.

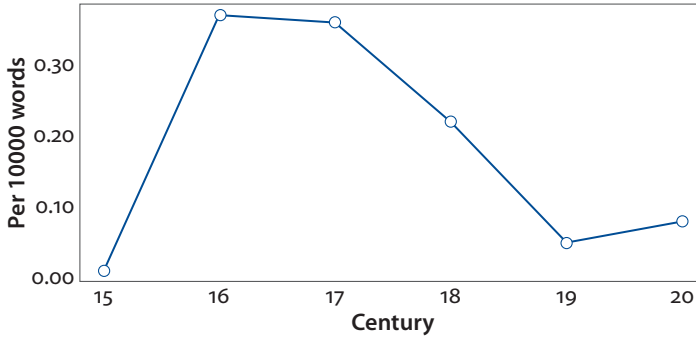


Figure 2. Frequency of "reduced cleft" constructions per 10000 words in the diachronic Spanish corpus analyzed in Pérez-Saldanya & Hualde (2021)

In (39) we provide an example of a reduced cleft from a text written towards the end of the 15th century or beginning of the 16th century. Notice that, in this example, the empty subject of the copula is co-referential with the noun phrase *una gran cosa* 'a big thing' and that it could be made explicit with a free relative clause such as *lo que he descubierto* 'what I have found out':

(39) Una gran cosa he descubyerto, y *es que* el señor duque d'Ater y capitán general de las montañas estava aquy encubyerto detrás del capitán Gordyano.
'I have found out a big thing; and *it is that* the Duke of Ater, Captain General of the mountains, was here, hidden behind Captain Gordyano.'
(Anonymous, *La corónica de Adramón*, ca. 1492, *Corde*)

5.2 Second change: From a reduced cleft to an inferential reduced cleft

Once the reduced cleft construction reaches a certain frequency, we start finding also the first examples that can be analyzed as inferential reduced clefts with an argumentative value. As it happens in many other changes related to the grammaticalization and conventionalization of inferences, in the case at hand as well the new construction arises from bridging contexts (Heine 2002: 84), where the interlocutor may make one of two different interpretations: either an etymological or "literal" interpretation, where the subject of the copulative verb is anaphorically recovered, or an "inferential" interpretation, in which the content of the empty subject is established through a discursive inference of causality, equivalent to 'the

reason for this' or 'the cause of this'. A possible example of this bridging context (which we take from Pérez Saldanya & Hualde 2021) is found in the dialogue in (40) from a 16th century novel:

(40) PANDULFO. ¿Qué fue, que tanta priessa hay?
what it.was that so.much rush there.is
'What has happened, so that you are in such a hurry?'
SIGERIL. *Es que* te llama nuestro amo.
is that to.you he.calls our master
'*Well*, our master is calling you.'
(Feliciano de Silva, *Segunda Celestina*, 1534, *Corde*)

In a possible reading of (40), Sigeril's construction with *es que* offers a reply to Pandulfo's question *¿qué fue?* 'what was it?, what has happened?'. In such a reading, the construction with *es que* in the reply by Sigeril is interpreted as a reduced cleft and the copula *es* has a null subject *pro* whose content is anaphorically recoverable from the preceding question. The reply is thus interpretable as *(Lo que fue) es que te llama nuestro amo* '(what has happened) is that our master is calling you'.

In a second reading, instead, the construction is interpreted as an inferential reduced cleft, in which the null subject is obtained through inference and can be paraphrased by 'the reason'; that is, 'the reason I am in such a hurry is that our master is calling you'.

Starting from examples that allow these two readings, the next stage is the conventionalization of discourse inferences, which gives rise to inferential reduced clefts. This is what happens in the following example taken from a satirical sonnet in *Don Quijote* (1605), containing a dialogue between El Cid's horse, *B(abieca)*, and Don Quijote's horse, *R(ocinante)*:

(41) B. ¿Es necedad amar?
is foolish love
R. No es gran prudencia.
not is big prudence
'B: Is it foolish to love? R: It's not too smart.'
B. Metafísico estáis.
metaphysical you.are
R. *Es que* no como.
is that not I.eat
'B: You're a philosopher. R: It is that I just don't eat.'
(Miguel de Cervantes, *El Quijote*, 1605, *CORDE*)

In this example, there is no possible antecedent that could be interpreted as co-referential with the subject of the copulative verb, but from the discourse context we may infer something like 'the reason'. That is, the reply *es que no como* is interpreted as 'The reason why I am a philosopher is that I don't eat'.[11]

5.3 Third change: From compound sentences to main clauses with the particle *es-que*

In the preceding section we have considered the origin of inferential reduced clefts with a null subject interpretable as 'the reason', 'what happens' or 'the thing'. Such constructions are also found in English (Declerck 1992), a language where *it's that* has not, however, become a discourse particle. It appears that an essential step in the evolution from *pro*+V+COMP to a particle has been the development of counter-argumentative values and other values that make it difficult or impossible to recover a subject for the copula.

The counter-argumentative construction arises in contexts of reply, where both an argumentative and a counter-argumentative interpretation are possible, as in the following example:

(42) A: Esta [espada] fué del tan celebrado don Fernando Cortés.
'A: This sword belonged to the famous Hernán Cortés.'
B: ¿Que ésta es? – dixo Andrenio–. ¡Cómo me alegro de verla! ¿Y es de azero?
'B: This one? – said Andrenio – I'm so happy to see it! And is it made of steel?'
A: ¿Pues de qué avía de ser?
'A: What else could it be made of?'
B: *Es que* yo avía oído dezir que era de caña.
'B: Well, [= the reason I asked is that] I heard it was made of cane.'
(Baltasar Gracián, *El Criticón*, 1651–1657, *CORDE*)

Notice that the construction with *es que* in speaker B's second turn can be interpreted as argumentative, justifying why this speaker asked whether the sword is made out of steel ('The reason why I asked if it is made of steel is that…'). But it can also be interpreted as counter-argumentative, introducing a justification against the criticism implicit in speaker A's rhetorical question ('What else could it be made of?').

---

11. As a reviewer reminds us, Rico (1997) has suggested that in this sonnet Cervantes is playing with the potential ambiguity of the word *metafísico*. Although in a first reading this word can be understood as equivalent to 'philosopher', there is also a second interpretation of this word as meaning 'without physical consistency, i.e., very skinny.'

In contexts of this type, which are argumentatively more complex, it is no longer clear what the potential subject of *es* could be or what the syntactic status of *es que* is. The identification of the potential subject becomes even more difficult when the value of the construction is purely counter-argumentative and becomes increasingly conventionalized through grammaticalization. The first purely counter-argumentative example documented in Pérez-Saldanya & Hualde (2021) is from the very beginning of the 19th century:

(43) D. DIEGO: No hay remedio… Usted no ha de dormir aquí.
'There is no remedy…You cannot sleep here.'
CALAMOCHA: *Es que* los caballos no están ahora para correr…
'*The thing is*, the horses cannot run now…'
(Leandro Fernández de Moratín, *El sí de las niñas*, 1801, *CORDE*)

In this example, the second speaker, Calamocha, introduces an argument against obeying Don Diego's command, justifying his difficulty to do it. It is at this point in the evolution that we can say with certainty that *es que* has already been reanalyzed as a discursive particle and that the construction has been conventionalized with a justificatory value, regardless of its argumentative orientation.

The fact that *es-que* is a particle is even more obvious when it starts being used with an expressive value, as an emphatic element and with other discourse functions that are typical of colloquial speech. The following example, extracted from a corpus of colloquial conversational speech, may serve as an illustration:

(44) [context: discussing a possible trip to Madrid]
es que me encantaría↑ me encantaría ir/ y nunca he tenido
is that to.me I.would.love to.me I.would.love go and never I.have had
oportunidad↑/ así como a Barcelona he ido muchísimas veces↑// a
opportunity as like to Barcelona I.have gone many times to
Madrid↑ no he ido en mi vida […] pero *es que* no↓ *es que* no↓/
Madrid not I.have gone in my life but is that no is that no.
¡joder! ¡que no!
f*ck. that no
'*Gee*, I would love it, I would love to go. And I have never had a chance. Just like I have been to Barcelona many times I haven't been to Madrid in my life. […] but, *well,* no, f*ck, no' (*Valesco*: 28–100–111)

In this usage, which starts being documented at the beginning of the 20th century, the justificatory character of *es que* is fuzzier than in other examples discussed above. The particle is used mainly at the beginning of sentences that express emotional

states of surprise, joy, complaint, rejection, etc. (Marín & Cuenca 2012: 79–82, Pérez Saldanya & Hualde 2021). Its function may also be simply that of an illative resource (Dufter 2008: 1775) or a strategy to seize or keep a speech turn (Bravo Cladera 2005: 176–179).

The diachronic evolution of the constructions with *es que* thus shows an increasing degree of grammaticalization, which is associated with progressively increasing (inter)subjectification (Traugott & Dasher 2002). In a first stage in this evolution, the reduced cleft acquires a certain frequency of usage. In a second stage, a reduced cleft construction that is used to specify an event is reinterpreted as an inferential reduced cleft, and more specifically as a construction that the speaker uses to justify what has just been said (Romera 2009: 159–161). In a third stage, the construction is reanalyzed as a main clause used to offer a reason against what the interlocutor just expressed or might be thinking. Subsequently, we find more complex discourse functions that may be described as emphatic or expressive.

6. Catalan and Portuguese

A similar diachronic process to the one we have outlined for Spanish in Section 5 has arguably taken place in the other Ibero-Romance languages as well. In those Ibero-Romance languages in contact with Spanish, most of whose speakers are in fact bilingual, such as Asturian, Galician, Aragonese and Catalan, it would seem sensible to think of a transfer from Spanish. Nevertheless, the fact is that in Catalan the justificatory use of *és que* is already well-attested in the 19th century, before bilingualism with Spanish had become widespread, as is shown in the example in (45), from a play:

(45) – Pepa: Calla, dona, y reposa.
be.quiet woman and relax
'Be quiet, woman, and relax.'
– Nuri: *Es que* no puch callar jo.
is that not can be.quiet I
'*It's just that* I cannot be quiet.' (Guimerà: 118, 1910)

Despite the absence of any widespread bilingualism with Spanish, the same justificatory particle also exists in Portuguese, as shown in (46), which repeats the examples given above in (34), by adding their version in the published Brazilian Portuguese translation of the novel (Puig 2003):

(46) a. Spanish original:
– Seguí un poco más.
'Go on a little more.'
– *Es que* con el sueño se me olvida la película, ¿qué te parece si la seguimos mañana? (Puig, Beso: 14)
'*Just that* I get sleepy and forget the film. What do you say we go on with it tomorrow? (Puig/Colchi: 8)

b. Portuguese translation:
– Continua um pouco mais.
– *É que* con o sonho me esqueço do filme, que acha você se continuarmos amanhã? (Puig/Rodríguez: 12)

We should point out, nevertheless, that some instances of *es que* in Puig's novel are left out in the Portuguese translation.

It is not implausible that we have parallel historical developments in different Ibero-Romance languages, instead of direct calquing. Other Romance languages such as Italian do not show this phenomenon, or not to the same extent, at least.

In Portuguese, as in Spanish, the first stage in the historical process would involve [pro *é que*] clauses in coordination and other contexts where the referent of the subject *pro* is easily recoverable, as in the 16th example in (47):

(47) Porem, eu quero saber de ti uma coisa e *é que*, quando nosso casamento vier a efeito, que sera logo, me digas… (Gonçalo Fernandes Trancoso, *Proveito*, c 1517–1596, *Corpus do Português: Gênero/Histórico*)
'However, I want to know one thing about you, and *it is that* when our marriage comes to have effect, which will be soon, you tell me…'

In certain contexts, these reduced clefts can be interpreted as inferentials with an argumentative value ('and I say this because', 'and this is so because'), as in (48):

(48) Para inteligência desta gravíssima e perigosa matéria, havemos de supor o que se não cuida; e *é que*, não só são talentos os dotes da natureza, os bens da fortuna e os dons particulares da graça, senão também os contrários, ou privações de tudo isto. (Antonio Vieira, *Sermões*, 1679–1695, *CHPTB*)
'For a proper understanding of this very grave and dangerous matter, we need to assume what one does not think; and *it is that*, not only are talents the gifts of nature, the favors of fortune and the particular presents of grace, but also their opposite, or lack of all of this.'

The argumentative value can be taken to be fully established in the following example:

(49) Meu paizinho, não se admire de ver este labirinto. *É que* nós vamos à ópera.
(José Daniel Rodrigues da Costa, *Entremezes de Cordel*, end of 18th c or beginning of 19th c, *CHPTB*)
'My good Father, do not be surprised to see this labyrinth. *It's just that* we are going to the opera.'

Also like in Spanish, the use of *é que* with a counter-argumentative value and functioning clearly as a particle is more recent:

(50) a. – Ri te, ri te, cá estou eu para chorar. *É que* tu não sabes o que é meu pae, em se lhe mettendo uma coisa de estas em a cabeça.
(Almeida Garrett, *Theatro*, 1845, *CHPTB*)
'Laugh, laugh, because I am about to cry. *The thing is that* you don't know what my father is like, when he gets one of these things in his head.'

b. – Como? Tão cedo, e já acabada a tarefa? – *É que* eu – observou Luzia, enleada – desejava sair hoje mais cedo… (Domingos Olímpio, *Luzia-Homem*, 1878, *Corpus do Português: Gênero/Histórico*)
'– What? So early and you have already finished the work?
– *it's just that*, I – said Luzia, confused – wanted to leave earlier today…'

Finally, like in Spanish, in Portuguese as well we find examples showing that the grammaticalization of *é que* has proceeded further ahead, becoming a particle with an essential emphatic function:

(51) a. Oh Ana Maria! O Windows 10 corre MUITO mais fluido e rápido que o El Capitain em o meu MacBook de 2010. *É que* nem tem comparação! Só não mudo completamente porque continuo a gostar de o OSx.
(*EUA querem que funcionários usem o Windows 10… em casa*, Portugal, 16–04–17, *Corpus do Português: Now*)
'Hi Ana Maria! Windows 10 runs better and faster than El Capitan on my 2010 McBook. *Gee*, there is no comparison! I am not switching only because I still like OSx.'

b. MG: em essa idade, eu não era tão conhecido. *É que* nem pensar.
(Miguel Guilherme e César Mourão falam de tudo Portugal, 15–08–08, *Corpus do Português: Now*)
'At that age, I was not so well known. *Gee*, not at all!'

For Catalan, we may assume a similar evolution, although it is hard to document all stages because of the sociolinguistic situation of this language during the Modern period (16th-18th centuries), when it was largely replaced by Spanish as a written language. In any event, in the 16th and 17th centuries we find examples of *és que* in reduced clefts, where the referent of the *pro* subject is identified in the preceding clause:

(52) Una cosa breument per conclusió d'aquest capítol vull advertir y *és que* per l'entrada y domini universal dels godos en tota Espanya se múdan totas las cosas d'ella. (Pere Gil, *Història moral de Catalunya*, 1600, *CIGCMod*)
'I want to point out one thing as a conclusion of this chapter, and *it is* that with the arrival of the Goths and their universal control of all of Spain, all things changed there.'

We have not found any clear examples of inferential reduced clefts with argumentative value in texts from the Modern period, but we may assume that they existed at that time, since in the 19th century we trace examples in which *és que* functions as a particle with both an argumentative (53a),(b) and a counter-argumentative value (53c):

(53) a. Ells li diuhen: Dona, ¿perqué ploras? Ella los digué: *Es que* sen han portat d'aquí á mon Senyor, y no sé ahont l'han posat.
(Josep Melcior Prat i Solà, Busanyà Ramon, *Lo sant evangeli de Jesu-Christ segons sant Joan*, 1832, *CTILC*)
'The say to her: Woman, why are you crying? And she says to them: *(The reason) is that* they have taken my lord away from here and I don't know where they have taken him.'

b. ¡Hi haurá una bella mes rara! *es que* no sen trobará un altre sobra la terra.
(Josep Robrenyo, *La union ó la tia Sacallona, en las fiestas de Barcelona*, 1833, *CTILC*)
'There will be a most rare fair lady!. Ø A comparable one will not be found on earth.'

c. – Rey, aquesta no es la resposta que espero.
– *Es que* no sè qué dirvos… (Antoni de Bofarull, *La orfaneta de Menargues ó Catalunya agonisant*, 1862, *CTILC*)
'– King, this is not the answer that I am hoping for.
– *Well*, I don't know what to tell you…'

Finally, like Spanish and Portuguese, Catalan has also developed purely emphatic uses of the particle *és que*:

(54) és una cosa que no… si no ho sents no saps lo que és
is a thing that no if not it you.feel not you.know PRON that is
\vull dir *és que* jo no te puc explicar lo angoixós que és
I.want say is that I not to.you I.can explain PRON scary that is
ofegar-te \ *és que* és *horrorós*\
drown-2sg.REFLX is that is horrible
'It is a thing that if you don't experience it yourself you don't know what it is. It is, I mean, well, I cannot explain it to you; how scary it is to drown. *Jesus*, it is horrible!' (*Corpus oral de conversa col·loquial, ap.* Marín & Cuenca 2012: 79–80)

7. Conclusions

We have distinguished four types of diachronically-related constructions with *es que*: *a*) in a first stage in its diachronic evolution, *es que* appears in pseudo-clefts and similar identificative copulative sentences. *b*) Subsequently, we find *es que* in reduced clefts, with a null subject coindexed with a noun phrase in the previous discourse. *c*) In a third stage, we find inferential reduced clefts, where there is no explicit referent, but the null subject of the copula can be interpreted contextually as 'the problem', 'what happens', etc. *d*) Finally, at a later stage, we find constructions where *es que* is a particle, not a V+COMP sequence, and there is no null subject *pro*.

Although these four types of *es que* constructions are all found in contemporary Spanish as well as in Catalan and Portuguese, they are diachronically ordered. The evolution is thus one where an empty subject *pro* first loses its coindexing and at a second stage is deleted, with restructuring of the surface sequence. As noted, this evolution has some similarities with the process of insubordination. In particular, utterances headed by the particle *es-que* have a similar pragmatic usage as insubordinate conditionals in present-day Spanish.

The historical data that we have reviewed here has led us to conclude that there has been a reanalysis process resulting in the development a discourse particle. An important aspect of this process has been the gradual extension of the use of *es que* from contexts where it has an argumentative function to counter-argumentative contexts, where the referent of a *pro* subject is more difficult to identify.

It is at the last stage of the grammaticalization that *es que* is reanalyzed by speakers as a particle, which in turn has facilitated its borrowing in colloquial Basque, in spite of the two languages having very different syntactic structures.

The diachronic process appears to have been identical in Spanish, Catalan and Portuguese. In all three Ibero-Romance languages, the use of *es/és/é que* with argumentative value is attested earlier than its counter-argumentative use, and its use as an emphatic marker is the last one to be attested in the corpus. It is hard to know to what extent one language influenced over the others through convergent but separate evolutions.

The historical trajectory outlined in the paper that leads to the formation of these particles can be fruitfully related to Evans' notion of insubordination. Both in Evans' insubordination phenomena and in the diachronic development studied here it is argued that it arises as a result of a diachronic deletion. The difference, however, lies in that in our proposal the final outcome is not exactly the syntactic independence of a formerly subordinate clause. In this respect, diachronic developments like the one leading to the formation of the discourse particle *es-que* may serve to also place the notion of insubordination within a larger context in historical syntax.

Acknowledgements

The first author wants to thank Ana Maria Martins for her help and information about Portuguese. We are grateful to Gemma Rigau and to two anonymous reviewers for many comments and suggestions that have undoubtedly improved the quality of this paper.

Funding

This research has been made possible in part thanks to a grant from the Spanish Ministry of Education to the first author for the project "Variación y cambio lingüístico en catalán. Una aproximación diacrónica según la Lingüística de Corpus" (MICINUN, Ref. PGC2018-099399-B-100371).

References

Bravo Cladera, Nadezhda. 2005. Y, pero, así que *y* es que. *Un estudio de su uso en las interacciones del español de jóvenes bilingües y unilingües*. Uppsala: Acta Universitatis Upsaliensis. Studia Romanica Upsaliensia 70.

Brinton, Laurel. 2014. The Extremes of Insubordination: Exclamatory *as if!*. *Journal of English Linguistics* 42(1), 93–113. https://doi.org/10.1177/0075424214521425

Company, Concepción. 2018. Four directionalities for grammaticalization: Evidence for new diachronic paths. *Journal of Historical Linguistics* 8 (3): 356–387. https://doi.org/10.1075/jhl.17032.com

Cuenca, Maria Josep. 2013. The fuzzy boundaries between discourse marking and modal marking. In *Discourse Markers and Modal Particles*, Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds), 191–216. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.234.08cue

Declerck, Renaat. 1988. *Studies in Copular Sentences. Cleft and Pseudo-Clefts*. Leuven: Leuven University Press / Foris Publication. https://doi.org/10.1515/9783110869330

Declerck, Renaat. 1992. The inferential *it is that* construction and its congeners. *Lingua* 87: 203–230. https://doi.org/10.1016/0024-3841(92)90008-7

Delahunty, Gerald P. 1990. Inferentials: The story of a forgotten evidential. *Kansas Working Papers in Linguistics* 15: 1–28. https://doi.org/10.17161/KWPL.1808.430

Delahunty, Gerald P. 2001. Discourse functions of inferential sentences. *Linguistics* 39: 517–545. https://doi.org/10.1515/ling.2001.022

Dufter, Andreas. 2008. Evolución pragmática de las oraciones hendidas en español: el papel de los usos no focalizadores. In *Actas del VII Congreso Internacional de Historia de la Lengua Española* 2, Concepción Company & José G. Moreno de Alba (eds), 1763–1780. Madrid: Arco Libros.

España, Margarita. 1996. Aspectos semántico-pragmáticos de la construcción "es que" en español. *Dicenda* 14: 129–147.

Evans, Nicholas. 2007. Insubordination and its uses. In *Finiteness: Theoretical and empirical foundations*, Irina Nikolaeva (ed), 366–431. Oxford: Oxford University Press.

Fernández Leborans, María Jesús. 1992. La oración del tipo *es que*. *Verba* 19: 223–239.

Fernández Leborans, María Jesús. 1999. La predicación: las oraciones copulativas. In *Gramática Descriptiva de la Lengua Española*, 2, Ignacio Bosque & Violeta Demonte (eds), 2357–2460. Madrid: Espasa.

Fuentes, Catalina. 1997. Los conectores en la lengua oral: "es que" como inductor de enunciado. *Verba* 24: 237–263.

Gras, Pedro. 2020. Construcciones gramaticales y ubicación discursiva: sobre la construcción independiente <si + indicativo>. In *Pragmática del español hablado: nuevos horizontes*, Adrián Cabedo & Antonio Hidalgo (eds), 211–226. València: Servei de Publicacions de la Universitat de València.

Gras, Pedro & Elvira-García, Wendy. 2021. The role of intonation in Construction Grammar: On prosodic constructions. *Journal of Pragmatics* 180: 232–247. https://doi.org/10.1016/j.pragma.2021.05.010

Harris, Alice & Campbell, Lyle. 1997. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Heine, Bernd. 2002. On the role of context in grammaticalization. In *New Reflections on Grammaticalization*, Ilse Wischer & Gabriele Diewald (eds), 83–10. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.08hei

Ibarra, Orreaga. 2008. Sobre estrategias discursivas del lenguaje de los jóvenes vascohablantes: aspectos pragmáticos y discursivos (conectores, marcadores). In *Gramatika jaietan: Patxi Goenagaren omenez*, Xabier Artiagoitia & Joseba Andoni Lakarra (eds), 395–412. Vitoria-Gasteiz: Universidad del País Vasco / Euskal Herriko Unibertsitatea.

Ibarra, Orreaga. 2020. Gipuzkoako euskaldun gazteen elkarrizketak eta whatsappak. In *Eibartik Zuberoara euskalkietan barrena: Koldo Zuazori gorazarre*, Iñaki Camino, Xabier Artiagoitia, Irantzu Epelde & Koldo Ulibarri (eds), 319–334. Bilbao: Universidad del País Vasco / Euskal Herriko Unibertsitatea.

Marín, Maria Josep & Cuenca, Maria Josep. 2012. De l'atribució a la modalitat: construccions amb *és que* en català oral. *Caplletra* 52: 65–94.

Mikkelsen, Line. 2019. Copular clauses. In *Semantics – Sentence and Information Structure*, Paul Portner, Claudia Maienborn & Klaus von Heusinger (eds), 250–280. Berlin/Boston: De Gruyter. https://doi.org/10.1515/9783110589863-007

Pérez Saldanya, Manuel & Hualde, José Ignacio. (2021). De las copulativas identificativas a las justificativas con *es que*. *Revista de Filología Española*. 101(2): 421-448. https://doi.org/10.3989/rfe.2021.014

Pérez Saldanya, Manuel & Hualde, José Ignacio. (In press). Estrategias de focalización II. Relativas libres o hendidas. In *Sintaxis histórica de la lengua española. Cuarta parte: Estructura argumental, estructura informativa y discurso. Tradiciones y soportes textuales*, Concepción Company (dir). México: Universidad Autónoma de México / Fondo de Cultura Económica.

Porroche, Margarita. 1998. Sobre algunos usos de *que*, *si* y *es que* como marcadores discursivos. In *Los marcadores del discurso. Teoría y análisis*, María A. Martín Zorraquino & Estrella Montolío (eds), 229–242. Madrid: Arco Libros.

Reig, Asela. 2011. The pragmatic meaning of the Spanish construction 'lo que pasa es que'. *Journal of Pragmatics* 43(5): 1435–1450. https://doi.org/10.1016/j.pragma.2010.10.015

Rico, Francisco. 1997. «Metafísico estáis» (y el sentido de los clásicos). *Boletín de la Real Academia Española* 77: 141–164.

Rizzi, Luiggi. 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar*, Liliane Haegeman (ed), 281–337. Dordrecht: Kluwer.

Rizzi, Luiggi & Bocci, Giuliano. 2017. Left Periphery of the Clause. In *The Wiley Blackwell Companion to Syntax*, Martin Everaert & Henk C. van Riemsdijk (eds), 1–30. New York: John Wiley & Sons.

Romera, Magdalena. 2009. The Multiple Origin of *es que* in Modern Spanish: Diachronic evidence. In *Current Trends in Diachronic Semantics and Pragmatics*, Maj-Britt Mosegaard Hansen & Jacqueline Visconti (eds), 151–168. Bingley: Emerald Group.

Roulet, Eddy. 2006. The description of text relation markers in the Geneva model of discourse organization. In *Approaches to Discourse Particles*, Kerstin Fischer (ed), 115–131. Amsterdam: Elsevier.

Siemund, Peter. 2018. *Speech Acts and Clause Types: English in a Cross-linguistic Context*, Oxford: Oxford University Press.

Traugott, Elisabeth C. & Dasher, Richard B. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Waltereit, Richard. 2018. Inferencing, reanalysis, and the history of the French *est-ce que* question. *Open Linguistics* 4: 56–69. https://doi.org/10.1515/opli-2018-0004

Corpus

CHPTB: Corpus Histórico do Português Tycho Brahe. <http://www.tycho.iel.unicamp.br/corpus/>.

CIGCMod: Josep Martines and Vicent Martines (ed), *Corpus Informatitzat de la Gramàtica del Català Modern*, ISIC-IVITRA, Universitat d'Alacant.

CORDE: Real Academia Española, *Corpus Diacrónico del Español* <www.rae.es>.

Corpus do português: Davis, Mark. 2006. *Corpus do português: Gênero/Histórico* <https://www.corpusdoportugues.org/hist-gen/>.

CREA: Real Academia Española, *Corpus de Referencia del Español Actual* <www.rae.es>.

CTILC: Institut d'Estudis Catalans, *Corpus Textual Informatitzat de la llengua catalana* <https://ctilc.iec.cat/>.

Guimerà, Angel. 1914. *Terra Baixa*, 5th ed. Barcelona: La Renaixensa.

Puig, Manuel. 1976. *El beso de la mujer araña*. Barcelona: Seix Barral.

Puig, Manuel. 1979. *Kiss of the Spider Woman*, translated from Spanish by Thomas Colchie. New York: Alfred A. Knopf.

Puig, Manuel. 2003. *O beijo da mulher-aranha*, translated by Gloria Rodríguez, 16th ed. Rio de Janeiro: José Olympio.

Valesco: Adrián Cabedo & Salvador Pons (eds): *Corpus Val.Es.Co 2.0*, <www.valesco.es>.

PART II

Syntactic analyses of discourse particles

CHAPTER 4

Agreeing complementizers may just be moody

Ömer Demirok and Balkız Öztürk
Boğaziçi University

This study investigates two discourse-related particles, *ya* and *şo*, in Laz, an endangered South Caucasian language. We argue that both *ya* and *şo* are *indexical shift* complementizers which can occur without an overt embedding verb, suggesting root complementizer behavior. However, when they appear embedded, the mood specification of the embedding verb determines which of the two will surface, suggestive of complementizer agreement in mood features. We show that, while *ya* and *şo* need to be semantically distinct in their root occurrences, there are compositionality challenges against the null hypothesis that *ya* and *şo* keep their meanings when embedded. As an alternative to a formal agreement account, we propose to semantically relate the embedded and root occurrences of these complementizers.

Keywords: root complementizers, indexical shift, imperative mood, complementizer agreement

1. Introduction

South Caucasian languages have discourse-related particles that have been descriptively labelled as *quotative* or *instructional* (Boeder, 2002).[1] In this paper, we focus on two such particles, namely *ya* and *şo,* in Laz, an endangered South Caucasian language primarily spoken in Turkey. Judging from the examples in (1), the particle *ya* may be labelled as quotative, and the particle *şo* would fall under what Boeder (2002) calls instructional. Nevertheless, in order to avoid prejudging an analysis, we will not be using these terms and directly refer to them and gloss them with their forms throughout our discussion.

(1) a. [oxori-s vore] ya.
house-LOC be.1SG YA
'She/He say/said: "I am home."'

1. Lacroix (2009) glosses both particles as DR for *discours rapporté*.

https://doi.org/10.1075/la.276.04dem

b. [oxori-s vore] şo!
house-LOC be.1SG ŞO
'Say: "I am home."!'

Given the translations provided for the sentences in (1), these particles may appear to be non-transparently inflected forms of a verb of saying in Laz. However, their distribution in the language challenges this initial hypothesis. Both particles may in fact be followed by an inflected verb of saying as shown in (2):

(2) a. Oxori-s vore ya t'k'-u.
house-LOC be.1SG YA say-PST.3SG
'She/he said: "I am home."'.
b. Oxori-s vore şo t'k'v-i!
house-LOC be.1SG ŞO say-IMP.2SG
'Say: "I am home."!'

The fact that *ya* and *şo* are able to occur with or without an overt verb raises the question whether verb-less strings as in (1) are derived by eliding the verb or not. We report important structural asymmetries between examples like those given in (1) and (2) and argue that they are *not* derived through ellipsis. If examples of the type given in (2) do not involve ellipsis, it becomes obvious that *ya* and *şo* have distinct meanings in root contexts. There will be two important questions that follow from this finding.

The first question concerns the syntactic category of *ya* and *şo.* Given that these particles appear to substitute for a main verb in the examples in (1), it may seem that they have an eventive core in their meanings (in particular a speech eventuality) and, perhaps, signal the presence of a VP at least. We provide empirical arguments against this idea and propose that *ya* and *şo* are *quotative complementizers* in both embedded and root contexts.

The second question concerns the relationship between the root and embedded occurrences of these quotative complementizers. The null hypothesis is that the meanings posited for *ya* and *şo* in root contexts directly extend to their embedded occurrences in examples such as (2). We show that implementing the null hypothesis leads to compositionality issues in embedded contexts, at best predicting incorrect truth conditions. We explore two distinct proposals to account for the morphological distinctness between *ya* and *şo*. If one submits to the idea that in the embedded occurrences of *ya* and *şo* there is no meaning difference between the two, their morphological distinctness in the embedded context could be analyzed as a result of agreement in a formal feature (in particular one that relates to *mood,* as we will discuss). Nevertheless, we argue for an alternative account that semantically relates the embedded and the root occurrences of these complementizers, hence preserving a meaning difference between *ya* and *şo* in embedded contexts.

The data that we will report from Laz and the questions they raise are in some ways paralleled in the investigations of various languages and constructions in the literature. As Güldemann (2008) reports, it is cross-linguistically common for verbs of saying to be omitted in the presence of quotative particles.[2] Hence, the main contrast in Laz between the examples in (1) and (2) above is not a rare phenomenon. However, as we will argue, a synchronic omission/ellipsis account is not tenable for Laz. Therefore, it turns out to be an important question how the root and the embedded occurrences of these formally distinct and meaningful quotative complementizers should be related in the compositional interpretation. Furthermore, as will be shown, quotative complementizers in Laz are special and morphologically distinct from the run-of-the-mill complementizer *na* in the language. They look like they introduce quotes in which *indexical shifting* is obligatory in their complements. However, they in fact introduce finite clauses that are *not* opaque to grammatical processes such as extraction.

Another question that has received attention in the literature is whether quotative particles have a verbal nature, providing a speech event visible to syntax. For example, Etxepare (2008) shows that *que* in Iberian Spanish has a use as a main clause complementizer, as illustrated in (3a), and argues that a speech eventuality is nevertheless visible to syntax/semantic interpretation, as can be detected by various tests such as modification by a quantificational adverb *siempre* 'always' quantifying over the speech eventualities, as shown in (3b):

(3) a. Oye, *que* el Barça ha ganado la Champions
listen that the Barça has won the Champions.League
'Listen, someone said that Barcelona has won the Champions League.'

b. Tú siempre *que* cuándo viene (Etxepare 2008: 36–41)
You always that when he-is-coming
'You are always saying "when is he coming?"'

We will argue that the root use of the quotative complementizers in Laz is different from Iberian Spanish *que* in that a speech event is not visible to syntax/semantic interpretation.

Finally, the relevant literature presents quite rich descriptions of the grammaticalization paths that lead to quotative complementizers (Lord 1993; Klamer 2000; Güldemann 2008, a.o.). While we do not have access to historical data, we briefly discuss possible historical origins of the quotative complementizers in Laz, highlighting the fact that speech verbs do not seem to be the source in Laz.

2. For similar empirical observations, see also Frajzyngier (1985) and Lord (1993).

This paper is organized as follows: in Section 2, we document the common functional and distributional properties of the particles *ya* and *şo* and argue that they are complementizers which can additionally function as root complementizers. In Section 3, we turn to their occurrences in embedded contexts and show that the hypothesis that they have identical meanings as their root occurrences fails. To accommodate their embedded occurrences, we flesh out an agreement account as well as an alternative account which preserves a meaning difference between the two. Section 4 concludes the paper.

2. The particles *ya* and *şo* are complementizers

In this section, we describe the common distributional and functional properties of the particles *ya* and *şo.* We argue that the particles *ya* and *şo* belong to the category of complementizers. This finds initial support from their embedded use in examples such as (5a) and (5b), where *ya* and *şo* appear to subordinate the matrix clause in (4):

(4) Mp'oli msk'va on.
Istanbul beautiful be.3SG
'Istanbul is beautiful.'

(5) a. Arte-k [Mp'oli msk'va on ya] t'k'-u
Arte-ERG Istanbul beautiful be.3SG YA say-PST.3SG
'Arte said: "Istanbul is beautiful."'

b. [Mp'oli msk'va on şo] t'k'v-i!
Istanbul beautiful be.3SG ŞO say-IMP.2SG
'Say: "Istanbul is beautiful."'

The goal of this section is to substantiate this claim and extend it to the root occurrences of these complementizers. Our claim regarding the embedded occurrences of these complementizers rests on two main observations: they cannot co-occur with other complementizers and they can embed finite clauses that are *not* syntactically opaque like quotations, even though both *ya* and *şo* make *indexical shift* in their scope possible (unlike the run-of-the-mill complementizer *na* which will be discussed shortly). As for the root occurrences of these complementizers, we argue that they are genuinely unembedded. Despite the translations offered for sentences that feature the root occurrences of *ya* and *şo,* all tests to identify a hidden VP embedding these clauses fail to detect such a VP, which provides evidence that *ya* and *şo* are functional categories that can introduce finite dependents without the help of an embedding verb. By the null hypothesis that embedded and root occurrences of *ya* and *şo* are of identical categories, we propose that they can act as

root complementizers. Furthermore, there is suggestive diachronic evidence that these complementizers have *not* followed a grammaticalization path from verbs of saying. Rather, it seems likely that they are historically related to manner deictic expressions, such as *thus* in the language, a common path of grammaticalization (cf. Güldemann 2002).

2.1 Evidence for the complementizer status of *ya* and *şo*

As was shown in (5) above, the particles *ya* and *şo* appear to subordinate finite clauses, which we take to indicate that they both function as complementizers. In this section, we bring evidence for this claim by situating *ya* and *şo* in the complementizer system of Laz. Laz has a run-of-the-mill complementizer *na* (akin to *that* in English), whose attachment behavior suggests that it is a proclitic. The proclitic complementizer *na* can embed all kinds of propositional attitudes as shown by the examples in (6):

(6) a. Arte-k [Mp'oli msk'va na=on] t'k'-u.
Arte-ERG Istanbul beautiful NA=be.3SG say-PST.3SG
'Arte said that Istanbul is beautiful.'
b. Arte-k [Mp'oli msk'va na=on] iduşun-am-s
Arte-ERG Istanbul beautiful NA=be.3SG think-IPFV-PRS.3SG
'Arte thinks that Istanbul is beautiful.'
c. Arte-k [Mp'oli msk'va na=on] moivar-u.
Arte-ERG Istanbul beautiful NA=be.3SG deny-PST.3SG
'Arte denied that Istanbul is beautiful.'

What is important for the current purposes is that the particles *ya* and *şo* cannot co-occur with *na*, as shown in the examples in (7) below. This suggests that they exhibit complementizer behavior in the language:

(7) a. Arte-k [Mp'oli msk'va (*na=)on ya] t'k'-u.
Arte-ERG Istanbul beautiful NA=be.3SG YA say-PST.3SG
'Arte said: "Istanbul is beautiful."'
b. [Mp'oli msk'va (*na=)on şo] t'k'v-i!
Istanbul beautiful NA=be.3SG ŞO say-IMP.2SG
'Say: "Istanbul is beautiful."'

On the other hand, *ya* and *şo* are much more restricted than *na. The* complementizer *na* appears in relative clauses as shown in (8), while *ya* or *şo* cannot do so (cf. 9):

(8) [Nana-şk'imi-k na=ç'-u] foga
mother-my-ERG NA=sew-PST.3SG dress
"the dress that my mother sewed"

(9) *[Nana-şk'imi-k ç'-u ya/şo] foga
mother-my-ERG sew-PST.3SG YA/ŞO dress
Intended: "the dress that my mother sewed"

Similarly, while the complementizer *na* appears in embedded questions as shown in (10) and (11), *ya* and *şo* cannot do so (cf. (12) and (13)):

(10) Şana-k [ham foga mi-k na=ç'-u] mi-ts'-u.
Şana-ERG this dress who-ERG NA=sew-PST.3SG 1SG.APPL-tell-PST.3SG
'Şana told me who sewed this dress.'

(11) [Ham foga mi-k na=ç'-u] mi-ts'v-i!
this dress who-ERG NA=sew-PST.3SG 1SG.APPL-tell-IMP.2SG
'Tell me who sewed this dress!'

(12) *Şana-k [ham foga mi-k ç'-u ya] mi-ts'-u.
Şana-ERG this dress who-ERG sew-PST.3SG YA 1SG.APPL-tell-PST.3SG
Intended: 'Şana told me who sewed this dress.'

(13) *[Ham foga mi-k ç'-u şo] mi-ts'v-i!
this dress who-ERG sew-PST.3SG ŞO 1SG.APPL-tell-IMP.2SG
Intended: 'Tell me who sewed this dress!'

We argue that (12) and (13) are unacceptable under the intended interpretations precisely because they are unable to embed questions. The declarative counterparts of the sentences in (12) and (13) are acceptable, as the control sentences in (14) and (15) demonstrate:

(14) Şana-k [ham foga Arte-k ç'-u ya] mi-ts'-u.
Şana-ERG this dress Arte-ERG sew-PST.3SG YA 1SG.APPL-tell-PST.3SG
'Şana told me Arte sewed this dress.'

(15) [Ham foga Arte-k ç'-u şo] mi-ts'v-i!
this dress Arte-ERG sew-PST.3SG ŞO 1SG.APPL-tell-IMP.2SG
'Tell me Arte sewed this dress!'

Despite these restrictions on *ya* and *şo,* there is robust evidence that they genuinely embed clauses which are syntactically non-opaque. As (16) and (17) illustrate, a *wh*-phrase can take matrix scope out of a clause that *ya* or *şo* embed. Note that both questions are construed as regular information-seeking questions:

(16) Şana-k [Arte *nak* ort'u ya] t'k'-u?
Şana-ERG Arte wherebe. PST.3SG YA say-PST.3SG
'Where did Şana say Arte was?'

(17) Şana-k [Arte *nak* ort'u şo] t'k'v-a-s?
Şana-ERG Arte where be.PST.3SG ŞO say-IMP-PRS.3SG
'Where should Şana say Arte was?'

Similarly, relativization out of a clause that *ya* or *şo* embed is licit, as shown in (18) and (19):

(18) [Şana-k [Arte-k e_1 dzir-u ya] na=t'k'-u] bere$_1$
Şana-ERG Arte-ERG see-PST.3SG YA NA=say-PST.3SG child
'the child who Şana said Arte saw'

(19) [Şana-k [Arte-k e_1 dzir-u şo] na=t'k'v-a-s] bere$_1$
Şana-ERG Arte-ERG see-PST.3SG ŞO NA=say-IMP-3SG child
'the child who Şana should say Arte saw'

So far, we have seen data in favor of categorizing *ya* and *şo* as complementizers in Laz, along with the run-of-the-mill complementizer *na*. Building on our earlier work on *ya* (Demirok and Öztürk 2015), in what follows we argue that *ya* and *şo* are not regular complementizers and are highly restricted in terms of which propositional attitudes they can embed. Furthermore, we argue that they have the unique property of requiring *indexical shift* in their complements.

The complementizers *ya* and *şo* have a restricted distribution, unlike *na*. They cannot embed all sorts of propositional attitudes, as evidenced by the limited set of attitude verbs under which they can occur. The complementizer *ya* is able to occur under the Laz equivalents of 'say', 'tell', and 'think', as illustrated in (20):

(20) a. Arte-k [Mp'oli msk'va on ya] t'k'-u.
Arte-ERG Istanbul beautiful be.3SG YA say-PST.3SG
'Arte said that Istanbul is beautiful.'

b. Arte-k [Mp'oli msk'va on ya] mi-ts'-u.
Arte-ERG Istanbul beautiful be.3SG YA 1SG.APPL-say-PST.3SG
'Arte told me that Istanbul is beautiful.'

c. Arte-k [Mp'oli msk'va on ya] iduşun-am-s.
Arte-ERG Istanbul beautiful be.3SG YA think-IPFV-PRS.3SG
'Arte thinks that Istanbul is beautiful.'

However, *ya* cannot co-occur with any other attitude predicate, as illustrated by the examples in (21). All of these predicates require *na* for embedding:

(21) a. *Arte-k [Mp'oli msk'va on ya] moivar-u.
Arte-ERG Istanbul beautiful be.3SG YA deny-PST.3SG
Intended: 'Arte denied that Istanbul is beautiful.'

b. *Arte-s [Mp'oli msk'va on ya] uşk'un.
Arte-DAT Istanbul beautiful be.3SG YA know. PRS.3SG
Intended: 'Arte knows that Istanbul is beautiful.'

c. *Arte-s [Mp'oli msk'va on ya] aceren.
Arte-DAT Istanbul beautiful be.3SG YA believe.PRS.3SG
Intended: 'Arte believes that Istanbul is beautiful.'

The complementizer *şo* is similar to *ya* but is even more restricted: it is only able to occur under the Laz equivalents of 'say', 'tell', but not 'think' (or other any predicate). This is shown in the examples in (22) below:

(22) a. Aşela-k [Mp'oli msk'va on şo] t'k'v-a-s!
Aşela-ERG Istanbul beautiful be.3SG ŞO say-IMP-3SG
'Aşela should say that Istanbul is beautiful!'
b. Aşela-s [Mp'oli msk'va on şo] u-ts'v-i!
Aşela-DAT Istanbul beautiful be.3SG ŞO 3SG.APPL-say-IMP.2SG
'Tell Aşela that Istanbul is beautiful!'
c. *Aşela-k [Mp'oli msk'va on şo] iduşun-a-s!
Aşela-ERG Istanbul beautiful be.3SG ŞO think-IMP-3SG
Intended: 'Aşela should think that Istanbul is beautiful!'

The restricted use of the complementizers *ya* and *şo* is related to a common function that they have in the language. Indexicals in the scope of *ya* and *şo* shift obligatorily. As is well-known, indexical shift, as seen in directly quoted speech reports, refers to the shift in deictic elements like personal pronouns (cf. Kaplan 1989; Schlenker 1999, 2003). To illustrate this point, the shift is obligatorily observed in the sentences in (23), where the embedded first person pronoun *ma* refers to Şana, not to the speaker who utters the sentences in (23):

(23) a. Şana-k [ma msk'va vore ya] t'k'-u.
Şana-ERG 1SG beautiful be.1SG YA say-PST.3SG
Lit: 'Şana said: I am beautiful.'
'Şana$_1$ said that she$_1$ is/*I am beautiful.'
b. Şana-k [ma msk'va vore ya] iduşun-am-s.
Şana-ERG 1SG beautiful be.1SG YA think-IPFV-PRS.3SG
Lit: 'Şana thinks: I am beautiful.'
'Şana$_1$ thinks that she$_1$ is/*I am beautiful.'

Clearly, this behavior has to be attributed to *ya* in that *na* does not allow indexical shift in its scope, as shown in (24). The same behavior extends to *şo,* as one can see in (25):

(24) Şana-k [ma msk'va na=vore] t'k'-u.
Şana-ERG 1SG beautiful NA=be.1SG say-PST.3SG
'Şana$_1$ said that I am/*she$_1$ is beautiful.'

(25) a. Şana-k [ma msk'va vore şo] t'k'v-a-s!
Şana-ERG 1SG beautiful be.1SG ŞO say-IMP-3SG
Lit: 'Şana should say: I am beautiful!'
'Şana$_1$ should say that she$_1$ is/*I am beautiful.'

b. [ma msk'va vore ṣo] t'k'v-i!
1SG beautiful be.1SG ṢO say-IMP.2SG
Lit: 'Say: I am beautiful!'
'Say that you are/*I am beautiful!'

We have seen evidence that clauses introduced by *ya* and *ṣo* are *not* syntactically opaque because they allow relativization, and because the *wh*-phrases they contain can take matrix scope. Then, it may seem suspicious that indexical shift is *obligatory* under *ya* and *ṣo*. As is well known, this behavior is associated with quotations, which are opaque domains for operations such as relativization, as evidenced by (26):

(26) *This is the dress$_1$ that Sue said: "I wanna buy e_1"

However, there is now ample cross-linguistic evidence that indexical shift is possible in clauses that are not quotes (Anand and Nevins 2014; Podobryaev 2014; Shklovsky and Sudo 2014; Deal 2020; a.o). This is also evidenced in Laz by the diagnostic criteria of non-opaqueness. Indexicals also obligatorily shift in demonstrably non-opaque clauses introduced by *ya* and *ṣo,* as illustrated by the sentences in (27) and (28):[3]

(27) a. Ṣana-k [ma nak vore ṣo] t'k'v-a-s?
Ṣana-ERG 1SG where be.1SG ṢO say-IMP-3SG
'Where should Ṣana$_1$ say that she$_1$ is/*I am?'
b. Ṣana-k [ma nak vore ya] t'k'-u?
Ṣana-ERG 1SG where be.1SG YA say-PST.3SG
'Where did Ṣana$_1$ say that she$_1$/*I was?'

(28) a. [Ṣana-k [ma e_1 bdziri ya] na=t'k'-u] bere$_1$
Ṣana-ERG 1SG see.PST.1SG YA NA=say-PST.3SG child
'the child who Ṣana$_1$ said she$_1$/*I saw'
b. [Ṣana-k [ma e_1 bdziri ṣo] na=t'k'v-a-s] bere$_1$
Ṣana-ERG 1SG see.PST.1SG ṢO NA=say-IMP-3SG child
'the child who Ṣana$_1$ should say she$_1$/*I saw'

3. An anonymous reviewer asks if the sentences in (27) also have declarative construals where it is a question that is being directly quoted. Indeed, our consultant has confirmed that this is possible for both *ya* and *ṣo*. Accordingly, (27a) can mean: 'Ṣana should say: "where am I?"' and (27b) can mean: 'Ṣana said: "where am I?".

To conclude, *ya* and *şo* are special complementizers, serving as devices that signal indexical shift.[4] For our purposes, what derives the distinctness of *ya* and *şo* is orthogonal to the question of how indexical shift is to be analyzed, for both complementizers exhibit obligatory indexical shift in their scope. However, it is important to observe that their morphological distinctness from *na* must be attributed to this indexical shift property shared by both. Moreover, the phenomenon of indexical shift is not available under all attitude verbs in Laz, which is line with the data available on indexical shift across languages. Typological surveys have shown that languages can be quite picky about which of their verbs allow indexical shift in their complements (Deal 2020). Hence, the restrictions on which verbs *ya* and *şo* can occur with may at least partially follow from possibly independent restrictions on which attitude predicates allow for indexical shift in their complements. In the next subsection, we extend our claims to the root occurrences of *ya* and *şo.*

2.2 Evidence for the complementizer status of *ya* and *şo*

As we have mentioned in the introduction, *ya* or *şo* can also appear without an embedding verb, as illustrated by the examples in (29):

(29) a. Arte-k [oxori-s vore] ya.
Arte-ERG house-LOC be.1SG YA
'Arte$_1$ said/says that he$_1$ is home.'
b. [noseri vore] şo!
smart be.1SG ŞO
'Say that you are smart!'

The interpretation of such sentences is as if there is still an embedding verb, as the translations into English provided suggest. Identifying whether examples of the type given in (29) involve verb ellipsis or not is crucial. If there is an operation of verb ellipsis, there are no new questions arising from that (except, of course, questions about ellipsis). We would simply have the two complementizers embedded under a verb; however, an ellipsis process (albeit a curiously special one) would elide the verb along with any morphology it bears, giving us strings without a main verb as in (29).

4. There are various proposals on how to derive indexical shift. One approach involves so-called *monster* operators argued to be part of the functional sequence in the left periphery of some clauses, see e.g. Shklovsky and Sudo (2014). Deal (2020) proposes that Laz in fact bundles such operators into complementizers, citing our earlier work on *ya* (Demirok and Öztürk 2015). An alternative view is defended by Sundaresan (2019), who takes indexical shift in some languages to be made possible by complementizers themselves. Both kinds of approaches are compatible with Laz, which has a morphologically distinguished class of complementizers for indexical shift.

If, on the other hand, there is no verb ellipsis, it seems inevitable to posit that *ya* or *şo* are meaningful and, moreover, that they have distinct meanings, as the translations make it clear. This would also raise the question of whether their meanings can be extended to their occurrences in embedded contexts.

In our earlier work on *ya* (Demirok et al. 2019), we provided structural evidence that points to the *absence* of ellipsis when *ya* appears without a main verb. Here, we also report parallel data on *şo* and argue that *ya* and *şo* exhibit identical behavior when they appear without a main verb.

First, questioning a constituent of a clause introduced by *ya* or *şo* is only possible if it is embedded by an overt verb (i.e. an overt verb of saying has to follow *ya* or *şo*). This is illustrated by the sentences in (30) below:

(30) a. Arte-k [*nak* vort'i] ya *(t'k'-u)?
Arte-ERG where be.PST.1SG YA say-PST.3SG
'Where did Arte$_1$ say that he$_1$ was?'

b. Arte-k [*nak* vort'i] şo *(t'k'v-a-s)?
Arte-ERG where be.PST.1SG ŞO say-IMP-3SG
'Where should Arte$_1$ say that he$_1$ was?'

Notably, there is evidence that it is not the prosodic requirements of matrix interrogative structures that preclude the verb omission here. As shown by the sentences in (31), a matrix question can be formed by questioning the matrix subject even when there is no overt verb following *ya* or *şo*. The only difference between the sentences in (30) and the sentences in (31) is a structural one, namely where the *wh*-extraction originates in:[5]

(31) a. *Mi-k* [oxori-s vort'i] ya (t'k'-u)?
who-ERG house-LOC be.PST.1SG YA say-PST.3SG
'Who$_1$ said that they$_1$ were home?'

b. *Mi-k* [oxori-s vort'i] şo (t'k'v-a-s)?
who-ERG house-LOC be.PST.1SG ŞO say-IMP-3SG
'Who$_1$ should say that they$_1$ were home?'

Second, if the verb appears in the abilitative form requiring an oblique dative subject, it cannot be omitted. It seems that the dative case on the subject does not suffice to recover the modal content expressed on the verb, as shown in (32):

5. Nevertheless, there is evidence independent of wh-extraction that root *ya* and *şo* do not necessarily introduce quotes. As reported in Demirok et al. (2019), *Berepek Lazi voret ya* 'Lit: The children say: "We are Laz".' is judged felicitous if each children utters *Lazi vore* "I am Laz", and none utters *Lazi voret* "We are Laz.". This is also true for *şo*.

(32) a. Arte-s [Lazi vore] ya *(a-zit'-u).
Arte-DAT Laz be.PRS.1SG YA ABIL-say-PST.3SG
'Arte$_1$ was able to say that he$_1$ is Laz.'
b. Arte-s [Lazi vore] şo *(a-zit'-a-s)!
Arte-DAT Laz be. PRS.1SG ŞO ABIL-tell-IMP-3SG
'Arte$_1$ should be able to say that he$_1$ is Laz!'

Third, a clause introduced by *ya* or *şo* cannot be further embedded without a main verb, as shown by the examples in (33):

(33) a. [Arte-k [oxori-s vort'i] şo *(t'k'v-a-s)] bgorum.
Arte-ERG house-LOC be.PST.1SG ŞO say-IMP-3SG want.IPFV.1SG
'I want Arte$_1$ to say that he$_1$ was home.'
b. [Arte-k [oxori-s vort'i] ya *(t'k'-u)] domats'onu.
Arte-ERG house-LOC be.PST.1SG YA say-PST-3SG think.PST.1SG
'I thought that Arte$_1$ said that he$_1$ was home.'

It is not clear how these asymmetries are to be accounted for. Nevertheless, they are puzzling under an account which takes some verbs to be elided by a mysterious ellipsis process. Ellipsis is never possible for verbs that embed clauses introduced by *na*. This is illustrated by the sentence in (34) below:

(34) Arte-k [Mp'oli msk'va na=on] *(t'k'v-u).
Arte-ERG house-LOC beautiful NA=be.3SG say-PST-3SG
'Arte said that Istanbul is beautiful.'

Based on these structural asymmetries, we argue that there is sufficient evidence to reject an ellipsis account. Notably, in the absence of ellipsis, it is inevitable to conclude that *ya* and *şo* are meaningful elements and, furthermore, that they have distinct meanings.

In what follows, we argue that despite what translations suggest, *ya* and *şo* do not signal a VP, and do not make a speech event visible to the compositional semantics.[6] In that sense, they differ from the root complementizer *que* in Iberian Spanish which Etxepare (2008) argues to signal the presence of a VP, i.e. a speech eventuality. We argue that all that *ya* and *şo* can do is to relate an individual, always

6. As an anonymous reviewer rightly points out, it is not always possible to talk about an isomorphism between events and verbs. Since this point is largely orthogonal to the discussion here, we make the simplifying assumption that a speech event being visible to the compositional interpretation entails the presence of a VP in the structure, and hence we use these terms interchangeably throughout.

marked ergative,[7] to a proposition (the finite dependent it introduces). No further modification is possible when the verb is omitted.

First, while an ergative 'subject' is licensed without a verb following *ya* or *şo,* as can be seen in the sentences in (31), adding an overt dative addressee/indirect object is only possible when there is an overt verb following *ya* or *şo,* as illustrated in (35) below:

(35) a. Arte-k *nana-s* [oxori-s vort'i] ya *(u-ts'-u).
Arte-ERG mother-DAT house-LOC be.PST.1SG YA APPL.3SG-tell-PST.3SG
'Arte$_1$ said to the mother that he$_1$ was home.'

b. Arte-k *nana-s* [oxori-s vort'i] şo *(u-ts'v-a-s)!
Arte-ERG mother-DAT house-LOC be.PST.1SG ŞO APPL.3SG-tell-IMP-3SG
'Arte$_1$ should say to the mother that he$_1$ was home!'

Second, an adverb of quantification cannot be used when the verb is omitted, as shown in (36):

(36) a. Arte-k *p'anda* [noseri vore] ya *(it'ur-s).
Arte-ERG always smart be.PRS.1SG YA say.IPFV-PRS.3SG
'Arte$_1$ always says that he$_1$ is smart.'

b. Arte-k *p'anda* [noseri vore] şo *(t'k'v-a-s)!
Arte-ERG always smart be.PRS.1SG ŞO say-IMP-3SG
'Arte$_1$ should always say that he$_1$ is smart!'

Third, temporal adverbs cannot be used when the verb is omitted, as shown in (37):

(37) a. Arte-k *ğoma* [noğa-şe vidare] ya *(t'k'-u).
Arte-ERG yesterday market-ALL go.FUT.1SG YA say-PST.3SG
'Yesterday, Arte$_1$ said that he$_1$ would go to the marketplace.'

b. Arte-k *oç'ume* [noğa-s v ort'i] şo *(t'k'v-a-s)!
Arte-ERG tomorrow market-LOC be.PST.3SG ŞO say-IMP-3SG
'Tomorrow, Arte$_1$ should say that he$_1$ was at the marketplace.'

7. An anonymous reviewer fairly asks how the ergative case is licensed here and whether it implies a functional or verbal head being in the structure. As we discuss below, there is no evidence for a VP. Hence, it is not clear how a functional projection such as *voice* can be present in the structure, licensing the ergative case. We do not make any definitive claims as to the source of the ergative case in Laz. However, we would like to point out that the relevant facts in Laz are also compatible with the ergative case being a dependent case (Baker and Bobaljik 2017), which would reflect a relation between two thematic entities rather than imply a relation with a head like *voice.* In short, in the absence of corroborating evidence, we reject the idea that the ergative case entails verbal structure.

Fourth, the lowest category of eventuality modifiers, namely manner adverbs, cannot be used when the verb is omitted, as shown in (38):[8]

(38) a. Şana-k *şurite* [kapça bgorum] ya *(t'k'-u).
Şana-ERG quietly anchovy want.IPFV.1SG YA say-PST.3SG
'Şana$_1$ quietly said that she$_1$ wants anchovies.'
b. Şana-k *şurite* [kapça bgorum] şo *(t'k'v-a-s)!
Şana-ERG quietly anchovy want.IPFV.1SG ŞO say-IMP-3SG
'Şana$_1$ should quietly say that she$_1$ wants anchovies.'

All of these tests point towards the conclusion that *ya* and *şo* do not by themselves provide a speech eventuality and further corroborate the claim that there is no verb ellipsis. As Etxepare (2008) argues for the root complementizer *que* in Iberian Spanish, the compatibility with adverbs of quantification, manner adverbs, addressee NPs, and temporal adverbs serves to diagnose a speech eventuality (and some higher projections) as being part of the structure. Hence, parallel constructions in Iberian Spanish and Laz clearly contrast on the same tests that Etxepare employs, and thus show us that *ya* and *şo* in Laz correspond to minimal structures, complementizers by hypothesis.

Finally, it should be mentioned that there is also no diachronic evidence for a speech eventuality being signaled by *ya* and *şo.* A grammaticalization path from speech eventualities to complementizers has been argued to be cross-linguistically attested (Klamer 2000). Nevertheless, there is no morphological relatedness of *ya* and *şo* to any of the attested speech verbs in Laz, as shown in (39):

(39) a. zit' 'say' – elsewhere allomorph
b. t'k'v 'say' – perfective allomorph
c. it'ur 'say' – imperfective allomorph
d. ts'v 'tell'

Rather, it seems that the complementizers *ya* and *şo* have their origin in manner deictic expressions such as 'thus', 'be as follows', 'that is to say', which have been

8. Upon reviewer recommendations, we decided to apply the tests Etxepare employs to detect a speech eventuality. To our surprise, in the tests that concern the acceptability of manner adverbs and temporal adverbs as modifiers of *ya* clauses that occur without an embedding verb, we elicited judgments different from what we report in Demirok et al. (2019). If there was no error on our part in eliciting the data for our earlier work, our primary consultant's grammar seems to have shifted towards the variant we report here. In order to ascertain the judgment, we have consulted three additional Laz native speakers, who have unanimously rejected the sentences in (35)–(38) when the main verb was omitted. We should also note that the judgments we report here regarding the acceptability of manner adverbs and temporal adverbs are consistent with the findings from the other two tests, namely the incompatibility with an addressee and adverbs of quantification.

shown to acquire functions of quotative markers and complementizers in the context of languages such as Shona, as discussed in detail in Güldemann (2002).

As one may check in (40b)–(c), *ya* and *şo* can still synchronically combine with the demonstratives in (40a) yielding manner deictic expression:

(40) a. ha- 'this'
hi- 'that'
b. ha-*şo* 'so, in this way'
hi-*şo* 'so, in that way'
muç'o [←?mu-*şo*] 'how, in what way'
c. ha-*ya* 'this'
hi-*ya* 'that/he/she/it'

However, a synchronic account that builds on the historical origins of *ya* and *şo* is beyond the scope of this paper. It should suffice for our current purposes that these complementizers are not historically linked to a speech eventuality.

3. On the morphological distinctness of *ya* and *şo*

In the previous section, we have tried to uncover the commonalities between *ya* and *şo.* We have shown that they are special quotative complementizers that can introduce finite clauses which are not opaque to grammatical processes such as extraction. Furthermore, we have argued that they can occur as main clause complementizers, without the help of an embedding verb. Once we admit this, *ya* and *şo*, as root complementizers, will need to have distinct meanings.

In the next subsection, we discuss our proposal for what *ya* and *şo* mean when they are root complementizers. In the subsection that follows, we will turn to embedded occurrences of *ya* and *şo.*

3.1 What do the roots complementizers in Laz mean?

Given examples such as (41), it seems that *şo* corresponds to an imperative verb of saying while *ya* corresponds to an indicative verb of saying, respectively:

(41) a. Arte-k [noseri vore] ya.
Arte-ERG smart be.1SG YA
'Arte$_1$ said/says that he$_1$ is smart.'
b. Arte-k [noseri vore] şo!
Arte-ERG smart be.1SG ŞO
'Arte$_1$ should say that he$_1$ is smart!'

However, recall from Section 2.2 that there is no evidence for an outer VP layer being available in sentences like the ones in (41). They do not allow any additional NP argument (e.g. addressee) or an event modifier (e.g. manner adverbs, temporal adverbs, adverbs of quantification). We take these facts to show that there is no node in their structure which these elements can (type-wise) combine with. In other words, even if *ya* and *şo* talked about speech eventualities, they would not make a predicate of events of type <*v,t*> available in syntax. What they *do* seem to do is relate an individual to a proposition, without the mediation of a predicate of events in the composition. Accordingly, we propose the approximate meaning for *ya* in (42a) and the approximate meaning for *şo* in (42b):

(42) a. $[[\text{ya}]]^c = \lambda p.\ \lambda x.$ *linguistic-production*(p) = 1 and *source*(p)=x
b. $[[\text{şo}]]^c = \lambda p.\ \lambda x.$ speaker in c WANT: [SAY(p)(x)]

Needless to say, the proposed meanings gloss over many important aspects of meaning calculation. For example, the indexical shift that takes places in the complement of *ya* and *şo* would need to be formally integrated with the semantics provided here. Also, WANT is obviously a placeholder for an accurate semantics for imperatives (see Kaufmann 2012, 2019; Portner 1997, 2004, a.o.) which we do not need to go into the details of for the purposes of the current discussion. Despite these tentative aspects of the proposal, there are correct predictions that we make.

As shown in (42a), *ya* does not introduce a speech eventuality in the compositional interpretation. Hence, we correctly predict that it is impossible for an addressee NP or any event modifier to be added to the structure. We make use of the concept of "linguistic production" (in the result/output sense) to accommodate the fact that root *ya* clauses can be understood to report inner speech/thoughts which are also a kind of linguistic content like speech (but unlike belief).[9] Recall that *ya* can be embedded under a verb of thinking but not under a verb of believing. Similarly, our consultant informs us that in the absence of an overt verb, too, an "inner speech/thought" construal is available. For example, the sentence in (43) is reported to be felicitous in a context where Arte is shy and never talks about himself but from his behavior/manners we can infer that "I am smart" is what he is thinking/saying to himself:

(43) Arte-k [noseri vore] ya.
Arte-ERG smart be.1SG YA
'Arte is saying/thinking: "I am smart."'

9. We adapt the notion of linguistic production from Demirok et al. (2019) where *linguistic production* is formalized as an event predicate that corresponds to the union of *think* and *say*.

We propose that *şo* bundles a modal layer in its meaning, for which we use WANT as a shorthand. Roughly, "speaker in c WANT: [SAY(p)(x)]" is meant to convey that in all worlds where the speaker's preferences are met, x says p. We tentatively choose to represent the inner proposition [SAY(p)(x)] as building out of a speech event. However, note that this speech eventuality is "buried" under a modal layer within the meaning of *şo*, correctly predicting the attested restrictions we observe for *ya*. In other words, there is a speech event but it does not correspond to any syntactic node that can combine with event modifiers. Our motivation to make use of a speech event in *şo* (as opposed to the notion of "linguistic production") is empirical. First, recall that *şo* can only be embedded under verbs of saying/telling, not thinking. Second, our consultant informs us that whoever utters (44) has a preference for there to be an utterance event where Arte says the sentence *Lazi vore* "I am Laz":

(44) Arte-k [Lazi vore] şo!
Arte-ERG Laz be.1SG ŞO
'Arte should say: "I am Laz."!'

In other words, Arte thinking "I am Laz" does not suffice for (44) to be utterable: Arte should say it.

While much detail is being left out in the brief discussion here, we believe that the meanings that we attribute to the root occurrences of *ya* and *şo* are on the right track and make correct empirical predictions. In the next subsection, we turn to the embedded occurrences of *ya* and *şo* and show that the null hypothesis, namely that embedded *ya* and *şo* should mean what root *ya* and *şo* mean, cannot work. We argue for a semantic account that preserves a meaning difference between the two in embedded contexts as an alternative to a formal agreement account, which does not.

3.2 Embedded *ya* and *şo*

The embedded uses of *ya* and *şo* place a restriction on the grammatical mood of the matrix attitude predicate that embeds them. It seems that *şo* can only occur with an attitude predicate marked with the *imperative* mood whereas *ya* is the elsewhere form, i.e. occurring with attitude predicates marked with the indicative mood.[10] This distribution is illustrated in the examples in (45) and (46) below:

10. For expository purposes, we take third person forms typically glossed as subjunctive or optative to be imperative forms. It seems to us that this label is essentially orthogonal to the claims and characterizations made in this paper.

(45) a. Arte-k [Mp'oli msk'va on ya] t'k'-u.
Arte-ERG Istanbul beautiful be.3SG YA say-PST.3SG
'Arte said that Istanbul is beautiful.'
b. Arte-k [Mp'oli msk'va on ya] iduşun-am-s.
Arte-ERG Istanbul beautiful be.3SG YA think-IPFV-PRS3SG
'Arte thinks that Istanbul is beautiful.'

(46) a. [Mp'oli msk'va on şo] t'k'v-i!
Istanbul beautiful be.3SG ŞO say-IMP.2SG
'Say that Istanbul is beautiful!'
b. Tanura-k [Mp'oli msk'va on şo] t'k'v-a-s!
Tanura-ERG Istanbul beautiful be.3SG ŞO say-IMP-3SG
'Tanura should say that Istanbul is beautiful!'[11]

As shown by the examples in (47), *ya* is ungrammatical under an imperative marked predicate and likewise *şo* is ungrammatical under an indicative marked predicate:

(47) a. *[Mp'oli msk'va on *ya*] t'k'v-i!
Istanbul beautiful be.3SG YA say-IMP.2SG
'Say that Istanbul is beautiful.'
b. *Arte-k [Mp'oli msk'va on *şo*] t'k'-u.
Arte-ERG Istanbul beautiful be.3SG ŞO say-PST.3SG
'Arte said that Istanbul is beautiful.'

From a functional perspective, the morphological distinctness between *ya* and *şo* in their embedded uses is superfluous. For example, an equally expressive, hypothetical, grammar is one where *şo* is only possible as a root complementizer while *ya* is used to signal indexical shift in *all* embedded environments. In this grammar, the sentence in (47a) would be grammatical while the sentences in (46) would be ungrammatical. However, Laz is different from this hypothetical grammar. In the meaning we gave to *şo*, repeated in (48) below, there is a modal layer which is supposed to be a shorthand for whatever semantics the imperative mood marked on verbs brings in. In addition, it embeds the meaning of a verb of saying. Essentially, *şo* is a meaning unit that bundles all the pieces of information that enters in the compositional interpretation of a verb of saying combining with certain functional projections on top of it (such as imperative mood). This means that the sentence in (46b), for example, doubles that information, having both *şo* and *t'k'v-a-s*:

(48) $[[\text{şo}]]^c = \lambda p.\, \lambda x.$ speaker in c WANT: [SAY(p)(x)]

11. We systematically use *should* in English to translate the third person imperative forms in Laz.

To put it simply, the sentence in (46b) is predicted to still mean what it means if we could simply remove *şo* from the picture. However, what is troubling is that the doubling of that information is not innocent. The null hypothesis is that morphemes have the same meaning in every context. But if we, by that null hypothesis, assume that *şo* has the meaning in (48) in its embedded use, things go wrong. For example, the sentence in (46b) ends up having the incorrect interpretation in (49).[12] A similar problem arises if we insist on interpreting embedded *ya* with the denotation we assigned to the root *ya*:

(49) speaker in c WANT:[SAY (speaker in c WANT: [SAY (*Istanbul is beautiful*) (Tanura?)])(Tanura)] = *the speaker wants Tanura to say that the speaker wants Tanura to say that Istabul is beautiful.*

Hence, the root meanings for *ya* and *şo* make incorrect predictions when we use them to interpret embedded *ya* and *şo.* Therefore, we reject the null hypothesis and argue, instead, for a semantic account that preserves a meaning difference between the two in embedded contexts as an alternative to a formal agreement account, which does not. In what follows, we first flesh out the agreement account and, then, turn to the alternative semantic account that connects the root and the embedded occurrences of the quotative complementizers.

3.2.1 *Are embedded* ya *and* şo *differentiated through agreement?*

As we have seen, the morphological distinctness of *ya* and *şo* in their embedded uses is superfluous. The data are, on the surface, readily compatible with an account where a formal mood feature, say [imperative], is at work in the morphological distinctness of *ya* and *şo* in embedded contexts. While this is doable with allomorphy conditioned by morphosyntactic features, the syntactic distance between the mood layer of a clause and the complementizer of the clause it embeds makes this untenable, for allomorphy is known to be subject to tight locality restrictions in general (Božič 2019). Such long distance co-variance effects are better handled by agreement. Notably, the agreement relation here cannot be the strictly downward Agree relation in Chomsky (2000, 2001). But it rather needs to be what is usually called Upwards or Reverse Agree (Zeijlstra 2012), given that the agreeing complementizer would need to receive its value from the mood layer of the *embedding* clause. This is perhaps unusual from the perspective of well-studied cases of complementizer agreement (e.g. in Germanic). But it is not unattested. See, for example, agreeing

12. For the types to work out, let us assume that the matrix subject binds a null pronoun in the embedded clause acting as the individual argument of *şo*.

complementizers in Lubukusu (Diercks 2013; Carstens 2016), which exhibit parallel behavior in terms of agreement directionality.[13]

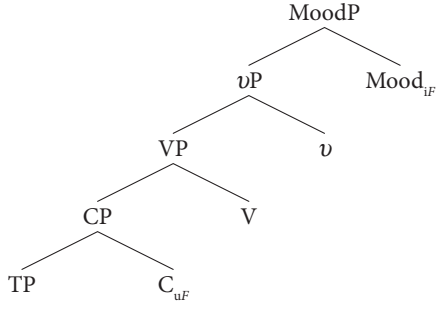
Notably, the kind of agreement that we would need to posit for Laz complementizers is not unprecedented. For example, Zeijlstra (2007) argues that in a sentence like (50), which is argued to feature what is called *modal concord*, the embedded modal *must* is not interpreted. Rather, it bears an uninterpretable mod(al) feature and has to agree with a higher element that carries the interpretable counterpart of this feature, which is the verb *demands* in this case:

(50) The general demands$_{[i\text{-MOD}]}$ that the troops must$_{[u\text{-MOD}]}$ leave

The kind of agreement relation posited here requires a configuration where the source of the agreement (Goal) c-commands the locus of the agreement (Probe).

Assuming an agreement operation that works upwards, we need a feature F whose interpretable variant is associated with the matrix mood projection and the uninterpretable variant is associated with the agreeing complementizer. A rough structure where the morphological co-variance between matrix mood and the complementizer is shown below in (51):

(51)



Positing an (agreement-related) uninterpretable mood feature is not a far-fetched idea language-internally. Verbal negation in Laz has mood-dependent forms, as shown by the examples in (52):

(52) a. Baba-sk'ani-k si haminapes *va* g-dzir-u.
father-your-ERG you around.here NEG 2SG-see-PST.3SG
"Your father didn't see you around here."

b. Baba-sk'ani-k si haminapes *mo* g-dzir-a-s.
father-your-ERG you around.here NEG.IMP 2SG-see-IMP-3SG
"Your father should not see you around here!"

13. See also Bjorkman and Zeijlstra (2019), who set out to settle the debate in favor of a view of grammar which universally establishes agreement in this direction, also addressing the problems with canonical phi-agreement configurations across languages.

However, this pattern may as well be analyzed as allomorphy, unlike the mood dependent complementizer forms which require a non-local conditioning context. We do not undertake the task of determining which analysis is correct for the data in (52).

The upwards agreement account proposed above provides a simple description of the basic facts about the distribution of embedded *ya* and *şo*. It is also a natural account in that the morphological distinction between *ya* and *şo* does indeed seem like a feature co-variance effect in embedded contexts.[14] Hence, we take this account seriously and evaluate it against further empirical data in what follows.

First, in Laz, the future marker *-ere* obligatorily attaches to a stem that is marked as subjunctive/imperative, as shown in (53). Future-marked verbs are not understood as commands or permissions, even though they require an imperative (alternative gloss: subjunctive) stem. Hence, it seems safe to say that the imperative marking is essentially vacuous in (53b):

(53) a. Arte-k oşkuri şkom-a-s!
Arte-ERG apple eat-IMP-3SG
'Arte should eat apples/the apple!'
b. Arte-k oşkuri şkom-a-s-*ere*.
Arte-ERG apple eat-IMP-3SG-FUT
'Arte will eat apples/the apple.'

Nevertheless, from the perspective of the agreement account, there is still a question. Can future-marked say/tell verbs embed clauses introduced by *şo*? The answer is no.

As shown in (54a) and (54b), under a future-marked verb of saying we can only see *ya*. Compare (54a) and (54c) to see the blocking effect of the future-marking *-ere*:

(54) a. *Aşela-k baba-muşi-s [oxoris vore *şo*] u-ts'v-a-s-ere.
Aşela-ERG father-her-DAT home be.1SG ŞO APPL.3SG-tell-IMP-3SG-FUT
'Aşela$_1$ will tell her father that she$_1$ is home.'
b. Aşela-k baba-muşi-s [oxoris vore *ya*]
Aşela-ERG father-her-DAT home be.1SG YA
u-ts'v-a-s-ere.
APPL.3SG-tell-IMP-3SG-FUT
'Aşela$_1$ will tell her father that she$_1$ is home.'

14. Regarding an anonymous reviewer's question, we would like to stress that the agreement account sketched here is aimed to address the morphological distinctness between *ya* and *şo* in *embedded* contexts only, where it appears as though they do not contribute to the meaning calculation. In root contexts, *ya* and *şo* have distinct meanings and they are both interpreted and are not agreeing with anything. See also Section 3.2.1 for a semantic account that we develop as a better alternative with wider empirical coverage.

c. Aşela-k baba-muşi-s [oxoris vore *şo*] u-ts'v-a-s!
Aşela-ERG father-her-DAT home be.1SG ŞO APPL.3SG-tell-IMP-3SG
'Aşela$_1$ should tell her father that she$_1$ is home!'

There is a legitimate question for the agreement account: if it is agreement with a formal feature that licenses *şo*, then how come we do not have that same morphosyntactic feature (being exponed) in future-marked forms which obligatorily attach to an imperative stem? A licit response to this question would resort to the distinction between uninterpretable and interpretable features (Chomsky 1995). If it is only interpretable features that can be the source of an agreement operation, then the imperative-marking under future could be the realization of an uninterpretable variant of the formal feature that normally licenses *şo*, explaining the lack of agreement in (54a) above.

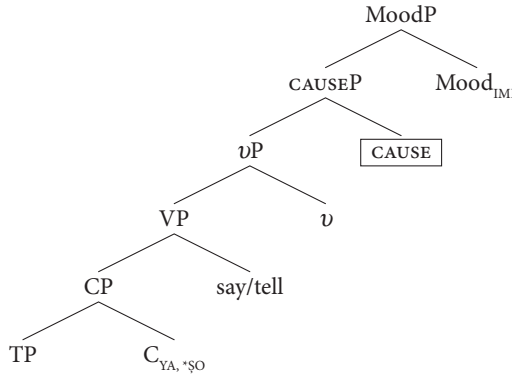
A more challenging observation for the agreement account comes from causativization. In Laz, the prefix *o-* and the suffix *-ap* are added to a verbal stem to causativize it. Causativization in this morphological frame can also target transitive predicates, including verbs that are compatible with *şo*. Interestingly, we observe that a causativized speech predicate bearing imperative inflection cannot embed a clause introduced by *şo*, as shown in (55a) and (56a). Instead, the complementizer *ya* appears despite imperative marking on the embedding predicate, as shown in (55b) and (56b):

(55) a. *Şana-s [noseri vore *şo*] o-zit'-ap-i!
Şana-DAT smart be.1SG ŞO CAUS-say-CAUS-IMP.2SG
Intended: 'Make Şana$_1$ say that she$_1$ is smart!'
b. Şana-s [noseri vore *ya*] o-zit'-ap-i!
Şana-DAT smart be.1SG YA CAUS-say-CAUS-IMP.2SG
'Make Şana$_1$ say that she$_1$ is smart!'

(56) a. *Arte-k Şana-s [noseri vore *şo*] o-zit'-ap-a-s!
Arte-ERG Şana-DAT smart be.1SG ŞO CAUS-say-CAUS-IMP-3SG
Intended: 'Arte should make Şana$_1$ say that she$_1$ is smart!'
b. Arte-k Şana-s [noseri vore *ya*] o-zit'-ap-a-s!
Arte-ERG Şana-DAT smart be.1SG YA CAUS-say-CAUS-IMP-3SG
'Arte should make Şana$_1$ say that she$_1$ is smart!'

This state of affairs is compatible with the idea that *ya* is the default form for the agreeing complementizer in a way that is similar to default third person singular forms that we observe with failed person-number agreement (Preminger 2014). From this perspective, *şo* would be the realization of the complementizer which successfully finds the imperative mood feature in the embedding clause. Then, it is possible to ascribe the blocking effect illustrated above to the additional layer of causative in the structure, as illustrated in (57) below:

(57)



However, it is not at all obvious why causative would block the agreement relation. A plausible hypothesis is that causative creates a syntactic domain (perhaps a *phase* in the sense of Chomsky (2000, 2001)) that the search space of the agreement operation cannot go beyond. We are not aware of any further data that supports or refutes this idea for Laz.

There is one last challenge for the agreement account, which we take to be beyond its reach. As shown in (58), neither *ya* nor *şo* can be in the scope of negation. By the logic of default realization with failed agreement (Preminger 2014), we would at least expect to see *ya* if negation is a blocker for agreement. This hypothesis is then a non-starter:

(58) a. *Tanura-k [noseri vore ya] va t'k'-u!
Tanura-ERG smart be.1SG YA NEG say-PST.3SG
Intended: 'Tanura did not say that he is smart!'
b. *Tanura-k [noseri vore şo] mo t'k'v-a-s!
Tanura-ERG smart be.1SG ŞO NEG.IMP say-IMP-3SG
Intended: 'Tanura should not say that he is smart!'

In what follows, we argue for an alternative semantic account that can also shed light on why neither *ya* nor *şo* is acceptable under negation.

3.2.2 *Are embedded* ya *and* şo *differentiated semantically?*

We have seen that the meanings we gave to root *ya* and *şo* are not usable in their embedded occurrences. As a first line of attack, we have sketched an agreement account under which there is simply no semantic difference between the two when they occur as embedded complementizers. While this account works, it also faces some challenges such as the causative layer blocking the putative agreement relationship and the failure of *ya* and *şo* to occur under negation. In this section, we sketch a semantic account that semantically relates root and embedded occurrences of *ya* and *şo* and explains the contrasts that we have subjected the agreement account to.

We propose an account where embedded *ya* and *şo* keep their original meanings but as not-at-issue content. In particular, we argue that *ya* and *şo* in embedded contexts both come with a semantic requirement that directly relates to their meanings in root contexts. We characterize these semantic requirements as a *post-supposition* (Brasoveanu 2013).[15] In other words, the semantic requirements embedded *ya* and *şo* bring in are definedness conditions that are necessarily evaluated late, once truth conditions are computed for the entire sentence.

To illustrate on concrete examples, our intuition is that the embedded occurrence of *şo* in (59a) requires that the sentence in question entail the sentence in (59b), and that the embedded occurrence of *ya* in (60a) requires that the sentence in question entails the sentence in (60b):

(59) a. Arte-k [noseri vore *şo*] t'k'v-a-s!
Arte-ERG smart be.1SG ŞO say-IMP-3SG
'Arte$_1$ should say that he$_1$ is smart!'
b. Arte-k noseri vore *şo*!
Arte-ERG smart be.1SG ŞO
'Arte$_1$ should say that he$_1$ is smart!'

(60) a. Şana-k [noseri vore *ya*] t'k'-u!
Şana-ERG smart be.1SG YA say-PST.3SG
'Şana$_1$ said that she$_1$ is smart!'
b. Şana-k noseri vore *ya*!
Şana-ERG smart be.1SG YA
'Şana$_1$ says/said/thinks that she$_1$ is smart!'

Recall that we gave the meanings in (61) to the root occurrences of *ya* and *şo*. Accordingly, the embedded occurrences of *ya* and *şo* will relate to their corresponding root meanings as in (62). Note that the embedded occurrences of *ya* and *şo* are truth-conditionally vacuous but come with definedness conditions that are evaluated late, once truth conditions are computed for the entire sentence.

(61) a. $[[\text{ya}_{\text{root}}]]^c = \lambda p.\ \lambda x.$ *linguistic-production*$(p) = 1$ and *source*$(p) = x$
b. $[[\text{şo}_{\text{root}}]]^c = \lambda p.\ \lambda x.$ speaker in c WANT: [SAY$(p)(x)$]

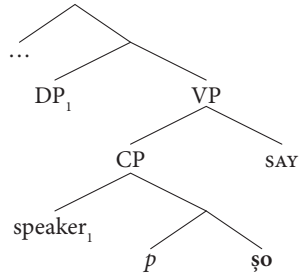
(62) for any sentence *S* that embeds *ya*/*şo*
a. $[[\text{ya}_{\text{embedded}}]]^c = \lambda p.\ \lambda x.$ p defined if and only if *S* entails $[[\text{ya}_{\text{root}}]]^c (p)(x)$.
b. $[[\text{şo}_{\text{embedded}}]]^c = \lambda p.\ \lambda x.$ p defined if and only if *S* entails $[[\text{şo}_{\text{root}}]]^c (p)(x)$.

Notably, the meanings in (62) requires that both $\text{ya}_{\text{embedded}}$ and $\text{şo}_{\text{embedded}}$ have access to the speaker *x*. Hence, for this proposal to work, a particular syntax is needed for where 'speaker' is represented syntactically in clauses embedded under

15. For the use of the term *semantic requirement*, see Szabolcsi (2015).

attitude predicates as in (62). For relevant discussion, see Akkuş and Hill (2021) and Portner et al. (2019) and the references therein.

(63)



The intuition behind our proposal will hopefully become clearer when we go through some examples and put this idea to work. Recall the fact that *şo* is not licensed with causativized speech predicates, as shown in (64). The semantic requirement we have posited for *şo* might help us understand why this is the case. Example (64), if it were grammatical, would be conveying that the speaker wants *Arte to make* Şana$_1$ say that she$_1$ is smart. This interpretation is distinct from the meaning requirement imposed by *şo* shown in (65). After all, if one has a preference for *y* to make *x* say *p*, that person does not necessarily have a preference for *x* to say *p*. Hence, the causative layer may be argued to be incompatible with the semantic requirement that *şo* imposes:

(64) *Arte-k Şana-s [noseri vore *şo*] o-zit'-ap-a-s!
Arte-ERG Şana-DAT smart be.1SG ŞO CAUS-say-CAUS-IMP-3SG
Intended: 'Arte should make Şana$_1$ say that she$_1$ is smart!'

(65) [[(64)]]^c is defined iff [[(64)]]^c entails that the speaker in c WANT: [SAY("I am smart")(Şana)]

On the other hand, following the same logic, the sentence in (66), for example, will not be precluded. Indeed, if one has a preference for *x* to tell y *p*, then that person has a preference for *x* say to *p*. Hence, the semantic requirement of *şo* is met in (66):

(66) Arte-k Şana-s [noseri vore *şo*] u-ts'v-a-s!
Arte-ERG Şana-DAT smart be.1SG ŞO APPL.3SG-tell-IMP-3SG
'Arte$_1$ should tell Şana say that he$_1$ is smart!'

The same logic can be applied to cases where there is no imperative marking on the matrix verb and cases where the imperative marking is not interpreted. Neither the formally imperative marked future form in (67a) nor the past tense marked form in (67b) has anything to do with speaker preferences. Hence, it could be argued that the semantic requirement of *şo* precludes such sentences that exhibit a mismatch in what the sentence actually says and what *şo* requires it to say:

(67) a. *Aşela-k [noseri vore *şo*] t'k'v-a-s-ere.
Aşela-ERG smart be.1SG ŞO say-IMP-3SG-FUT
'Aşela$_1$ will say that she$_1$ is smart.'
b. *Aşela-k [noseri vore *şo*] t'k'-u.
Aşela-ERG smart be.1SG ŞO say-PST.3SG
'Aşela$_1$ said that she$_1$ is smart.'

Recall that in cases where *şo* fails to get licensed, *ya* is licensed. This can be attributed to the fact that ya$_{embedded}$ ends up having a much weaker semantic requirement. It merely requires that the embedded speaker, which we syntactically represent, is the *source* of a proposition p such that p is the syntactic complement/first semantic argument of ya$_{embedded}$ and it is a *linguistic-production*. For example, in (68) the semantic requirement that ya$_{embedded}$ brings in is that the sentence in (68) entails that "I am smart" is a *linguistic-production* and that its *source* is Tanura. This is clearly met:

(68) Tanura-k Şana-s [noseri vore *ya*] u-ts'-u!
Tanura-ERG Şana-DAT smart be.1SG YA APPL.3SG-tell-PST.3SG
'Tanura$_1$ told Şana that he$_1$ is smart!'

An issue that we should acknowledge here is the fact that *ya* seems like the elsewhere of the two complementizers. For example, it is licensed under a causativized speech predicate when *şo* is not licensed, as shown in (69). However, if it is licensed here, then we may expect it to be licensed in (70), as well, being in free variation with *şo* in embedded contexts. Yet, this is not what we observe:

(69) Arte-k Şana-s [noseri vore *ya*] o-zit'-ap-a-s!
Arte-ERG Şana-DAT smart be.1SG YA CAUS-say-CAUS-IMP-3SG
'Arte should make Şana$_1$ say that she$_1$ is smart!'

(70) *Arte-k Şana-s [noseri vore *ya*] u-ts'v-a-s!
Arte-ERG Şana-DAT smart be.1SG YA APPL.3SG-tell-IMP-3SG
Intended: 'Arte$_1$ should tell Şana say that he$_1$ is smart!'

Therefore, it seems necessary to stipulate an extrinsic ordering between *ya* and *şo* in embedded environments: choose *şo* over *ya* unless *şo* is not licensed.

Importantly, there is also an environment where neither is licensed, as we have seen in the previous section. Neither *ya* nor *şo* can be in the scope of negation, as shown by the examples repeated in (71). Our semantic account for embedded *ya* and *şo* correctly predicts that *ya* and *şo* will not be licensed under negation, for the semantic requirements they introduce in the meaning calculation contradict what the sentences mean.[16] For example, *şo* in (71b) requires that the sentence in (71b)

16. Etxepare (2008, 2010, 2013) shows that the root *que* in Iberian Spanish exhibits restrictions concerning negation, disjunction, and quantification and argues that Krifka's (2001, 2003)

entail that the speaker *has* a preference for Tanura to say he is smart, contrary to fact. Similarly, *ya* in (71a) requires that the sentence in (71a) entail that "I am smart" is a *linguistic-production* whose *source* is Tanura, again contrary to fact:

(71) a. *Tanura-k [noseri vore ya] va t'k'-u!
Tanura-ERG smart be.1SG YA NEG say-PST.3SG
Intended: 'Tanura did not say that he is smart!'
b. *Tanura-k [noseri vore ṣo] mo t'k'v-a-s!
Tanura-ERG smart be.1SG ṢO NEG.IMP say-IMP-3SG
Intended: 'Tanura should not say that he is smart!'

To conclude, in this section we have presented an alternative to a formal agreement account and argued that it is superior with respect to empirical coverage. Moreover, we believe it to be generally desirable to semantically relate seemingly semantically vacuous occurrences of morphemes to their semantically non-vacuous occurrences, whenever this is possible. We hope to have demonstrated that this can be achieved in the case of *ya* and *ṣo*.

4. Final remarks

In this paper, we have investigated two quotative particles *ya* and *ṣo* in Laz and argued that they are special complementizers that introduce clauses that are not syntactically opaque like quotes and yet require indexical shift. Both can function as root complementizers, occurring without an embedding verb. We have argued that, when *ya* and *ṣo* function as root complementizers, they exhibit interesting restrictions that can be explained if root *ya* and *ṣo* do not occur under a covert/elided verb or make a speech eventuality available to the compositional interpretation. All diagnostic tests to detect a speech eventuality in root clauses headed by *ya* and *ṣo* point towards the absence of a speech eventuality available to the compositional interpretation. With this finding at hand, we have justified giving distinct meanings to *ya* and *ṣo* in root contexts. In the second part of the paper, we have explored the consequences of this move for their occurrences in embedded contexts. We have argued that the null hypothesis that they have the exact same semantic contribution in root and embedded contexts makes incorrect predictions. Given that their morphological differentiation in embedded contexts is dependent on the mood of the embedding verb, we have considered an account where the distinctness of *ya*

proposal that certain logical operations are not defined for speech acts explains these restrictions. While Etxepare's insightful explanations in terms of Krifka's semantics may in part carry over to root *ya* and *ṣo* in Laz (even though there are differences), it is not yet clear to us what it could offer for the embedded *ya* and *ṣo.* We leave a detailed comparison to future work.

and *şo* is not semantically-driven but is a reflex of an agreement relation with an interpretable mood feature in the embedding clause. We have shown that such an account faces some challenges (even though nothing conclusively shows that the agreement account is wrong). As an alternative to this formal agreement account, we have also sketched a semantic account where there is a systematic relationship between the meanings of root and embedded occurrences of these complementizers. We have shown that this semantic account is superior to the agreement account in terms of its empirical coverage. Furthermore, we think that it is generally desirable to relate the seemingly semantically vacuous occurrence of a morpheme to its semantically non-vacuous occurrence, which is achievable in the case of quotative complementizers in Laz.

There are also further issues that we have to leave to future work: a theory of indexical shift integrated with the proposals made here, further embedding configurations, speaker preferences shifting to addressee preferences in questions that feature *şo*, the grammaticalization path that *ya* and *şo* have followed, among others.

Acknowledgments

We would like to thank the organizers of the International Workshop on Discourse Particles who made this volume possible, and the two anonymous reviewers of this contribution as well as the audience at the workshop for their helpful and insightful feedback. We also gratefully acknowledge our primary language consultant for Laz, Ismail Bucaklişi, and three additional Laz speakers, whom we consulted to verify part of the reported data. This work builds on our earlier research on the various aspects of the complementizer system of Laz which we presented at SLE 48 and NELS 49. We also thank the audiences in these venues, and most importantly Deniz Özyıldız, who was our co-author at NELS 49. All remaining errors are solely ours.

References

Akkuş, Faruk & Hill, Virginia. 2021. Overt speakers in syntax. *Glossa* 6(1): 1–33. https://doi.org/10.5334/gjgl.1286

Anand, Pranav & Nevins, Andrew. 2004. Shifty operators in changing contexts. In *Proceedings of Semantics and Linguistic Theory (SALT)* 14, Robert B. Young (ed), 20–37. Ithaca: CLC Publications. https://doi.org/10.3765/salt.v14i0.2913

Baker, Mark C., & Bobalijk, Jonathan. 2017. On Inherent and Dependent Theories of Ergative Case. In *The Oxford Handbook of Ergativity*, Jessica Coon, Diane Massam, & Lisa De Mena Travis (eds), 113–134. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198739371.013.5

Bjorkman, Bronwyn M. & Zeijlstra, Hedde. 2019. Checking up on (ϕ-) Agree. *Linguistic Inquiry* 50(3): 527–569. https://doi.org/10.1162/ling_a_00319

Boeder, Winfried. 2002. Speech and thought representation in the Kartvelian (South Caucasian) languages. In *Reported Discourse: A meeting ground of different domains* [Typological Studies in Language 52], Tom Güldemann & Manfred von Roncador (eds), 3–48. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.52.04boe

Božič, Jurij. 2019. Constraining long-distance allomorphy. *The Linguistic Review* 36(3): 485–505. https://doi.org/10.1515/tlr-2019-2031

Brasoveanu, Adrian. 2013. Modified numerals as post-suppositions. *Journal of Semantics* 30(2): 155–209. https://doi.org/10.1093/jos/ffs003

Carstens, Vicki. 2016. Delayed valuation: a reanalysis of goal features, "upwards" complementizer agreement, and the mechanics of Case. *Syntax* 19(1): 1–42. https://doi.org/10.1111/synt.12116

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

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays in honor of Howard Lasnik*, Roger Martin, David Michaels, Juan Uriagereka & Samuel Jay Keyser (eds), 89–155. Cambridge MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, Michael Kenstowicz (ed), 1–52. Cambridge MA: MIT Press. https://doi.org/10.7551/mitpress/4056.003.0004

Deal, Amy Rose. 2020. *A Theory of Indexical Shift*. Cambridge MA: MIT Press. https://doi.org/10.7551/mitpress/12374.001.0001

Demirok, Ömer & Öztürk, Balkız. 2015. The logophoric complementizer in Laz. *Dilbilim Araştırmaları Dergisi [Journal of Linguistics Research]* 26(2): 45–69. https://doi.org/10.18492/dad.08998

Demirok, Ömer, Özyıldız, Deniz, & Öztürk, Balkız. 2019. Complementizers with attitude. In *Proceedings of North East Linguistic Society (NELS)* 49, Maggie Baird & Jonathan Pesetsky (eds), 213–222. Amherst MA: GLSA.

Diercks, Michael. 2013. Lubukusu complementizer agreement as a logophoric relation. *Natural Language & Linguistic Theory* 31(2): 257–407. https://doi.org/10.1007/s11049-013-9187-7

Etxepare, Ricardo. 2008. On quotative constructions in Iberian Spanish. In *Pragmatics of Clause Combining*, Ritva Laury (ed), 35–78. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.80.04etx

Etxepare, Ricardo. 2010. From hearsay evidentiality to samesaying relations. *Lingua* 120(3), 604–627. https://doi.org/10.1016/j.lingua.2008.07.009

Etxepare, Ricardo. 2013. Quotative expansions. In *Romance Languages and Linguistic Theory 2011: Selected Papers from 'Going Romance' Utrecht 2011*, Sergio Baauw, Frank Drijkoningen, Luisa Meroni & Manuela Pinto (eds), 93–124. Amsterdam: John Benjamins. https://doi.org/10.1075/rllt.5.05etx

Frajzyngier, Zygmunt. 1985. Logophoric Systems in Chadic. *Journal of African Languages and Linguistics* 7(1): 23–38. https://doi.org/10.1515/jall.1985.7.1.23

Güldemann, Tom. 2002. When 'say' is not say: the functional versatility of the Bantu quotative marker *ti* with special reference to Shona. In *Reported discourse: a meeting ground for different linguistic domains*, Tom Güldemann & Manfred von Roncador (eds.), 253–288. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.52.18gul

Güldemann, Tom. 2008. *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*. Berlin: de Gruyter. https://doi.org/10.1515/9783110211450

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: Oxford University Press.

Kaufmann, Magdalena. 2012. *Interpreting imperatives*. Berlin: Springer. https://doi.org/10.1007/978-94-007-2269-9

Kaufmann, Magdalena. 2019. Fine-tuning natural language imperatives. *Journal of Logic and Computation* 29(3): 321–348. https://doi.org/10.1093/logcom/exw009

Klamer, Marian. 2000. How report verbs become quote markers and complementisers. *Lingua* 110(2): 69–98. https://doi.org/10.1016/S0024-3841(99)00032-7

Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9(1): 1–40. https://doi.org/10.1023/A:1017903702063

Krifka, Manfred. 2003. Quantifiers in questions. *Korean Journal of English Language and Linguistics* 3(4): 449–526.

Lacroix, René. 2009. Description du dialecte laze d'Arhavi (caucasique du sud, Turquie): grammaire et textes. PhD dissertation, Université Lumière Lyon 2. <http://theses.univ-lyon2.fr/documents/lyon2/2009/lacroix_r/pdfAmont/lacroix_r_these.pdf>

Lord, Carol. 1993. *Historical Change in Serial Verb Constructions*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.26

Podobryaev, Alexander. 2014. Persons, imposters, and monsters. PhD dissertation, Massachusetts Institute of Technology.

Portner, Paul. 1997. The semantics of mood, complementation, and conversational force. *Natural Language Semantics* 5(2): 167–212. https://doi.org/10.1023/A:1008280630142

Portner, Paul. 2004. The semantics of imperatives within a theory of clause types. In *Semantics and Linguistic Theory (SALT) 14*, Robert B. Young (ed), 235–252. Ithaca: LSA and CLC Publications. https://doi.org/10.3765/salt.v14i0.2907

Portner, Paul, Pak, Miok & Zanuttini, Raffaella. 2019. The speaker-addressee relation at the syntax-semantics interface. *Language* 95(1): 1–36. https://doi.org/10.1353/lan.2019.0008

Preminger, Omer. 2014. *Agreement and Its Failures*. Cambridge MA: MIT Press. https://doi.org/10.7551/mitpress/9780262027403.001.0001

Schlenker, Philippe. 1999. Propositional attitudes and indexicality. PhD dissertation, Massachusetts Institute of Technology.

Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26(1): 29–120. https://doi.org/10.1023/A:1022225203544

Shklovsky, Kirill, & Sudo, Yasutada. 2014. The syntax of monsters. *Linguistic Inquiry* 45(3): 381–402. https://doi.org/10.1162/LING_a_00160

Sundaresan, Sandhya. 2019. A New Theory of Indexical Shift. Ms, University of Leipzig. <https://www.dropbox.com/s/dw5c9w51wn3l3xg/sundaresan_indexicalshift.pdf?dl=0> (30 September 2021)

Szabolcsi, Anna. 2015. What do quantifier particles do?. *Linguistics and Philosophy* 38(2): 159–204. https://doi.org/10.1007/s10988-015-9166-z

Zeijlstra, Hedde. 2007. Modal Concord. In *Proceedings of Semantics and Linguistic Theory (SALT)* 17, Masayuki Gibson & Tova Friedman (eds), 317–332. Ithaca: CLC Publications. https://doi.org/10.3765/salt.v17i0.2961

Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29(3): 317–332. https://doi.org/10.1515/tlr-2012-0017

CHAPTER 5

Outer particles vs tag particles

A distinction in homophony

Aitor Lizardi Ituarte
University of the Basque Country (UPV/EHU)

This article aims to draw a syntactic analysis that accounts for the differing properties of two sets of discourse particles: Outer (or Final) Particles (OutPs) and Tag Particles (TagPs). In recent years, researchers have built a syntactic model to accommodate OutPs (Haegeman 2014; Wiltschko & Heim 2016; a.o.), but TagPs have received little attention. In order to highlight the different nature of these two sets of particles, I will focus on two segmentally homophonous Basque particles: *ala*$_{OutP}$ and *ala*$_{Tag}$. After discussing their distinct prosodic, syntactic and pragmatic properties, I will argue that these particles show structural differences: while OutPs merge in the right periphery of the clause, TagPs are intransitive X^0s and head their own Speech Act Phrase (SAP).

Keywords: discourse particles, syntax, pragmatics

1. Introduction

In the last years, linguistic research has taken a great step forward in the syntactic and semantic/pragmatic understanding of Discourse Particles (DisP). Descriptions and analyses of different Modal or Inner Particles (InnPs) have expanded our knowledge of DisPs located in the Tense Phrase (TP) field. The progress on the characterization of Final or Outer Particles (OutPs), i.e. DisPs that appear after or before a clause, has also been remarkable.

However, the literature on the characterization of DisPs has often neglected a relevant set of particles: Tag particles (TagPs). Some isolate attempt aside (e.g. Wiltschko & Heim 2016), TagPs have not attracted the interest of researchers working on DisPs, despite their distinctive prosodic, syntactic and pragmatic features.

This article wants to contribute to fill in this gap by proposing a first syntactic analysis that distinguishes OutPs from TagPs. With this in mind, I will analyze two Basque particles, *ala*$_{OutP}$ and *ala*$_{Tag}$, in detail. Based on their different prosodic, syntactic and pragmatic properties, I will conclude that OutPs are transitive Speech Act

https://doi.org/10.1075/la.276.05itu

heads, whereas TagPs are intransitive Speech Act heads; more specifically, OutPs take a complement, the proposition, whereas TagPs do not take any complement.

This article is structured as follows. In the section 2, I present the different types of discourse particles, the motivation for the classification I adopt, and I sketch a general introduction to Basque discourse particles. In Section 3, I discuss in detail the properties of the Basque particles ala_{OutP} and ala_{Tag}. In Section 4, I offer an analysis of TagPs, which relies on the distinction between OutPs and TagPs, previously outlined in Section 3. The main conclusions of this work are summarized in Section 5.

2. Introduction to discourse particles

2.1 Syntactic models for discourse particles: An overview

The term Discourse Particle (DisP) does not have a uniform definition across the literature. Here, I will take DisPs to be elements that "have the function of fitting the propositional content of a sentence to the context of speech by giving an utterance its specific 'shade' (Hartmann 1998: 660) or, alternatively, by imposing restrictions on appropriate contexts for a given utterance" (Zimmermann 2011: 2012–2013). In other words, DisPs are elements that relate a sentence to the discourse, by potentially providing the sentence with contextual background, such as discourse participants' (speaker's and addressee's) epistemic states and the Common Ground of the conversation at the utterance time.

This definition is broad enough to accommodate the two main types of DisPs discussed in the literature: Inner (or Modal) Particles (InnP) and Outer (or Final) Particles (OutP).[1] The description and analysis of the former have a long tradition in Germanic linguistics (Aijmer 1996; Schelfhout *et al.* 2005; Scherf 2017, a.o.), especially in German (Abraham 1991; Coniglio 2008; Bayer & Obenauer 2011; Struckmeier 2014), but similar particles have also been reported in other languages, such as Basque (Haddican 2008; Monforte 2020). Regarding OutPs, the literature is typologically more diverse, ranging from Romance (Munaro & Poletto 2002; Corr 2016) and Germanic (Haegeman 2014) to different East Asian languages like Mandarin (Li 2006; Kuong 2008; Paul & Pan 2017) or Japanese (Kuwabara 2013).

1. I avoid the terms "Modal Particles" and "Final Particles", as they are not accurate enough. The expression "Modal Particle" links them to modal verbs, but they are different both in their distribution and their pragmatic function, so I refer to them as "Inner Particles". Regarding "Final Particles", some of them may appear also in initial position. Each position seems to be related to different pragmatic functions (Haselow 2012), but since both positions have the function of regulating discourse, I will categorize them as "Outer Particles".

The most remarkable difference between both particle sets concerns their distribution. InnPs surface close to the inflected verb (1). In contrast, OutPs are generally located on the left or right edges of the sentence (2):

(1) David ist ja ein Zombie.
David is InnP a Zombie
'David is a Zombie (as you know)'. (German, Gutzmann 2015: 218)

(2) Xiáofú dú-guo zhè běn shū ma?
Xiaofu read-EXP this CL book OutP
'Did Xiaofu read this book?' (Mandarin, Li 2006: 9)

Another difference between InnPs and OutPs is that, in general terms, InnPs may appear in embedded clauses (3); in contrast, OutPs cannot (4):

(3) Tom erinnerte Ulf, [dass es ja Erdbeeren gäbe].
Tom reminded Ulf that it InnP Strawberries give.SBJV
'Tom reminded Ulf that there would be strawberries.'
(Zimmermann 2011: 2023)

(4) *[Ākiū lái ma] méi yǒu guānxi.
Akiu come OutP NEG have relation
'Whether or not Akiu comes doesn't matter.' (Paul & Pan 2017: 19)

Based on these properties, most analyses have proposed that InnPs occupy a high position in the TP layer (Abraham 1991; Bayer & Obenauer 2011; Struckmeier 2014; Scherf 2017). Some properties of InnPs suggest that they interact with the left periphery, which may be challenging for such an analysis: clause type conditions their grammaticality, and some InnPs interact with the illocutionary force of the proposition (Gutzmann 2015). Proposing a left peripheral merging position for InnPs is problematic too, as it cannot account for their distribution. Thus, I consider that Bayer & Obenauer (2011) suggest an appropriate account for this fact, by proposing that InnPs merge in the TP field but are in an Agree relation with the Force head.

Regarding OutPs, most syntactic work follows two general lines of analysis.[2] The earliest works on Romance or Germanic (e.g. Munaro & Poletto 2002) and

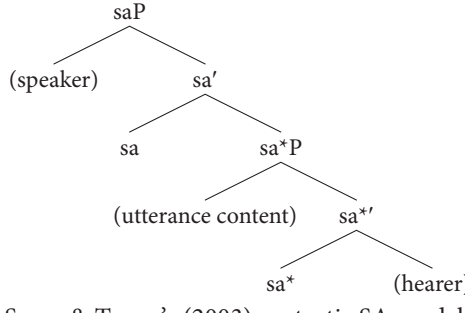
2. The two lines are the split-CP and the neoperformative hypotheses. The difference between them lies on a theoretical level. The former does not commit to a structured and in principle universal syntactic layer for discourse phenomena, which results in different language-specific proposals of sets of phrases, depending on the distribution of the particles. On the contrary, the neoperformative hypothesis departs from the view that discursive phenomena have a dedicated syntactic layer that is universally valid. In fact, they are not contradictory lines of analysis. I thank the anonymous reviewer who pointed out the need for clarifying this point.

the Asian language tradition (Li 2006, Kuong 2008, Kuwabara 2013, Paul & Pan 2017, a.o.) base their analysis on the split-CP hypothesis by Rizzi (1997).[3] Many researchers on OutPs have followed Rizzi's path and have claimed that the left periphery of the clause is even more complex than Rizzi's original proposal. Different analyses have been put forth in the literature based on a variety of OutPs and languages, which have led to various proposals on the specific functional projections involved and the hierarchy these projections give rise to. Even for the same language, Mandarin, and similar particles, Li (2006) and Paul & Pan (2017) offer alternative structures for CP:

(5) a. DiscourseP > DegreeP > ForceP > EvaluativeP > MoodP > FinP (Li2006)
b. AttitudeP > ForceP > LowCP (Paul & Pan 2017)

The second trend of analysis is based on some form of the neo performative hypothesis. Recent papers (Speas & Tenny 2003; Hill 2007a-b; Haegeman 2014; Wiltschko & Heim 2016; Corr 2016, a.o.) have supported the idea that we need a syntactic layer that accommodates those grammatical elements that link the sentence to discourse, i.e. a syntactic speech act layer. Two are the main models proposed. The first approach, by Speas & Tenny (2003), encodes the speech participants (the speaker and the addressee) in syntax (6). Hill (2007a, b) adapts this model to account for sentential adverbs and vocatives:

(6)



Speas & Tenny's (2003) syntactic SA model (adapted)

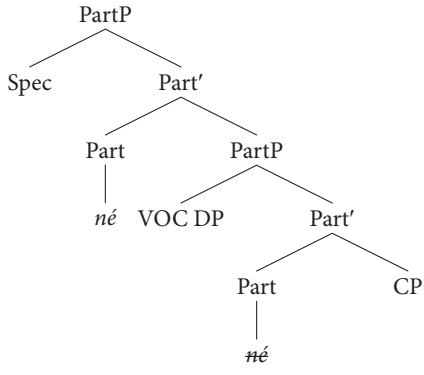
Haegeman (2014) proposes an alternative structure, trying to find a syntactic accommodation for DisPs and vocatives. She notes the different type of contribution and distribution of discourse particles and vocatives, and proposes that the speech act (SA) layer is divided into two 'shells'. The higher shell is 'dynamic and directional' and has the function of relating "the utterance to an addressee as the one for

3. Rizzi (1997) splits up the multifunctional Complementizer Phrase (CP) into four phrases, each with its precise function:

(i) CP = [$_{\text{ForceP}}$ [$_{\text{TopP*}}$ [$_{\text{FocP}}$ [$_{\text{TopP*}}$ [$_{\text{FinP}}$ [$_{\text{TP}}$…]]]]]] (Rizzi 1997: 297)

whom the utterance is intended". The lower shell, on the contrary, is 'stative' and more 'attitudinal' (Haegeman 2014: 135):

(7)



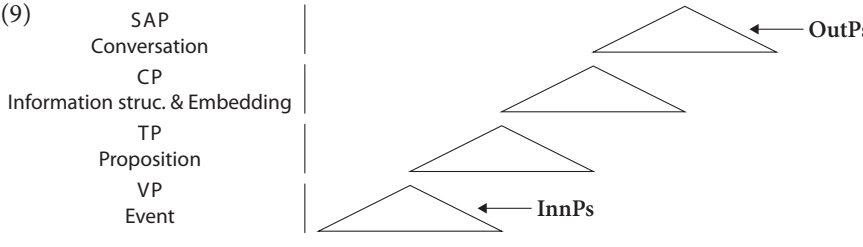
Wiltschko (Wiltschko & Heim 2016; Wiltschko 2017, a.o.) proposes another alternative structure that combines some feature of the previous ones. She splits the speech act layer into an attitudinal lower part, GroundingP, and a higher part, the ResponseP, that has a 'call on addressee' function, but she also considers it necessary to divide each part to account for the speech participants' point of view (8):

(8) [ResPAdd [ResPSp [GroundPAdd [GroundPSp [CP …]]]]] (Wiltschko 2017)

Following the neoperformative model, I assume that sentences are pragmatic objects too, and have a specific syntactic layer that regulates their discursive dimension and accommodates the grammatical elements that encode pragmatic functions. The internal structure of the SA layer goes beyond the scope of this paper.

In this paper I adopt the sentence structure in (9), where I have pointed out the function of each clausal layer and the merging position of InnPs and OutPs:[4]

(9)



4. (9) shows Wiltschko & Heim's (2016) structure with the merging points of different discourse particles. It does not show the specific sentence structure of any language, and therefore I do not commit to any head-final or head-initial configuration for Basque. I thank an anonymous reviewer for pointing out this potential misunderstanding.

However, there is also another particle set that has not received much attention in the literature: Tag particles (TagPs). In a broad sense, TagPs are sentence final expressions that ask the addressee for the confirmation of the sentence content. The two English examples in (10) illustrate two types of TagPs found crosslinguistically, an invariable particle (10a) or a varying reduced clause (10b):

(10) a. You live in London, *right*? / She lives in London, *right*?
b. You live in London, *don't you*? / She has lived in London, *hasn't she*?

In this paper, I will focus on the first type and will use the term TagP to refer to invariable TagPs.

From a syntactic point of view, TagPs have not been analyzed in depth. One of the major syntactic analyses of TagPs is that developed in Wiltschko & Heim (2016), who analyze TagPs (in their terminology, *confirmationals*) and defend that they merge in GroundP. Such an analysis raises some questions on the nature of OutPs and TagPs: should we understand OutPs and TagPs as different particle sets, or are TagPs a subtype of OutPs? If we assume that OutPs and TagPs are different linguistic objects, do they still merge in the same phrase? Should we make a syntactic distinction between them and, if so, how?

Based on prosodic, syntactic and pragmatic evidence, in this article I will argue that OutPs and TagPs are different particle sets, and that this difference has a syntactic reflex.

2.2 General overview of Basque particles

In order to make the difference between OutPs and TagPs clear, I will analyze these particles in Basque in detail. This subsection will provide some basic information on the discourse particle system of this language.

Basque is a relatively free word-order SOV language, but constituents may move for information structure reasons, in which case the verbal complex (the main verb and the auxiliary) must be adjacent to the focalized element, as shown in (11):

(11) a. (Top_1) (Top_n) – XP_{Foc}-*V-T* – (given_1) (given_n)
b. Atzo zurekin pozik ikusi nuen Mikel.
yesterday.TOP you.with.TOP happy.FOC seen AUX Mikel.GIV
'Yesterday, I saw Mikel HAPPY with you.'

Basque has both InnPs and OutPs. Basque InnPs are evidential-epistemic particles that contribute to the propositional content and/or modulate the illocutionary force of the utterance. OutPs act in a different dimension, as they regulate the discursive function of the utterance within the conversation. The most common InnPs

are *omen, ei*, *ote* and *al* (Zubeldia 2010; Etxepare & Uria 2016; Monforte 2020). Among the most used OutPs, we find *ba, e*, *gero* and *ala* (Lizardi Ituarte 2022). In the following examples, I will focus on the hearsay evidential InnP *ei* and the OutP *ba*, which generally marks the utterance as a response to a previous intervention.

I use the following criteria to establish the division between InnPs and OutPs:

i. InnPs appear before the auxiliary and phonologically attached to it,[5] whereas OutPs may only appear on the left or right edge of the sentence (or exceptionally attach to T in sentences with focus) (12):[6]

(12) a. (*Ei) Miren etorri (ei) da (*ei).
InnP Miren come InnP AUX InnP
'Miren has come [reportedly].'

b. (Ba) Miren (*ba) etorri (*ba) da (ba).
OutP Miren OutP come OutP AUX OutP
'[So] Miren has come.'

ii. InnPs may appear in embedded clauses, whereas OutPs cannot:

(13) a. Jon etorriko ei denez, ni ez naiz etorriko.
Jon come.FUT InnP AUX.COMP I NEG AUX come.FUT
'Since Jon will [reportedly] come, I will not come.'

b. Jonek [(*ba) etorriko zarela (*ba)] esan dit.
Jon.ERG OutP come.FUT AUX.COMP OutP say AUX
'Jon has told me that you will come.'

iii. InnPs cannot appear with the response particles *yes* and *no*; OutPs can:

(14) a. (*Ei) bai (*ei)
InnP yes InnP
'Yes [reportedly].'

b. (Ba) bai (ba).
OutP yes OutP
'[So] yes.'

---

5. The particles *omen* and *ote* are InnPs only in central-western varieties. In eastern dialects, they have a freer distribution and may appear with response particles (*Bai ote? / Bai omen*) (Etxepare & Uria 2016; Monforte 2020), which suggests that they merge higher in the structure and that they should be considered another type of particles, probably OutPs. In this case, we are dealing with a clear dialectal division.

6. Phonological processes occur between *ba* and the auxiliary (*ez dakit ba* > *eztakipa* 'I don't know'), which suggests that *ba*, and OutPs by extension, are phonologically attached to the tensed verb.

Basque also has TagPs. In the following example, I will use the TagP *ezta*:

(15) Mikel etorri da, ezta?
Mikel come AUX TagP
'Mikel has come, hasn't he?'

Particles of different sets may combine in a sentence, as long as they are semantically-pragmatically compatible.

In the following sections, I will analyze the differences between OutPs and TagPs in Basque in depth.

3. Appearances are deceptive

Basque has two particles that are homophonous at the segmental level: the OutP *ala* and the TagP *ala*. I will refer to the former as ala_{OutP} and to the latter as ala_{Tag}. They have the same form as the disjunction *ala* 'or' used in alternative questions, as in (16):

(16) Zure bila joango da ala zeure kasa etorriko zara?
your pick.up go.FUT AUX ala_{or} your own come.FUT AUX
'Will (s)he pick you up or will you come on your own?' (Goenaga 2009: 382)

Both of them are mostly used in western and central dialects, but ala_{OutP} is geographically more extended than ala_{Tag}. Each of them conveys a different pragmatic meaning. With ala_{OutP}, the speaker indicates that she has inferred the proposition (either from a previous utterance or from contextual evidence) and expresses surprise and unexpectedness, due to the fact that her inferred expectation and the reality do not agree. On the contrary, when the speaker uses ala_{Tag}, she commits to the proposition with a high degree of certainty, but asks the addressee to confirm it, assuming that the addressee has more direct evidence than herself. (17) provides examples of both uses. The meaning of ala_{OutP} is paraphrased between square brackets:

(17) a. Miren etorri da ala?
Miren come AUX OutP
'[So] Miren has come? [I did not expect it!]'
b. Miren etorri da, ala?
Miren come AUX TagP
'Miren has come, hasn't she?

Ala$_{OutP}$ is only grammatical in interrogative sentences but it is incompatible with interrogative InnPs, such as *al* or *ote*:[7]

(18) a. Miren etorri (*al) da ala?
Miren come InnP AUX OutP
'[So] Miren has come? [I did not expect it!]'
b. Miren etorri (*ote) da ala?
Miren come InnP AUX OutP
'[So] [I wonder] Miren has come? [I did not expect it!]'

The reader may wonder whether we should really regard *ala*$_{OutP}$ and *ala*$_{Tag}$ as two different particles rather than a single one with different functions. In order to answer this question, next I will compare their prosodic, syntactic and pragmatic properties in detail.

3.1 Intonation

Ala$_{OutP}$ and *ala*$_{Tag}$ differ considerably in their phonological properties. Here, I will focus on two aspects: i. the intonation contour, and ii. the phonological integration between the clause and the particle.

Figures 1 and 2 show the intonation contour of utterances with *ala*$_{OutP}$ and *ala*$_{Tag}$ respectively, using the same sentence content: *Mirenek egingo du* 'Miren will do it'.

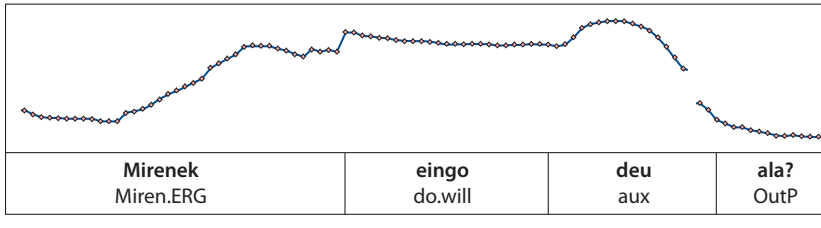


Figure 1. Intonation contour of an *ala*$_{OutP}$ utterance

7. The incompatibility of *ote* with *ala*$_{OutP}$ may be due to semantic-pragmatic reasons. *Ala*$_{OutP}$ expresses an inference, which is incompatible with a conjectural question that *ote* introduces. Regarding *al*, it is unclear why it is incompatible with *ala*$_{OutP}$; further research is needed to clarify this question. However, this incompatibility is surely related to the fact that InnPs are ungrammatical in alternative questions, which use the disjunctive conjunction *ala* (Goenaga 2009).

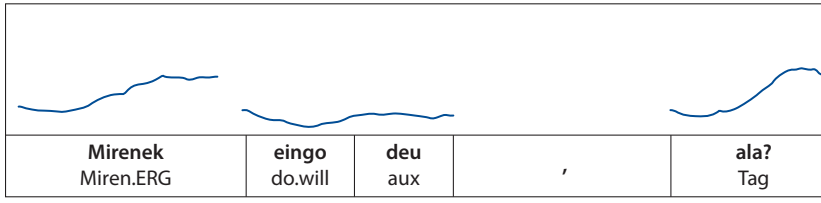


Figure 2. Intonation contour of an *ala*$_{Tag}$ utterance

There is a remarkable difference in the intonation contour of these two utterances. In the *ala*$_{OutP}$ utterance, the intonation starts low, rises on the subject, and keeps high until the auxiliary, where it falls until the end of the sentence. In contrast, in *ala*$_{Tag}$ utterances there is a little initial rise that falls until the particle, where there is an abrupt rise. Additionally, in the *ala*$_{Tag}$ utterance, there may be a pause between the auxiliary and *ala*, represented by a comma in Figure 2, which is not grammatical in the case of *ala*$_{OutP}$ utterances.

Another difference between both utterance types is the phonological integration of the particle with preceding material. In the utterance in Figure 1, if we get rid of *ala*$_{OutP}$ maintaining the intonation contour, the result is an ungrammatical sentence. A Basque regular yes/no question cannot have its intonation peak on the auxiliary, and the final fall cannot be that marked. In other words, an *ala*$_{OutP}$ utterance is not divisible; the particle is intonationally integrated in the clause. In contrast, in an *ala*$_{Tag}$ utterance, if we were to remove *ala*$_{Tag}$ from the material that precedes it, the result would be a regular declarative sentence. Let us compare Figure 2 to Figure 3, where I show a declarative sentence that contains the same sentence content as Figure 2.

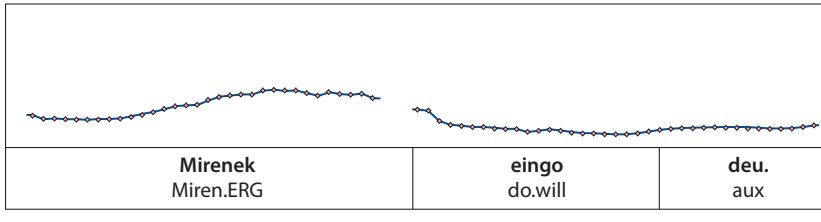


Figure 3. Intonation contour of a declarative sentence

As we can observe, the intonation patterns of declarative sentences and that of the sentence content of *ala*$_{Tag}$ utterances are very similar. In fact, as shown in Figure 2, *ala*$_{Tag}$ utterances seem to be composed by two clearly divisible parts: a declarative part on the sentence content, with a declarative-like intonation contour, and the TagP, which has a rising intonation typical of questions. In addition, the speaker can make a clear pause between the sentence content and the TagP (represented

by a comma in Figure 2), whereas a pause between the sentence content and the OutP leads to ungrammaticality in *ala*$_{OutP}$ utterances. This is strong evidence that we need to distinguish these two particles and their preceding sentence content on phonological grounds.

To sum up, the intonational properties of both types of utterances differ significantly. *Ala*$_{OutP}$ utterances show a question-like pattern, whereas *ala*$_{Tag}$ utterances may be clearly divided into a declarative part in the sentence content and a question part associated with the particle. Moreover, they also differ in the integration of the particle in the clause: in *ala*$_{OutP}$ utterances, the particle is completely integrated within the clause, whilst in the case of *ala*$_{Tag}$ the particle is intonationally separated from the clause.[8]

3.2 Word order and syntax

These two utterance types also differ in their syntactic properties and trigger differences in word order. Here I will mention two of them: i. their ability to license Negative Polarity Items (NPIs), and ii. their ability to license InnPs that are grammatical only in declarative sentences.

Let us first focus on the scope of the interrogative operator, looking at the grammaticality of NPIs. Basque, as well as English, has a set of NPIs: *inor* 'anyone', *inon* 'anywhere', etc. They are not possible within affirmative declarative sentences, but they are licensed in negative sentences, questions and conditional sentences. (19) shows an instance of a question, where the NPI *inork* is licensed:

(19) Inork egin du?
anyone.ERG done AUX
'Has anyone done it?'

If we depart from the idea that *ala*$_{OutP}$ and *ala*$_{Tag}$ utterances ask for the answer of the addressee and that they are therefore questions, we would expect NPIs to be grammatical in both utterance types. Nevertheless, the data reject this prediction:

8. It is interesting to compare the phonological properties of these two *ala* particles with that of alternative questions containing *ala*. In alternative questions, *ala* is more integrated to the second disjunct than to the first one. This can be seen, for instance, in that the intonation peak of the whole utterance occurs at the end of the first disjunct, before *ala* (Goenaga 2009: 391). Compared to the particles analyzed in this article, disjunctive *ala* is less integrated to the previous sentence, which makes it closer to *ala*$_{Tag}$ utterances, and the intonation contour of alternative questions is closer to *ala*$_{OutP}$ utterances. In any case, the relation between the disjunctive *ala* and the particles *ala*$_{OutP}$ and *ala*$_{Tag}$ deserves further research.

(20) a. Inork egin du ala?
anyone.ERG done AUX OutP
'Has anyone done it [I infer]?'
b. *Inork egin du, ala?
anyone.ERG done AUX TagP
'Anyone has done it, hasn't she?'

(20) demonstrates that NPIs are possible in *ala*$_{OutP}$ utterances, but they result in ungrammaticality in *ala*$_{Tag}$ questions. These results suggest that the scope of the question operator is sentential in *ala*$_{OutP}$ utterances, whereas in sentences with *ala*$_{Tag}$ it scopes over the TagP, but does not scope over the sentence content.

This conclusion is fully consistent with the intonational properties presented in subsection 3.2. From Figure 2, I concluded that the intonation contour of TagP questions is divisible into a declarative-type part on the sentence content and a final rising on the particle typical of questions. The behaviour of NPIs supports this approach: NPIs are ungrammatical when there is no interrogative intonation, which suggests that the part of the utterance that precedes *ala*$_{Tag}$ utterances is not a real question, but a declarative.

Let us now move on to the compatibility of *ala*$_{OutP}$ and *ala*$_{Tag}$ with evidential InnPs.[9] The evidential InnPs *omen* and *ei* are grammatical in declarative sentences but not in interrogative sentences, as (21) shows. When we try to add these evidential particles to the sentences under analysis, we observe that they are ungrammatical in *ala*$_{OutP}$ utterances (22a), whereas they are compatible with *ala*$_{Tag}$ (22b):

(21) a. Miren etorri ei da.
Miren come InnP AUX
'[Reportedly,] Miren has come.'
b. *Miren etorri ei da?
Miren come InnP AUX
'[Reportedly,] has Miren come?'

(22) a. *Miren etorri ei da ala?
Miren come InnP AUX OutP
'[Reportedly,] has Miren come [I infer]?' (intended)
b. Miren etorri ei da, ala?
Miren come InnP AUX TagP
'[Reportedly,] Miren has come, hasn't she?'

To recap, the different behaviour in licensing NPIs and evidential InnPs fits very well with the phonological results. *Ala*$_{OutP}$ has a stronger integration in the clause than *ala*$_{Tag}$. All these properties lead us to conclude that *ala*$_{OutP}$ utterances behave as interrogatives, while *ala*$_{Tag}$ utterances have the properties of declarative sentences.

9. I would like to thank Ricardo Etxepare for bringing this paradigm to my attention.

3.3 Pragmatics

In addition to phonology and syntax, these two *ala* utterances also differ in their pragmatic characterization. In this subsection I will make a first formal characterization of ala_{OutP} and ala_{Tag} utterances. For that purpose, I will use Malamud & Stephenson's (2015) framework, who among other things address a pragmatic analysis of English tag questions.

Malamud & Stephenson (2015) base their proposal on Farkas & Bruce's (2009) dialogical model. Farkas & Bruce (2009) defend that we need to define at least the following discursive notions to formally characterize discourse.

a. The Speech Participants. Speaker (S) and addressee (A).
b. The Common Ground (CG). The set of propositions that all speech participants are committed to. s_x represents the CG state in a given point of the conversation.
c. The participants' Discourse Commitments (DC). The set of propositions that each speech participant is committed to.
d. The Table (T). When a speech participant makes a conversational move, she places the sentence on the conversational 'table'. The table records what is 'at issue' in the conversation (Farkas & Bruce 2009: 87) and, consequently, what must be solved. The elements in the table form a 'stack'.
e. The Projected CG (CG*). Each conversational move projects a set of future CGs in which the issue is decided. An assertion projects a future CG that contains *p*, and a question projects a future CG in which the question is answered.
f. Projected DCs (DC_X*), for each participant (Malamud & Stephenson 2015).

Let us start formalising the impact of assertions in discourse.[10] When S asserts *Mirenek egingo du* 'Miren will do it', she makes the following discursive movements: (i) she adds *p* to DC_S*; (ii) she adds *p* to DC_S; (iii) she places the sentence on T; and (iv) she projects a single future CG in which *p* is added to the CG. This is shown in (23):

(23) Effects of an assertion in discourse
  i. $DC_{S,o}* = DC_{S,i}* \bigcup \{p\}$[11]
  ii. $DC_{S,o} = DC_{S,i} \bigcup \{p\}$
  iii. $T_o = push(\langle$'*Mirenek egingo du*'[decl.];$\{p\}\rangle, T_i)$[12]
  iv. $CG_o* = CG_i* \bigcup \{p\}$

10. We will assume that T is empty before the utterance time.

11. Subscripts *i* and *o* represent *input* and *output*, respectively.

12. *Push* represents the new stack obtained by adding item *e* to the top of the stack *T* (Farkas & Bruce 2009: 90).

Now there is an item on the table. The next step is to empty T: A must decide whether to accept *p* and add it to his DC_A, and, consequently, to the CG, or to go against the CG* and reject it, which would mean that *p* would only belong to DC_S but neither to DC_A nor to the CG.

The effect of polar questions in discourse is slightly different. Uttering the polar question *Mirenek egingo du?* 'Will Miren do it?' brings the following two moves into the conversation: (i) she places the sentence in the CG*; (ii) she projects a DC_S* in which she commits to a possible answer {*p*, ¬*p*}; and (iii) S projects a CG* that contains an answer to the question. This is illustrated in (24):

(24) Effects of a polar question in discourse
   i. $T_o = push(\langle$'*Mirenek egingo du*'[int.];$\{p\}\rangle, T_i)$
   ii. $DC_{S,o}{}^* = DC_{S,i}{}^* \bigcup \{p, \neg p\}$
   iii. $CG_o{}^* = CG_i{}^* \bigcup \{p, \neg p\}$

In the context stage that we get after the move in (24), the addressee must solve the item on T by committing to *p* or to ¬*p*.

Now, let us move to characterize the different pragmatic import of these *ala* utterances formally within this framework. At the beginning of Section 3, I mentioned that ala_{OutP} conveys expressive content such as surprise and unexpectedness, and that it expresses the idea that S inferred the sentence content from A or by contextual evidence. This expressive dimension cannot be directly accounted for in this model, because in this system the only possible status for propositions is projected or effective commitment. Commitment implies a public compromise that S accepts the truth of *p*, so that she cannot contradict *p* later on in the conversation. Notions such as implication or expressiveness do not fit into that definition. In order to accommodate such effects, I propose that the schema by Farkas & Bruce (2009) and Malamud & Stephenson (2015) must be expanded. In addition to DC and DC*, I propose another category, the Doxastic State (DS), defined in (25):

(25) *Doxastic State (DS)*: discourse participants' knowledge at the current situational context and conversational exchange.

The DS regulates the epistemic background of each participant and contains both the items in a participant's DC and items towards which the participant is not publicly committed. With the addition of this category, we can now account for the contribution of ala_{OutP}.

Let us account for the pragmatic contribution of the following utterance: *Mirenek egingo du ala?* 'Will Miren do it ala_{OutP}?'. First of all, this sentence cannot be uttered out of the blue. It needs a previous background in which the previous

utterance (q),[13] i.e. the topmost element of the stack on T, entails the denotation of p, i.e. that Miren will do it; and, additionally, p must have been added to DS_S. When S utters p in that input context, she does the following operations: (i) she places the utterance p on the top of the stack on T; (ii) S projects a future CG that will contain the answer of the question, $\{p, \neg p\}$; (iii) DS_A is updated with p; (iv) DS_A is updated with the item 'S has inferred p from q'; and (v) DS_A gets updated with the item 'S is surprised by p'. This is formally summarized in (26):

(26) Effects of an ala_{OutP} utterance
Input context
i. $top(T_i) = (\{q \,/\, q \vDash p\})$
ii. $DS_{S,o} = DS_{S,i} \bigcup \{p\}$
Output context
i. $T_o = push(\{p, \neg p\}, T_i)$
ii. $CG^{\star} = \{s_i \bigcup \{p\} \vee s_i \bigcup \{\neg p\}\}$
iii. $DS_{A,o} = DS_{A,i} \bigcup \{p\}$
iv. $DS_{A,o} = DS_{A,i} \bigcup$ (S infers p from q)
v. $DS_{A,o} = DS_A \bigcup$ (S is surprised and unexpected about p)

Essentially, (26) represents the same speech act that we have in regular polar questions, with slight modifications: a previous background is necessary, it projects a future CG that includes either p or $\neg p$, and it adds p and expressive meaning to DS_A. Thus, ala_{OutP}'s contribution is to be understood as a modification of a questioning speech act.

The case of ala_{Tag} utterances is different. They produce the same effect as reverse polarity tags, as described in Malamud & Stephenson (2015). Unlike in ala_{OutP} utterances, in this case S tries to commit to the proposition. The difference with plain declaratives is that, in ala_{Tag} utterances, S accepts A as a higher authority to decide on p's validity and that S's commitment depends on A's answer, i.e. S adds p to $DC_S^{\star}$. In all other respects, an ala_{Tag} utterance has the effect of a declarative in discourse. Thus, a sentence such as *Mirenek egingo du, ala?* 'Miren will do it, won't she?' has the following effect on discourse: i. S projects that she will commit to p; ii. S puts the sentence and its denotation on T, and iii. S adds p to $CG^{\star}$. These operations are summarized in (27):

---

13. In cases where the inference refers to contextual evidence, the only thing that would change is that q would not represent an utterance, but it would represent some abstract information received by the speaker through contextual evidence.

(27) Effects of an ala_{Tag} utterance
  i. $DC_{S,o} = DC_{S,i} \bigcup \{p\}$
  ii. $T_o = push(\{p\}, T_i)$
  iii. $CG^{\star} = \{s_i \bigcup \{p\}\}$

The main difference between ala_{OutP} and ala_{Tag} utterances is that ala_{OutP} utterances modify an interrogative sentence adding expressive content to participants' DS, whereas ala_{Tag} utterances merge declarative and interrogative properties into a single utterance.

The pragmatic analysis of ala_{OutP} and ala_{Tag} utterances shows that they have different effects on discourse. The most remarkable feature is that in the case of the former, ala_{OutP} is responsible for modifying a questioning speech act, whereas the latter is a merger of asserting and questioning speech acts.

3.4 Interim summary

In this section, I have compared the intonational, syntactic and pragmatic properties of *ala* utterances discussed in the paper. The data show that the particle ala_{OutP} is intonationally more integrated to the sentence content, appears in sentences that are syntactically questions, and pragmatically modifies the sentence content. By contrast, ala_{Tag} utterances are intonationally divided into a declarative sentence and a questioned particle, the particle is less integrated to the sentence content, the sentence content has syntactic properties of declaratives clauses, and these utterances are pragmatically combinations of declaratives and questions. In Table 1, I summarize the properties discussed so far.

Table 1. Differences between ala_{OutP} and ala_{Tag} utterances

ala	**OutP**	**TagP**
Intonation	Integrated.	Possible break. Divisible declarative and interrogative parts.
Illocutionary force	Question.	Assertion & question.
Meaning	Modify *p*. Expressive.	Weakly assert *p* / ask for {*p*}.

Ala_{OutP} and ala_{Tag} differ in their prosodic, syntactic and pragmatic properties. In other words, they have no common properties in any linguistic aspect. Therefore, not an isolate differing property, but the combination of many leads me to consider ala_{OutP} and ala_{Tag} as two different particles.

In this section, I have demonstrated the different nature of these two *ala* utterances. These data lead us to think that these differences are not only due to the lexical contribution of each particle, but are also rooted in two underlying syntactic

structures that combine the sentence content and the particle in a different way. In the next section, I will propose distinct syntactic analyses for these different structures.

4. Accommodating tagPs in syntax

In Section 2, I presented different syntactic models that accommodate discourse particles in syntax. I assume that InnPs merge in a position in the TP field, while OutPs are located in the SA field. In Section 3, I showed that *ala*$_{OutP}$ and *ala*$_{Tag}$ utterances differ in phonological, syntactic and pragmatic aspects. I consider these differences a consequence of different syntactic structures. In this section, I will propose a syntactic analysis for tag questions, which builds on the hypothesis that they differ structurally from regular questions with OutPs.

With that goal, I will first discuss the contribution and limits of Wiltschko & Heim's analysis of TagPs, the only work that tries to account syntactically for this particle type. Next, I will present my proposal, which attempts to address the shortcomings of previous analyses. Finally, I will present two implications of this analysis that reinforce my proposal.

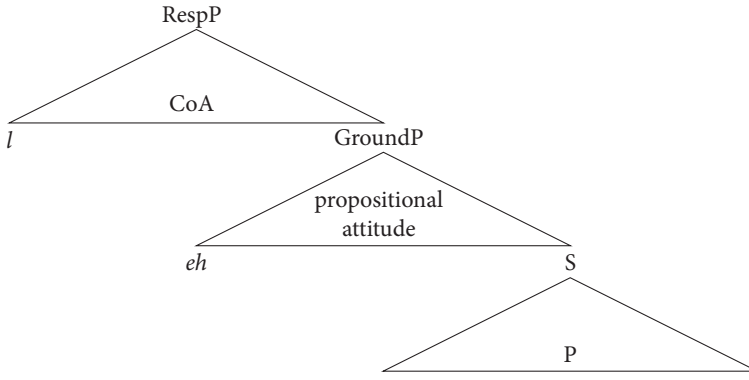
4.1 Previous accounts of the syntax of tagPs

From the data presented in Section 3, I conclude that a syntactic analysis of TagPs must take two aspects into account. First, it must be a model that allows integrating discourse related phenomena in syntax. This means adopting one of the different neoperformative models that have been summarized in Section 2. Second, it must consider the prosodic, syntactic, and pragmatic differences between OutPs and TagPs seen in Section 3.

As far as I know, the only syntactic analysis of TagPs so far has been the one presented by Wiltschcko & Heim (2016), who specifically analyze the syntax of the TagP (aka *confirmational*) *eh*, from Canadian English.

Regarding the first point, they do assume a conversational layer in syntax for discursive elements like particles. According to their model, the conversational layer is structured as a lower layer that regulates the grounding process of the conversation, "where S's propositional attitude is encoded" (Wiltschcko & Heim 2016: 328), and a higher response layer, which encodes *call on addressee*, i.e. what S wants A to do with the proposition. Based on this structure, they argue that *eh* merges on the Grounding layer, whereas the rising contour on the particle is associated with the Response layer:

(28)



However, this proposal does not account for the second point that every analysis of TagPs should take into consideration, i.e. the distinction between OutPs and TagPs. Tag utterances are not speech modifiers like OutPs, but they are combinations of assertions and questions. Therefore, TagPs have a lighter syntactic integration than OutPs with the sentence content, which means that OutPs and TagPs cannot occupy the same syntactic phrase. If OutPs merge on GroundP, TagPs should in any case merge in a higher position. However, there is only another projection above GroundP, namely ResponseP. ResponseP regulates the speech act of the sentence, marked by the intonation, but that is not the function of TagPs. Consequently, there is no additional projection to accommodate TagPs, and the structure proposed by Wiltschko & Heim (2016) is too small to account for a different analysis of OutPs and Tags. Moreover, integrating TagPs below a single ResponseP raises an additional question: if tag utterances are combinations of assertions and questions, how can we account for the assertion part of the sentence?

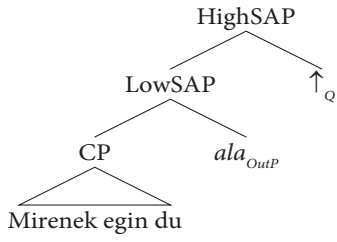
4.2 Proposal

In a footnote, Wiltschko & Heim (2016) point out that a reviewer asked them whether *confirmationals* (aka TagPs) may be analyzed as bi-clausal structures, as they involve the combination of two speech acts: "it is however less economic than the present proposal in that it would require an obligatory process of deleting the complement of one of the operators because only one of the two p-structures […] can be spelled out" (Wiltschko & Heim 2016: 328). While this is a legitimate answer, I would like to argue that economy cannot be a pretext to avoid the complexity of a phenomenon. Their proposal does not account for the double speech act nature of tag questions and omits the distinction between OutPs and TagPs. In order to avoid these shortcomings, next I will present an alternative proposal that represents these structures in all their complexity.

Before I present my analysis, I will first briefly introduce the SA structure that I adopt. For the sake of simplicity, here I will assume a bi-phrasal SA domain in syntax, in line with Haegeman (2014) and Corr (2016). These works agree that the lower phrase is linked to the sentence content and relates the speech participants to the proposition, whereas the higher layer is 'directional', relating the utterance to the addressee.

Once we adopt this model, the position of *ala*$_{OutP}$ is clear: it merges on the lower SA head and the intonation, responsible for the inquisitive nature of the clause, and is associated with the high SA phrase:

(29)

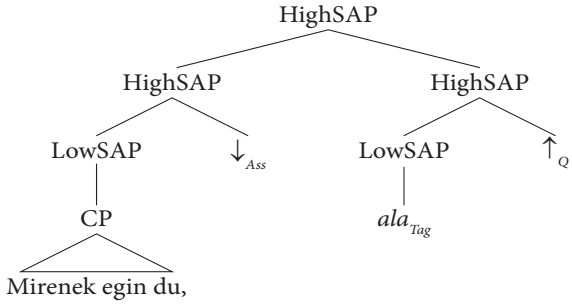


In (29),[14] we can see that *ala*$_{OutP}$ is a transitive SA phrase that takes the sentence content as its complement. Thus, *ala*$_{OutP}$ is part of the clause, it heads the LowSAP and modifies the unique speech act of the clause, i.e. questioning, reached by the questioning intonation, merged in the head of HighSAP.

With respect to tag utterances, I propose that they are complex SAPs that result from the combination of a declarative SAP, the sentence content, and a questioning SAP, the TagP. TagPs merge on the lower SAP, they do not take any complement, i.e. they are intransitive heads, and are dominated by the HighSAP head, which triggers the question intonation. The sentence content has its own SAP, with assertion intonation. Like other functional intransitive heads, TagPs are anaphoric elements, which take the sentence content as their referent. Finally, both SAP structures are adjoined, forming a more complex SA object that includes the whole utterance. (30) illustrates this structure:

14. An anonymous reviewer has warned that this structure goes against FOFC. It would be true if I adopted a specific sentence structure model for Basque, i.e. Ortiz de Urbina (1999), but I have not done so. Moreover, I agree with Paul & Pan (2017) when they claim that discourse particles "are likened to 'extra-metrical' elements in phonology, i.e. elements not counting for rules, in this case the FOFC." This follows from the fact that there is a great tendency, both in OV and VO languages, for discourse particles to be sentence final.

(30)



The attachment of both SAPs into a greater SAP is necessary, because the TagP is not usually an independent element. Notice that there is no possibility of merging both SAPs via coordination:

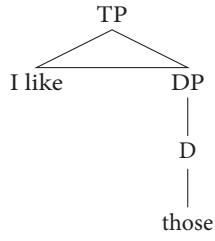
(31) Mirenek egingo du, (*eta / *edo / *baina) ala?
Miren.ERG do.FUT AUX and or but TagP
'Miren will do it, (*and/or/but) won't she?'

This analysis solves the problems posited in Section 2, as well as the shortcomings of Wiltschko & Heim's (2016) proposal. Firstly, it marks a distinction between OutPs and TagPs: the former are part of the main SA phrase, the latter form their own SA phrase. Secondly, it accounts for the double speech act nature of tag questions, since tag questions have two connected speech acts. Finally, this proposal has the advantage of avoiding the deletion of a second questioned sentence as Wiltschko & Heim's (2016) envision as alternative, because the TagP is an anaphoric element that takes the sentence content as its referent.

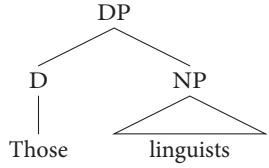
TagPs are intransitive SA heads, similar to demonstratives and personal pronouns, which are considered intransitive D heads (Abney 1987), and the polarity particles *yes* and *no* (Krifka 2013; Holmberg 2015). All these elements are anaphoric, in the same way that TagPs are.

Regarding the Determiner Phrase (DP) domain, Abney (1987) argued that articles, demonstratives and personal pronouns are D heads. Following Abney (1987), articles are transitive heads that take the NP as their complement, personal pronouns are intransitive D heads, and demonstratives may be transitive or intransitive. In (32a), *those* is an intransitive demonstrative, in (32b) *those* is a transitive demonstrative, and in (32c), *we* is an intransitive personal pronoun. In all three cases, the D heads *those* and *we* take their referent from discourse, i.e. they are anaphoric:

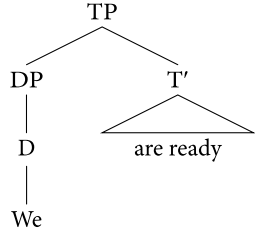
(32) a.



b.



c.



Abney (1987: 180, adapted)

Let us now focus on demonstratives and personal pronouns, since they may be intransitive heads and need a referent. Personal pronouns take their referent from the situational context, as their function is to refer to speech participants, and intransitive demonstratives reach the referent from a previous DP,[15] as they are anaphoric elements.

Polarity particles such as *yes* and *no* have been proposed to be another example of intransitive functional heads. Recent research considers them propositional anaphors (Krifka 2013), because they need a previous propositional referent to confirm or reverse its polarity. There are two main approaches to these elements. Krifka (2013) considers German *ja* 'yes' and *nein* 'no' intransitive TP heads. Holmberg (2015) considers that *yes* and *no* are specifiers of FocP that agree with the Pol head of the clause, in which the propositional content is present but phonologically omitted.[16] In both cases, *yes* and *no* (and their language specific counterparts) are anaphors: they need a previous utterance to refer to.

---

15. They can also refer to a non-verbal contextual referent, but, in any case, they need a contextual referent to function in the conversation.

16. This analysis corresponds to polarity answering systems like English or Basque. For the sake of simplicity, we will not address here the precise analysis of truth based answering systems, as it is not an important issue for the purpose of this article.

I propose that TagPs are comparable to elements like intransitive determiners and polarity particles. They are anaphoric intransitive Speech Act heads, they refer anaphorically to a previous speech act.

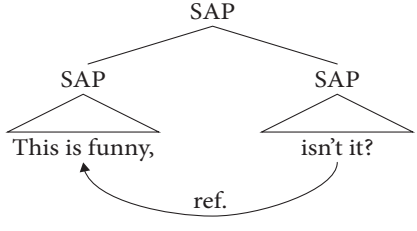
The previous discussion on these particles leads me to defend that, although segmentally homophonous, *ala*$_{OutP}$ and *ala*$_{Tag}$ are different particles, they might have emerged from different grammaticalization paths of the disjunctive conjunction *ala* 'or', which is a very common source for interrogative particles[17] (Heine & Kuteva 2003; Bailey 2015). Both *ala* particles present alternatives: *ala*$_{OutP}$ presents alternatives to the inferred proposition; with *ala*$_{Tag}$ S asserts *p* and asks A to confirm *p* or to present an alternative.

The idea of different grammaticalization paths is supported by the different meanings of both utterances. The meaning difference between *ala*$_{OutP}$ and *ala*$_{Tag}$ cannot only be explained by different syntactic structures. If *ala*$_{OutP}$ and *ala*$_{Tag}$ were the same element differing only in their syntax, we would expect that *ala*$_{Tag}$ would also convey expressive meaning like surprise and inference, but this is not the case. Therefore, I consider that we need two lexically different *ala*s.

Before I finish this section, I would like to discuss some further implications of this analysis that support my proposal.

The first implication is the following: if an intransitive SAP may attach to another SAP that contains the sentence content, we could expect a similar possibility to exist also for other structures. I argue that tags like English 'isn't it' should be analyzed precisely along these lines. These tags are SAPs with a richer internal structure, consisting of two SAPs, the second of which attaches to the first one, as in (33). This first SAP corresponds to the main sentence.

(33)



The second implication is related to the attachment process of the particle. It is true that TagPs are usually attached to the main sentence, but they may also appear alone,[18] as long as the referent sentence is inferable from discourse. In (34), I present some literary examples of the Basque TagP *ezta?* without its referent sentence:

17. According an anonimous reviewer, the grammaticalization source of the disjunctive conjunction *ala* might be the demonstrative *hala* 'so, like that'.

18. An anonymous reviewer asks whether all TagPs may appear on their own. According to my intuitions, all Basque TagPs may appear alone.

(34) a. Ezta, Jaime?
TagP Jaime
'Isn't it, Jaime?' (Mintegi 2002: 131)
b. Ezta? - esan zuen Harryk.
TagP said AUX Harry.ERG
'Isn't it? – said Harry.' (Rowling 2002: 244)
c. Ezta? - isiltasun sakona.
TagP silence deep
'Isn't it? – deep silence.' (Morillo 2003: 30)

The first implication shows that the attachment of different SAs is also present in other more complex structures. I refer to structures such as (35):

(35) Miren etorri da, ez da hala?
Miren come AUX NEG AUX so
'Miren has come, isn't it so?'

In (35), *ez da hala* is not a TagP, but a whole SAP consisting of the negation *ez*, the verb *da* 'is' and the adverb *hala* 'so', and it is attached to its main SAP, *Miren etorri da* 'Miren has come'. The second implication demonstrates that TagPs are not necessarily attached to their referent, which reinforces the idea that they are intransitive heads.

5. Conclusions

In this article, I presented a proposal for the syntactic accommodation of TagPs. I compared two segmentally homophonous Basque particles: ala_{OutP} and ala_{Tag}. Each particle involves different intonational, syntactic and pragmatic properties, which affect the syntactic translation of these utterance types. This difference raises the issue of the necessity of a syntactic distinction between OutPs and TagPs, an issue that no previous work has addressed. I have shown that ala_{OutP} is phonologically and syntactically more integrated in the clause than ala_{Tag}, and that ala_{OutP} modifies the pragmatic component of the questioning speech act, whereas ala_{Tag} utterances are in fact a combination of two different speech acts, i.e. asserting the proposition and asking for its confirmation.

Contra Wiltschko & Heim's (2016), who do not distinguish between OutPs and TagPs, I have made a syntactic proposal that accounts for this difference. I have suggested that OutPs are transitive SA heads that take the sentence content as a complement and modify their speech act, in line with the analyses of OutPs in other languages (Haegeman 2014; Corr 2016). On the other hand, TagPs are intransitive SA heads that form their own SA domain. The TagP may be attached

to the main sentence, but this is not a necessary condition, as the referent sentence may be understood from the context.

This proposal accounts in a better way for the properties of both OutPs and TagPs, explains their intonational, syntactic and pragmatic differences, and presents a framework in which every particle type can be accommodated in the clausal syntax.

Funding

This research has been partially supported by the project VASTRUD PGC2018-096870-B-I00), funded by the Ministry of Science, Innovation and Universities (MiCIU) / Spanish Research Agency (AEI) and the European Regional Development Fund (FEDER, UE), the HiTT Research group IT1396-19 (Basque Government), the HiTT research group GIU18/221 (University of the Basque Country, UPV/EHU), and the project BIM ANR-17-0011-BIM (Agence National de la Recherche, French Government).

References

Abney, Steven. 1987. The English Noun Phrase in its Sentential Aspect. PhD dissertation, MIT.

Abraham, Werner. 1991. Discourse particles in German: How does their illocutive force come about. In *Discourse particles: descriptive and theoretical investigations on the logical, syntactic, and pragmatic properties of discourse particles in German*, Werner Abraham (ed), 203–252. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.12.08abr

Aijmer, Karin. 1996. Swedish modal particles in a contrastive perspective. *Language Sciences* 18(1–2): 393–427. https://doi.org/10.1016/0388-0001(96)00027-7

Bailey, Laura. 2015. Word order and the syntax of question particles. In *Final Particles*, Sylvie Hancil, Alexander Haselow & Margje Post (eds), 406–427. Berlin: De Gruyter. https://doi.org/10.1515/9783110375572-017

Bayer, Joseph & 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

Coniglio, Marco. 2008. Modal particles in Italian. *Working Papers in Linguistics* 18: 91–129.

Corr, Alice. 2016. Ibero-Romance and the syntax of the utterance. PhD dissertation, University of Cambridge.

Etxepare, Ricardo & Uria, Larraitz. 2016. Microsyntactic variation in the Basque hearsay evidential. In *Microparameters in the Grammar of Basque*, Beatriz Fernández & Jon Ortiz de Urbina (eds), 265–288. Amsterdam: John Benjamins. https://doi.org/10.1075/lfab.13.10etx

Farkas, Donka & Bruce, Kim. 2009. On reacting to assertions and polar questions. *Journal of Semantics* 27: 81–118. https://doi.org/10.1093/jos/ffp010

Goenaga, Patxi. 2009. Hauta-galderez. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"-International Journal of Basque linguistics and Philology* 43: 381–404.

Gutzmann, Daniel. 2015. *Use Conditional Meaning*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198723820.001.0001

Haddican, Bill. 2008. Euskal perpausaren oinarrizko espez-buru-osagarri hurrenkeraren aldeko argudio batzuk. In *Antisimetriaren hipotesia vs. buru parametroa: euskararen oinarrizko hurrenkera ezbaian*, Iñigo Arteatx, Xabier Artiagoitia & Arantzazu Elordieta (eds), 65–92. Bilbo: UPV/EHUren Argitalpen Zerbitzua.

Haegeman, Liliane. 2014. West Flemish verb-based discourse particles and the articulation of the Speech Act layer. *Studia Linguistica* 68(1): 116–139. https://doi.org/10.1111/stul.12023

Hartmann, Dietrich. 1998. Particles. In *Concise Encyclopedia of Pragmatics*, Jacob L. Mey (ed), 657–663. Amsterdam: Elsevier.

Haselow, Alexander. 2012. Subjectivity, intersubjectivity and the negotiation of the common ground in spoken discourse: Final particles in English. *Language & Communication* 32(3): 182–204. https://doi.org/10.1016/j.langcom.2012.04.008

Heine, Bernd & Kuteva, Tania. 2003. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Hill, Virginia. 2007a. Vocatives and the pragmatics-syntax interface. *Lingua* 117(12): 2077–2105. https://doi.org/10.1016/j.lingua.2007.01.002

Hill, Virginia. 2007b. Romanian adverbs and the pragmatic field. *The Linguistic Review* 24 (1): 61–86. https://doi.org/10.1515/TLR.2007.003

Holmberg, Anders. 2015. *The syntax of yes and no*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198701859.001.0001

Krifka, Manfred. 2013. Response particles as propositional anaphors. In *Proceedings of Semantics and Linguistic Theory (SALT)* 23, Todd Snider (ed), 1–18. https://doi.org/10.3765/salt.v23i0.2676

Kuong, Io-Kei Joaquim. 2008. Yes/no question particles revisited: The grammatical functions of *mo4*, *me1*, and *maa3*. In *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, Marjorie K. M. Chan and Hana Kang (eds), 715–733. Columbus: The Ohio State University.

Kuwabara, Kazuki. 2013. Peripheral effects in Japanese questions and the fine structure of CP. *Lingua* 126: 92–119. https://doi.org/10.1016/j.lingua.2012.12.004

Li, Boya. 2006. Chinese Final Particles and the Syntax of Left Periphery. PhD dissertation, Leiden University.

Lizardi Ituarte, Aitor. 2022. On intonationally integrated complex speech acts: the Basque discourse particle *ba* in interrogatives. In *Proceedings of Chicago Linguistic Society* 56. Chicago: The University of Chicago.

Malamud, Sophia & Stephenson, Tamina. 2015. Three ways to avoid commitments: declarative force in the conversational scoreboard. *Journal of Semantics* 32(2): 275–311. https://doi.org/10.1093/jos/ffu002

Monforte, Sergio. 2020. Galderetako *-a*, *al* eta *ote* partikulak euskaraz: sintaxia, mikroaldakortasuna eta interpretazioa. PhD dissertation, University of the Basque Country UPV/EHU.

Munaro, Nicola & Poletto, Cecilia. 2002. Ways of clausal typing. *Rivista di Grammatica Generativa* 27: 87–105.

Ortiz de Urbina, Jon. 1999. Focus Phrases, Force Phrases and Left Heads in Basque. In *Grammatical Analyses in Basque and Romance Linguistics Papers in honor of Mario Saltarelli*, Jon A. Franco, Alazne Landa & Juan Martín (eds), 179–194. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.187.11urb

Paul, Waltraud & Pan Victor J. 2017. What you see is what you get: Chinese sentence-final particles as head-final complementisers. In *Discourse particles. Formal approaches to their syntax and semantics*, Joseph Bayer & Volker Struckmeier (eds), 49–77. Berlin: De Gruyter.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, Liliane Haegeman (ed), 281–337. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7

Schelfhout, Carla, Coppen, Peter-Arno, Oostdijk, Nelleke & van der Silk, Frans. 2005. Interjections in Dutch: a corpus-based approach. *Nederlandistiek.nl*. 05.05.

Scherf, Nathalie. 2017. The syntax of Swedish modal particles. In *Discourse Particles: Formal Approaches to their Syntax and Semantics*, Joseph Bayer & Volker Struckmeier (eds), 78–99. Berlin: De Gruyter.

Speas, Peggy & Tenny, Carol. 2003. Configurational properties of point of view roles. In *Asymmetry in Grammar*, Anna Maria Di Sciullo (ed), 315–344. Amsterdam: John Benjamins. https://doi.org/10.1075/la.57.15spe

Struckmeier, Volker. 2014. Ja doch wohl C? Modal Particles in German as C-related elements. *Studia Lingüística* 68(1): 16–48. https://doi.org/10.1111/stul.12019

Wiltschko, Martina & Heim, Johannes. 2016. The syntax of confirmationals. A neo-performative analysis. In *Outside the Clause. Form and Function of Extra-Clausal Constituent*, Gunther Kaltenbock, Evelien Keizer and Arne Lohmann (eds), 303–340. Amsterdam: John Benjamins.

Wiltschko, Martina. 2017. Ergative constellations in the structure of Speech Acts. In *The Oxford Handbook of Ergativity*, Jessica Coon, Diane Massam & Lisa De Mena Travis (eds), 419–446. Oxford: Oxford University Press.

Zimmermann, Malte. 2011. Discourse particles. In *Semantics: An International Handbook of Natural Language Meaning*, Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds), 2011–2038. Berlin: De Gruyter.

Zubeldia, Larraitz. 2010. *Omen* partikularen azterketa semantikoa eta pragmatikoa. PhD dissertation, University of the Basque Country UPV/EHU.

Corpus references

Mintegi, Migel Anjel. 2002. *Denak du bere prezioa*. Bilbao: Ibaizabal.

Morillo, Fernando 2003. *Ortzadarra sutan*. Donostia: Elkar.

Rowling, Joanne K. 2002. *Harry Potter eta Azkabango presoa*. Translated by Iñaki Mendiguren. Donostia: Elkarlanean.

CHAPTER 6

Anchoring primary and secondary interjections to the context

Nicola Munaro
Ca' Foscari University of Venice

On the basis of empirical evidence from various Italian dialects, I argue that primary and secondary interjections lexicalize different functional heads which are computed syntactically at the edge of the clause. Secondary interjections should be clearly distinguished from primary ones; only secondary interjections lexicalizing a SpeechAct° head represent autonomous speech acts and are prosodically and syntactically independent from the co-occurring clause, which they can attract to their specifier position, raising eventually to the adjacent head Speaker° in order to achieve the necessary spatio-temporal contextual anchoring. Primary interjections, which can co-occur with secondary ones and surface clause-initially, lexicalize arguably the highest functional head Speaker°, interacting in interesting ways with lower projections and with the overt realization of the complementizer in Force.

Keywords: interjection, speech act, Force, contextual anchoring

1. Introduction

Adopting a cartographic approach to the functional layout of the left periphery, in this article I argue for the necessity of a syntactic encoding of information pertaining to the interface between utterance and discourse within the highest layer of clause structure, above what is usually referred to as the C-domain.[1]

---

1. A preliminary version of this material was presented at the *International Workshop on Discourse Particles* organized by the University of the Basque Country (UPV/EHU) and held online on October 8th 2020. I wish to thank the participants of this event, two anonymous reviewers and Guglielmo Cinque for helpful comments and suggestions on various aspects of this work, although the responsibility for what I claim is entirely mine. Thanks are also due to Marco Fioratti for kindly collecting the data from Alto Polesano. For a detailed discussion of closely related issues the reader is referred to Munaro (2019).

https://doi.org/10.1075/la.276.06mun

Within the cartographic project (cf. Cinque & Rizzi 2010) the sequence of functional projections which constitute the structural layout of the clause are associated to formal features ultimately responsible for the crosslinguistically parametrized lexicalization of specific functional heads as well as for the displacement of lexical material. The simplified functional sequence of the left periphery, the so-called CP-layer – as proposed by Rizzi (1997) – is the following:

(1) Force >>> Topic >>> Focus >>> Topic >>> Finiteness

Capitalizing on this seminal proposal, a series of studies (cf. Speas & Tenny 2003; Sigurđsson 2004; Bianchi 2006; Baker 2008; Giorgi 2010; Miyagawa 2012; Haegeman & Hill 2013; Haegeman 2014; Hinterhölzl & Munaro 2015; Zanuttini 2017 among others) have developed an innovative view of different aspects of the interface properties characterizing the relationship between utterance and discourse, enriching this basic sequence with additional functional positions where the relevant interpretive factors can be codified. As to the nature of the interface between utterance and discourse, I will submit that, much in the spirit of the cartographic approach, this relation is encoded in the feature inventory of single functional projections of the left-peripheral functional spine, namely in the structural area devoted to connecting the propositional content to the context, that is, to the linking with the conversational or situational background.

Building on the above mentioned studies on the syntactization of logophoric and conversational features, I will sketch here a formal account of the syntactic and interpretive properties of primary and secondary interjections based on empirical evidence from some Italo-Romance dialects. From a descriptive point of view, secondary (or lexical) interjections should be clearly distinguished from primary (or non-lexical) ones and be split into three categories, depending on whether they must, they can or they cannot be prosodically and syntactically integrated with the associated clause; according to the formal analysis proposed here, the degree of integration depends on the merge position of the interjection, which is in turn strictly connected to its discourse linking properties. On the other hand, primary interjections are argued to lexicalize the highest functional head of the left periphery, to which interjections must raise in order to achieve the appropriate spatio-temporal anchoring.

The article is structured as follows: in Section 2 I analyze, on the basis of evidence from some North-Eastern Italian dialects, the interpretive and distributional properties of discourse-linked secondary interjections, which must be integrated with the associated clause; in Section 3, based again on evidence from the same Italo-Romance varieties, I discuss the defining properties of hybrid secondary interjections, which can, but need not, be integrated with the co-occurring clause; in Section 4 I examine non-integrated secondary interjections attested across

Venetan dialects and sketch a formalization of the process of contextual anchoring; in Section 5 I describe the main features of primary interjections on the basis of relevant evidence from some upper Southern Italian dialects; finally, Section 6 contains some concluding remarks.

2. Discourse-linked secondary interjections

The first class includes the lexical interjections that must be integrated with the associated clause and are intrinsically discourse-linked, in the sense that they can only be used to reply to a previous utterance in the discourse situation.

2.1 Emilian dialects

These interjections can be exemplified with the clause initial locution *as capés* in Modenese, like in (2a), or with the interjection *eterché* in Reggiano, like in (2b); these two elements, prosodically integrated with the following propositional clause, clearly have the function to connect the utterance to a previous question or contrary statement of the addressee:

(2) a. Modena
As capés chi m'an invidé!
Interjection that me-have invited
'Of course they invited me!'
b. Reggio Emilia
Eterché se m'an invidé!
Interjection if me-have invited
'They invited me indeed!'

Furthermore, in Modenese the complex clause initial interjection *mo vaca* is obligatorily followed by the complementizer *se* and is clearly prosodically integrated with the rest of the clause:

(3) Modena
a. *Mo vaca* s' l'è gnù èlt!
Interjection if he=is become tall
'He has become tall indeed!'
b. *Mo vaca* s' l'è bèl!
Interjection if he=is handsome
'He is handsome indeed!'

As discussed in detail by Alessandrini (2012), no lexical element can intervene between *mo* and *vaca* (which is predicted by the structural representation in (7) below).[2] The same holds in Reggiano for the complex interjection *mo deg*, which is obligatorily followed by the complementizer *che* and can not be split by intervening lexical material:[3]

(4) Reggio Emilia
a. *Mo deg* c' l'è bèl!
Interjection that he=is handsome
'He is handsome indeed!'

---

2. So for example an overtly realized subject like *Luigi* can appear either in clause final position, like in (ia), or immediately after the complementizer *se*, like in (ib), but not after the particle *mo*, as witnessed by the ungrammaticality of (ic):

(i) Modena
a. *Mo vaca* s' l'è èlt, Luigi!
Interjection if he=is tall, Luigi
b. *Mo vaca* se Luigi l'è èlt!
Interjection if Luigi he=is tall
c. **Mo* Luigi *vaca* s' l'è èlt!
Interj. Luigi interj. if he=is tall
'Luigi is tall indeed!'

3. Moreover, as pointed out by Alessandrini (2012), a topicalized constituent like *a so/to sure-la* cannot intervene between the complex interjection and the complementizer *se/che*, which strongly suggests that the complementizer occupies the head Force° rather than a lower head of the CP layer if, as proposed by Rizzi (1997), Topic projections are located lower than Force in the clausal spine:

(i) Modena
a. *Mo vaca* *se* a so surela Mario al gh'à telefunè!
Interjection if to his sister Mario he=her=has phoned
b. **Mo vaca* a so surela *se* Mario al gh'à telefunè!
Interjection to his sister if Mario he=her=has phoned
'Mario has called his sister indeed!'

(ii) Reggio Emilia
a. *Mo deg* *che* a to surela a gh'ò regalè un bel leber!
Interjection that to your sister I=her=have given an interesting book
b. **Mo deg* a to surela *ch'*a gh'ò regalè un bel leber!
Interjection to your sister that I=her=have given an interesting book
'I have given an interesting book to your sister indeed!'

Based on this evidence, Alessandrini (2012) proposes that interjection and complementizer occupy respectively the specifier and the head of the left-peripheral projection ForceP, as represented in (iii):

(iii) a. [ForceP *Mo vaca* [Force° s'] [FinP l'è èlt!]]
b. [ForceP *Mo deg* [Force° c'] [FinP l'è èlt!]]

b. *Mo deg* c' al stodia!
Interjection that he=studies
'He studies a lot indeed!'

Notice that in exclamative contexts the lexical items *vaca/deg* can also appear in sentence initial position without the particle *mo*:

(5) a. Modena
Vaca s' l'èra elegant!
Int. if he=was elegant
b. Reggio Emilia
Deg c' l'èra elegant!
Int. that he=was elegant
'How elegant he was!'

A very plausible structural analysis for the examples in (5) is the one reported in (6), which is perfectly compatible with the exclamative reading, by which the speaker states that the degree of elegance of the person referred to is situated beyond an expected threshold:

(6) a. [$_{\text{ForceP}}$ *Vaca* [$_{\text{Force°}}$ s'] [$_{\text{FinP}}$ l'èra elegant!]]
b. [$_{\text{ForceP}}$ *Deg* [$_{\text{Force°}}$ c'] [$_{\text{FinP}}$ l'èra elegant!]]

The exclamative reading is in fact generally taken to be linked to the activation of ForceP, the functional projection encoding clause typing features (cf. Rizzi (1997), Benincà (2001)).[4] However, the structural representation in (6) raises an obvious question concerning the position of the discourse particle *mo* in examples like (3)–(4), where the presence of *mo* before *vaca/deg* adds an evaluative shade, in the sense that it requires as linguistic antecedent an utterance of the addressee towards which the speaker expresses his/her emotionally salient reply. Building on Munaro & Poletto (2008) and Poletto & Zanuttini (2010), who have proposed that the semantics of *mo* – a grammaticalized temporal adverb – contains an evaluative component related to the speaker's point of view, Hinterhölzl & Munaro (2015) suggest that *mo* lexicalizes (the head EvalS° of) an Eval(uative)S(peaker) projection immediately dominating ForceP;[5] capitalizing on this analysis, I will suggest that

4. Moreover, this proposal captures in terms of spec-head agreement the selectional link between the element filling the specifier of ForceP and the type of complementizer lexicalizing the head Force°.

5. In particular, Hinterhölzl & Munaro (2015) propose that modal particles in exclamations and special questions function as evidential markers interacting with the evaluative component to derive the diverse expressive meanings. In order to provide an adequate syntactic account of the distribution of these particles, they propose that their semantic impact requires the syntactic representation of separate evaluative and evidential phrases pertaining to speaker and hearer as well as the syntactic representation of the speech act operator.

the particle *mo* appearing in the complex interjections *mo vaca* / *mo deg* is first merged as the head EvalS° of the projection EvalSP, and propose the following structural representation, where *mo* lexicalizes EvalS°, while *vaca/deg* occupy the specifier position of ForceP:

(7) a. [$_{EvalSP}$ [$_{EvalS°}$ *Mo*] [$_{ForceP}$ *vaca* [$_{Force°}$ s'] [$_{FinP}$ l'è èlt!]]]
b. [$_{EvalSP}$ [$_{EvalS°}$ *Mo*] [$_{ForceP}$ *deg* [$_{Force°}$ c'] [$_{FinP}$ l'è èlt!]]]

The impossibility for a left-dislocated constituent to intervene between *vaca/deg* and the complementizers *se/che* suggests that the latter can be plausibly argued to occupy the head Force°. Notice that, given this rigid hierarchy of functional categories, it is correctly predicted that the lexical elements hosted in them are not interchangeable.

2.2 Venetan dialects

In this section I report some examples of discourse-linked interjections attested in Alto Polesano, a dialectal variety spoken in the extreme South-Western part of the region Veneto (at the border with Emilia). Also in this group of dialects a clause initial discourse-linked interjection like *altroché* or *anca massa* must be prosodically integrated with the co-occurring clause and followed respectively by the complementizers *se/che*, as exemplified in (8):

(8) a. Badia Polesine
Altroché se Marco el ga passà l' esame!
Interjection if Mark he=has passed the exam
b. Lendinara
Anca massa che Marco el ga passà l' esame!
Interjection that Mark he=has passed the exam
'Mark has passed the exam indeed!'

In analogy with the proposal put forth above, the structural analysis I propose for these examples is the following, where the interjection lexicalizes (alternatively the head or the specifier of) the EvalSP projection, while the complementizers occupy the head Force°:

(9) a. [$_{EvalSP}$ [$_{EvalS°}$ *Altroché*] [$_{ForceP}$ [$_{Force°}$ se] [$_{FinP}$ Marco el ga passà l'esame!]]]
b. [$_{EvalSP}$ *Anca massa* [$_{EvalS°}$] [$_{ForceP}$ [$_{Force°}$ che] [$_{FinP}$ Marco el ga passà l'esame!]]]

Interestingly, unlike the interjections that will be discussed in the next sections, this kind of interjections cannot surface in final position after the propositional clause, as witnessed by the ungrammaticality of the following examples:

(10) a. Badia Polesine
*Marco el ga passà l' esame! *Altroché*!
Mark he=has passed the exam! Interjection!
b. Castelguglielmo
*Marco el ga passà l' esame! *Anca massa*!
Mark he=has passed the exam! Interjection!

In the next section I will propose a tentative account of this peculiar distributional restriction of discourse-linked interjections, pointing out here simply that this property is a distinctive feature of interjections of this first class.[6]

---

6. An anonymous reviewer points out that in Margherotto, a variety of Venetian, the examples in (10) are both grammatical, as well as the following ones from standard Italian containing the interjections *eccome* and *altroché*, which to my ear must instead be integrated with the associated clause:

(i) a. *Eccome*/*Altroché* se Gianni ha passato l' esame!
Interjection if John has passed the exam!
b. **Eccome*!/*Altroché*! Gianni ha passato l' esame!
Interjection John has passed the exam
c. *Gianni ha passato l' esame! *Eccome*!/*Altroché*!
John has passed the exam Interjection
'John did pass the exam indeed!'

In (ia) the clause initial interjection is obligatorily followed by the complementizer *se* and is clearly prosodically integrated with the rest of the clause, as witnessed by the ungrammaticality of (ib-c). On the basis of their compositional nature – both *eccome* and *altroché* being transparently decomposable into *e* + *come* and *altro* + *che* – I will assume that these interjections are (at least originally) maximal projections from the categorial point of view, hence presumably occupy a specifier position:

(ii) [EvalSP *Eccome/Altroché* [ForceP [Force° se] [FinP Gianni ha passato l'esame!]]]

Interestingly, the sequence in (ia) is typically uttered in response to a previous question or to a contrary statement, and is used to underline the speaker's commitment in stating the relevant propositional content, that is, to express emphatically his personal stance, which leads me to hypothesize that these two interjections lexicalize the functional projection EvalSP, as represented in (ii). Most likely, in the course of time *eccome* and *altroché* are being reanalyzed as filling the head EvalS°, as a consequence of a well attested diachronic process of specifier to head reanalysis taking place within the CP layer (along the lines of the proposal put forth by van Gelderen (2004)); the interjection is still fully integrated with the associated clause, and does not represent an independent illocutionary act. On the specifier to head reanalysis process within the left-periphery the reader is referred also to Willis (2007). It should be pointed out here that the word order in (ic) with *eccome* following the associated clause is grammatical, provided there is no intonational break between the clause and the interjection:

(iii) Gianni ha passato l' esame *eccome*!
John has passed the exam interjection
'John has passed the exam indeed!'

3. A hybrid class of secondary interjections

The second class includes the interjections that can (but need not) be integrated with the associated clause, that is, the ones that are only optionally linked to a discourse antecedent.

3.1 Emilian dialects

The second category of interjections is exemplified by *sorbla* in Bolognese and *madosca* in Reggiano; the interjection can either be followed by *se*, like in (11a)/(12a), or be prosodically and syntactically independent, in which case it usually precedes the associated clause, as exemplified in (11b)/(12b):

(11) Bologna
  a. *Sorbla* s' l'è gnù èlt!
     Interjection if he=is become tall
     'He has become tall indeed!'
  b. *Sorbla*! L'è gnù propria èlt!
     Interjection he=is become really tall
     'He has become really tall!'

(12) Reggio Emilia
  a. *Madosca* s' l'è gnù èlt!
     Interjection if he=is become tall
     'He has become tall indeed!'
  b. *Madosca*! L'è gnù propria èlt!
     Interjection he=is become really tall
     'He has become really tall!'

The structural representation proposed for these examples is reported in (13a)–(13b); as a first working hypothesis, we can assume that this type of interjections are structurally ambiguous in the sense that they can be analyzed by the speakers

This might suggest that once the reanalysis of *eccome* from specifier to head is completed, the interjection, lexicalizing the head EvalS°, can become an attractor for the associated clause, which can raise to the specifier of EvalSP, yielding the sequence in (iii). I will submit that a similar diachronic process can account for the grammaticality of examples like (10) in Venetan varieties like Margherotto.

either as the head EvalS°, like in (13a), or as the head of a contiguous SpeechAct projection, like in (13b):[7]

(13) a. [$_{EvalSP}$ [$_{EvalS°}$ *Sorbla/Madosca*] [$_{ForceP}$ [$_{Force°}$ se] [$_{FinP}$ l'è gnù èlt!]]]
b. [$_{SpeechAct1P}$ [$_{SA°}$ *Sorbla!/Madosca*!] [$_{EvalSP}$ [$_{ForceP}$ [$_{FinP}$ l'è gnù propria èlt!]]]]

As to the degree of syntactic independence of SpeechAct1P with respect to the associated clause, although the status of extra-clausal and parenthetical constituents is still highly controversial, in the spirit of Haegeman & Hill (2013) I will entertain that, despite codifying an autonomous linguistic act, it is still part of the clausal spine of the extended clausal projection; in particular, I would assume that the associated clause may be endowed with its own SpeechAct1P, as it must be somehow anchored in the discourse as well. The merge of the lexical interjection inside the head SpeechAct1° results in an independent speech act, which accounts for the prosodic non-integration of the interjection – separated from the associated clause by a clear intonational break – as well as for the different pragmatic conditions associated to (11b)/(12b), which do not require a discourse background and can be freely uttered out of the blue.

3.2 Venetan dialects

The distributional pattern of the second class of interjections reported from the Emilian dialects in the previous section finds again a striking confirmation in the data collected from Alto Polesano, where interjections like *madona* and *ostrega* display similar properties: they can be prosodically integrated and appear in sentence initial position followed by the complementizer *se*, like in the examples in (14a)/(15a), or be prosodically and structurally independent, in which case they can either precede, like in (14b)/(15b), or follow, like in (14c)/(15c), the associated clause:

(14) Badia Polesine
a. *Madona* se Marco el ga passà l' esame!
Interjection if Mark he=has passed the exam
'Mark did pass the exam indeed!'
b. *Madona*! Marco el ga passà l' esame!
Interjection! Mark he=has passed the exam
'Surprisingly, Mark has passed the exam!'

7. Crosslinguistic evidence for the postulation of a very high Speech Act projection dominating ForceP within the left periphery of the clause is provided by Munaro (2010), Haegeman & Hill (2013), and Haegeman (2014) among others; for a closely related formal proposal see also Coniglio & Zegrean (2012).

c. Marco el ga passà l' esame! *Madona*!
Mark he=has passed the exam! Interjection!
'Mark has passed the exam, surprisingly!'

(15) Lendinara
a. *Ostrega* se Marco el ga passà l' esame!
Interjection if Mark he=has passed the exam
'Mark did pass the exam indeed!'
b. *Ostrega*! Marco el ga passà l' esame!
Interjection! Mark he=has passed the exam
'Surprisingly, Mark has passed the exam!'
c. Marco el ga passà l' esame! *Ostrega*!
Mark he=has passed the exam! Interjection!
'Mark has passed the exam, surprisingly!'

Interestingly, as shown by the glosses, the interpretation of the examples in (a) differs from the one of (b)–(c): while the examples in (a) can only be uttered in response to a question about Marco's passing the exam (and the speaker already knows that Marco has passed the exam, which conforms to his expectation), in (b)–(c) the speaker rather conveys emphatically his surprise about Marco's passing the exam (that is, he learns in that precise moment that Marco has passed the exam and wants to underline that this fact is contrary to his expectation).

This interpretive asymmetry is tentatively captured by the different structural representation proposed for the examples in (14a–c) and (15a–c), which is reported in (16a–c):

(16) a. [$_{EvalSP}$ [$_{EvalS°}$ *Madona/Ostrega*] [$_{ForceP}$ [$_{Force°}$ se] [$_{FinP}$ Marco el ga passà l'esame!]]]
b. [$_{SpeechAct1P}$ [$_{SA°}$ *Madona/Ostrega*!] [$_{EvalSP}$ [$_{ForceP}$ [$_{FinP}$ Marco el ga passà l'esame!]]]]
c. [$_{SpeechAct1P}$ [$_{FinP}$ Marco el ga passà l'esame!]$_{x}$ [$_{SA°}$ *Madona/Ostrega*!] [$_{EvalSP}$ [$_{ForceP}$ t_{x}]]]

In (16a) the clause initial interjection lexicalizes the head EvalS° and is linked to the associated clause by the complementizer *se* located in Force°, while in (16b) it lexicalizes the head SpeechAct1° and the complementizer is not overtly realized; finally, in (16c) the entire propositional clause raises to the specifier of SpeechAct1°. As represented in (16c), I assume that the clause final position of the interjection can be derived from the raising of the nuclear propositional clause FinP to the specifier of SpeechAct1P, possibly in obeyance to a criterial requirement to the effect that the interjection and the associated clause must enter, at some level of

representation, a spec-head agreement configuration (cf. also Munaro & Poletto (2008), Munaro (2010)).[8]

4. Non-integrated secondary interjections and contextual anchoring

Let us turn finally to the third type of interjections, namely the ones that cannot be integrated with the associated clause, and do not need any linguistic antecedent in the speech situation.

4.1 Venetan dialects

This third class of interjections can be exemplified with the lexical item *ciò*, which is attested across the Venetan dialects. Besides appearing in isolation, it can be associated to a clause either in sentence initial position, like in (17b), or in sentence final position, like in (17c), as witnessed by the following examples from Alto Polesano:

(17) Badia Polesine
a. **Ciò* che Marco el ga passà l' esame!
Interjection that Mark he=has passed the exam
'Mark did pass the exam indeed!'
b. *Ciò*! Marco el ga passà l' esame!
Interjection! Mark he=has passed the exam
'Surprisingly, Mark has passed the exam!'

---

8. This second category of interjections is exemplified by items like *caspita* or *accidenti* in standard Italian; they can either be followed by *se*, like in (ia), or be prosodically and syntactically independent, in which case they can either precede or follow the associated clause, as exemplified in (ib) and (ic) respectively:

(i) a. *Caspita*/*Accidenti* se Gianni ha passato l' esame!
Interjection if John has passed the exam!
'John did pass the exam indeed!'
b. *Caspita*!/*Accidenti*! Gianni ha passato l' esame!
Interjection John has passed the exam!
'Surprisingly, John has passed the exam!'
c. Gianni ha passato l' esame! *Caspita*!/*Accidenti*!
John has passed the exam! Interjection
'John has passed the exam, surprisingly!'

The structural representation proposed for the examples in (ia)–(ic) corresponds to the one reported in (16) in the main text.

c. Marco el ga passà l' esame! *Ciò*!
Mark he=has passed the exam! Interjection!
'Mark has passed the exam, surprisingly!'

Interestingly, as shown by the ungrammaticality of (17a), in this variety *ciò* cannot be followed by the complementizer *che*, which means, under the present approach, that it cannot undergo a process of prosodic and structural integration with the co-occurring clause;[9] this restriction can be explained by the hypothesis that this kind of interjections, unlike the ones discussed in the previous sections, are generated as lexicalizations of a higher SpeechAct2° head, from where they can attract the associated clause to their specifier, as represented in (18):[10]

(18) a. [$_{\text{SpeechAct2P}}$ [$_{\text{SA°}}$ *Ciò*!$_{x}$] [$_{\text{SpeechAct1P}}$ [$_{\text{EvalSP}}$ [$_{\text{ForceP}}$ [$_{\text{FinP}}$
Marco el ga passà l'esame!]]]]]
b. [$_{\text{SpeechAct2P}}$ [$_{\text{FinP}}$ Marco el ga passà l'esame!]$_{x}$ [$_{\text{SA°}}$ *Ciò*!] [$_{\text{SpeechAct1P}}$ [$_{\text{EvalSP}}$
[$_{\text{ForceP}}$ t_{x}]]]]

---

9. In fact, some speakers of other varieties of Alto Polesano report that the presence of the complementizer *che* after *ciò* is grammatical, as long as *ciò* is preceded by the particle *eh*, like in the following example:

(i) Lendinara
Eh ciò che Marco el ga passà l' esame!
Int Int that Mark he=has passed the exam
'Obviously, Mark has passed the exam!'

Uttering (i), the speaker underlines that the fact that Marco has passed the exam conforms to his expectations. Similarly, in other Venetan varieties *ciò* can occur with or without a following complementizer, but the pragmatic import of the two alternatives differs:

(ii) a. *Ciò* che ghe vago!
Interjection that there=go
'I will surely go there!'
b. *Ciò*! Vago via!
Interjection! Go away
'Hey listen, I am leaving!'

The utterance in (iia) is used pragmatically to convey a sense of surprise by the speaker at the question of the interlocutor, hence is used to reply to a previous utterance, while in (iib) *ciò* has the pragmatic function of drawing the attention of the interlocutor to what is being said. I will assume that in examples like (i) and (iia) *ciò* should be analyzed as a lexical item which is still syntactically and prosodically integrated with the associated clause (hence presumably, under the present analysis, lexicalizing the head EvalS°), while in (iib) it displays the distinctive properties of an interjection (lexicalizing a SpeechAct head), as witnessed by the strong intonational break intervening between *ciò* and the rest of the clause.

10. An anonymous reviewer points out that a possible argument in favour of this syntactic analysis comes from the prosodic contour of the utterance, since before sentence final *ciò* there is no intonational break, but rather a rising tone, followed by a falling tone on *ciò*.

Still following Haegeman & Hill (2013), I postulate a recursive SpeechAct layer, articulated in a higher SA2°, lexicalized by the non-integrated secondary interjection and encoding the setting up of the discourse layer (with an 'attention seeking' attitude of the speaker), and a lower SA1° with a 'bonding' function, encoding the consolidation of the discourse relation and anchoring the associated clause to the discourse.[11]

As anticipated above, the non-integrated version of the interjection *ciò* is attested throughout the Veneto region, like in the following examples from central Venetan:

(19) Central Venetan
a. *Ciò*! Che bela casa che la se ga comprà!
Interjection! What nice house that she=herself=has bought
b. Che bela casa che la se ga comprà, *ciò*!
What nice house that she=herself=has bought, interjection
'She bought herself a really nice house!'

Hence, for the Example (19b) I propose the following simplified structural representation:

(20) $[_{\text{SpeechAct2P}}\ [_{\text{ForceP}}$ Che bela casa che la se ga comprà$]_x\ [_{\text{SA}^\circ}$ *ciò*!$]\ t_x\ ]$

As discussed in Del Gobbo, Munaro & Poletto (2015), the crossdialectal distribution of *ciò* in Venetan strongly suggests a derivational cline according to which this discourse marker was originally a specifier which has been eventually reanalyzed by the speakers as a discourse-related head of the left periphery (arguably a SpeechAct head), and has become then in the central Veneto area an attractor for the associated clause, giving rise to the sequence in (19b), represented structurally as in (20).[12]

---

11. In this perspective, the raising of the associated clause – a SpeechActP – to the specifier of the higher SA2° would be plausibly driven by a sort of SpeechAct criterial condition.

12. As for standard Italian, we find interjections like *toh* and *però*, expressing slight surprise and mirativity respectively, both of which can either precede or follow the associated clause, and, if clause initial, cannot be followed by a complementizer:

(i) a. *Toh*! Maria ha dimenticato le chiavi!
Interjection! Mary has forgotten the keys!
b. Maria ha dimenticato le chiavi! *Toh*!
Mary has forgotten the keys! Interjection
c. **Toh* se/che Maria ha dimenticato le chiavi!
Interjection if/that Mary has forgotten the keys!
'Look, Mary has forgotten the keys!'

(ii) a. *Però*! Gianni sta ancora studiando!
Interjection! John is still studying!
b. Gianni sta ancora studiando! *Però*!
John is still. studying! Interjection
c. **Però* se/che Gianni sta ancora studiando!
Interjection if/that John is still studying!
'Surprisingly, John is still studying!'

4.2 On contextual anchoring

Interjections have an intrinsically deictic nature in the sense that they can only be uttered in the presence of the mental state they give vent to; their use entails therefore a crucial reference to the speaker's spatio-temporal coordinates, namely to the default indexical values (i.e. *hic et nunc*) defining the logophoric center represented by the speaker.[13]

Elaborating on Poggi (1988), Munaro (2010) tries to discompose the interpretive import of secondary interjections, pointing out that the Italian examples in (21) and (22) can be rephrased as in (23a) and (23b) respectively:

(21) *Toh*! [Maria ha dimenticato le chiavi!]
Interjection! [Mary has forgotten the keys!]

(22) *Però*! [Gianni sta ancora studiando!]
Interjection! [John is still studying!]

(23) a. *This event* [= the fact that Mary forgot the keys] arouses in me a slight surprise
b. I am positively struck by *this event* [= the fact that John is still studying]

The rephrased versions in (23) reveal the informational content of the interjection, which includes (at least) two components: the mental state of the speaker, which is conventionally codified by the interjection, and a deictic expression, typically a demonstrative, that refers to the event of the external world that is the source of that mental state (which can be made explicit by the clause associated to the interjection, otherwise it must be recovered from the linguistic or situational context).

Interestingly, as we have seen above, only secondary interjections belonging to the second and third class can be uttered in isolation in out of the blue contexts; this property can be derived by the hypothesis that only lexical interjections occupying

---

I surmise that this kind of interjections are first merged as heads of the SpeechAct2 projection, giving rise to the basic word order in (ia)/(iia); as for the reverse word order in (ib)/(iib), as suggested above, the clause final position of the interjection can be derived again from the optional fronting of the nuclear clause FinP to the specifier of SpeechAct2P, as represented in (iii):

(iii) [SpeechAct2P [FinP Gianni sta ancora studiando!]$_x$ [SA° *Toh/Però*!] [ForceP t$_x$]]

Although clause fronting to the specifier of SpeechActP has no evident interpretive effects, it can be pointed out that the interjection in clause final position is perceived as a sort of afterthought, optionally added by the speaker in order to spell out the speaker's mental attitude towards the propositional content of the associated clause.

13. On the ontological nature of interjections, and, more specifically, on the spatio-temporal anchoring of the associated event as well as on the question whether an interjection can exhaustively constitute a complete speech act, I refer the reader to Menza (2006).

a SpeechAct° head, after attracting the associated clause into the corresponding specifier, can reach by head movement the head of the adjacent Speaker projection where, according to Giorgi (2010, 2012), the speaker's logophoric coordinates are codified:[14]

(24) $[_{\text{SpeakerP}}\ [_{\text{Sp°}}\ \textit{Sorbla!Ostrega!/Ciò!}_{x}]\ [_{\text{SpeechActP}}\ [_{\text{FinP}}\ \emptyset]_{y}\ [_{\text{SA°}}\ t_{x}]\ [_{\text{ForceP}}\ t_{y}\ ]]]$

I submit that only the raising of the interjection to the next head Speaker° can provide the appropriate contextual anchoring of the utterance through the selection of the (default) spatio-temporal coordinates so that the interjection will be spelled-out in isolation, without the phonetic realization of the associated clause. Under this analysis, only after the interjection has reached the next higher head Speaker° can the appropriate spatio-temporal anchoring of the utterance come about.[15]

5. On the contextual anchoring of primary interjections

Let us turn next to primary, i.e. non-lexical interjections which, unlike secondary ones, are not etymologically related to a lexical item and are used to emphatically express the speaker's strong emotional reaction to a linguistic or extra-linguistic event that is manifest in the speech situation.

As witnessed by the following minimal pair from standard Italian, primary vocalic and crosslinguistically attested interjections (such as *ah/eh/ih/oh/uh*) appear preferably in clause initial position (although there seems to be some degree of variation among speakers in the acceptability of these structures):

---

14. The reader is referred to Ritter & Wiltschko (2014) on the existence of a unique syncretic head encoding spatial and temporal anchoring. Notice that if the raising of the interjection takes place by head movement, we expect it to be subject to the head movement constraint, so that ungrammaticality should arise if another lexical head intervenes between the initial position and the final landing site, as discussed for example in Footnote 18 below.

15. In other words, only in that case can take place the deictic reference to the event of the external world that is the source of the speaker's mental state, which allows for the non-realization of the associated clause. As for the secondary interjections of the first class, they can only be uttered in isolation in response to a previous utterance of the interlocutor: not being merged as SpeechAct heads, they are correctly predicted to require a linguistic antecedent in the discourse with which they presumably form a unique syntactic object, arguably in virtue of being endowed with an *edge*-feature in the sense of Chomsky (2008); in fact Munaro (2010) argues that minimal syntactic elements like interjections, short answers and particles do have an edge-feature and tries to unify the syntactic analysis of interjections and particles with the one of prepositions as attractors formulated by Kayne (2002).

(25) a. [Ah/Eh/Ih/Oh/Uh], Gianni ha passato l' esame!
b. * Gianni ha passato l' esame, [ah/eh/ih/oh/uh]!
[Interjection] John has passed the exam [Interjection]
'John has passed the exam indeed!'

Moreover, they can co-occur with a secondary interjection, obligatorily preceding it:

(26) a. [Ah/Eh/Ih/Oh/Uh], eccome/altroché/accidenti/caspita se Gianni ha passato l'esame!
b. *Eccome/Altroché/Accidenti/Caspita se Gianni ha passato l' esame,
[Interjection] interjection if John has passed the exam
[ah/eh/ih/oh/uh]!
[Interjection]
'John has passed the exam indeed!'

On the basis of their distributional and interpretive properties, it is extremely tempting to analyze primary interjections as the lexicalization of the highest head Speaker°. Under the plausible assumption that the specifier of SpeakerP is occupied by the speaker's contextual spatio-temporal coordinates, and is therefore inaccessible to the fronting of the associated clause, we straightforwardly derive in this way the ungrammaticality of (25b) and (26b).

As discussed in detail by Colasanti & Silvestri (2019), the insertion of a non-lexical interjection within the non-propositional area of the clause located above Force interacts in various ways with the embedded syntactic structure. More precisely, their research reveals the existence of a tight relation between the Speech Act layer and the Force layer, showing that in some Southern Italian dialects jussive, concessive, and optative matrix clauses may be introduced by complementizers whose insertion is strictly dependent on the utterance of speech act material at the outset of the sentence.[16]

In particular, Colasanti & Silvestri take into account, among others, a type of matrix clauses that in both the upper Southern Italian dialects and the extreme Southern Italian dialects display a structurally consistent pattern, i.e. optative clauses, where the speaker expresses his/her desire for the realization of a (mainly counterfactual) state of things which is beyond his/her control. The context provided to their informants was a situation whereby the speaker is extremely upset

16. According to them, in these dialects optatives, as jussives and concessives, provide a complex picture based on the different degrees of pragmatic stance conveyed by the presence or the absence of the initial complementizer. They also observe that the morphological mood expresses the irrealis modality of the sentence, and that in the unmarked word order of the clause the subject is placed in postverbal position.

about the hearer's behaviour to the point of imprecating curses on him. They notice that in all the varieties considered, the complementizer can be optionally inserted at the outset of an optative clause with negative orientation, as exemplified in (27a) with a Northern Abruzzese variety and in (27b) with a Northern Calabrian dialect:

(27) a. Teramo
(Chə/*Ca) tə pozzənə accidə!
(That) you-can kill
'May they kill you!'
b. Santa Maria del Cedro
(Chə/*Ca) tə pəgghjissa nu lampə!
(That) you-stoke a lightning
'May lightning strike you!'

They point out that informants are strikingly consistent in providing the interpretation of the clause introduced by the complementizer, in the sense that its insertion corresponds to a stronger stance of the curse, hence to a deeper emotional involvement of the speaker.

An interesting revealing property of optative clauses becomes evident when the utterances in (27) are realized with an initial interjection; the insertion of a primary interjection within the non-propositional area of the clause interacts in interesting ways with the embedded syntactic structure, triggering the spell-out of lower heads, among which Force°. Colasanti & Silvestri observe that the presence of a clause initial non-lexical interjection triggers the mandatory realization of a complementizer, as witnessed by the following minimal pairs from two upper Southern Italian dialects:[17]

(28) Santa Maria Capua Vetere
a. (Ca) tə putesserə accidə!
(That) you-could kill
b. Ih *(ca) tə putesserə accidə!
Interjection *(that) you-could kill
'May they kill you!'

17. The same restriction holds, according to Colasanti & Silvestri, for Extreme Southern Italian dialects, as exemplified in (i) with a Southern Calabrian variety:

(i) Melito
a. (Chi) *(mi) ti mangianu i cani!
(That) *(mi) you-eat the dogs
b. Ih *(chi) mi ti mangianu i cani!
Interjection *(that) mi you-eat the dogs
'May the dogs devour you!'

(29) Teramo
a. (Chə/*Ca) tə pozzənə accidə!
(That) you-can kill
b. Ih *(chə) tə pozzənə accidə!
Interjection *(that) you-can kill
'May they kill you!'

The same holds for the following optative examples from Northern Campanian; the example in (30a) is used in a context in which Mario is not behaving properly and the speaker wants him to behave, while the example in (30b) is used by the speaker to curse the hearer:

(30) Santa Maria Capua Vetere
a. Oh *(ca) Màriə facesse u brave!
Interjection *(that) Mario did the good
'Mario had better behave!'
b. Ih *(ca) te pozza piglià ne colpə!
Interjection *(that) you-can take a blow
'May you be hit!'

According to the analysis of Colasanti & Silvestri, in both cases the complementizer *ca* lexicalizes the head Force° and carries a [+speech] feature that has to be checked with the primary interjection situated beyond the Force layer.

Under the present approach, the upper Southern Italian dialects where the presence of the interjection in optative clauses triggers the obligatory presence of a complementizer provide empirical evidence that primary interjections may raise to Speaker° from a lower position.

As discussed by Colasanti & Silvestri, it can be argued that in these structures the interjection and the complementizer form one prosodic unit, which is confirmed by the elliptical use of the optative that speakers may resort to in order to mitigate the invective; the utterance reported in (31), from a Northern Calabrian variety, includes only the interjection and the complementizer, which display prosodic and pragmatic autonomy:

(31) Santa Maria del Cedro
Ih-chə…!
Interjection-that…

We can reasonably assume that, while secondary interjections of the second and third class may raise by head movement to Speaker°, raising of primary interjections to Speaker° must obtain in order to achieve the deictic reference to the event of the external world that is the source of the speaker's mental state. The present

approach offers a straightforward account for the elliptical structure in (31); as represented in (32b), on its way up to Speaker°, from SpeechAct° the interjection may attract the associated FinP to its specifier, stranding the complementizer in Force°; as we have postulated above, the final step, namely the raising of the interjection to Speaker°, results in the logophoric anchoring of the utterance and in the phonetic deletion of the associated clause, so that only the prosodic unit formed by interjection and complementizer will be spelled-out, as in (32a):[18]

(32) Santa Maria del Cedro
a. Ih-chə…!
b. $[_{\text{SpeakerP}}\ [_{\text{Sp°}}\ \text{Ih}_x][_{\text{SpeechActP}}\ [_{\text{FinP}}\ \text{Ø}]_y\ [_{\text{SA°}}\ t_x][_{\text{ForceP}}\ [_{\text{Force°}}\ \text{chə}]\ t_y]]$

The empirical evidence discussed by Colasanti & Silvestri (2019) reveals a certain degree of consistency in the behaviour of jussive, concessive and optative matrix clauses with respect to different factors of internal variation, like the optional overt realization of the complementizer, the utterance of initial interjections and the differences in the interpretive outcomes based on these elements; in particular, it turns out that the insertion of speech-act related material, such as interjections, triggers the obligatory realization of the complementizer, thus revealing interesting interactions between the Force layer and higher left-peripheral functional layers, which deserve further investigation.

6. Conclusions

On the basis of empirical evidence from Northern Italian dialects, I have proposed that, with respect to the degree of integration with the associated clause, secondary interjections can be split basically into three categories, depending on whether they must, they can or they cannot be integrated with the associated clause; the degree of integration with the co-occurring clause depends on the merge position of the interjection. I have suggested that interjections lexicalize different functional heads which are computed syntactically at the edge of the clause.

---

18. As already noted, the prosodic contour of the utterance reveals that the interjection and the complementizer form an unbreakable speech unit and prosodically weld together, as no further element, such as a vocative, can intervene between them:

(i) Santa Maria del Cedro

Ih	(*Marì)	chə	tə pəgghjissa	nu	lampə!	
Interjection	(*Mary)	that	you-stoke	a	lightening	

'May lightning strike you!'

Only interjections lexicalizing (by external Merge) a SpeechAct° head represent autonomous linguistic acts, and are as such prosodically and syntactically independent from the associated clause, if this is present; from this position they can attract the associated clause to the corresponding specifier position and raise to the adjacent head Speaker° in order to provide the necessary contextual anchoring by entering a local relation with the speaker's coordinates. Interjections lexicalizing the lower projection EvalSP are instead intrinsically discourse-linked, in the sense that they can only be used to reply to a previous utterance in the discourse situation.

Finally, upper Southern Italian dialects provide empirical evidence that non-lexical interjections lexicalize the highest functional head Speaker°, to which they must raise in order to achieve the necessary spatio-temporal anchoring, interacting in interesting ways with lower projections and with the overt realization of the complementizer in Force°.

References

Alessandrini, Elisa. 2012. Proprietà sintattiche della frase esclamativa nei dialetti emiliani. Master thesis, Ca' Foscari University – Venice.

Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge, Mass.: Cambridge University Press. https://doi.org/10.1017/CBO9780511619830

Benincà, Paola. 2001. The position of Topic and Focus in the left periphery. In *Current Studies in Italian Syntax – Essays offered to Lorenzo Renzi*, Guglielmo Cinque & Giampaolo Salvi (eds), 39–64. Amsterdam: Elsevier.

Bianchi, Valentina. 2006. On the syntax of personal arguments. *Lingua* 116: 2023–2067. https://doi.org/10.1016/j.lingua.2005.05.002

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory. Essays in honor of Jean-Roger Vergnaud*, Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds), 133–166. Cambridge MA: MIT Press. https://doi.org/10.7551/mitpress/9780262062787.003.0007

Cinque, Guglielmo & Rizzi. Luigi. 2010. The cartography of syntactic structures. In *The Oxford Handbook of Linguistic Analysis*, Bernd Heine & Heiko Narrog (eds), 51–65. Oxford/New York: Oxford University Press.

Colasanti, Valentina & Silvestri, Giuseppina. 2019. Matrix complementizers in Italo-Romance. In *Italian Dialectology at the Interfaces*, Silvio Cruschina, Adam Ledgeway & Eva-Maria Remberger (eds), 155–183. Amsterdam: John Benjamins. https://doi.org/10.1075/la.251.08col

Coniglio, Marco & Zegrean, Iulia. 2012. Splitting up Force: Evidence from discourse particles. In *Main clause phenomena: New horizons*, Lobke Aelbrecht, Liliane Haegeman & Rachel Nye (eds), 229–256. Amsterdam: John Benjamins. https://doi.org/10.1075/la.190.10con

Del Gobbo, Francesca, Munaro, Nicola & Poletto, Cecilia. 2015. On sentential particles: A cross-linguistic study. In *Final Particles*, Sylvie Hancil, Alexander Haselow, Margje Post (eds), 359–386. Berlin: de Gruyter. https://doi.org/10.1515/9783110375572-015

van Gelderen, Elly. 2004. Economy, innovation, and prescriptivism: from spec to head and head to head". *Journal of Comparative Germanic Linguistics* 7: 59–98. https://doi.org/10.1023/B:JCOM.0000003601.53603.b2

Giorgi, Alessandra. 2010. *About the speaker: Towards a syntax of indexicality*. Oxford/New York: Oxford University Press.

Giorgi, Alessandra. 2012. The theory of syntax and the representation of indexicality. In *Functional Heads*, Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds), 42–54. Oxford/New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199746736.003.0003

Haegeman, Liliane. 2014. West-flemish verb-based discourse markers and the articulation of the speech act layer. *Studia Linguistica* 68: 116–139. https://doi.org/10.1111/stul.12023

Haegeman, Liliane & Hill, Virginia. 2013. The syntacticization of discourse. In *Syntax and its limits*, Raffaella Folli, Christina Sevdali & Robert Truswell (eds), 370–390. Oxford/New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199683239.003.0018

Hinterhölzl, Roland & Munaro, Nicola. 2015. On the interpretation of modal particles in non-assertive speech acts in German and Bellunese. In *Discourse-oriented syntax*, Josef Bayer, Roland Hinterhölzl & Andreas Trotzke (eds), 41–70. Amsterdam: John Benjamins. https://doi.org/10.1075/la.226.03hin

Kayne, Richard. 2002. Prepositional complementizers as attractors. In *Parameters and Universals*, 282–323. Oxford/New York: Oxford University Press.

Menza, Salvatore. 2006. *Il paraverbo. L'interiezione come sottoclasse del verbo*. Alessandria: Edizioni dell'Orso.

Miyagawa, Shigeru. 2012. Agreements that occur mainly in the main clause. In *Main clause phenomena: New horizons*, Lobke Aelbrecht, Liliane Haegeman & Rachel Nye (eds), 79–112. Amsterdam: John Benjamins. https://doi.org/10.1075/la.190.04miy

Munaro, Nicola. 2010. On the edge feature of particles, interjections, and short answers. In *Edges, heads and projections – Interface properties*, Anna Maria Di Sciullo & Virginia Hill (eds), 67–85. Amsterdam: John Benjamins. https://doi.org/10.1075/la.156.07mun

Munaro, Nicola. 2019. On the syntactic encoding of lexical interjections in Italo-Romance. In *Italian Dialectology at the Interfaces*, Silvio Cruschina, Adam Ledgeway & Eva-Maria Remberger (eds), 185–202. Amsterdam: John Benjamins. https://doi.org/10.1075/la.251.09mun

Munaro, Nicola & Poletto, Cecilia. 2008. Sentential particles and clausal typing in Venetan dialects. In *Dislocated elements in discourse: Syntactic, semantic and pragmatic perspectives*, Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds), 173–199. London/New York: Routledge.

Poggi, Isabella. 1988. Le interiezioni. In *Grande grammatica italiana di consultazione*, Lorenzo Renzi, Giampaolo Salvi & Anna Cardinaletti (eds), vol.3, 403–425. Bologna: Il Mulino.

Poletto, Cecilia & Zanuttini, Raffaella. 2010. Sentential particles and remnant movement. In *Mapping the left periphery*, Paola Benincà & Nicola Munaro (eds), 201–227. Oxford/New York: Oxford University Press.

Ritter, Elizabeth & Wiltschko, Martina. 2014. The composition of INFL. *Natural Language and Linguistic Theory* 32: 1331–1386. https://doi.org/10.1007/s11049-014-9248-6

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, Liliane Haegeman (ed), 281–337. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7

Sigurđsson, Halldor. 2004. The syntax of person, tense, and speech features. *Rivista di Linguistica – Italian Journal of Linguistics* 16: 219–251.

Speas, Peggy & Tenny, Carol. 2003. Configurational properties of point of view roles. In *Asymmetry in grammar*, Anna Maria Di Sciullo (ed), 315–344. Amsterdam: John Benjamins. https://doi.org/10.1075/la.57.15spe

Willis, David. 2007. Specifier-to-head reanalysis in the complementizer domain: evidence from Welsh. *Transactions of the Philological Society* 105(3): 432–480. https://doi.org/10.1111/j.1467-968X.2007.00194.x

Zanuttini, Raffaella. 2017. Presentatives and the Left Periphery. In *Elements of comparative syntax: Theory and description*, Enoch Aboh, Eric Haeberli, Genoveva Puskàs & Manuela Schönenberger (eds), 221–256. Boston/Berlin: de Gruyter. https://doi.org/10.1515/9781501504037-008

CHAPTER 7

Sentence-final particles in Mandarin Chinese

Syntax, semantics and acquisition

Waltraud Paul and Shanshan Yan
CNRS-EHESS-INALCO, CRLAO, Paris / School of Chinese as a second language, Peking University

Sentence-final particles (SFPs) in Mandarin Chinese realize the heads of three projections in the rigidly ordered head-final CP 'Low CP < ForceP < AttitudeP'. Only the highest projection AttitudeP encodes discourse-related properties, whereas ForceP encodes the sentence-type (interrogative, imperative). Low Cs interact with properties of the TP-internal extended verbal projection and are obligatory when acting as (non-default) anchors. They play an important role in determining the temporal interpretation and finiteness in Mandarin Chinese and can therefore no longer be neglected by studies addressing these issues. There is no evidence for an "incremental" acquisition "up the tree" of the different projections in the split CP nor for the acquisition of TP prior to CP, as postulated by the cartographic approach.

Keywords: split CP, sentence type, finiteness, tense, aspect

1. Introduction

Mandarin Chinese not only features SFPs linked to discourse (as would be expected from its alleged "discourse-oriented" nature), but also SFPs encoding the sentence-type (interrogative, imperative etc.) as well as a set of SFPs that interact with the properties of the TP-internal extended verbal projection and in certain cases are obligatory. SFPs all occupy a position in the (right) sentence periphery (CP) and are construed with the entire clause, leading to a transparent syntax/semantics mapping in terms of scope relations. More precisely, SFPs realize heads in a three-layered split CP in the spirit of Rizzi (1997):[1]

1. Given the complex nature of finiteness in Chinese, Rizzi's (1997) Finiteness Phrase is replaced by CLowP as the lowest level in the Chinese split CP.

https://doi.org/10.1075/la.276.07pau

(1) Split CP in Mandarin Chinese (Paul 2009):
[Attitude-CP [Force-CP [ClowP [TP NP V NP] Clow°] Force°] Attitude°]]]

Rizzi (1997) demonstrated in great detail that the sentence periphery above TP, the sentence proper, does not consist of a single CP hosting e.g. the fronted *wh*-phrase (and the "dummy" verb *do*, in the absence of an auxiliary verb) in English sentences such as *[$_{CP}$ What$_i$ [$_{C'}$ [$_{C°}$ did] [$_{TP}$ he buy t_i]]]?*. On the contrary, the sentence periphery is "split up", i.e. divided into numerous subprojections displaying a rigid order, among them projections for topic phrases and focus phrases. As for the heads present in the left periphery, i.e. complementizers, he likewise argued that they are of different types and hence occur in different projections within the split CP. Complementizers indicating the type of clause (declarative "force", interrogative "force" etc., e.g. *that, whether* in English; *che* in Italian) head the projection ForceP *preceding* the topic and focus projections; by contrast, prepositional complementizers in Romance such as Italian *di* introducing infinitivals realize the head of FinitenessP, a projection immediately above TP and *below* topic and focus projections:

(2) a. Penso (*a Gianni) che, a Gianni, gli dovrei parlare
think.1SG to Gianni that to Gianni him should speak
'I think that to Gianni, I should speak to him.'
b. Penso, a Gianni, di (*a Gianni) dovergli parlare
think.1SG to Gianni that to Gianni him.should speak
'I think, to Gianni, 'of' to have to speak to him.' [sic]
(Rizzi 1997: 304, [61], [62])

Subsequent studies of mostly Romance and Germanic languages extended this approach to *matrix* clauses and analyzed as different types of complementizers those items at the sentence periphery that had so far been called "particles", for want of a precise categorial status (cf. among others Munaro and Poletto 2002). Importantly, these studies also provided evidence for the existence of an additional, discourse-related projection *above* ForceP, equivalent in function to the projection labelled *AttitudeP* by Paul (2009) in (1) for Chinese (cf. a.o. Benincà 2001 for Romance languages, Haegeman 2014 and Haegeman and Hill 2013 for West-Flemish):

(3) DiscourseP > ForceP > FiniteP > TP
(Split CP for Germanic and Romance languages)

(Note that [3] concentrates on the projections within the split CP that are exclusively realized by heads, to the exclusion of topic and focus phrases.) The hierarchy in (3) thus extends Rizzi's (1997) original hierarchy where the highest projection

had been ForceP. Comparing (3) with (1) for Chinese, we see that they only differ in the directionality, head-initial for Rizzi (1997), head-final in Chinese.[2]

The split CP for Chinese in (1) in fact recasts into modern terms the observations by the eminent Chinese scholar Zhu (1982: 207–213). He identified three distributional classes of SFPs whose relative order is fixed. The SFPs belonging to the first class, SFP_1, occur nearest to the sentence proper (TP) and are said to "express tense/aspect"; they comprise SFPs such as *le* and *láizhe* (cf. [6a] below) and realize LowCP in (1). The SFPs of the second class, SFP_2, to the right of the position for SFP_1, convey notions such as *yes/no* question (*ma*) and imperative (*ba*) (cf. [5a] and [7a] below) and thus illustrate the ForceP in (1). The third, "outermost" class of SFP_3, finally, is explicitly stated to be different from the two other classes, because it involves the speaker's attitude or feelings (hence the label AttitudeP in [1]); SFPs belonging to this class are e.g. *a*, *ou* etc. (cf. [7a], [8a] below). Zhu (1982: 208) emphasizes that co-occurring SFPs belong to hierarchically different levels, while SFPs of the same class are mutually exclusive, such as e.g. *le* and *láizhe*, which both belong to the innermost class, SFP_1 (cf. [6b] below).

(4) $[[[[_S$…..] $SFP_1]$ $SFP_2]$ $SFP_3]$

The ordering restrictions underlying the configuration in (4) are illustrated below:

(5) a. [$_{CP2}$ [$_{CP1}$ [$_{TP}$ Tā bù chōu yān] le] ma]?
3SG NEG inhale cigarette SFP1 SFP2
'Does he no longer smoke?'
b. *[$_{CP1}$ [$_{CP2}$ [$_{TP}$ Tā bù chōu yān] ma] le]?
3SG NEG inhale cigarette SFP2 SFP1

(6) a. [$_{CP1}$ [$_{TP}$ Wǒ chī wǎnfàn] le / láizhe].
1SG eat dinner SFP1 / SFP1
'I (just) had dinner.'
b. *[$_{CP1}$ [$_{TP}$ Wǒ chī wǎnfàn] {le láizhe} / {láizhe le}]
1SG eat dinner SFP1 SFP1 / SFP1 SFP1

(7) a. [$_{CP3}$ [$_{CP2}$ [$_{TP}$ Jìnlái] b'ou (=ba+ou)]]!
enter SFP(fusion) SFP2+SFP3
'Hurry, come in!' (Zhu 1982: 212; bracketing added)
b. *[$_{CP2}$ [$_{CP3}$ [$_{TP}$ Jìnlái] ou] ba]!
enter SFP3 SFP2

2. Conjunctions in adverbial clauses also instantiate Cs and project a head-initial CP: *rúguǒ* 'if', *jíshǐ* 'even if', *jiùsuàn* 'even though', *jìrán* 'since', *suīrán* 'although', *yīnwèi* 'because' (cf. Pan & Paul 2018: 147). By contrast, Chinese lacks an embedding C equivalent to English *that* (cf. Paul 2015: 305 for further discussion).

(8) a. [$_{CP3}$ [$_{CP1}$ [$_{TP}$ Bù zǎo] l'ou (= le + ou)]]!
NEG early SFP (fusion) SFP1+SFP3
'Hey, it's already late!'
b. *[$_{CP1}$ [$_{CP3}$ [$_{TP}$ Bù zǎo] ou le !
NEG early SFP3 SFP1

Starting with examples (8a) and (7a), an SFP_3 such as *ou*, which expresses the speaker's impatience, must follow the SFP_2 *ba* (expressing a "softened" imperative) in (7a) and the SFP_1 *le* in (8a). Since it consists of a single vowel, *ou* fuses phonetically with the preceding SFP into a single syllable.[3] Likewise, the innermost SFP_1 *le* must always precede SFP_2 such as the interrogative *ma* (cf. [5a]), as shown by the unacceptability of the opposite order in (5b). (8a) further illustrates that Zhu (1982) basically uses the same reasoning in order to determine the relative order of SFPs as the cartographic approach does when establishing the hierarchy of functional projections.[4] Since for semantic reasons it is rather difficult to construe and find sequences where all the three classes co-occur, Zhu (1982: 208) applies the notion of *transitivity*: if a given SFP *A* is shown to precede the SFP *B* and the SFP *B* precedes the SFP *C*, then necessarily *A* likewise must precede *C*. This same notion of transitivity also underlies Zhu's (1982: 208) statement that the relative order always holds, i.e. also when a given SFP position remains empty, as in the combination of the SFP_1 *le* with the SFP_3 *ou* in (8a).

The article is organized as follows. Section 2 provides the general background for the analysis of SFPs as C-heads in Chinese. No exhaustive presentation is intended nor possible here. (For systematic overviews, cf. Paul 2015, Chapter 7; Pan 2015, 2019a; Paul & Pan 2017). Instead, we choose to highlight some aspects of SFPs which have either not received enough attention so far or have not been sufficiently spelt out. Section 2.1 on ClowP argues that the low Cs *le,* ne_1 and *láizhe* contribute to the temporal interpretation, but do *not* encode tense or aspect themselves. They are thus *not* on a par with the TP-*internal* aspectual heads, as also evidenced by their co-occurrence. Section 2.2 turns to the SFPs in ForceP realizing sentence types such as interrogative and imperative. Given that these SFPs have been extensively studied, claims in the literature that Chinese lacks such Force heads (cf. Li 2006; Bailey 2015; Del Gobbo et al. 2015) are very surprising and straightforwardly invalidated by the data. Section 2.3 examines the third, highest layer in the split CP,

3. This phonetic fusion only applies to spelt-out SFPs in the correct linear order (cf. the unacceptability of [7b], [8b]). It is limited to SFP_3 simply because only the latter consist of a single vowel.

4. For example, Cinque's (1999: 41) complete hierarchy of the functional projections hosting adverbs relies on the stepwise application of the relative order established for a given pair of adverbs, the full hierarchy not being attested.

i.e. AttitudeP. Section 3 briefly addresses the issue of acquisition. Importantly, by the age of two years, children have basically acquired all three types of SFPs, including Attitude heads, which are subject to rather complex semantic-pragmatic constraints. These data challenge Friedmann et al.'s (2021) hypothesis of an incremental step-by-step acquisition, both with respect to the relative order of TP and CP and the order within the periphery. Section 4 concludes the article and emphasizes the importance of the low Cs for future studies on tense and finiteness in Chinese.

2. The three-layered CP in Chinese: Overview and some in-depth case studies

The analysis of SFPs as complementizers goes back to Lee (1986) who was the first to claim C-head status for the *yes/no*-question particle *ma*. The analysis of *ma* as C became the standard analysis and was confirmed by subsequent studies, which also introduced another C, i.e. *ne* (cf. Cheng 1991). Tang (1989: 541) extended the C-analysis to SFPs in general. The architecture of the Chinese sentence periphery was developed in more detail within Rizzi's (1997) split CP approach by Paul (2005) and subsequent work, where an additional projection AttitudeP *above* Rizzi's ForceP was motivated (cf. Paul 2009, 2014).

(9) The split root CP (based on Paul 2014; cf. Pan 2015, 2019a; Paul & Pan 2017 for a more fine-grained picture)

C_1 (Low C)	**C_2 (Force)**	**C_3 (Attitude)**
le currently relevant state	ba_{Imp} (advisative *ba*)	*a* softening $láizhe_3$ what did you just say?
$láizhe_1$ recent past	$ba_{Qconfirmation}$	*ei* gentle reminder
	ma_2 yes/no question	*ou* impatience, surprise
ne_1 continuing sit.		ma_3 dogmatic assertion
		zhene intensifier ne_3 exaggeration

*N.B. The semantic values indicated for each SFP can give a rough approximation only.

Importantly, there are several cases of homonymy between low C-heads and Attitude-heads, hence our indexing the low C-heads ne_1 and $láizhe_1$ with 1 and the Attitude-heads ne_3 and $láizhe_3$ with 3. The two force heads *ba*, advisative *ba* (softening the imperative) and the question confirmation *ba*, are likewise homonyms, but can be told apart by the associated semantics and sentence intonation. Finally, the *yes/no*-question force head *ma* and the dogmatic assertion attitude head *ma* (where the speaker insists on her/his opinion) can be easily distinguished by the

resulting interpretation and different intonational contours. Although homonymy is a widespread phenomenon in Mandarin Chinese, the homonymy between SFPs belonging to different projections has led to quite a confusion in the literature and must be carefully controlled for. Finally, the table in (9) foremost captures the relative hierarchy between SFPs and is not meant to imply that they can all co-occur, given the semantic constraints observed for each SFP further discussed below.

2.1 Low CP

2.1.1 *The C-heads* láizhe$_1$, ne$_1$, le *as "expressing tense" (Zhu 1982)*

Zhu (1982: 208) characterized the three SFPs *ne$_1$, le*, *láizhe$_1$* as "expressing tense", based on the different interpretations obtained in the triple below, where the (bare) lexical predicate *xià yǔ* 'fall rain' remains constant and only the SFPs vary. Note that the SFPs are obligatory here, a fact not explicitly mentioned by Zhu (1982).[5]

(10) a. [$_{ClowP}$ [$_{TP}$ Xià yǔ] ne. (Zhu 1982: 209)
fall rain CLOW
'It's (still) raining.' (Zhu's comment: It was raining before.)

b. Xià yǔ le.
fall rain CLOW
'(Look), it's raining (now).' (Zhu's comment: It didn't rain before.)

c. Xià yǔ láizhe.
fall rain CLOW
'It (just) rained.' (Zhu's comment: It rained a moment ago.)

On the basis of these examples, Zhu (1982: 209) proposes the following interpretative values for the three SFPs. *Láizhe* indicates that the event has occurred in the recent past. *Le* signals that the situation at hand is (conceived of as) new.[6] *Ne$_1$* expresses a continuing situation or state. Importantly, this is not tantamount to postulating *tense* as a *verbal* category for Chinese. Instead, these characterizations rather attempt to capture the semantic import of the SFPs, which is also reflected in the constraints observed for the type of TP each SFP can select, to be examined in detail further below.

As reflected in Zhu's comments and the translations provided, the low Cs *inter alia* differ in whether or not the event held in the past, and in whether the event

5. Unlike statives, activity predicates must bear aspect markers for a non-habitual, episodic reading (cf. Kong 1994, Sun 2014). Otherwise, a low C is required (cf. Paul 2018).

6. "Conceived of as new" refers to the subjective perception of the speaker, i.e. (10b) is also compatible with a situation where it might have rained before, but where the speaker notices it only at this moment (hence Li & Thompson's [1981: 238] label "currently relevant state").

holds at speech time or not. While these two values are obvious in the translations of (10a) and (10b), i.e. 'It's (still) raining.' and '(Look), it's raining (now).', (10c) with the low C *láizhe* requires a further comment. More precisely, for *láizhe* the default interpretation is that the event no longer holds at the speech time, as evidenced by the following mini-dialogue (but cf. [15] below):

(11) A: Wàimiàn dì zěnme shī le?
outside ground how humid CLOW
'How come the ground is humid outside?'
B: Wàimiàn xià yǔ láizhe.
outside fall rain CLOW
'It (just) rained outside.'

The second triple provided by Zhu (1982) with *mén kāi* 'door open' as lexical material allows us to further sharpen the differences between the three low Cs. We elaborate on (i) the interaction of low Cs with TP-internal properties as illustrated by certain incompatibilities, (ii) the difference between the low Cs and aspect as a verbal category, and (iii) thus make more precise the import and role of the SFPs.

We again start with the low C *ne* as the most straightforward case:

(12) Mén kāi -zhe ne. Nǐ wèishénme qiāo mén?
door open-IMP CLOW 2SG why knock door
'The door is (standing) open. Why do you knock?'

As stated for *ne* in (10) above, the eventuality holding now, i.e. state of standing open, already held in the past, as explicitly mentioned by Zhu (1982: 209).

This contrasts with *le* in (13), as shown by Zhu's comment enclosed in square brackets:

(13) Mén kāi le. [Yuánlái guān-zhe.] Wǒmen jìnqù ba.
door open CLOW before close-IMP 1PL enter FORCE
'The door is open (now). [It was closed before.] Let's go in.'

As expected, *(mén) yuánlái guān-zhe* '(the door) was closed before' would be unfelicitous as a continuation for (12) *Mén kāi-zhe ne* 'The door is (still) open'.

Turning now to *láizhe*, we first illustrate (in a mini-dialogue) the default reading where the eventuality holding in the past no longer holds at the speech time.

Context: A says that B just went to C's office for some documents, but couldn't get in.

(14) C answers:
Shì ma? Mén kāi-zhe láizhe, tā wèishénme jìn-bù-lái?
be FORCE door open-IMP CLOW 3SG why enter-NEG-come
'Is that so? The door was open, (so) why couldn't he get in?'

While in the default reading the eventuality holding in the past no longer holds at the speech time, it is not excluded for the eventuality to still hold at speech time; as a result, both continuations in (15) are felicitous:

(15) Mén kāi -zhe láizhe, xiànzài guān-shàng -le / xiànzài yě
door open-IMP CLOW now close-ascend-PERF/ now also
kāi -zhe ne.
open-IMP CLOW
'The door was (standing) open, now it's closed/ and it's still open now.'

To wrap up our results so far, the crucial differences between the three low Cs consist in whether the event holds at speech time or not and whether the event held before or not. For *ne*$_1$, the event still holds at speech time and likewise held in the past. For *le*, the event holds at speech time and did not hold in the past. For *láizhe*, the event held in the past and may or may not hold at speech time (with the latter as default case).

2.1.2 *The selectional properties of the low Cs*

Against this backdrop, we now examine the selectional properties of each low C, which at the same time highlight that the low Cs themselves can *not* be analyzed as aspect markers (*pace* a.o. Zhang 2019).[7]

As demonstrated by Yan (2018: § 3.2.2, § 7.2.2), *ne*$_1$ only allows for atelic activity predicates in the TP-complement (cf. [18a-c]), and excludes telic (cf. [16]) as well as stative predicates (i.e. stative verbs in [17b] and adjectives in [17a]):

(16) a. *[$_{\text{ClowP}}$ [$_{\text{TP}}$ Xiǎo Wáng *líkāi* gōngchǎng] ne].
Xiao Wang leave factory CLOW
'Xiao Wang is leaving the factory.'
b. *Xiǎo Wáng *xuéhuì* hànyǔ ne.
Xiao Wang acquire Chinese CLOW
'Xiao Wang is acquiring/mastering Chinese.'

7. Zhang (2019) concentrates on the finiteness issue in *complement* clauses and claims that when a complement clause does not allow for a low C (her "sentence-final aspect particles"), then it is automatically non-finite. This claim is too simplistic and incorrect and not borne out by the data. Many matrix clauses are perfectly grammatical without any SFPs (cf. [18c], [19b], [26a]), and likewise acceptable in the same form as complement clauses. Furthermore, the SFPs *le*, *ne*, *láizhe* are mechanically tested by Zhang in all kinds of sentences, without their associated semantic constraints being taken into account. Accordingly, in many cases, the SFP in question is simply incompatible with the TP-internal predicate in both the complement clause and the corresponding matrix clause; accordingly, this incompatibility cannot be taken as a diagnostic for the non-finite status of the clause at hand.

(17) a. Tā hěn cōngmíng (*ne).
3SG very intelligent CLOW
'She is bright.'
b. Tā hěn xǐhuān shùxué (*ne).
3SG very like mathematics CLOW
'She likes mathematics.'

(18) a. [$_{ClowP}$[$_{TP}$ Xiǎo Wáng *cānguān* gōngchǎng] ne].
Xiao Wang visit factory CLOW
'Xiao Wang is visiting the factory.'
b. Xiǎo Wáng *xuéxí* hànyǔ ne.
Xiao Wang learn Chinese CLOW
'Xiao Wang is learning Chinese.'
c. Tā zhèng zài tiē-zhe biāoyǔ (ne). (Zhu 1982: 210)
3SG just PROGR paste-IMP poster CLOW
'He is pasting posters right now.'

The presence of the progressive aspect auxiliary *zài* in (18c) demonstrates that aspect as a verbal category is distinct from the low C *ne*. Importantly, *ne* in (18c) is optional, confirming our point just made. This contrasts with (18a)–(b) and with Zhu's (1982) (10a) above where *ne* is obligatory, given the bare nature of the activity predicate.

With respect to the alleged status of low Cs as aspect markers, the low C *le* has caused quite some confusion, due to its homonymy with the perfective aspect verb suffix *-le* . Although the distinctness of the SFP *le* and the aspectual suffix *-le* was established a long time ago (cf. a.o. Chao 1968: 246; Teng 1973; Chan 1980; Li & Thompson 1981: 296; Zhu 1982), claims that both items instantiate one and the same category regularly make their reappearance in the literature (like the famous Loch Ness monster). This seems particularly futile given the many cases provided in the literature where the aspectual suffix *-le* and the SFP *le* co-occur (cf. Paul 2015: 276–277 for further discussion):

(19) a. Wǒ zài zhèr zhù -le wǔ nián le. (Zhu 1982: 209)
1SG at here live -PERF 5 year CLOW
'I have been living here for five years now.'
b. Wǒ zài zhèr zhù -le wǔ nián.
1SG at here live -PERF 5 year
'I lived here for five years.'

Zhu's example (19a) nicely illustrates both the very common co-occurrence of the aspectual suffix *-le* and the SFP *le* and the semantic import of *le*. Given that *le* relates the event to the speech time, (19a) with *le* unambiguously states that my living here

still obtains at the speech time. By contrast, as pointed out by Zhu (1982: 209), (19b) without the SFP *le* implies my no longer living here.[8]

(20) illustrates that the meaning of a sentence with an SFP is derived in a clearly compositional way, with *le* as C having scope over the entire TP:

(20) [$_{\text{ClowP}}$ [$_{\text{TopP}}$ Nà [$_{\text{Top'}}$[$_{\text{TP}}$ wǒ jiù bù děng tā] le]]].
in.that.case 1SG then NEG wait 3SG CLOW
'In that case I will no longer wait for him.'

Le signals that the proposition 'I won't wait for him' obtains at the speech time (in the absence of any other reference time), which leads to 'I will no longer wait for him'.

Finally, when an explicit reference time (different from the speech time) is provided ('as soon as I rang the bell'), *le* relates the event to that time:

(21) [$_{\text{CPlow}}$ [$_{\text{TopP}}$ [$_{\text{TP}}$ Wǒ yī àn mén-líng] [$_{\text{Top'}}$ [$_{\text{TP}}$ tā jiù lái kāi
1SG once ring door-bell 3SG then come open
mén] le]]]
door CLOW
'As soon as I rang the door bell, he came and opened the door.'
(slightly modified example from Chao 1968: 799)

Unlike *le*, *láizhe* by default indicates that the event that held in the past no longer holds at speech time. Accordingly, *le* in (22a) is unacceptable, because *gāngcái* 'just now, a moment ago' explicitly locates the event in the past, whereas *le* relates the very same event to the speech time. This is not the case for *láizhe*, hence its acceptability.

(22) a. Tā gāngcái hái zài bàngōngshì láizhe /*le.
3SG just.now still at office CLOW / CLOW
'He was in his office just now.' (Paul & Pan 2017: 58, [24])
b. [$_{\text{ClowP}}$[$_{\text{TP}}$ Xiǎo Wáng cānguān /* líkāi gōngchǎng] láizhe].
Xiao Wang visit / leave factory CLOW
'Xiao Wang visited/left the factory.'
c. Tā yǐqián xǐhuān wǒ láizhe.
3SG before like 1SG CLOW
'She liked me before.'
d. Tā qùnián hěn pàng láizhe.
3SG last.year very fat CLOW
'He was overweight last year.'

8. For some speakers (19b) might be ambiguous and then in addition also has the reading in (19a).

As shown in (22b)–(d), *láizhe*, like *ne*, excludes telic activity predicates (cf. [22b]), but is compatible with stative predicates (cf. [22c]–[d]).

The approximate, "shorthand" characterization of *láizhe* as "recent past" makes it very tempting to consider it as a genuine tense marker. However, there are at least two pieces of evidence challenging this idea. First, in a sentence with *láizhe*, the presence of (past) temporal adverbs and of aspect is in fact preferred:

(23) a. Wàimiàn gāngcái xià-guò yǔ (láizhe).
outside just fall-EXP rain CLOW
'It just rained a moment ago.'

b. Tā {zuìjìn/shàng ge yuè} qù-guò gùgōng (láizhe).
3SG recently/last CL month go-EXP imperial.palace CLOW
'She went to the imperial palace recently/ a month ago.'

This shows that the temporal interpretation of the event is based on the TP-internal material, a fact confirmed by the optionality of *láizhe* in (23a)–(b). Also note that "recent past" is a flexible notion and not limited to adverbs such as *gāngcái* 'just now' and *zuìjìn* 'recently'; instead, what counts as "recent past" also depends on the speaker's judgement of the immediacy of the event at hand, as witnessed by the acceptability of 'last month' in (23b) (cf. Song 1981: 272).

Secondly, in addition to the "recent past" feature, *láizhe* asserts that the event has taken place (cf. Song 1981: 275; Lü 2000: 348–349). As a result, *láizhe* is incompatible with a TP whose predicate is negated. This well-known observation from the literature was confirmed by the acceptability judgement test with eight native speakers (average age around 28 years) who uniformly rejected negation in *láizhe* sentences (cf. the first clause in [25b]) and fully accepted the assertion strengthening function of *láizhe* in [24b]):[9]

(24) A: Nǐ shì bù shì qù kàn diànyǐng le?
2SG be NEG be go watch movie CLOW
'You went to the movies, didn't you?'
(Literally: 'Is it the case or not that you went to the movies?')

B: Wǒ zài jiā zuò zuòyè láizhe, méi qù kàn diànyǐng.
1SG at home do homework CLOW NEG go watch movie
'(In fact) I did my homework at home, I didn't go to the movies.'

(25) A: Nǐ shì bù shì zài jiā zuò zuòyè le?
2SG be NEG be at home do homework CLOW
'You did your homework at home, didn't you?'

9. This observation seemed worthwhile checking, given that the relevant literature is from 40 years ago (the first edition of Lü [2000] dating back to 1980).

B: *Wǒ méi zài jiā zuò zuòyè láizhe. (Wǒ qù
1SG NEG at home do homework CLOW 1SG go
kàn -le diànyǐng.)
watch-PERF movie
'(In fact) I didn't do my homework at home. (I went to the movies.)'

In B's response in (24b), *láizhe* strengthens the assertion and thereby corrects A's wrong assumption. In (25b), the first clause is unacceptable, due to the conflict between the assertion strengthening component of *láizhe* and the negated predicate.

2.1.3 *Low CP: Interim summary*

The low Cs *láizhe*$_1$, *le* and *ne*$_1$ all interact with TP-internal material and impose constraints on the properties of the extended verbal projection, including its aktionsart. However, they clearly occupy a TP-external position in the left periphery (contra Erlewine (2017); cf. Pan (2019b) for a critical review).

The low Cs themselves do neither encode aspect nor tense, as evidenced *inter alia* by their co-occurrence with aspect markers on the verb. Instead, low Cs indicate whether the event in question holds at the speech time and/or whether it held before. In joint ongoing work with Gillian Ramchand (University of Tromsø), this is taken to indicate that – roughly speaking – the low Cs are overt versions of the (non-default) anchor, the default anchor being NOW, a moment. Low Cs thus contribute to the finiteness of a sentence in the absence of aspect markers, which either turn the event into a state (able to be true at a moment) or indicate a temporal precedence relation (as in the case of *-le*).[10] The temporal interpretation of a sentence obtains as resulting from the interaction between the properties of the TP-internal predicate (bare or not bare, state or non-state), on the one hand, and the precise nature of (the stative intermediate reference situation introduced by) the low C.

2.2 SFPs realizing ForceP

2.2.1 *The* yes/no*-question Force head* ma$_2$

The SFP *ma*$_2$ indicating the *yes/no* question status of a sentence (cf. [26b]) was the first SFP to be analyzed as a complementizer (cf. Lee 1986; Tang 1989: 540):

10. Assertability in a root context is the most language general and theory neutral definition of *finiteness*. Furthermore, habituals and negation of events can be likened to states, as widely assumed in the literature.

(26) a. Tā huì shuō zhōngwén.
3SG. can speak Chinese
'He can speak Chinese.'
b. [$_{CPforce}$ [$_{TP}$ Tā huì shuō zhōngwén] ma]?
3SG can speak Chinese FORCE
'Can he speak Chinese?'

Since *ma* turns a declarative sentence into a *yes/no*-question, it must have scope over the entire sentence, whence the analysis of *ma* as a C-head taking a clausal complement (TP or ClowP, cf. [28] below). The complement status of TP and the head status of *ma* are confirmed by the fact that *ma* imposes selectional restrictions: it can only select a non-interrogative TP and is therefore incompatible with *wh*-questions (cf. [27a]) and TP-internal *yes/no*-questions in the 'A-not-A' form (cf. [27b]).

(27) a. [$_{TP}$ Nǐ wèn-le *shéi*] (*ma)?
2SG ask-PERF who FORCE
'Whom did you ask?'
b. [$_{TP}$ Tā *dǒng bù dǒng* wèntí] (*ma)?
3SG understand NEG understand problem FORCE
'Does he understand the problem?'

The Force head status of *ma* is confirmed by its position above, i.e. to the right of, low Cs such as *le* (cf. [28a]) and *ne* (cf. [28b]):

(28) a. [$_{ForceP}$ [$_{ClowP}$ [$_{TP}$ Tā bù chōu yān] le] ma]?
3SG NEG inhale cigarette CLOW FORCE
'Does he no longer smoke?'
b. [$_{ForceP}$ [$_{ClowP}$ [$_{TP}$ Tā hái méi zǒu] ne$_1$] ma] ?!
3SG still NEG leave CLOW FORCE
'Hasn't he left yet?!'
c. Kuài le, kuài le, tā ná -wán shū mǎshàng jiù zǒu.
quick CLOW quick CLOW 3SG take-finish book immediately then leave
'Almost there, almost there (i.e. 'He's nearly ready to leave), he finishes taking his books and then leaves at once.'

As indicated by the question-plus-exclamation mark, (28b) requires an angry or surprised intonation and can be continued by another speaker's uttering (28c).

2.2.2 *The Force head* ba$_{Qconfirmation}$*: confirmation request or conjecture*

A *yes/no*-question with *ba*$_{Qconfirmation}$ is not neutral, but implies the speaker's expectation to receive a positive answer to her/his request:

(29) Jīntiān xīngqīsān ba? (Zhu 1982: 211)
today Wednesday FORCE
'It is Wednesday today, correct?'

(30) Nǐ xiànzài míngbái le ba? (Yang-Drocourt 2007: 312)
2SG now understand CLOW FORCE
'You understand now, don't you?'

It is this component of confirmation request which explains why *ba*$_{\text{Qconfirmation}}$ is incompatible with *wh* questions and *yes/no* question in the 'A-not-A' form, both being genuine information seeking questions.

(31) a. *Shéi míngbái ba?
who understand FORCE
b. *Nǐ míngbái bù míngbái ba?
2SG understand NEG understand FORCE

Lü (2000: 57) provides neat minimal pairs where either both *ba*$_{\text{Qconfirmation}}$ and *ma* are possible (*modulo* the associated meaning differences) or where only *ba*$_{\text{Qconfirmation}}$ is acceptable:

(32) a. Zhèi zuò fángzi shì [xīn gài de] ma?
this CL house be new build SUB FORCE
'Is this house a new one?
b. Zhèi zuò fángzi shì [xīn gài de] ba?
this CL house be new build SUB FORCE
'This house is a new one, isn't it?'

While (32a) with *ma* is a genuine information request, this is not the case for (32b) where a positive answer is expected. Accordingly, only *ba*$_{\text{Qconfirmation}}$, but not *ma* is compatible with adverbs such as *dàgài* 'probably', *yěxǔ* 'perhaps', *shuōbùdìng* 'possibly, perhaps':

(33) [Tā dàgài yǐjīng zǒu-le] ba / *ma?
3SG probably already leave-PERF FORCE/ FORCE
'She has already left, I guess?'

(34) [Xiànzài shuōbùdìng yǐjīng guò -le shí'èr diǎn] le {ba /*ma}?
now perhaps already pass-PERF 12 o'clock CLOW FORCE/FORCE
'It might very well be past twelve o'clock now?'

When *ba*$_{\text{Qconfirmation}}$ occurs with declarative sentences, its conjecturing component results in a weakening of the assertion (cf. Hu Mingyang 1981: 416):

(35) Nǐ tīngcuò -le ba.
2SG mishear-PERF FORCE
'You must have misheard.'/ 'You probably have misheard.'

Being a Force head, *ba* – be it as question confirmation or conjecture – occurs above low Cs such as *le* and *ne*$_1$:

(36) a. [$_{\text{ForceP}}$ [$_{\text{ClowP}}$ [$_{\text{TP}}$ Sānshí nián qián hái méi yǒu shǔbiāo] ne$_1$] ba].
30 year ago still NEG exist mouse CLOW FORCE
'Thirty years ago there probably didn't exist any computer mouse yet.'
(Paul & Pan 2017: 67, [50])

b. [$_{\text{ForceP}}$ [$_{\text{ClowP}}$ [$_{\text{TP}}$ Èrshí nián qián fāmíng-le shǔbiāo]
20 year ago invent-PERF mouse
{le /*ne$_1$ }] ba].
CLOW/ CLOW FORCE
'Twenty years ago they had probably invented the computer mouse.'

(36a) is read in a low contour, in accordance with *ba* expressing a conjecture rather than asking for a confirmation. The default stress lies on *sānshí nián qián* '30 years ago', but stressing *méi yǒu* 'not exist' or even *shǔbiāo* 'mouse' might also be possible. The analysis of *ne* as a low C here is confirmed by its unacceptability in (36b) with the telic predicate *fāmíng-le* 'have invented', given that *ne* requires atelic predicates in its TP-complement (cf. [16] – [18] above).

2.2.3 *The Force head* ba$_{Imp}$: *advice or suggestion*

The SFP *ba*$_{\text{Imp}}$ is called "advisative" by Chao (1968: 807) because of its "softening" effect. Accordingly, an imperative containing *ba*$_{\text{Imp}}$ is understood as less harsh an order than the corresponding imperative sentence without *ba*$_{\text{Imp}}$ (also cf. Hu 1981: 416):

(37) a. [Kuài diǎnr zǒu] ba!
quick a.bit go FORCE
'Better hurry up and go!' (Chao 1968: 807)

b. [[[Bié chàng] le] ba]!
NEG sing CLOW FORCE
'Better stop singing.' (Hu 1981: 416)

Again, the rigid ordering with respect to the low C *le* (cf. [37b] above) and the Attitude head *ou* (cf. [38] below) confirms the status of *ba*$_{\text{Imp}}$ as a Force head in the second CP-layer:

(38) Zǒu b'ou [= ba + ou] (Zhu 1982: 208)
go SFP(fusion) FORCE+ATT
'You better go!'

Finally, (39) below illustrates the role the semantic (in)compatibility plays, in addition to the hierarchy in the split CP. While the order 'low C < Force' in itself is correct, the sequence '*ne*$_1$ *ba*$_{\text{advis}}$' is ruled out for semantic reasons: an event in the future whence neither holding in the past nor ongoing in the present is incompatible with the low C *ne*$_1$, while this type of event is fine for the imperative advisative *ba*. This conflict leads to an unacceptable sentence:

(39) *[$_{\text{ForceP}}$ [$_{\text{ClowP}}$ [$_{\text{TP}}$ Wǒmen qù diànyǐngyuàn kàn diànyǐng] ne] ba]!
1PL go movie.theatre watch movie CLOW FORCE

This semantic incompatibility is a very robust phenomenon confirmed by all of the eighteen native speaker participants, who stated they could *not* assemble a correct sentence by using *all* of the lexical items provided, as required in this sentence assembly task.

2.2.4 *ForceP: Interim summary*

The observations in this section, mainly based on widely used grammars and grammar manuals, straightforwardly invalidate Li's (2006: 171) claim that the clause-typing heads always remain covert in Mandarin and Cantonese (whereas they may be realized overtly in Wenzhou). Given the semantically transparent and extremely well-documented *yes/no* question Force head *ma*$_2$ this is a very surprising statement. In particular, Li (2006: 64–65) doesn't see that there are two SFPs *ma*, the *yes/no* question Force head *ma*$_2$ and the Attitude head *ma*$_3$ (cf. Section 3 immediately below), despite the well-established difference between the two (cf. a.o. Chao 1968: 801). Both Del Gobbo et al. (2015) and Bailey (2015) adopt Li's (2006) incorrect claim that Chinese has no SFPs realizing Force such as imperative and interrogation. While Del Gobbo et al. (2015: 378) see this as a parallel with sentential particles in Romance, Bailey (2015: 420) considers it a general characteristic of final question particles in VO languages that they are in fact markers of "something other than interrogative force".

2.3 Attitude Phrase

The SFPs instantiating AttitudeP involve both speaker and hearer, via the speaker's assumptions concerning the beliefs of the hearer. Again, Chinese is not unique in this respect, given that e.g. Japanese (cf. Endo 2007: 175–198) as well as Romance and Germanic languages likewise display particles in the sentence periphery encoding properties of the speaker-hearer interaction. Examining Romanian and West-Flemish, Haegeman and Hill (2013) postulate the projection DiscourseP, equivalent in function to AttitudeP in Chinese. Importantly, the characteristics of

SFPs realizing DiscourseP established by Haegeman and Hill (2013) also hold for Attitude SFPs in Chinese.

First, AttitudeP does not concern nor affect the truth value of the proposition at hand. This contrasts with the SFPs instantiating ForceP, where as we have seen $ba_{\text{Qconfirmation}}$ conveys the speaker's belief that the proposition is true, and *ma* is a request as to the truth value of the proposition (*yes/no*). It is correct that an SFP such as the advisative ba_{Imp} also conveys the speaker's (friendly) attitude, but at the same time this SFP is linked to a particular sentence type, i.e. the imperative. Furthermore, its status as Force head is confirmed by its obligatorily preceding Attitude SFPs such as *ou* (cf. [7a] above). As for low C, *láizhe* 'recent past' was shown to be incompatible with TP-internal negation, due to its event assertion feature (cf. Section 2.1.2 above). Attitude SFPs are thus fundamentally distinct from both low C and Force heads, an observation already made by Zhu (1982: 208), although not further elaborated upon.

Second, Attitude SFPs indicate the speaker's commitment to the sentence content; they are interactional and imply the obligatory presence of a hearer (hence infelicitous in broadcasts).

Third, Attitude SFPs are deictic, i.e. they are directly correlated with the speech act, but do not require a preceding utterance as "trigger". Finally, Haegeman and Hill (2013) concede that it is difficult to determine exactly the interpretive properties of Attitude SFPs, even though their semantic import is clearly discernible when comparing sentences with and without them. This leads to the fourth characteristic, which is the "optionality" of Attitude heads. A *caveat* is necessary here, though; if one wants to signal the discourse function associated with a particular Attitude SFP, then the presence of this SFP is evidently required.

2.3.1 *The Attitude heads* ne$_3$ *and* bale

Note first of all that *ne*$_3$ is *not* a "wh-question particle", i.e. it is *not* a Force head indicating the sentence-type (*pace* Cheng 1991), a fact again well-documented in the literature (cf. a.o. Hu 1981: 418; Paris 1981: 389; Li and Thompson 1981: 305; Lin 1984: 220; also cf. Pan and Paul 2016). In other words, in a *wh*-question (cf. [40]) or in an 'A-*not*-A' polar question (cf. [41]), the Attitude head *ne*$_3$ is not obligatory, for the simple reason that *ne*$_3$ does not encode the interrogative force. However, if one wants to signal the discourse function associated with *ne*$_3$, which *inter alia* is to solicit the co-speaker's attention, rendered here by "listen, and you…", it is evidently obligatory (cf. a.o. Wu 2005; Li 2006; Pan 2011):

(40) a. Nǐ zuì xǐhuān hē nǎ ge páizi de déguó píjiǔ?
2SG most like drink which CL brand SUB German beer
'Which brand of German beer do you like most?'

b. Nǐ zuì xǐhuān hē nǎ ge páizi de déguó píjiǔ ne?
2SG most like drink which CL brand SUB German beer ATT
'Listen, and you, which brand of German beer do you like most?'

(41) a. Tā huì bù huì shuō bāfālìyàyǔ?
3SG can NEG can speak Bavarian
'Can he speak Bavarian?'
b. Tā huì bù huì shuō bāfālìyàyǔ ne?
3SG can NEG can speak Bavarian ATT
'And he, can he speak Bavarian?'

Ne$_3$ also occurs in rhetorical questions:

(42) Wǒ zěnme bù jìde ne?! (Zhou & Shen 2006: 121)
1SG how NEG remember ATT
'How would I not remember [it]?!'

Being an Attitude head, *ne*$_3$ can naturally also combine with a non-interrogative complement, further invalidating its alleged status as a "clause typer" for *wh*-questions. It then expresses an exclamation/exaggeration or conveys a boasting tone (cf. [43]) and is obligatory in the presence of the speaker-oriented emphatic adverb *kě* 'really' (cf. [44]):

(43) [$_{TP}$ Tā huì kāi fēijī] ne! (Zhu 1982: 213)
3SG can drive airplane ATT
'(Imagine) he can fly an airplane!'

(44) Déguó yǔyánxuéjiā kě duō *(ne)!
German linguist really many ATT
'There really are a lot of German linguists!' (Paul & Pan 2017: 55, [14])

Zhu (1982: 213) also provides the neat minimal pair below (slightly changed) where *ne*$_3$ alternates with *bàle*, the latter being paraphrasable as 'that's all there is to it' and having the effect of "downplaying", which is exactly the opposite of the boasting tone mediated by *ne*$_3$:

(45) Tāmen yào wǔbǎi kuài qián ne! Bù shì ge xiǎo shùmù!
3SG want 500 CL money ATT NEG be CL small sum
'They want (as much as) 500 dollars! That's not a small sum!'

(46) Tāmen yào wǔbǎi kuài qián bàle! Méi yǒu shénme liǎobùqǐ!
3SG want 500 CL money ATT NEG have what extraordinary
'They (only) want 500 dollars! That's nothing extraordinary!'

Finally, *ne*$_3$ as an Attitude head can co-occur with low Cs such as *le* and must follow them:

(47) [$_{AttP}$ [$_{ClowP}$ [$_{TP}$ Shànghǎi de gōngyuán kě dà] le] ne]!
Shanghai SUB park really big CLOW ATT
'Shanghai's parks are really extremely big!'

The semantics of *ne*$_3$ and its being able to co-occur with low Cs clearly distinguishes it from the low C *ne*$_1$ and warrants the status of the former as an Attitude head. This result also challenges a uniform analysis of *ne*$_1$ and *ne*$_3$ (cf. a.o. Hu 1981; Wu 2005).[11]

2.3.2 *The Attitude head* ma

The Attitude head *ma* (henceforth *ma*$_{Att}$) implies that the speaker presupposes the hearer *not* to be up to date and provides a correction of the hearer's belief, conveying something like 'this is self-evident', 'you should know' (cf. Chao's 1968: 801 term "dogmatic assertion"):

(48) Tā bù shì Lǎolǐ ma? Ràng tā jìnlái ma$_{Att}$.
3SG NEG be Laoli FORCE let 3SG come.in ATT
'Isn't that Laoli? Let him come in. (Why do I have to tell you?)' (Lü 2000: 375)

(49) Wǒ shuō jīntiān shì xīngqīsān ma$_{Att}$! Nǐ shuō bù shì!
1SG say today be Wednesday ATT 2SG say NEG be
'I say it's Wedndesday today! You say it isn't!' (Zhu 1982: 213)

The Attitude head *ma*$_{Att}$ is clearly distinct from the Force head *ma* encoding *yes/no*-questions, as generally acknowledged in the literature (cf. a.o. Chao 1968: 800–801; Zhu 1982: 211–213; Lü 2000: 375–376) and nicely illustrated by (48) with both

---

11. Constant (2014) goes even further and incorrectly conflates the Attitude head *ne*$_3$ with the topic head *ne* (cf. [i]). More precisely, both are claimed to be instances of the *contrastive topic* (CT) and realizations of the CT operator (distinct from the "aspectual" SFP *ne*$_1$) (p. 438). Cf. Paul (2014; 2015, Chapter 6) for showing that the allegedly inherent contrastive value of the topic head *ne* is not borne out by the data; *inter alia*, Top° *ne* can host sentential adverbs (where [± contrast] does not apply) in its specifier, and Top° *ne* can co-occur in the same sentence with *ne*$_3$:

(i) [$_{TopP}$ Quèshí/qíshí [$_{Top'}$ [$_{Top°}$ ne][$_{TP}$ tā de nénglì shì bǐ wǒ qiáng]]]
indeed/ in.fact TOP 3SG SUB ability be compared.to 1SG strong
'Indeed/In fact, his abilities *are* greater than mine.'

(ii) [$_{AttP}$ [$_{ClowP}$ [$_{TopP}$ Nǐ [$_{Top'}$ [$_{Top°}$ ne] [$_{TP}$ nǐ wèn shéi] le] ne]?
2SG TOP 2SG ask who CLOW ATT
'And you, whom have you asked?'

Constant's (2014) conflation also glosses completely over the well-known fact that TopP is head-*initial*, while AttP is head-*final*, and over the associated consequences. Furthermore, Wei and Li (2018: 206) highlight an important tonal difference between the Top° *ne* and the Att° ne$_3$, which holds regardless of the tone of the preceding syllable; while the Top° *ne* is always pronounced in a high tone, the Att° *ne*$_3$ is always pronounced in a low tone.

SFPs in successive sentences. This invalidates Li (2006: 64–65) who postulates a single *ma* "mark[ing] a high degree of the strength of the assertive or directive force".[12]

2.3.3 *The Attitude head* láizhe

As shown by Pan (2019a: 109), in addition to the low C *láizhe* 'recent past', there is also an Attitude head *láizhe*:

(50) [$_{AttP}$ [$_{TP}$ Tāmen liǎ shénme shíhou jié hūn] láizhe] ?
3PL two what time join marriage ATT
'By the way, when will they get married?'[13]
(Pan 2019a: 109, [40a]; bracketing simplified)

While Pan (2019a: 109) renders this *láizhe* as 'by the way', we think it rather refers to a former state of knowledge and accordingly can be paraphrased as 'What did you say' or 'What was + sentence', indicating that the speaker no longer exactly recalls the sentence or prior knowledge. This is particularly neat in (51a) below:

(51) a. [Nǐ xìng shénme] láizhe? (Chao 1968: 810)
2SG call what ATT
'What (did you just say) is your family name?'
'What was your family name?' (I forgot.)
(NOT: What was your family name before?')
b. [Wǒ yǐqián xìng Zhāng] (láizhe), xiànzài xìng Bāo.
1SG before call Zhang CLOW now call Bao
'My family name used to be Zhang, now it's Bao.'

This "interlocutive" *láizhe* is clearly different from the low C *láizhe*; accordingly, (51a) is not a question concerning somebody's family name in the past. The latter requires a temporal adverb such as *yǐqián* 'before, earlier' as in (51b), where the low C *láizhe* is optional.

---

12. Li (2006) also glosses over the intonational difference observed by Chao (1968: 801). While the Force had *ma* is associated with a fairly high intonation ending in a slight drawl, the Attitude head *ma* is always short.

13. Implementing the generalization in Paul (2015: 285) that only low Cs may occur in non-root clauses, to the exclusion of Force and Attitude heads, Pan (2019a: 109) substantiates the Attitude head status of *láizhe* in (50) by showing its unacceptability in embedded contexts such as relative clauses:

(i) *[DP [AttP [TP Tāmen liǎ shénme shíhou jié hūn] láizhe] de wèntí]
3PL two what time join marriage ATT SUB question
bù qīngchǔ
NEG clear.
('The question (*by the way) when will they get married is not really clear.')
(Pan 2019a: 109, [40b])

Importantly, being an Attitude head, the interlocutive *láizhe* can co-occur with, i.e. follow the low C *le*, something completely excluded for the low C 'recent past' *láizhe* (cf. [6b] above):

(52) [$_{AttP}$ [$_{CLOWP}$ [$_{TP}$ Xiǎo Wáng qù nǎr mǎi cài] le] láizhe]? Wǒ
Xiao Wang go where buy vegetables CLOW ATT 1SG
xiǎng-bù- qǐlái le.
think-NEG-rise CLOW
'Where was it/did you say that Xiao Wang went buying groceries? I (simply) cannot recall.'

As to be expected, the inverse order '*láizhe le*' is ruled out. This also holds for those speakers who only marginally accept (52); they categorically reject '*láizhe le*', given that there is no Attitude head *le* and that accordingly, '*láizhe le*' can only be parsed as the illegitimate sequence of two low Cs. Finally, note that the examples illustrating the interlocutive *láizhe* are all questions, showing that they involve the hearer, either a real or an imaginary one, because (52) can also be a self-directed question or musing.

2.3.4 *The Attitude head* a

The SFP *a* has rather complicated morphophonemics depending on the preceding word, which is often reflected in different transliterations: *ia*, *(u)a*, *(n)a*, *(ng)a* etc. (cf. Chao 1968: 803; Zhu 1982: 212; Yang-Drocourt 2007: 192–195 for detailed discussion). For ease of exposition, we gloss over these phonological alternations and use the transliteration *a* throughout.

The SFP *a* is rather ubiquitous and occurs with all kinds of sentence types (declaratives, questions, imperatives, exclamatives), which makes its semantic characterization very difficult. Scholars agree that *a* conveys the personal implication of the speaker and has a general softening effect; the different interpretations observed for *a* are then due to the different sentence types it combines with (cf. a.o. Chao 1968: 803–806; Zhu 1982: 212; Li & Thompson 1981: 313–317). For example, Chao (1968: 804) observes that a question with the SFP *a* is less blunt than one without it, an effect which can be paraphrased as 'by the way' or 'excuse me' etc.

(53) Nǐ míngtiān chūqù bù chūqù a?
2SG tomorrow go.out. NEG go.out ATT
'(By the way) are you going out tomorrow?'

Likewise, an imperative with the SFP *a* has less the flavour of a command than an imperative without it (though according to Chao [1968: 804] the softening effect with *a* is less strong than with the advisative *ba*$_{Imp}$ discussed above):

(54) Shuō a, bié hàipà a!
say ATT NEG be.afraid ATT
'Say it, don't be afraid!

In an exclamative, *a* expresses the emotion of the speaker which, depending on the sentence meaning, can be anger, astonishment, enthusiasm etc.:

(55) Nǐ kàn a, biànhuà duō dà a! (Yang-Drocourt 2007: 311)
2SG see ATT change much big ATT
'Look, how much everything has changed!'

Finally, *a* is also compatible with rhetorical questions (cf. Pan 2015: 855, [66]):

(56) Shéi bù xǐhuān chī tílāmǐsū a?!
who NEG like eat tiramisu ATT
'Oh, who doesn't like tiramisu?! = Everyone likes tiramisu!'

2.3.5 *Wrap-up*

The strict ordering observed by Zhu (1982, ch. 16) for the three classes of SFPs can be easily recast as a split CP à la Rizzi, *modulo* the addition of the projection AttitudeP (absent from Rizzi's original hierarchy) above ForceP. Importantly, studies on Romance and Germanic languages within Rizzi's split CP approach independently argue for the necessity of such a speaker/hearer related Discourse projection.

SFPs are clearly heads, because they impose selectional restrictions on their clausal complement (such as declarative or interrogative sentence type). Low Cs as (non-default) anchor interact with the properties of the extended verbal projection; depending on the (non-) bare nature of the predicate and its aktionsart (state vs non-state), they may be obligatory to make the sentence finite.

The detailed study of *ne* and *láizhe* has illustrated several problems encountered in the analysis of SFPs in general, among them the homonymy between C-elements instantiating different projections (here LowCP and AttitudeP). The SFPs ba_{Imp} and $ba_{\text{Qconfirmation}}$ reveal yet another analytic difficulty, namely the homonymy between SFPs belonging to the same projection, in this case ForceP.

The decision to be made for homonymous SFPs is further complicated by the interaction between the properties of each SFP, the sentence meaning itself, the sentence intonation and the context, all of which contribute to the interpretation obtained. As a consequence, it is not always easy to pin down the meaning component provided by the SFP itself. Besides, the use of SFPs, especially those realizing AttitudeP, is also subject to individual and regional differences which still remain to be elucidated. (In general, Northern speakers seem to use SFPs more frequently

than Southern speakers.) These caveats notwithstanding, it is evident that SFPs are an integral part of the syntax and as such subject to syntactic constraints, the most visible being the hierarchy of the different projections reflected in their rigid order.

3. The cartographic approach and the acquisition of SFPs

Having recast Zhu's (1982) three SFP-classes into a split CP à la Rizzi (1997), one might wonder whether another tenet of the cartographic approach likewise holds for Chinese, *viz.* that acquisition is "incremental" and proceeds layer by layer "up the tree", i.e. first the TP and then the periphery (cf. Friedmann et al. 2021: § 5.2.1). The left periphery in Hebrew is said to be acquired in two steps "defining two zones: first a lower LP zone including Fin, Mod, and Q and then a higher LP zone that includes Force, Int, and Top". Note that neither a head nor an entire zone can be "skipped". At the same time, the authors concede that the two zones do not form natural classes characterizable by a common feature, "because it includes topics, force markers (embedding markers), and operators such as yes/no operators (in embedded questions), relative clause operators, and *why*".

Although we concentrate on SFPs here only and abstract away from the acquisition of Topic Phrase (included in Friedmann et al.'s study), the Chinese data might nevertheless shed some light on their hypothesis. In particular, we will see that low C, Force head and Attitude heads are more or less acquired simultaneously. Since Chinese is a *pro*-drop language, on the surface a TP often simply consists of a bare VP. It is therefore difficult to decide whether indeed the full structure of the TP (including aspect, auxiliaries and negation) is acquired before the CP. Given that the child data also include utterances consisting of NPs plus the Attitude head *a* (cf. [65c] below), it is equally plausible to postulate the simultaneous acquisition of TP *and* CP (something which has to be assumed for the acquisition of Germanic V2 languages with V-to-C movement in any case).

Tao (2012) sets the onset for the productive use of SFPs at the age of 01;07 and reports the spontaneous use of the following SFPs before the age of two years by the four children examined: the low C ne_1 and the Attitude ne_3, the Force head *ma* (*yes/no* question) and the dogmatic assertion Attitude head *ma*, the two Force heads *ba*, i.e. advisative *ba* in imperatives and the confirmation request question *ba*.

(57)–(61) show a sample of the sentences produced by the children (cf. Tao 2012: 29–34):

(57) Chī táng ba. (ZTX 01; 08; 18)
eat candy FORCE
'(Let me) have some candy.'

(58) XXX, wǒ zài nǎr ne? (CY 01; 11; 17;)
XXX 1SG be.at where ATT
'XXX, where am I?' (playing hide and seek with the interviewer XXX)

(59) a. Adult: Béng wán'r le!
NEG play CLOW
'Do no longer play!' = 'Stop playing.'
b. Child: Hē shuǐ ne. (SJQ 01; 07; 16)
drink water CLOW
'I'm drinking water.' (i.e. the child is not playing with her water cup as implied by the adult's admonition.)

(60) a. Adult: Nǐ kàn bù shì huǒchái.
2SG see NEG be match
'You see, it is not the matches (that set the fire).'
b. Child: Shì huǒchái ma. (SJQ 01/ 10; 22)
be match ATT
'It is the matches (that set the fire).'

(61) a. Adult: Bù xǐhuān chī táng, shì bù shì?
NEG like eat candy be NEG be
'You don't like to eat candies, do you?'
b. Child: Chī táng ma! (ZTX 01; 08; 24)
eat candy ATT
'(Naturally) I eat candies!'

The children distinguish between the low C ne_1 in (59b) and the Attitude head ne_3 in (58), with both sentences exactly corresponding to the target adult grammar. Their mastering of the Attitude head *ma* is particularly noteworthy, because this implies their contradicting the previous adult utterance and insisting on their own utterance as the only truthful statement.

Guo (2016) and Peng (2016) report similar results from three Beijing Mandarin-speaking children aged between 1;3 and 3;1; in general, SFPs are acquired by the age of 2 years. For reasons of space, we only illustrate phenomena not included in Tao Yu's sample.

(62) (ZZC 1; 10; 13)
Méi diàn. Méi diàn la! [= le + a]!
NEG electricity NEG electricity SFP-fusion CLOW + ATT
'There is no electricity. There is no electricity anymore!'(Peng 2016: 118–119)

According to Peng (2016), this fused form of the low C *le* and the Attitude head *a* is the first SFP acquired by the children in her sample. Given that the same children also produce sentences with the low C *le* (cf. [64a] below), they clearly know the difference between the two.

(63) a. Māma zài zhè ma? (WYF 1; 10; 16)
mum be.at here FORCE
'Is mum here?'
b. Hǎo chǒu a ! (WYF 1; 09; 15)
so be.ugly ATT
'(This) is so ugly!'

(64) Wǒ zhǎodào māma le. (ZZC 2; 00; 21)
1SG find mum CLOW
'I have found mum.'

(65) a. Child: Yǒu jīmù. (ZZC 1; 10; 20)
exist toy.bricks
'There are toy bricks.'
b. Father: Jīmù
toy.bricks
c. Child: Jīmù a!
toy.bricks ATT
'Ha, toy bricks!'

(63) illustrates the Force *yes/no* question head *ma*, distinct from the homonymous Attitude head *ma* in (60b) and (61b) above. ZZC's sentence of the form 'NP *a*' is rather intriguing, because unlike her/his father in (65b), ZZC does not use a one-word utterance, but adds *a*. (Note that ZZC starts out with the complete sentence [65a]). Whether (65c) indeed represents an AttP (whose TP complement consists of an NP only) or rather an NP followed by the homonymous pause particle *a* must be left open here.

To summarize, by the age of 2–3 years, children have in principle acquired all the three types of SFPs, with their often very subtle semantics/pragmatics and the selectional restrictions on their complements. There is no real evidence for the "first TP, then CP" incremental acquisition hypothesis postulated by Friedmann et al. (2021). Upon further reflection, this is in fact a desirable result, because TP and CP delimit each other; for the child to acquire the TP projection requires the knowledge about the periphery beyond.

4. Conclusion

SFPs in Mandarin Chinese have been demonstrated to be complementizers and to realize the heads of three projections in the rigidly ordered head-final CP 'Low CP < ForceP < AttitudeP'. Importantly, this split CP only exists in root contexts, whereas in non-root contexts at most one C is allowed, if at all. More precisely, C-elements acceptable in non-root contexts are restricted to low C (*láizhe*$_1$, *le*, *ne*$_1$), to the exclusion of the Force and Attitude heads.

All SFPs display a complex feature make-up (among them the specification for [± root]), thus challenging Huang, Li & Li's (2009: 35) view that such complex feature bundles are a characteristic of functional categories in Indo-European languages, but do not exist in Chinese.

Importantly, low Cs, by virtue of their acting as (non-default) anchors, have been shown to play an important role in determining the temporal interpretation and finiteness in Chinese, and can therefore no longer be neglected by studies addressing these issues, as has been the case so far (cf. a.o. Sun 2014; Lin 2012).

Acknowledgments

This is a thoroughly revised version of our invited talk at the *International workshop on discourse particles* organized by the University of the Basque Country (UPV/EHU) at Vitoria-Gasteiz in October 2020. We would like to thank the two anonymous reviewers for comments and suggestions. We are very grateful to the editors for their swift and efficient work on this volume and their constant good humour in these difficult times.

References

Bailey, Laura. 2015. Word order and the syntax of question particles. In *Final Particles* [Trends in Linguistics. Studies and Monographs Series 284], Sylvie Hancil, Alexander Haselow & Margje Post (eds), 407–26. Berlin: De Gruyter. https://doi.org/10.1515/9783110375572-017

Benincà, Paola. 2001. The position of topic and focus in the left periphery. In *Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi* [North-Holland Linguistic Series: Linguistic Variations 59], Guglielmo Cinque & Giampaolo Salvi (eds), 39–64. Amsterdam: Elsevier.

Chan, Marjorie K. M. 1980. Temporal reference in Mandarin Chinese: An analytical-semantic approach to the study of the morphemes *le*, *zài*, *-zhe*, and *ne*. *Journal of the Chinese Language Teachers' Association* 15(3): 33–79.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Los Angeles: California University Press.

Cheng, Lisa Lai-Shen. 1991. On the Typology of *Wh*-questions. PhD dissertation, MIT.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads. A Cross-linguistic Perspective*. Oxford: Oxford University Press.

Constant, Noah. 2014. Contrastive topic: Meaning and Realizations. PhD dissertation, University of Massachusetts at Amherst. <https://scholarworks.umass.edu/dissertations_2/171>

Del Gobbo, Francesca, Munaro, Nicola & Cecilia, Poletto. 2015. On sentential particles: a cross-linguistic study. In *Final Particles* [Trends in Linguistics. Studies and Monographs Series 284], Sylvie Hancil, Alexander Haselow & Margje Post (eds), 359–86. Berlin: De Gruyter. https://doi.org/10.1515/9783110375572-015

Endo, Yoshio. 2007. *Locality and Information Structure: A Cartographic Approach to Japanese*. Amsterdam: John Benjamins. https://doi.org/10.1075/la.116

Erlewine, Michael Yoshitaka. 2017. Low sentence-final particles in Mandarin Chinese and the Final-over-Final Constraint. *Journal of East Asian Linguistics* 26(1): 37–75. https://doi.org/10.1007/s10831-016-9150-9

Friedmann, Naama, Belletti, Adriana & Rizzi, Luigi. 2021. Growing trees: The acquisition of the left periphery. *Glossa: A Journal of General Linguistics*. 6(1: 131): 1–38. https://doi.org/10.16995/glossa.5877

Guo, Xiuli. 2016. Hànyǔ értóng zǎoqī yíwènjù huòdé yánjiū [A Study on the acquisition of early interrogative sentences in Chinese children]. PhD dissertation, Chinese Academy of Social Sciences.

Haegeman, Liliane. 2014. West-Flemish verb-based discourse markers and the articulation of the Speech Act layer. *Studia Linguistica* 68 (1): 116–139. https://doi.org/10.1111/stul.12023

Haegeman, Liliane & Hill, Virginia. 2013. The syntacticization of discourse. In *Syntax and Its Limits*, Raffaella Folli, Christina Sevdali & Robert Truswell (eds), 370–390. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199683239.003.0018

Hu, Mingyang. 1981. Běijīnghuà de yǔqì zhǔcí hé tàncí [Mood particles and interjections in the Beijing dialect]. *Zhongguo Yuwen* 5: 347–350; 6: 416–423.

Huang, C.-T. James, Li, Audrey, Y.-H & Li, Yafei. 2009. *The Syntax of Chinese*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139166935

Kong, Lingda. 1994. Yǐngxiǎng hànyǔ jùzi zìzú de yǔyán xíngshì [Linguistic forms that affect sentence completeness in Chinese]. *Zhongguo Yuwen* 6: 434–440.

Lee, Thomas Hun-tak. 1986. Studies on Quantification in Chinese. PhD dissertation, University of California.

Li, Boya. 2006. Chinese Final Particles and the Syntax of the Periphery. PhD dissertation, University of Leiden.

Li, Charles N. & Thompson, Sandra A. 1981. *Mandarin Chinese. A Functional Reference Grammar*. Los Angeles: California University Press.

Lin, Jo-Wang. 2012. Tenselessness. In *The Oxford Handbook of Tense and Aspect*, Robert Binnick (eds), 669–95. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195381979.013.0023

Lin, William C. 1984. What does the Mandarin particle *ne* communicate? *Cahiers de Linguistique – Asie Orientale* 13 (2): 217–40. <https://www.persee.fr/doc/clao_0153-3320_1984_num_13_2_1157> https://doi.org/10.1163/19606028-90000293

Lü, Shuxiang (ed). 2000. *Xiàndài hànyǔ bābáicí* [800 words of modern Mandarin]. Beijing: Shangwu.

Munaro, Nicola & Poletto, Cecilia. 2002. Ways of clausal typing. *Rivista di Grammatica Generativa* 27: 87–105.

Pan, Victor Junnan. 2011. *Interrogatives et Quantification en Chinois Mandarin: Une Approche Générative*. Rennes: Presses Universitaires de Rennes.

Pan, Victor Junnan. 2015. Mandarin peripheral construals at the syntax-discourse interface. *The Linguistic Review* 32(4): 819–68. https://doi.org/10.1515/tlr-2015-0005

Pan, Victor Junnan. 2019a. *Architecture of the Periphery in Chinese. Cartography and Minimalism*. London: Routledge. https://doi.org/10.4324/9781315158228

Pan, Victor Junnan. 2019b. Derivation of the apparent narrow scope of sentence-final particles in Chinese: A reply to Erlewine (2017). *Studies in Chinese Linguistics* 39(2): 99–126. https://doi.org/10.2478/scl-2018-0004

Pan, Victor Junnan. & Paul, Waltraud. 2016. Why Chinese SFPs are neither optional nor disjunctors. *Lingua* 170: 23–34. https://doi.org/10.1016/j.lingua.2015.10.005

Pan, Victor Junnan & Paul, Waltraud. 2018. The syntax of complex sentences in Mandarin Chinese: A comprehensive overview with analyses. *Linguistic Analysis* 42, 1–2: 63–161. <https://www.linguisticanalysis.com/volumes-volume-42-issue-1-2/>

Paris, Marie-Claude. 1981. *Problèmes de Syntaxe et de Sémantique en Linguistique Chinoise*. Paris: Collège de France.

Paul, Waltraud. 2005. Low IP area and left periphery in Mandarin Chinese. *Recherches linguistiques de Vincennes* 33: 111–134. https://doi.org/10.4000/rlv.1303

Paul, Waltraud. 2009. Consistent disharmony: Sentence-final particles in Chinese. Unpublished ms., Paris. <http://lodel.ehess.fr/crlao/docannexe.php?id=1323>

Paul, Waltraud. 2014. Why particles are not particular: Sentence-final particles in Chinese as heads of a split CP. *Studia Linguistica* 68 (1): 77–115. https://doi.org/10.1111/stul.12020

Paul, Waltraud. 2015. *New Perspectives on Chinese Syntax*. Berlin: De Gruyter. https://doi.org/10.1515/9783110338775

Paul, Waltraud. 2018. The encoding of finiteness in Chinese. *10th Conference of the European Association of Chinese Linguistics*, Università degli Studi di Milano & U. degli Studi di Milano-Bicocca, Milan, 28–29 September 2018. <http://crlao.ehess.fr/docannexe/file/2174/handout_short_version.pdf>

Paul, Waltraud & Pan, Victor Junnan. 2017. What you see is what you get: Chinese SFPs as head-final complementisers. In *Discourse Particles. Formal Approaches to Their Syntax and Semantics*, Josef Bayer & Volker Struckmeier (eds), 49–77. Berlin: De Gruyter.

Peng, Lulu. 2016. *Értóng jùfǎ jiégòu de huòdé* [The acquisition of syntactic structure]. PhD dissertation, Chinese Academy of Social Sciences.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*. Handbook in Generative Syntax, Liliane Haegeman (eds), 281–337. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7

Song, Yuzhu. 1981. Guānyú shíjiān zhùcí *de* hé *láizhe* [On the tense particles *de* and *laizhe*]. *Zhongguo Yuwen* 4, 271–276.

Sun, Hongyuan. 2014. Temporal Construals of Bare Predicates in Mandarin Chinese. PhD dissertation, Leiden University. <https://openaccess.leidenuniv.nl/handle/1887/29991>

Tang, Ting-chi. 1989. *Hànyǔ cífǎ jùfǎ xùjí* [Studies on Chinese Morphology and Syntax], vol. 2. Taipei: Student Book Co.

Tao, Yu. 2012. *The Acquisition of Mandarin SFPs*. MA dissertation, Tsinghua University.

Teng, Shou-hsin. 1973. Negation and Aspects in Chinese. *Journal of Chinese Linguistics* 1 (1): 14–37.

Wei, Haley Wei & Li, Yen-Hui Audrey. 2018. Adverbial clauses in Mandarin Chinese, Part 1: Preverbal adverbial PPs and clauses. *Linguistic Analysis* 42(1–2): 163–234. <https://www.linguisticanalysis.com/volumes-volume-42-issue-1-2/>

Wu, Guo. 2005. The discourse function of the Chinese particle *ne* in statements. *Journal of the Chinese Language Teachers' Association* 40 (1): 47–81.

Yan, Shanshan. 2018. Chinese Sentence-final Particles and Their Behaviours in English Speakers' L2 Chinese. PhD dissertation, University of Cambridge. Downloadable at: <https://www.repository.cam.ac.uk/handle/1810/275336>

Yang-Drocourt, Zhitang. 2007. *Parlons Chinois*. Paris: L'Harmattan.

Zhang, Niina Ning. 2019. Sentence-final aspect particles as finite markers in Mandarin Chinese. *Linguistics* 57(5): 967–1023. https://doi.org/10.1515/ling-2019-0020

Zhou, Shihong & Shen, Li. 2006. *Ne*$_2$ de gōngnéng, yòngfǎ jí zài duiwài hànyǔ jiàoxué zhōng de yìngduì cèlüè [The functions and usages of *ne*$_2$ and instructional strategy in teaching Chinese as a foreign language]. *Shijie hanyu jiaoxue* 2: 119–128.

Zhu, Dexi. 1982. *Yǔfǎ jiǎngyì* [On grammar]. Beijing: Shangwu Press.

PART III

The semantic-pragmatics of discourse particles

CHAPTER 8

Meaning and use of the Basque particle *bide*

Kepa Korta and Larraitz Zubeldia
University of the Basque Country (UPV/EHU)

In this work we study the meaning and use of the Basque particle *bide*. We contend that, uttering a *bide*-sentence, the speaker asserts the proposition she would assert had she uttered a sentence without *bide*, but conveys additional evidential and doxastic information. She conveys that the evidence for the belief she expresses is indirect and that she is not absolutely certain on its truth – although it seems that the latter is sometimes cancellable. To put it in speech-act theoretic terms, *bide* is an illocutionary force indicator with no contribution to the propositional content of the speech act, which imposes to the assertion a certain preparatory condition and, perhaps, a constraint on the degree of strength of the belief expressed.

Keywords: doxastic dimension, evidential dimension, illocutionary force, propositional content

1. Introduction

Languages have various means (verbs, adverbs, enclitic elements, etc.) to express the speaker's source of knowledge – an evidential element – and/or her stance towards the truth of the proposition expressed – an epistemic or, better said, doxastic element, since it is related to believing rather than knowing. Characteristically, Basque has particles with meanings involving those dimensions: in particular, *omen*, *ei*, *bide*, *ote* and *al*.[1] These particles are all grouped together syntactically – they appear in the verbal complex, before the conjugated verb, in their canonical use (Monforte 2018).[2]

1. Gascon, a Romance language that has been and is still in contact with Eastern Basque, has a similar system of particles (see, among others, Pusch 2002, 2007; Morin 2008; Marcus 2010 and Suïls & Ribes 2015). Marcus (2010) argues that the system of Basque particles is the source of the Gascon system. Thanks to the second referee for this remark.

2. Although some classical grammarians considered these elements to be adverbs (e.g., Larramendi 1729 and van Eys 1873), adverbial locutions (Gèze 1875), or even verbal prefixes (Azkue 1905–1906), nowadays they are unanimously categorized as particles. We follow that trend, first,

https://doi.org/10.1075/la.276.08kor

(1) Euri-a ari *omen/bide* du
rain-DET PROG PRT/PRT AUX
'It's *omen/bide* raining.'[3]

(2) Euri-a ari *ote/al* du?
rain-DET PROG PRT/PRT AUX
'Is it *ote/al* raining?'[4]

They have been mainly studied from a syntactic point of view (e.g. Mujika 1988; Haddican 2004; Etxepare 2010; Etxepare & Uria 2016 and Monforte 2018, 2020a, b), with comparatively little research on their meaning and use (Jendraschek 2003; Carretero & Cid-Abasolo 2014; Garmendia 2014; Korta & Zubeldia 2014, 2015, 2016).

In this paper, we add to the semantic and pragmatic studies on *bide*. After providing a brief description of its morphosyntactic characteristics (Section 2), we study its meaning and use (Section 3). We argue that *bide* is related to both the evidential and the doxastic dimensions (see Faller 2002 for a similar conclusion about the Cuzco Quechua enclitic *-chá*). We consider *bide* an illocutionary force indicator, not contributing to the propositional content of the utterance, and we offer a speech-act theoretic account of it. The evidential aspect is part of the preparatory condition of the speech act – that is, the *bide*-assertion presupposes that the speaker's evidence is indirect. The doxastic aspect is a combination of two kinds of meaning. On the one hand, *bide*-assertions express the speaker's belief on the truth of its propositional content, exactly as their bare counterparts do. On the other hand, an expression of incomplete certainty is typically present. We offer two alternative accounts of this latter doxastic aspect: either it is a semantic non-truth-conditional meaning that constrains the degree of strength of the speaker's belief (or a *conventional* implicature, using Grice's category); or it is a pragmatic aspect of meaning, a generalized conversational implicature (GCI), cancellable in particular contexts (Section 4). Finally, we draw conclusions about *bide* and its comparison with *omen* (Section 5). We supplement our conceptual discussion with corpus data.[5]

---

because of their canonical invariable syntactic position within the verbal complex. And, second, and most importantly, for the lack of an intrinsic lexical meaning outside the sentence. To use the terms of medieval logicians, they are syncategorematic terms.

3. There is no direct equivalent of *bide* in English, but it could be approximately rendered as *apparently*.

4. There is no direct equivalent of *ote* and *al* in English, but they could be approximately rendered as *I wonder* (Monforte 2019) and *by chance* (Euskaltzaindia 2020), respectively.

5. We used the corpus *Ereduzko Prosa Gaur (EPG)* [*Contemporary Reference Prose*] from the University of the Basque Country (UPV/EHU): https://www.ehu.eus/en/web/eins/ereduzko-prosa-gaur-epg-

In previous works (Korta & Zubeldia 2015, 2016), we proposed that both the doxastic and the evidential aspects are part of the meaning of the *bide*-sentence, and we suggested that the doxastic dimension has some priority over the evidential, especially compared to the case of *omen*. Here, we revise the relation between the two dimensions, and suggest that it is the evidential dimension which has some priority over the doxastic. In those previous works we also proposed that *bide* is an illocutionary force indicator. However, in our speech-act theoretic account, we only considered the doxastic dimension. Here, we also situate the evidential dimension in its place.

2. Morphosyntactic characteristics of *bide*

The particle *bide* has traditionally been employed most usually in Eastern dialects (Lapurdian and Low Navarrese), and also in High Navarrese and Gipuzkoan or central dialects (according to the *Orotariko Euskal Hiztegia* 'Basque General Dictionary', Euskaltzaindia 2020). Nowadays, its spoken ordinary use is limited to the Eastern dialects. In standard Basque it is used in formal language. In this work, we focus on this contemporary standard use.

Bide, in its canonical standard use, is attached to conjugated verbs as part of the verbal complex, and it appears preceding the conjugated verb. In affirmative sentences with synthetic verbs it goes just before the verb (3), and with periphrastic verbs it appears between the main verb and the auxiliary (4).[6]

(3) Lantoki-a-n *bide* da-go.
workplace-DET-LOC PRT 3SG.PRS-be
'(S)he *bide* is at her/his workplace.'

(4) Egunero joa-ten *bide* da lantoki-ra.
everyday go-PROG PRT AUX workplace-ADL
'(S)he *bide* goes to her/his workplace everyday.'

In negative sentences with synthetic verbs it is also placed before the inflected verb (5). In contrast, with periphrastic verbs it appears before the auxiliary verb, with the auxiliary verb moved ahead of the main verb (6).

---

6. There are cases where *bide* appears in some other slots in the sentence; for instance, before the auxiliary verb when this is fronted, the main verb remaining behind:

(i) Hare-k *bide* zion Goitino-ri mikroba bikoitz-a pasa-tu.
(s)he-ERG PRT AUX Goitino-DAT microbe double-DET pass-PTCP
(*Herria*, 04–12–2003, p. 3, without signature)

'It was (s)he who *bide* passed to Goitino a double microbe.'

(5) Ez *bide* da-go lantoki-a-n.
not PRT 3SG.PRS-be workplace-DET-LOC
'(S)he is not *bide* at her/his workplace.'

(6) Ez *bide* da egunero joa-ten lantoki-ra.
not PRT AUX everyday go-PROG workplace.DET-ADL
'(S)he does not *bide* go everyday to her/his workplace.'

Besides, it typically appears in declarative sentences, both in main sentences (as in Examples (3)–(6)) and subordinate sentences, as in the following example:

(7) Oso poliki hazi-ta-ko tumore-a *bide* zue-la
very slowly grow-PTCP-REL tumour-DET PRT 3SG.ERG.PST.have.PST-COMP
buru-ko-a, esan zion emakume-a-ri (Zurutuza 2002: 27)
head-of-DET say.PFV AUX woman-DET-DAT
'That the one (that she had) in her head *bide* was a tumour grown very slowly, (s)he told the woman.'

However, it can also appear in interrogative sentences:

(8) Zergatik iragar-tzen *bide* didazu gure arteko gutun bidezko
why announce-IPFV PRT AUX our.GEN between letter by.means
harreman oro hauste-ra d-oa-la? (Muñoz 1997: 426)[7]
relation all breaking-ADL 3SG.PRS-go-COMP
'Why do you *bide* announce to me that all of our pen-relation is going to break?'

In those cases, *bide* is taken as equivalent to the question particle *ote*, which is translated as *acaso* 'perhaps' into Spanish (see Euskaltzaindia 2020). We leave aside this non-canonical use of *bide* here.

7. This is a translation into Basque (made by Jon Muñoz) of the following passage from Choderlos de Laclos (1782: 375–376): "*pourquoi semblez-vous m'annoncer que toute correspondance va être rompue entre nous ?*".

3. *Bide*: Evidential and doxastic meaning

Bide has been associated to two dimensions in Basque linguistics:

1. To an evidential dimension, as an inferential particle (Novia de Salcedo 1887; Jendraschek 2003; de Rijk 2008; Alcázar 2009 and Boye 2012, among others). In our view, uttering a *bide*-sentence the speaker asserts that *p* and presupposes that she has indirect evidence for its truth.
2. To a doxastic dimension, that is, to the expression of a certain degree of belief or certainty on the truth of the proposition (see, for instance, van Eys 1873; Euskaltzaindia 1987; Orpustan 1993 and Elhuyar-Elkar 1994). As we see it, uttering a *bide*-sentence the speaker expresses that she believes, but is not entirely certain, that *p*.

We take it that the evidential aspect and the doxastic aspect are both present in a *bide*-sentence and, hence, an example such as (9) could be glossed as (10) or as (11):[8]

(9) Asko mugi-tzen *bide* naiz lo-tan na-go-ela.
a.lot move-IPFV PRT AUX sleep-LOC 1SG.PRS-be-COMP
'I *bide* move a lot while sleeping.' (Etxebarria 2003: 111)[9]

(10) I have indirect evidence that I move a lot while sleeping [evidential dimension].

(11) I believe with a high but not absolute degree of certainty that I move a lot while sleeping [doxastic dimension].

Regarding the evidential dimension, the speaker when uttering a *bide*-sentence expresses that she has indirect evidence for the truth of the proposition. By indirect evidence we mean that she has arrived at the belief that *p* not by perception but by reasoning – including not only (deductive, inductive or abductive) inference but also speculation (see Faller 2002: 175)–, as in Examples (12) and (13), or by testimony or hearsay (reported evidence), as in (14).[10] It is reasonable to conclude, then, that *bide* is an inferential evidential particle.

(12) Egarritu-a *bide* zara, engoitik, hainbertze denbora oihuka kalakan.
thirsty-DET PRT AUX already so.much time.DET shouting talking
'You are *bide* already thirsty, so much time talking shouting.'
(Bidegain 2002: 195)

8. See also Korta & Zubeldia (2015, 2016).

9. Slightly adapted example.

10. In some cases, it seems that the report by some other person is the base for the inference ("report-based" inference by Marín-Arrese 2015).

(13) Jon, ipurdi-a hara eta hona astin-tzen, dantzan ari *bide* zen.
Jon bottom-DET there and here shake-IPFV dancing PROG PRT AUX
'Jon, moving his bottom here and there, was *bide* dancing.'
(Iturriaga 1999: 226)

(14) Gau-ez, ostera bai, ager-tzen *bide* dira itsaspeko-ak
night-INS however yes appear-PROG PRT AUX submarine-DET.PL
azal-e-ra, Gregory-k kontatu-ko du-en legez.
surface-DET-ADL Gregory-ERG narrate-PROSP AUX-COMP as
'On the contrary, at night, yes, the submarines *bide* appear in the surface, as Gregory will narrate.' (Jimenez 2003: 187)

There is a strong argument to claim that the evidential dimension belongs to the meaning of *bide*-sentences: the evidential dimension is not cancellable. As a rule, the result of adding perception (something like "and/but I have seen it") as the grounds for asserting a *bide*-sentence is an infelicitous utterance, as the following examples show:[11]

(15) #Lantoki-a-n *bide* da-go. Han ikus-i dut.
workplace-DET-LOC PRT 3SG.PRS-be there see-PTCP AUX
'(S)he is *bide* at her/his workplace. I saw her/him there.'

(16) #Ez *bide* da-go lantoki-a-n. Taberna-n ikus-i dut.
not PRT 3SG.PRS-be workplace-DET-LOC bar.DET-LOC see-PTCP AUX
'(S)he is not *bide* at her/his workplace. I saw her/him at the bar.'

As for the doxastic dimension, we think, first, that by uttering a *bide*-sentence the speaker expresses her belief on the proposition expressed, that is, her belief on the proposition she would have expressed by the same sentence without *bide*. Infelicitous utterances are produced when we add something like "but I don't believe it" after a *bide*-utterance as in (17), "but I believe it hasn't" as in (18), or "I believe neither it has nor it hasn't" as in (19). They cause something very much like Moore's paradox ("The cat is on the mat, but I don't believe it is").

(17) #Euri-a egin *bide* du, baina ez dut uste.
rain-DET do.PFV PRT AUX but not AUX believe
'It has *bide* rained, but I don't believe it has.'

---

11. Adding "and/but I have seen it" to the *bide*-utterance in (9) is plain nonsense, of course, but it is probably due to the impossibility of having direct evidence of anything while asleep (watching it later on recorded images does not count as *direct* evidence).

(i) #Asko mugi-tzen **bide** naiz lo-tan na-go-ela. Eta ikus-i dut.
a.lot move-PROG PRT AUX sleep-LOC 1SG.ABS.PRS-be-COMP and see-PTCP AUX
'I *bide* move a lot while sleeping. And I saw it'.

(18) #Euri-a egin *bide* du, baina uste dut ez du-ela egin.
rain-DET do.PFV PRT AUX but believe AUX not AUX-COMP do.PFV
'It has *bide* rained, but I believe it hasn't.'

(19) #Euri-a egin *bide* du, baina ez dakit zer pentsa-tu,
rain-DET do.PFV PRT AUX but not know.1SG.ERG what think-PTCP
egin du-ela ala ez du-ela egin.
do.PFV AUX-COMP or not AUX-COMP do.PFV
'It has *bide* rained, but I don't know what to think that it has rained or it hasn't.'

Second, we think that *bide* is incompatible with the speaker's complete certainty on the truth of *p*. Infelicitous utterances are produced when the speaker adds an expression of complete certainty on the truth of the proposition as a continuation of a *bide*-utterance.[12]

(20) #Euri-a egin *bide* du, eta/baina erabat ziur na-go euri-a
rain-DET do.PFV PRT AUX and/but totally certain 1SG.PRS-be rain-DET
egin du-ela.
do.PFV AUX-COMP
'It has *bide* rained, and/but I am absolutely certain that it has rained.'

It seems that the equivalents of *bide* in Basque, according to the main dictionaries, behave similarly in both respects; for example, the adverbs *antza* (21), *nonbait* (22), and *sobrare* (23) 'apparently'; the verbs *behar* 'must' (24), *iduri du*, *badirudi* 'it seems' (25) and the future or participle + future ((*etorri*) *izango da*) (26), among others. First, regarding lack of belief, disbelief, and suspension of (dis)belief:[13]

(21) #Euri-a egin du *antza*, baina ez dut uste
rain-DET do.PFV AUX adverb but not AUX believe
It has *antza* rained, but I don't believe it has.'

(22) #Euri-a egin du *nonbait*, baina uste dut ez du-ela egin.
rain-DET do.PFV AUX adverb but believe AUX not AUX-COMP do.PFV
'It has *nonbait* rained, but I believe it hasn't.'

(23) #Euri-a egin du *sobrare*, baina ez dakit zer pentsa-tu,
rain-DET do.PFV AUX adverb but not know.1SG.ERG what think-PTCP
egin du-ela ala ez du-ela egin
do.PFV AUX-COMP or not AUX-COMP do.PFV
'It has *sobrare* rained, but I don't know what to think, that it has rained or it hasn't.'

12. But see below.

13. For the sake of brevity, we only give examples of adverbs in this case, and of verbs in the other. However, all elements behave similarly in both respects.

And, second, with respect to complete certainty:

(24) #Etxe-a-n *behar du*, argi-a piztu-ta bait-ago, eta/baina
home-DET-LOC must AUX light-DET switch.on-PTCP COMP-is and/but
han da-go.
there 3SG.PRS-be
'(S)he must be at home, as the light is switched on, and/but (s)he is there.'

(25) #*Badirudi/iduri du* euri-a egin du-ela, eta/baina egin du.
it.seems rain-DET do.PFV AUX-COMP eta/but do.PFV AUX
'It seems that it has rained, and/but it has.'

(26) #Euri-a *egin izan-go du*, eta/baina egin du.
rain-DET do.PFV do-PROSP AUX and/but do.PFV AUX
'It must have rained, and/but it has.'[14]

To sum up, we can initially conclude, then, that for any declarative sentence *S*, and its corresponding propositional content *p*,

(27) a. Uttering S_{bide}, the speaker asserts that *p*, the same proposition that she would assert by uttering *S*. Or, in other words, the truth-conditions of (an utterance of) S_{bide} are the same as the truth-conditions of (an utterance of) *S*; hence[15]

---

14. Again, there is no direct equivalent of the verb form future or participle + future ((*etorri*) *izango da*) in English. We provide *must* + past perfect as an approximate equivalent.

15. The results of the application of the assent/dissent test (see, among others, Faller 2006; Papafragou 2006; Matthewson et al. 2007; Murray 2010 and Matthewson 2012) to *bide*-utterances support this assumption. When a *bide*-utterance such as (i) is accepted, rejected or challenged, the proposition *p* – and not (also) p_{bide} – is the target.

(i) Euri-a egin *bide* du.
rain-DET do.PFV PRT AUX
'It has bide rained.'
a. Bai, asko egin du.
yes a.lot do.PFV AUX
'Yes, it has rained a lot!'
b. Ez da egia, kale-a garbi-tzen du-en makina
not AUX true.DET street-DET clean-PROG AUX-REL machine.DET
pasa-tu da.
pass-PFV AUX
'That's not true, the machine that cleans the street has passed through.'
c. Bai? Ez ote da kale-a garbi-tzen du-en makina pasa-tu?
yes? not PRT AUX street-DET clean-PROG AUX-REL machine.DET pass-PFV
'Are you sure? Wouldn't it be the machine that cleans the street that passed through?'

The utterances created by accepting, rejecting or challenging the evidential and doxastic dimensions are extremely artificial and even hard to formulate.

b. Uttering S_{bide}, the speaker expresses the belief that *p* is true, the same type of belief that she would express had she uttered just *S*; but
c. Uttering S_{bide}, and not just *S*, the speaker conveys that she has only indirect evidence for the truth of *p*; and
d. Uttering S_{bide}, she adds an upper limit to her degree of certainty (lower than entire certainty) on the truth of *p*.

We think they are fairly uncontroversial assumptions about the semantics of *bide*-sentences, except for (27d). In a few examples in the corpus, *bide* appears with adverbial phrases that express complete certainty (for example, *dudarik gabe* and *ezbairik gabe* 'no doubt'), as in the following example:

(28) *Dudarik gabe* jende guztiz zorrotz eta bertute guztiz bortitzeko
without.doubt people completely strict and virtue completely aggressive
horie-i on iduri *bide* zaie beren izen-ak
those-DAT good think.PTCP PRT AUX.DAT their.GEN name-DET.PL
ezagut-u-ak izan daiteze-n. (Muñoz 1997: 119)[17]
know-PTCP-DET.PL be AUX-SBJV
'No doubt, those completely strict people of completely aggressive virtue *bide* think it is good their names be well-known.'

Here the certainty upper limit we attribute to *bide* seems to be cancelled by the adverbial phrase *dudarik gabe* 'no doubt'. If our attribution (27d) were correct, some sort of infelicity would arise from the clash between not *absolutely certain* (from *bide*) and *no doubt* (from the adverbial phrase). It is clear that the adverbial cannot be interpreted as a sort of correction of the *bide*-sentence, since it comes earlier and, most importantly, has wide scope over it. So, either this kind of examples are, in fact, infelicitous, or we have to revise our thesis (27d) about the lack of complete certainty expressed by *bide*-utterances.[17]

In addition, some linguists and grammarians (among others, Lafitte 1979 [1962]; Altuna 1979; Minaberry et al. 1978; and Kintana & Xarritton 1997) provide the adverbial phrases/adverbs *sin duda* 'no doubt' in Spanish, and *sans doute* 'no doubt' and *sûrement* 'surely' in French as equivalents of *bide*. If so, instead of a clash, some redundancy should be expected in examples like the ones above. We are not sure about that. Cases like (28) are exceptions to the canonical standard

16. This is a translation into Basque (made by Jon Muñoz) of the following passage from Choderlos de Laclos (1782: 106): "*Sans doute ces gens si sévères, et d'une vertu si rigide, consentent à être nommés*".

17. As it can be noted, the use of *bide* in the translation into Basque of the Example (28) could be a result of a literal translation. It might also be that it is simply an instance of sloppy writing (thanks to the second referee for this remark). This is why we are dubious about this kind of examples, taking into account, in addition, that they are scarce.

usage in contemporary Basque, due perhaps to the influence of Eastern dialects (if not infelicitous cases), which might not exclude as a rule the speaker's complete certainty on the truth of the proposition expressed. We leave it as an open issue whether element (27d) is part of the semantics of *bide* (as suggested by its canonical use in standard contemporary Basque) or part of pragmatics (as possibly in its Eastern dialectal use). More questionnaires and experiments are needed in order to settle the issue, but for the moment we will formulate an alternative account to the picture proposed above.

Leaving aside assumption (27d) above, the only special point about the semantics of an utterance of S_{bide} with respect to an utterance of *S* is the evidential dimension as expressed in (27c). Not having direct evidence on the truth of a proposition might be taken, in general, as a reason for not being entirely certain about it. The speaker, when using a *bide*-sentence, could be taken to express that she has only indirect evidence for *p* and, consequently, in general, she cannot confidently assert the truth of the inferred utterance. The upper limit to the certainty of the speaker can be a GCI, cancellable in particular contexts. Thus, the utterance of a *bide*-sentence would be compatible with the expression of absolute certainty about the truth of *p*, as in (28) above. In Gricean terms (Grice 1975), we could say that it is generally inferred assuming that the speaker is observing the Cooperative Principle and the second maxim of quantity: "Do not make your contribution more informative than is required". In particular circumstances, the degree of certainty can be present or not. In neo-Gricean terms, the GCI could be the default result of the Third (M) Heuristic: "What's said in an abnormal way, isn't normal; or Marked message indicates marked situation" (Levinson 2000: 33).[18]

Reasons of theoretical economy would favour this option, taking the doxastic upper-bound limit as a pragmatic (secondary) meaning derived from its purely evidential semantic meaning, together with principles (or heuristics) governing communication. We are not entirely convinced, though.

First, because the case of particles with both evidential and doxastic semantic meanings is not entirely unusual. The Cuzco Quechua enclitic *-chá* (Faller 2002) and the inferential evidential *k'a* in St'át'imcets (Matthewson et al. 2007), for example, involve both dimensions too. Faller (2002: 177) argues that *-chá* encodes that the speaker arrived at his or her statement by reasoning, and also that the speaker is less than 100% certain that the proposition expressed is true. On the other hand, Matthewson et al. (2007) analyze evidentials in St'át'imcets as epistemic modals. They take modal evidentials to be elements that perform both evidential and modal

18. In Gricean terms, we can also take it to be a conventional implicature generated by the conventional meaning of the particle *bide*. This would be an adequate Gricean option, if we took examples like (20) as cases that show hard cancellability.

functions concurrently; that is to say, they are epistemic modals with an extra restriction about the source of the evidence. To give an example, "an utterance of the form *k'a p* asserts that according to the speaker's knowledge state, p is possibly or necessarily true, and presupposes that the speaker has inferential indirect evidence for p" (Matthewson 2011: 335).

But, second, and most importantly, we are not completely satisfied because it seems to us that the doxastic aspect is not cancellable or a mere default. If it were, adding after a *bide*-utterance something like "but I don't believe it" or even "I'm absolutely certain" would be perfectly possible, but we think that it is not, as we showed earlier by Examples (17) and (20).

That is why we leave the issue open, and we provisionally contemplate both options when we account for the meaning of *bide* in speech-act theoretic terms in what follows.

4. A speech-act theoretical account of *bide*

In speech-act theoretic terms, (27a) and (27b) above mean that *bide* is an illocutionary force indicator with no contribution to the truth-conditional content of the assertion.[19] Take a bare assertion such as (29) and a *bide*-utterance such as (30), both produced in Donostia today.

(29) Euri-a ari du.
rain-DET PROG AUX
'It is raining.'

(30) Euri-a ari *bide* du.
rain-DET PROG PRT AUX
'It is *bide* raining.'

Both are assertions with no difference in their direction of fit (words to world), conditions of satisfaction (that *p* represents a real state of affairs) or sincerity conditions (belief that their common propositional content *p* is true) (see Table 1). When a speaker utters either *S* or S_{bide}, and the world is such that the state of affairs represented by *p* does not exist, *p* is false, and both speech acts are not satisfied. When the speaker utters either *S* or S_{bide} without believing that its propositional content *p* is true, she is making an insincere assertion in both cases.

19. Faller (2002: 177) reaches a similar conclusion about the Cuzco Quechua enclitic *-chá*. She considers two possible ways of analysing the semantics of the enclitic: (i) as a possibility epistemic modal, within possible worlds semantics or (ii) as an illocutionary operator, with the analysis of its evidential meaning as a sincerity condition (Faller 2002, p. 179). She opts for an illocutionary force analysis for *-chá*.

Table 1. Comparison of illocutionary forces of *S* and S_{bide}

	***S* (Euria ari du [It is raining])**	**S_{bide} (Euria ari *bide* du [It is *bide* raining])**
direction of fit	words to world	words to world
condition of satisfaction	the speech act is satisfied if and only if *p* (it is raining in Donostia today) is true	the speech act is satisfied if and only if *p* (it is raining in Donostia today) is true
sincerity condition	the speaker believes that *p* (it is raining in Donostia today)	the speaker believes that *p* (it is raining in Donostia today)
preparatory condition	the speaker has some evidence for the truth of *p*	the speaker's evidence for the truth of *p* is indirect

Now, there are some differences too. First, when a speaker makes an assertion, it is presupposed that she has some evidence for the truth of the propositional content of her speech act; this is called the preparatory condition of the assertion (Searle 1992 [1969]: 65–67). When the assertion is made by uttering S_{bide}, we contend that the presupposition is more specific: the speaker's evidence is indirect, that is to say, inferential or reported (see Table 1).[20]

And, perhaps, second, if we follow our initial intuitions – assumption (d) above–, there is a difference in another component of the illocutionary force: a difference in what Searle and Vanderveken (1985: 15) and Vanderveken (1990: 119–121) call the *degree of strength* of the mental state the speaker expresses; in this case, the degree of strength of the belief that *p*. Numerically represented, if we are on the right track, *bide* requires the degree of strength to be in the interval 0.5 to 1 – strictly higher than 0.5 and strictly lower than 1.[21]

---

20. Our proposal shows some similarities to von Fintel & Gillies' (2010) proposal for English *must*. They argue that the *must*-utterance is as strong as its bare counterpart, and that the verb, in addition to the epistemic modal meaning (quantificational force or strong necessity), presupposes an indirect inference: it signals that the proposition was held via inference, and not known directly nor through reliable reports.

21. An explanatory note on a relative notion of belief might be helpful at this point. According to an absolute notion of belief, an agent believes a proposition *p* or she doesn't. If she does, she doesn't believe its negation, and vice versa. She might also suspend (dis)belief, and believe neither *p* nor not-*p*. If we are on the right track, however, according to the canonical standard use of *bide*, the sincere speaker expresses a belief with a lack of complete certitude, which requires a relative notion of belief, in which the agent (dis)believes a proposition, with various degrees of certainty. An obvious way to represent such a notion is with a scale from 0 to 1, where 1 stands for absolutely certain belief, 0 for absolutely certain disbelief (belief on not-*p*), and 0.5 for suspension of (dis)belief. This relative doxastic notion should not be confused with (alethic) modal notions

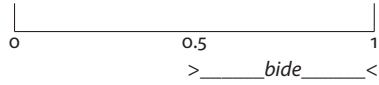


Figure 1. The degree of strength of the belief that *p* in the case of *bide*.[22]

As we said above, however, this modulation of the degree of strength of the speaker's belief could be taken to be an implicature generated by *bide*-utterances by default, without any particular context in mind (a GCI), or by the conventional meaning of the particle (a conventional implicature), as suggested by cases where the speaker expresses complete certainty. This option would also explain naturally away some equivalents amounting to entire certainty given to *bide* in some grammars and dictionaries. In that case, *bide* would not incorporate this particularity about the degree of strength of the mental state expressed by the speaker, but just the indirect evidential dimension. Further work is needed to go one way or the other.

5. Conclusions

In this work we have argued that the Basque particle *bide* encodes indirect evidential information and, perhaps, a specific doxastic information about the degree of strength of the mental state expressed. A *bide*-utterance and its bare counterpart assert the same thing, that is to say, *bide* does not add any constituent to the proposition expressed by an utterance, but contributes the presupposition that the speaker's evidence for the truth of such a proposition is indirect as well as the information that the speaker is not entirely certain about its truth. The apparent incompatibility of a *bide*-utterance with the explicit expression of the speaker's complete certainty makes it very plausible to treat this doxastic element as part of the semantics of *bide*; either as an element of the illocutionary force of *bide*-assertions (the degree of strength of the mental state expressed) or a conventional implicature generated by *bide*.

---

of possibility and necessity. On the other hand, in bare assertions, the degree of strength of the illocutionary act is signaled by other means, such as intonation and stress, producing utterances with different illocutionary forces from bald statements to suggestions. We thank one of the referees for raising this issue.

22. This relative notion of belief is, on the one hand, reminiscent of the *credence function* proposed by Davis et al. (2007), which is based on the concepts of *Lewisian quality* (Lewis 1976) and *quality threshold* (Potts 2006). On the other hand, it relates to the concept of *degrees of strength* (Wolf 2015) by which assertions are performed. We thank one of the referees for pointing this out to us.

However, there are a few cases where the *bide*-utterance seems compatible with the speaker's expression of complete certainty. This opens the possibility of locating the doxastic element at the pragmatic side of *bide*-utterances: as a GCI, that would be generally present, or by default; and cancellable in particular contexts. In addition, the equivalents given by some dictionaries and grammars are related to complete certainty. Taking into account that these equivalents collect the use of Eastern writers or are given by Eastern grammarians and lexicographers, and that it is in these dialects that *bide* is nowadays used in ordinary speech, this option deserves some credit and further research is called for. Especially, some questionnaires and experiments are needed to collect the spoken use of the particle to decide whether we were right on our initial thesis or the second option explains the use of *bide* best.

Finally, it is worth comparing these conclusions about *bide* with our findings about the reportative particle *omen*. In our view, a first difference is that *pace* Euskaltzaindia (1987: 515), Larramendi (1886 [1729]: 225), Lafitte (1979 [1962]: 162) and Wilbur (1981) among others, the doxastic dimension is totally absent from the semantic meaning of the *omen*-sentence. Uttering an *omen*-sentence, the speaker asserts that the reported proposition was stated by someone other than the speaker. *Omen* does not have any other semantic meaning. That is why *omen*-assertions take the whole spectrum of doxastic attitude towards the truth of the proposition expressed by the bare assertion (without *omen*), from absolute belief (1) to absolute disbelief (0), or any degree in between. *Bide*-assertions, however, always express belief, though, most often – if not always – with no complete certainty. Then, for instance, the Example (20) above with *bide* is infelicitous (as for the first proposal, at least); in contrast, it would be felicitous if we used *omen* instead.

(31) Euri-a egin *omen* du, eta erabat ziur na-go euri-a
rain-DET do.PFV PRT AUX and absolutely certain 1SG.PRS-be rain-DET
egin du-ela.
do.PFV AUX-COMP
'It is said that it has rained, and I am absolutely certain that it has rained.'

Similarly, the following *omen*-utterance is felicitous:

(32) Euri-a ari *omen* du, eta erabat ziur na-go ez
rain-DET PROG PRT AUX and absolutely certain 1SG.PRS-be not
du-ela ari.
AUX-COMP PROG
'It is said that it is raining, and I am absolutely certain that it is not.'

However, if we used *bide* rather than *omen*, an infelicitous utterance would result:

(33) #Euri-a ari *bide* du, eta erabat ziur na-go ez
rain-DET PROG PRT AUX and absolutely certain 1SG.PRS-be not
du-ela ari.
AUX-COMP PROG
'It is *bide* raining, and I am absolutely certain that it is not.'

The doxastic element in *omen*-utterances, if present, belongs always to pragmatics: it is a GCI, cancellable in particular contexts, and absent, in fact, in many examples. If we lean to the second (Eastern) pragmatic option about the doxastic information conveyed by *bide*-utterances, there would be some similarity about our accounts of the doxastic dimension associated with both particles. In both cases, the evidential meaning would be part of the semantic meaning (truth-conditional, in the case of *omen*; non-truth-conditional, in the case of *bide*), whereas the doxastic dimension would be an implicature, specifically a GCI. However, there are two important differences. One, *omen* covers all the doxastic spectrum while *bide* covers only positive belief. And, two, most importantly, the object of the doxastic belief under consideration here is not the propositional content of *omen*-assertions, but the propositional content of its bare counterpart. In other words, *omen* contributes to the proposition expressed by an *omen*-assertion; then, the proposition *p* expressed by the bare utterance without *omen* and the proposition p_{omen} expressed by an *omen*-utterance are two different propositions. Both are statements, but they are different.[23] In contrast, *bide* does not contribute to the proposition expressed by a *bide*-assertion – it is just an illocutionary force indicator.[24]

Further work is needed to decide among the theoretical options available for the correct account of the meaning and use of *bide*, to provide solid empirical support to it and, if possible, to aim at a unified semantic-pragmatic theory of Basque evidential and doxastic (or epistemic) particles, which may lead us to distinguish (or not) between a canonical standard use and an Eastern dialectal one.

Further work is also needed to crosslinguistically locate *bide* with respect to other indirect evidential particles.

---

23. Our evidence comes, on the one hand, from experimental results based on the assent/dissent test and on the other hand, on the scope test (see Korta & Zubeldia 2014 and Zubeldia 2015).

24. *Omen* also encodes some illocutionary force information: it indicates that the illocutionary point of the *omen*-utterances is always assertive. It indicates, as well, that the reported utterance is a statement; that is to say, an utterance of a declarative sentence – whatever their illocutionary point. See Caudal (2009) for a similar conclusion, for discourse connectives and tense systems that often contribute to both the propositional content and the illocutionary force.

Acknowledgements

We would like to thank the audience of the International Conference on Stance, (Inter)Subjectivity, Identity in Discourse 2020 (Madrid), the International Workshop on Discourse Particles (Vitoria-Gasteiz) and the Internal Seminar of the LAT group (ILCLI, Donostia), where earlier versions of this paper were presented. We also thank the anonymous reviewers for their helpful comments and suggestions. Needless to say, the responsibility of any mistake is ours.

Funding

This work was partially supported by grants from the Basque Government (IT1032–16) and the Spanish Ministry of Science and Innovation (PID2019–106078GB-I00, funded by MCIN/AEI/10.13039/501100011033).

Abbreviations used

1	first person
2	second person
3	third person
ABS	absolutive
ADL	adlative
AUX	auxiliary verb
COMP	complementizer
DAT	dative
DET	determiner
ERG	ergative
GEN	genitive
IPFV	imperfective
LOC	locative
PFV	perfective
PL	plural
PROG	progressive
PROSP	prospective
PRS	present
PRT	particle
PST	past
PTCP	participle
REL	relative
SBJV	subjunctive
SG	singular

References

Alcázar, Asier. 2009. Subjectification in Basque evidential particles. Manuscript.

Altuna, Francisco María. 1979. *Etxepareren hiztegia. Lexicon dechepariano*. Bilbao: Mensajero.

Azkue, Resurreccion Maria. 1905–1906. *Diccionario vasco-español-francés*. Bilbao: author's direction; Paris: Paul Geuthner (1984 edition, *Diccionario vasco- español-francés*. Bilbao: Euskaltzaindia).

Boye, Kasper. 2012. *Epistemic meaning: A Crosslinguistic and Functional-Cognitive study*. Berlin: de Gruyter. https://doi.org/10.1515/9783110219036

Carretero, Martha & Cid-Abasolo, Karlos. 2014. The interaction between epistemic modality and evidentiality: a semantic-pragmatic analysis of Basque *behar* and *omen/ei*. *The International Conference on Evidentiality and Modality in European Languages* 2014 (*EMEL'14*).

Caudal, Patrick. 2009. Tenses, connectives and narration(s). *Cahiers Chronos* 25: 69–89.
Davis, Christopher, Potts, Christopher & Speas, Margaret. 2007. The pragmatic values of evidential sentences. *SALT* XVII, Tova Friedman & Masayuki Gibson (eds), 71–88. Ithaca, NY: Cornell University. https://doi.org/10.3765/salt.v17i0.2966
Elhuyar-Elkar. 1994. *Euskal hiztegi modernoa*. Donostia: Elhuyar-Elkar.
Etxepare, Ricardo. 2010. *Omen* bariazioan. In *Euskara eta euskarak: Aldakortasun sintaktikoa aztergai*, Beatriz Fernandez, Pablo Albizu & Ricardo Etxepare (eds), 85–112. Bilbao: UPV/EHU.
Etxepare, Ricardo & Uria, Larraitz. 2016. Microsyntactic variation in the Basque hearsay evidential. In *Microparameters in the grammar of Basque*, Beatriz Fernández & Jon Ortiz de Urbina (eds), 265–288. Amsterdam: John Benjamins. https://doi.org/10.1075/lfab.13.10etx
Euskaltzaindia [The Royal Academy of the Basque language]. 1987. *Euskal Gramatika: Lehen Urratsak II*. Bilbao: Euskaltzaindia.
Euskaltzaindia [The Royal Academy of the Basque language]. 2020. *Orotariko Euskal Hiztegia*. 7th edition. [Basque General Dictionary]. Available online (in Basque) in: https://www.euskaltzaindia.eus/index.php?option=com_oehberria&task=bilaketa&Itemid=413&lang=eu.
Faller, Martina. 2002. Semantics and pragmatics of evidentials in Cuzco Quechua. PhD dissertation, Stanford University.
Faller, Martina. 2006. Evidentiality below and above speech acts. Manchester: University of Manchester, manuscript. Online: http://personalpages.manchester.ac.uk/staff/martina.t.faller/documents/Evidentiality.Above.Below.pdf
Garmendia, Joana. 2014. *Ote*: hiztunaren ziurtasunik eza. *Gogoa* 12-13: 7–26.
Gèze, Louis. 1875. *Eléments de grammaire basque: dialecte souletin, suivis d'un vocabulaire basque-français & français-basque*. Bayonne: Lamaignère.
Grice, Paul. 1975. Logic and conversation. In *The logic of grammar*, Donald Davidson & Gilbert Harman (eds), 64–75. Encino: Dickenson. Printed also in *Syntax and semantics 3: Speech acts*, Peter Cole & Jerry L. Morgan (eds), 1975, 41–58. New York: Academic Press. Reprinted in *Studies in the way of words*, Paul Grice, 1989, 22–40. Cambridge (MA): Harvard University Press.
Haddican, William. 2004. Sentence polarity and word order in Basque. *The linguistic review* 21: 87–124. https://doi.org/10.1515/tlir.2004.004
Jendraschek, Gerd. 2003. *La Modalité Épistémique en Basque*. Muenchen: Lincom Europa.
Kintana, Xabier & Xarritton, Piarres. 1997. *Hiztegia: euskara-frantsesa, frantsesa-euskara. Dictionnaire: basque-français, français-basque*. Donostia: Elkar.
Korta, Kepa & Zubeldia, Larraitz. 2014. The contribution of evidentials to utterance content: Evidence from the Basque reportative particle *omen*. *Language* 90–2: 389–423. https://doi.org/10.1353/lan.2014.0024
Korta, Kepa & Zubeldia, Larraitz. 2015. The evidential and doxastic dimensions of *bide*. *Belgian Journal of Linguistics* 29, special issue, *Evidentiality and the Semantics-Pragmatics interface*, Juana Marín-Arrese & Bert Cornillie (eds), 41–59. Amsterdam: John Benjamins. https://doi.org/10.1075/bjl.29.02kor. ISSN 0774-5141 | E-ISSN 1569-9676.
Korta, Kepa & Zubeldia, Larraitz. 2016. *Bide* partikula: eduki doxastikoa eta ebidentziala. *Gogoa* 15: 11–29. https://doi.org/10.1387/gogoa.17255
Lafitte, Pierre. 1979 [1962]. *Grammaire basque (Navarro-Labourdin littéraire)*. Donostia: Elkar.
Larramendi, Manuel de. 1729. *El impossible vencido. Arte de la lengua bascongada*. Salamanca: Antonio Joseph Villagordo Alcaráz (1886 edition, *El impossible vencido. Arte de la lengua bascongada*. Donostia: Hijos de I. Ramon Baroja).

Levinson, Stephen. 2000. *Presumptive Meanings*. Cambridge MA: MIT Press/Bradford Books. https://doi.org/10.7551/mitpress/5526.001.0001

Lewis, David. 1976. Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review* 85: 297–315. Reprinted In *Philosophical Papers*, Vol. 2, David Lewis, 1986, 133–156. New York: Oxford University Press. https://doi.org/10.2307/2184045

Marcus, Nicole Elise. 2010. The Gascon Énonciatif System: Past, Present, and Future. A study of language contact, change, endangerment, and maintenance. PhD dissertation, University of California.

Marín-Arrese, Juana I. 2015. Epistemicity and stance: A crosslinguistic study of epistemic stance strategies in journalistic discourse in English and Spanish. *Discourse Studies* 17 (2): 210–225. https://doi.org/10.1177/1461445614564523

Matthewson, Lisa. 2011. On apparently non-modal evidentials. *Empirical Issues in Syntax and Semantics* 8, Olivier Bonami & Patricia Cabredo Hofherr (eds), 333–357. http://www.cssp.cnrs.fr/eiss8

Matthewson, Lisa. 2012. Evidence about evidentials: Where fieldwork meets theory. In *Empirical approaches to linguistic theory: Studies in meaning and structure*, Britta Stolterfoht & Sam Featherston (eds), 85–114. Berlin: De Gruyter. https://doi.org/10.1515/9781614510888.85

Matthewson, Lisa, Davis, Henry & Rullmann, Hotze. 2007. Evidentials as epistemic modals: Evidence from St'át'imcets. *Linguistic Variation Yearbook 2007*, Jeroen Van Craenenbroeck (ed), 201–254. Amsterdam: John Benjamins. https://doi.org/10.1075/livy.7.07mat

Minaberry, Melle, Lafitte, Pierre & Michelena, Manuel. 1978. *Grammaire basque pour tous*. I. *Eléments non verbaux*. Hendaia: Haize Garbia.

Monforte, Sergio. 2018. Question particles in Basque. *Isogloss* 4/1: 29–53. https://doi.org/10.5565/rev/isogloss.48

Monforte, Sergio. 2019. What microvariation can show us: An analysis of the Basque discourse particle *ote*. In *Proceedings of the Workshop "Clause Typing and the Syntax- to-Discourse Relation in Head-Final Languages*", Arbeitspapier 130. Fachbereich Linguistik, Josef Bayer & Yvonne Viesel (eds), 175–191. Konstanz: Universität Konstanz.

Monforte, Sergio. 2020a. Modal particles in Basque. Two cases of interaction between *ote* and information structure. In *Information-Structural Perspectives on Discourse Particles*, Pierre-Yves Modicom & Olivier Duplâtre (eds), 277–300. Amsterdam: John Benjamins.

Monforte, Sergio. 2020b. Diachronical hypotheses accounting for synchronic variation: the case of the Basque particle *ote*. In *Fontes Linguae Vasconum 50 urte. Ekarpen berriak euskararen ikerketari*, Ekaitz Santazilia, Dorota Krajewska, Eneko Zuloaga & Borja Ariztimuño (eds), 437–451. Iruñea: Government of Navarre. https://doi.org/10.35462/fontes50urte.28

Morin, Annick. 2008. The syntax of Gascon clause-type particles. *Linguistica atlantica* 29: 137–155.

Mujika, Jose Antonio. 1988. Partículas modales de la flexión verbal. *Anuario del Seminario de Filología Vasca 'Julio de Urquijo'* XXII-2: 463–478.

Murray, Sarah. 2010. Evidentiality and the structure of speech acts. New Brunswick, NJ: Rutgers University dissertation.

Novia de Salcedo, Pedro. 1887. *Diccionario Etimológico del Idioma Bascongado*. Tolosa: Eusebio López.

Orpustan, Jean-Baptiste. 1993. *Oihenarten Hiztegia. Lexique Basque des Proverbes et Poésies d'Oyhenart Traduit en Français et Espagnol / par Jean-Baptiste Orpustan*. Baigorri: Izpegi.

Papafragou, Anna. 2006. Epistemic modality and truth conditions. *Lingua* 116: 1688–1702. https://doi.org/10.1016/j.lingua.2005.05.009

Potts, Christopher. 2006. Clausal Implicatures via General Pragmatic Pressures. In *Proceedings of Logic and Engineering of Natural Language Semantics 2006*, Eric McCready (ed), 65–79. Tokyo: Japan Society for the Promotion of Science.

Pusch, Claus Dieter. 2002. Preverbal modal particles in Gascony Occitan. In *Belgian Journal of Linguistics* 16, special issue, *Particles*, Ton van der Wouden, Ad Foolen & Piet Van de Craen (eds), 105–118. Amsterdam: John Benjamins. https://doi.org/10.1075/bjl.16.08pus

Pusch, Claus Dieter. 2007. Is there evidence for evidentiality in Gascony Occitan? *Rivista di Linguistica* 19.1: 91–108.

Rijk, Rudolf P. G. de. 2008. *Standard Basque: a Progressive Grammar*. Vol. 1: The grammar. Cambridge MA: MIT Press.

Searle, John R. 1992 [1969]. *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Searle, John R. & Vanderveken, Daniel. 1985. *Foundations of illocutionary logic*. Cambridge: Cambridge University Press.

Suïls, Jordi & Ribes, Salomé. 2015. Modality markers in Gascon, between grammar and stylistic variation. *eHumanista/IVITRA* 8: 544–576.

Vanderveken, Daniel. 1990. *Meaning and speech acts*, Vol. 1. *Principles of language use*. Cambridge: Cambridge University Press.

van Eys, Willem J. 1873. *Dictionnaire Basque-Français*. Paris: Maisonneuve; London: Williams and Norgate.

von Fintel, Kai & Gillies, Anthony S. 2010. *Must* … stay … strong! *Natural Language Semantics* 18: 351–383. https://doi.org/10.1007/s11050-010-9058-2

Wilbur, Terence H. 1981. Basque syntax. In *Euskalarien nazioarteko jardunaldiak* [Iker bilduma 1], 169–186. Bilbao: Euskaltzaindia.

Wolf, Lavi. 2015. Degrees of assertion. PhD dissertation, Senate of Ben-Gurion University of the Negev.

Zubeldia, Larraitz. 2015. *'Omen' partikularen azterketa semantikoa eta pragmatikoa*. Bilbao: UEU.

Corpus

Bidegain, Eneko. 2002. *Anbroxio*. Donostia: Elkar.

Ereduzko prosa gaur, The Basque Institute of UPV/EHU, http://www.ehu.es/euskara-orria/euskara/ereduzkoa/.

Etxebarria, Monika. 2003. *Goragalea*. Donostia: Alberdania-Elkar (Basque translation of Sartre, Jean-Paul. 1938. *La Nausée*. Paris: Éditions Gallimard).

Herria. 2003. Jean Pierre Goytino ameriketan euskal kazetari. 14–12–2003, p. 3, without signature.

Iturriaga, Unai. 1999. *Berandu da gelditzeko*. Zarautz: Susa.

Jimenez, Edorta. 2003. *Hemingway eta euskaldunak zerbitzu sekretuetan*. Zarautz: Susa.

Muñoz, Jon. 1997. *Harreman arriskutsuak*. Euba: Ibaizabal (Basque translation of Choderlos de Laclos, Pierre. 1782. *Les liaisons dangereuses*. Paris: Durand Neveu).

Newspaper library of *Berria* (from the web page http://hemeroteka.berria.info and *Ereduzko prosa gaur*).

Zurutuza, Imanol. 2002. *Errautsera arteko sugarra*. Donostia: Elkar.

CHAPTER 9

Three German discourse particles as speech act modifiers

Johannes Schneider
Universität Leipzig

This work attempts to reduce the properties of three German discourse particles (DPs), *ja*, *nicht* and *etwa*, to the basic building blocks of a formal discourse model (Farkas & Bruce 2010). We propose a definition of DPs as speech act modifiers that restricts the space of allowed variation of their meanings, arguing against previous approaches in terms of speaker attitudes. Speech acts modified by *ja* update the *Common Ground* (*CG*) directly; previous characterizations of the epistemic status of the proposition arise as descriptions of common justifications for such an imposed *CG* update. *Etwa* and *nicht* turn open polar questions with two default resolutions into questions with only one unmarked resolution; epistemic or bouletic attitudes arise as frequent connotations.

Keywords: German discourse particles, *ja*, *nicht*, *etwa*, Farkas and Bruce's discourse model

1. Introduction

In this paper we attempt to express the contributions of three German discourse particles by exclusively using the basic building blocks of a current formal model of discourse, Farkas & Bruce (2010). The primary motivation behind this is to be able to propose sufficiently formal definitions as answers to two questions about (a subclass of) German discourse particles: what they contribute to an utterance and how to delimit the space of possible meanings they can assume. Below, we go into detail about the motivation and background assumptions of this approach.

We start with the central question what discourse particles (DPs) mean. This question has received a range of answers in the previous literature; an influential one is the proposal that DPs specify or modify the attitude of the speaker towards the proposition or utterance in which the particle appears (see e.g. Egg & Mursell 2016 for a short overview). Although there rarely is a proposal as to which attitudes this can or cannot encompass, at least for the three particles discussed here this

https://doi.org/10.1075/la.276.09sch

usually involves epistemic notions such as knowledge, evidentiality, expectation or strength of belief. We specifically argue against such a view, for a number of reasons. The implications of e.g. a pre-theoretic description of *ja* as an epistemic particle marking the proposition as 'given, obvious or uncontroversial' (Kaufmann & Kaufmann 2012) are not innocent. Under such a view it is surprising that this epistemic particle can co-occur with a wide variety of items, expressing sometimes apparently contradictory epistemic states or attitudes, ranging from uncertainty to complete certainty:

(1) Peter ist *ja* offensichtlich/ sicherlich/ wahrscheinlich/ vermutlich/ vielleicht/ eventuell/ möglicherweise zu Hause.
'Peter is JA obviously/ for sure/ probably/ presumably/ maybe/ perhaps/ possibly at home.'

Unless the illocutionary or discourse-related aspects of the above-mentioned attitude label are not more precisely specified, it is unclear why DPs and standard epistemic or attitude-modifying adverbials do not behave alike, calling into question the justification for the existence of a distinct category DP; generally however, DPs are assumed to differ from these adverbs (e.g. Bayer & Obenauer 2011). Under this view, the property that DPs cannot be modified is surprising, as well. If a particle e.g., expresses something as obvious, or indicates heightened interest in an answer (Csipak & Zobel 2014 for *denn*) there is no a priori reason why gradations on such scales cannot be expressed via modifiers or many more related DPs. Therefore, we assume throughout this paper that any such attitude-related notions are epiphenomenal and merely frequent connotations. Let us give this heuristic the moniker *Vulcan Hypothesis*: the meanings of discourse particles do not express the attitudes, intentions, desires, beliefs or any other epistemic states of the speaker. We show below for each particle that any such proposed attitude is either a descriptive term that does not predict the particle's occurrence or a pragmatic inference that can be cancelled in appropriate counterexamples. We take this to mean that these notions are not at the core of DPs' meanings.

Instead, the line of research that we want to pursue is that DPs *modify the speech act in which they occur*. This idea goes back to at least Jacobs (1991) where he proposes that DPs modify speech act operators, resulting in a speech act with distinct properties. To make this precise, we make use of the Farkas & Bruce (2010) (F&B) model to formulate hypotheses about the effect of DPs. In short, the F&B model represents discourse as a game where utterances or speech acts are moves that have a precisely defined effect on the playing board. The playing board consists of five components (such as the *Table* or the *Common Ground*), each designed to capture discourse-related phenomena (such as the *QUD* or mutually agreed on

propositions). For each speech act, e.g. an assertion, there is an associated speech act operator that defines what effect the speech act has on the various components of the playing board. We propose that DPs take speech acts as arguments and modify them by applying a *single change* to the default effects that the speech act has on the playing board. To foreshadow, our proposal for *ja* changes the effect its speech act has on the *CG* (parallel to a proposal about non-restrictive relative clauses in the F&B model), while *etwa* and *nicht* modify the *Projected Set* for polar questions (parallel to the effect of assertions on this component). We elaborate in the respective sections that recasting previous proposals like this either lets previous proposals arise as frequent inferences or that it formalizes descriptive notions with the tools of the F&B model. This approach potentially selects a subset of what has previously been classified as discourse or modal particles. This could be used for a more fine-grained distinction between functions of different classes of particles (e.g., *ruhig* or stressed *ja* as indicators of force in imperatives, justifying the more traditional term modal particle).

This conception of DPs also allows us to propose an answer to the second question, the delimitation of potential meanings that DPs can and cannot assume. As indicated above, if the potential meanings of DPs amount to the space of possible attitudes or epistemic states, the fact that there are so few of them and that they cannot be modified is unexpected. However, it is an underappreciated fact that DPs *are a closed class* of a small number of functional elements (e.g. Gutzmann 2016); we take this observation to be in conflict with their *de facto* treatment as close relatives to other attitude indicating elements. With the above assumptions, however, DPs' status as a small, closed class receives a straightforward explanation. If DPs perform a *single* modification of a speech act, and if there are only finitely many speech acts that modify finitely many elementary discourse components, that can each only express a restricted number of states, we should expect only a small, finite number of (reasonable) DPs. The space of potential meanings for DPs is then the number of single modifications of a speech act's effect on those discourse components. No modifications of the F&B model itself are proposed, i.e. all effects make use only of the pre-existing variation in the discourse components that are independently necessary to capture the behaviour of default speech acts such as assertions and polar questions. In other words, the attempt to reduce the contribution of DPs to these general components is an attempt to reduce their allowed variation. Furthermore, since this variability amounts to binary or at least discrete choices, modification or scale-like gradability is unexpected. With a restriction to theoretical concepts that are independently necessary to describe language-universal phenomena such as the effect of speech acts, DPs in German are also no longer a language-specific curiosity but can be viewed as overt and lexicalized grammatical reflexes of aspects of these

effects that are otherwise expressed covertly or through different non-lexical means. This approach is similar in spirit to Rudin's (2020) treatment of the contribution of rising (and falling) intonation in the F&B model. He proposes that rising (or falling) intonation takes utterances as arguments, modifying the speech act to indicate absence (or presence) of speaker commitment (i.e. the effect of a speech act on the *Discourse Commitment List* component of the model). There as well as here, different grammatical reflexes contribute piecewise to the overall speech act effect.

The chapter is organized as follows: first, we introduce Farkas & Bruce's (2010) discourse model. Subsequently we discuss the particles *ja*, *nicht* and *etwa*, in that order. In each section we briefly summarize previous proposals for the respective particle. Guided by the expectations from the above assumptions about what DPs can and cannot mean, we discern which (aspects of the) proposals are unlikely to be their core contribution; relevant counterexamples are then provided. Then, we propose definitions of how these particles modify the speech act they occur in and elaborate how previously proposed notions arise as frequent pragmatic inferences from the underlying meaning. The last section discusses the overall results in light of the above assumptions and concludes.

2. The discourse model

The aim of the framework by Farkas & Bruce (2010) (F&B) is to build a dynamic model of discourse that captures the effect that questions and assertions, as well as reactions to these, have on the state of discourse. They propose several components that are necessary for a description of the grammatical properties of discourse: when a speaker utters e.g. an assertion, this assertion is put on the *Table*. What is on the *Table* is up to discussion, i.e. this is the place where *Questions Under Discussion* (*QUD*) are recorded. If both participants agree on its truth, the proposition is added to the *Common Ground* (*CG*), the set of mutual commitments. Making the preliminary *Table* stage necessary captures the view of assertions as proposals, in contrast to direct *CG* updates, as put forward in Stalnaker (1978). That is, an assertion is a *proposal* to change the *CG*; however, this change will only occur once the hearer agrees with what is said.

The component where the public individual beliefs of participant *X* are recorded is called the set of *Discourse Commitments* (DC_X). Something counts as a public belief if the speaker has explicitly committed herself to the proposition (e.g. by virtue of asserting it). Private, non-publicized views are not subsumed under this notion. As an example, if A asserts *p, p* is put on the Table, but also recorded in DC_A, as something that A has committed herself to.

There is also a component that anticipates the state of the future *CG*, called the *Projected Set* (*PS*). When an assertion is put forward a future *CG* is projected that contains the existing *CG* and the proposition that has been placed on the *Table*. This indicates the form of the future *CG* if the current move receives its standard resolution (e.g. the hearer agrees). This makes it possible to express that there are default reactions to certain speech acts. Informally, this captures the anticipations of the speaker and distinguishes unmarked from marked reactions. For example, the unmarked or default response to an assertion is its acceptance, as it fits the predictions of the *PS* where the conversation is steering. Conversational states are represented graphically as in Figure 1 (Farkas & Bruce 2010: 89):

A	**Table**	**B**
DC_A	S	DC_B
Common Ground CG	**Projected Set PS**	

Figure 1.

Speech acts are conceptualized as moves that have a defined effect on the above playing board. We illustrate the model with the effect that a declarative sentence has:

(2) A: Eva is a string theorist.

We assume an idealized empty initial context state before the utterance of the declarative. Standard declaratives (which are assigned the sentential feature [D]) result in the context state depicted in Figure 2:

A	**Table**	**B**
p	⟨'*Eva is a string theorist*'[D];{*p*}⟩	
$\mathbf{CG_2 = CG_1}$	$\mathbf{PS_2 = \{CG_1 \cup \{p\}\}}$	

Figure 2.

F&B (2010: 92) propose speech act operators for sentence types; the assertion operator takes a declarative sentence as argument and works as a function from input conversation state K_i (the state before the utterance) to output states K_o (the resulting state):

(3) **A**(S[D], a, K_i) = K_o such that
i. $DC_{a,o} = DC_{a,i} \cup \{p\}$
ii. $T_o = push([S[D];\{p\}], T_i)$
iii. $PS_o = PS_i \cup \{p\}$

In English: an assertion *A*, that takes sentence *S* with declarative feature [D] as argument, uttered by speaker *a*, changes the input context state K_i to the output context state K_o in the following way:

i. the proposition *p* denoted by that sentence is added to the initial discourse commitment list *DC* by speaker *A* $DC_{a,i}$, leading to the resulting output $DC_{a,o}$
ii. the syntactic object *S*[D] plus its denotation *p* are put on top of the stack on the input *Table* T_i resulting in the output *Table* T_o where the issue whether *p* is on the top of the *Table*
iii. the input *PS* is expanded by adding *p* to each possible *CG* in it, discarding any inconsistent set.

This is the effect of standard assertions. Default polar questions such as (4) have the effect depicted in Figure 3:

(4) Is Eva a string theorist?

A	**Table**	**B**
	⟨'*E. is a string theorist*'[I];{*p*,¬*p*}⟩	
CG_1	**PS_2 = {CG_1 ∪ {p}, CG_1 ∪ {¬p}}**	

Figure 3.

The syntactic object with an interrogative marker is placed on the *Table*, together with its denotation *{p,¬p}*. The *PS* computes both possible futures of the conversation, namely the *CG* enriched with the information that *p* holds, and the *CG* enriched with the information that ¬*p* holds. F&B call such a context state *inquisitive* (in contrast to a state that results after an assertion) since the *PS* sees multiple possible futures for the *CG*. This constitutes the case of a standard polar interrogative. Parallel to assertions, a polar question operator *PQ* is assumed. The definition of a standard polar interrogative is as follows, formalizing the two effects described above (Farkas & Bruce 2010: 95):

(5) **PQ**(*S*[I], K_i) = K_o such that
  i. $T_o = push([S[I]; \{p,\neg p\}], T_i)$
  ii. $PS_o = PS_i \cup \{p,\neg p\}$

These are the basic moves in F&B's framework. There are also moves that *react* to such assertions and questions, specifically confirming or reversing an assertion, and confirming or reversing a polar question. To exemplify, an assertion confirmation operator *AC* is defined as follows (Farkas & Bruce 2010: 98):

(6) **AC**
   a. *Input context conditions:*
      i. $top(T_i) = [S[\mathrm{D}]; \{p\}]$
      ii. p in $DC_{a,i}$
   b. *Change*:
      $\mathbf{AC}(b, K_i) = K_o$ where $DC_{b,o} = DC_{b,i} \cup \{p\}$

In contrast to assertions, an *AC* places requirements on the previous context state: on top of the stack on the *Table* is the declarative sentence *S*[D] with its denotation *p*, and this *p* is in the *DC* of the other discourse participant (in this case *A*, assuming she is the one who uttered (2). Then, for discourse participant *B* to confirm this sentence that is on the *Table* in the input context state K_i means to add the proposition *p* to his own *DC*. Such a confirmation can look like this:

(7) B: (Yes,) She is (a string theorist).

Note that a sentence that confirms an assertion can have the exact same semantics as the initial sentence that raised the issue on the *Table*. However, addressing an issue has a distinct discourse function which is why another operator is necessary. At this point, the issue *p* has already been raised, so reacting to it does not raise it again, but comments on the issue. Confirmation, therefore, does not place something on the *Table* but just adds the denotation already existing on the *Table* to its author's *DC*; congruent with the core idea of this model, addressing something on the *Table* only adds something to the speaker's *DC*, but not to the *CG* directly. In case a proposition is present in every participant's *DC*, the auxiliary move *M*' removes the issue *p* from the *Table* and adds it to the *CG*. We leave out the definitions for complementary moves such as total denial of an assertion or both possible standard reactions to a question as they are not relevant to this paper.

3. *Ja*

Thurmair (1989) argues that *ja* is a *CG*-marker, indicating that the proposition is known to both speaker and hearer. This certainly accounts for one of its most standard uses:

(8) *Two friends are at a party; both know that Speaker came by car. When everybody starts to leave, Speaker offers Hearer*:
Ich bin *ja* mit dem Auto hier, ich könnte dich also mitnehmen.
'I came JA here by car, so I could give you a ride.'

Here, the speaker uses *ja* to indicate that both interlocutors know the proposition and uses this sentence as a segue to his offer. Döring (2016) models this in F&B by stating that an assertion with *ja* is only defined if the core proposition is part of the *CG*. However, Lindner (1991) points out the use of *ja* in surprise contexts (already acknowledged by Thurmair (1989)) as problematic for a *CG* definition (roughly her Example 13):

(9) *B is climbing a tree, followed by A. Suddenly, A notices:*
Du hast *ja* ein Loch im Ärmel!
'You've got JA a hole in your sleeve!'

In this case, the hearer is ignorant of the information and even the speaker did not know right up until her utterance. For this reason, Lindner weakens the definition to "indicating that the state of affairs in question is not controversial" and that the speaker has "evidence – observational or derived from shared knowledge – that a particular state of affairs pertains" (Lindner 1991: 171). The speaker "assumes at *t* that there is no proposition *q* […] that is contradictory to the proposition […] i.e. the speaker assumes that the addressee will not contradict him/her" (ibidem: 173).

Although the idea of *ja* as a *CG* marker is persistent, most authors acknowledge the difficulties raised by this example. Kratzer (1999: 1) defines *ja* as being "appropriate in a context *c* if the proposition […] is a fact of w_c which – for all the speaker knows – might already be known to the addressee". Kratzer & Matthewson (2009) also note that *ja* does not require specific hearer knowledge but rather that "the speaker in *c* takes [the proposition] to be firmly established and, therefore, doesn't consider the question [*whether p*] to be an issue for inquiry in *c* or after *c*." (ibidem: 13). Grosz (2014) claims that *ja*'s meaning is a proper subset of the DP *doch*'s meaning in that both share an uncontroversiality presupposition, whereas *doch* also possesses a contrast presupposition. His definition of the uncontroversiality component is the following, arguing that both particles do not impose restrictions on the hearer's knowledge state (ibidem: 163):

(10) The speaker in *c* takes *p* to be firmly established in w_c and, therefore, assumes that it is safe to discard $\neg p$ as a possible answer to the question of whether *p* or $\neg p$ holds in w_c

Kaufmann & Kaufmann (2012) call *ja* and *doch* epistemic particles, saying that it is "widely agreed" that they "commit the speaker to the belief that *p* is in some sense given, obvious, or uncontroversial" (ibidem: 210). While noting that *ja* is in most cases used to indicate old information, they acknowledge that it can also be used to impart obvious information (Kaufmann & Kaufmann 2012: 211, Example 5):

(11) Du kannst *ja* zum Arzt gehen.
≈ 'You can [obviously] go to the doctor.'

So, the felicity conditions of *ja* must be widened to "include information that is readily available to anyone seeking it" (Kaufmann & Kaufmann 2012: 211). They also mention surprise examples, noting that these cases "may be amenable to an extension of that notion to include the mere absence of conflicting information" (ibidem: 212). Their formal definition for *ja* is as follows (ibidem: 212):

(12) *normally* in a situation like *c*, any rational agent whose go is to find out whether *p*, does find out whether *p* (from information already available or in the immediate surroundings).

To summarize, the clearly discourse-related definition of *ja* as a *CG* marker is challenged by surprise examples. Hearer knowledge appears to be irrelevant for the meaning of *ja*. This includes the notorious use of *ja* in discourse/topic-initial sentences where the information is new to the hearer and has to be believed at face value:

(13) Ich war *ja* gestern in der Innenstadt, und rate wen ich da getroffen habe – den Peter!
'I was JA downtown yesterday, and guess who I met there – Peter!'

Sometimes, *ja* simply marks an obvious conclusion:

(14) A: My boyfriend broke up with me.
B: Das ist *ja* blöd.
B: 'That's JA a bummer.'

The problem with the cited 'uncontroversiality' concept that supplants the previous *CG*-oriented proposals is that it is a descriptive term, not a theoretical concept and, hence, makes no predictions as to what situations fall under this property (similarly, what precisely counts as 'firmly established' or how much leeway is given to the idea of 'finding out'). Even the general intuition about *ja* as denoting an epistemic state of the proposition being somehow 'given, obvious or uncontroversial' is not without problems. As mentioned, *ja* can co-occur felicitously with elements that express uncertainty:

(15) Peter ist *ja* wahrscheinlich/ vermutlich/ vielleicht/ eventuell/ möglicherweise zu Hause
'Peter is JA probably/ presumably/ maybe/ perhaps/ possibly at home.'

Whereas combinations of more traditional attitude-oriented adverbs lead to infelicity:

(16) ??/# Juan ist (offensichtlich/ sicherlich) (wahrscheinlich/ vermutlich/ vielleicht) zu Hause.
??/# 'Juan is (obviously/for sure) (probably/presumably/maybe) at home.'

A reviewer points out that an interpretation of 'It is known that Peter is maybe at home' is conceivable, avoiding potential contradictions. However, a *CG*-interpretation of *ja* is already ruled out for independent reasons above. The examples are used to illustrate that even if one retreats to a speaker-attitude account, such an epistemic stance does not behave like any other known attitude, leading to further difficulties. While a more discourse-related version of obviousness is not logically excluded, this is not the path we want to pursue. Under the principles we assumed in the first section, we expect any attitude-related aspects in the use of *ja* to be pragmatic inferences, with its actual contribution to be found elsewhere.

There is in fact one robust, discourse-related property of *ja*, namely its inability to occur in direct answers to questions:

(17) A: Wie heißen Sie?
A: 'What's your name?'
B: #Ich heiße *ja* Joe.
B: #'My name is JA Joe.'

(18) A: Where's Peter?
B: #Er ist *ja* zu Hause.
B: #'He is JA at home.'

This is difficult to explain for attitude-related accounts of *ja* since attitude-modifiers have no problem occurring in answers ('*Peter is obviously/for sure at home.*' as response to the above question). An intuitive understanding of uncontroversiality also would not suffice since epistemic authority over one's own name can usually be assumed. Asking for a person's name could also be performed as conversation starter even if the other person's name is visible on his name tag. Still, the use of *ja* in the answer is excluded. The possibility of being able to find out the information is, therefore, not relevant for the use of *ja*. The above ban persists even if the inquirer knows the answer already (e.g., in teacher-student contexts and rhetorical questions) and, hence, cannot be due to pragmatic infelicity (a *CG* account, if it were not ruled out already by surprise examples, would need to provide independent reasons why *ja* is still infelicitous in said contexts).

We turn now to our proposal about how speech acts are modified by *ja*. The ban from appearing in answers to questions indicates that whatever the function of *ja* is prevents it from interacting with the *QUD*. Since we assume that DPs modify the effect of speech acts on the discourse components we conclude that the *Table*

and the *CG* for which the Table serves as an intermediate stop are the components relevant for *ja*. We propose the following: *ja* modifies a move by adding to its effects that it places its denotation into the *CG* directly (this idea is similar to the one expressed in Viesel 2015). Remember the assertion operator:

(19) **A**(S[D], a, K_i) = K_o such that
i. $DC_{a,o} = DC_{a,i} \cup \{p\}$
ii. $T_o = push([S[D];\{p\}], T_i)$
iii. $PS_o = PS_i \cup \{p\}$

Thus, a *ja*-assertion has the following effect:

(20) **ja**(A(S[D], a, K_i)) = K_o such that
i. $DC_{a,o} = DC_{a,i} \cup \{p\}$
ii. $T_o = push([S[D];\{p\}], T_i)$
iii. $PS_o = PS_i \cup \{p\}$
iv. $CG_o = CG_i \cup \{p\}$

Note that in default cases of the F&B model the *PS* can be automatically calculated from the content of the *Table* and is, therefore, redundant information added to moves for expository reasons. Specifying the *DC* and the *Table* in this case is mostly redundant as well since the *CG* is (at least) the union of all shared public commitments and any issue resolved by an element of the *CG* is removed from the *Table*. Therefore, the only non-trivial effect of a *ja*-assertion is to place the proposition into the *CG* (we come back later to the question why we only add the *CG* effect instead of letting it replace the *Table* effect).

This formalization achieves several things. Let us first derive the one robust empirical property of *ja*, its non-occurrence in answers. One of the crucial assumptions F&B make in their model is that assertions are *proposals* to update the *CG*. This is hard-coded into the model by the existence of the *Table*, and the fact that *Table* and *CG* are two distinct components. They emphasize this proposal nature, saying that "[c]haracterizing ordinary assertion as proposing additions to the *CG*, rather than actually changing it, is necessary in order to make room for the large variety of conversational moves that react to assertions" (Farkas & Bruce 2010: 82). In order to raise or address an issue, one has to take the route via the *Table*. It is crucial to note that a negative or positive answer to a question does not remove the issue from the *Table* but only via (explicit or implicit) agreement by the originator of the question. That predicts that one place where an effect as proposed here for *ja* is not allowed to take place is in answers to questions since it violates fundamental assumptions of the model's discourse components. This prediction is borne out. There is an exception to the above ban, however, namely in the presence of attitude-indicating elements:

(21) A: Where's Peter?
B: Vielleicht ist er *ja* zu Hause.
B: 'Maybe he is JA at home.'

(22) A: Is Lubos innocent?
B: Offensichtlich *ja* nicht!
B: 'Obviously JA not!'

The crucial question in this case is what is placed on the *Table* in the B sentences of the above examples (an issue already mentioned in Farkas 2007). The assumption that *Attit(p)* (i.e., the attitude the speaker expresses towards the proposition) is placed on the *Table* would predict that what is under discussion is the speaker's attitude and not the core proposition. This is clearly not how these sentences are interpreted. A *no*-answer to a sentence *Attit(p)* commits the speaker to $\neg p$, not to a position about the hearer's epistemic stances. It is, therefore, reasonable to assume that in these cases *p* alone is placed on the *Table*. Döring (2016) proposes that discourse commitments in general are stored directly in the *CG*, noting that "commitments do not have to be negotiated on the table: If a speaker commits to a proposition, this commitment usually is not questioned. So, the commitment can immediately be added to the common ground" (Döring 2016: 34). If we take the above elements to specify (properties of) the speaker's commitments this would mean that *Attit(p)* is placed into the *CG* directly. This is in any case a reasonable assumption since speaker attitudes are not under discussion and cannot be reacted to via normal (non-meta) linguistic means. We can reasonably assume that *ja* takes scope over these attitude-modifying elements, as can be inferred e.g., from their unmarked word order:

(23) Peter ist *ja* offensichtlich/ vermutlich/ vielleicht/ wahrscheinlich/ sicherlich zu Hause.
'Peter is JA obviously/ presumably/ maybe/ probably/ surely at home.'

An account (that is beyond the scope of this paper) of how epistemically modified propositions evade the *Table* (while making the core proposition at issue) together with the assumption that *ja* exerts its effect on the maximal element in its scope predicts that the conditions for the occurrence of *ja* in these cases are trivially satisfied. This also constitutes one reason for letting *ja* merely add the *CG* effect instead of replacing the *Table* function.

To sum up, the above proposal, together with independently necessary assumptions about attitude-modifying adverbials, correctly predicts the behaviour of *ja* in question-answer pairs. It is also transparent how previous proposals about the epistemic status of the proposition arise as cancellable pragmatic inferences. To

impose a *CG*-update requires some form of justification. The most uncontroversial way is if *p* is in the *CG* already which leads to the common use of *ja* as an apparent *CG* marker. (Note that under a strict reading of our definition, a DP as *CG* marker is not possible since it would place restrictions on preceding contexts, not just modify a speech act effect.) Direct visual evidence or obvious conclusions can serve as other justifications. However, no single justification is a necessary condition in every utterance, a problem that the current approach avoids. Instead, the above proposal formalizes notions such as 'uncontroversial' or 'firmly established'.

A remaining question is what constitutes felicitous contexts for *ja* if raising or directly addressing an issue necessarily has to proceed via the *Table*. In the following answer by a teacher to the question '*What are we going to do today?*' by her class, waiting outside the gym:

(24) Es regnet *ja*, wir werden also heute nicht rausgehen.
'It rains JA, so we won't go outside today.'

The clause containing *ja* does not address the *QUD* directly, providing only justification for the following clause. It is those utterances that do not interact directly with the *QUD* that are potential contexts for *ja*. For this reason, *ja* often appears in side remarks and sentences that add background information or set the stage for the main issue. In the following all-new, discourse-initial utterance, the *ja*-utterance simply serves as segue for the actual issue the speaker wants to raise:

(25) Ich war *ja* gestern in der Innenstadt, und rate, wen ich da getroffen habe – den Peter!
'I was JA downtown yesterday, and guess who I met there – Peter!'

A sole assertion with *ja* without a follow-up contribution by the speaker is infelicitous. A hearer will expect a follow-up statement as the main point.

(26) A: #Ich war *ja* gestern beim Friseur. *[End of speaker turn]*
A: #'Yesterday, I was JA at the barber's.'
B: …so?

That is, as long as *ja* does not raise the main issue or resolves an existing issue directly, it is allowed to occur. That some clauses do not interact with the *QUD* is a known effect for other phenomena such as appositives. A recent account of appositive content by AnderBois et al. (2015) for dynamic semantic frameworks such as F&B analyses e.g. non-restrictive relative clauses as imposed *CG*-updates, in contrast to the proposal nature of at-issue content.

(27) Andrew Wiles, who proved Fermat's Last Theorem, received the Fields Medal.

In their analysis, the content of the matrix clause is placed on the *Table* as a proposal to update the *CG* while the relative clause updates the *CG* directly. This is parallel to the effect proposed for *ja* here. This presents the possibility of *ja* as an overt appositive content marker. While investigating this possibility in depth is beyond the scope of this paper, we mention here a few parallels between *ja* and non-restrictive relative clauses.

First, their account predicts that *ja* is able to occur in every appositive construction that is syntactically able to host a DP. For non-restrictive relative clauses this appears to be true at first glance; we are not aware of counterexamples. A second issue are possible reactions to utterances with *ja*. As a first approximation, different linguistic means than a negative answer are necessary to refute non-at-issue content, e.g. for (27):

(28) B: Wait a minute, I thought he proved the Riemann Conjecture.

However, both '*No*' and '*Wait a minute!*' are possible reactions to *ja*-utterances:

(29) A: Peter war *ja* letztens im Park Sanssouci und -
A: 'Peter has JA been to Park Sanssouci recently and –'
B: Wait a second, I thought he's been to Park Babelsberg.
B′: No, he hasn't.

While this appears problematic for the non-at-issue nature of *ja*-utterances at first, an intriguing parallel can be found here, as well, as discussed by AnderBois et al (2015). Refuting content of a non-restrictive relative clause directly becomes possible in an utterance-final position, despite their effect as direct *CG*-updates (their Example 47/49, adapted):

(30) A: His husband, who had prostate cancer, was being treated at the Dominican Hospital.
B: ??No, he had lung cancer.
B′: No, he was being treated at the Stanford Hospital.

vs.

(31) A: He took care of his husband, who had prostate cancer.
B: No, he had lung cancer.
B′: No, he took care of his brother.

This indicates that linear adjacency to the challenged material is a relevant factor in determining possible responses. Since *ja* can modify full utterances, i.e. matrix clauses that are not embedded in contrast to relative clauses, directly adjacent responses are always possible in those cases, and a similarly extended range of

reactions is expected (it might also be another reason to let the *ja*-assertion affect the *Table* at least notationally).

To conclude, even though the current proposal does not depend on an exact parallelism, effects like the one proposed here for *ja* are known and required for other phenomena that also display additional similarities.

4. *Etwa & nicht*

4.1 *Nicht*

We now turn to the particles *etwa* and *nicht* which we argue, following Gieselman & Caponigro (2013), represent two sides of the same coin. We start with *nicht.* The idea of *nicht* as a DP follows Thurmair's (1989) original classification. As Thurmair describes, the sentences below invariably convey that the speaker thinks that the positive proposition is true or strongly suspects a positive answer:

(32) Ist mein Baby *nicht* wunderschön?
'Isn't my baby beautiful?'

(33) Hat Peter *nicht* Marie eingeladen?
'Didn't Peter invite Mary?'

(32) is in essence a rhetorical question, only allowing for a positive answer. (33) expresses the expectation of the speaker that she thought Peter invited Mary. In these examples, *nicht* does not express propositional negation like its homophonous counterpart, i.e., a rough paraphrase of (33) is "*Peter invited Mary, am I right?*" and not "*Peter did not invite Mary, am I right?*".

Elements that are homophonous to negation that have a similar meaning as this particular German *nicht* have also been observed in English and other languages under the label high/preposed negation (Ladd 1981; Romero & Han 2004; more recently Goodhue 2019 and AnderBois 2019). While the present proposal is independent of whether it also applies to other languages, the semantic properties of high negation proposed in the literature are highly parallel to the ones of *nicht*. We, therefore, treat it as the same phenomenon as null hypothesis and only discuss properties from the previous literature where *nicht* behaves equivalently. We start by providing an overview of the literature and then elaborate on the reasons why we depart from some of the assumptions of previous approaches.

The literature that we review here briefly (Romero & Han 2004; Goodhue 2019; AnderBois 2019) shares certain basic assumptions even though the proposals differ in detail. First, all authors assume that high negation is still a form of negation. The

difference in meaning comes about since it is structurally higher than standard negation. The proposals differ more strongly in how exactly this leads to the new meaning contribution. However, a second shared assumption is that the meaning of high negation questions comes about via pragmatic inference/implicature. In Romero & Han (2004), high negation gives rise to an epistemic VERUM operator which the negation scopes over. In their approach, this leads to unbalanced question partitions. Reasoning via Gricean Maximes and principles of economy leads the hearer to assume that such a sentence is only felicitous when resolving an epistemic conflict. The pronounced content of such questions, "*Are you not sure that we should add to CG that p?*", asks the hearer for reasons to doubt *p*; therefore, *p* must be the speaker's belief (or bias) and ¬*p* the hearer's. Goodhue's (2019) account is quite similar, except there is an epistemic operator *O* instead, and the pragmatic reasoning involves conversational principles of utility. AnderBois' (2019) approach uses two-tiered semantics that distinguishes between main and projected issues (*QUD*s) that utterances can raise/resolve and steer the conversation towards. High negation questions are cast as raising the same issue as default polar questions, while not projecting a secondary issue towards which it steers the conversation. Reasoning over the utility of projected issues leads to the bias contribution: default polar questions steer the conversation towards a positive resolution; since high negation questions do not, they give more importance to the negative answer. Such a move is felicitous where there is a tension between the speaker's prior belief and new evidence. A bias for a positive answer arises as a default preference for maintaining one's beliefs.

The last shared assumption is that the core contribution of higher negation (whether it is framed as positive belief, bias, answer expectation etc.) is invariably present in all occurrences.

However, as far as we can see the last two assumptions are a contradiction in terms. Since the above accounts rely on pragmatic reasoning, the question arises why this meaning contribution cannot be cancelled if it is a conversational implicature. The pragmatic reasoning from the compositional object to the pragmatic inference is also by no means inevitable in all three accounts presented here. In Romero & Han's (2004) (and similarly, Goodhue's 2019) account, the denotation of a high negation question is to ask the addressee if she is sure or not to add *p* to the *CG*, whereas the resulting inference is one of speaker bias. Even if such an inference were to arise, there is no reason why it cannot be cancelled. A bias for positive or negative resolution arises in default polar questions with positive/negative polarity. This notion of bias, however, is markedly different from the one in high negation questions, as all authors agree: the former can be cancelled, the latter cannot. The same criticism applies to AnderBois' (2019) approach: there is no reason why a

question that does not privilege any form of resolution is not understood as the ultimate open question. Even if it were not, it should in any case be even easier to cancel such a bias than in polar questions, contrary to fact. We, therefore, conclude that a lexical account of its contribution is necessary. We take the invariable question bias that results if high negation appears in a sentence to *be* its meaning.

Let us address an at first sight unwanted consequence. It is clear that by this classification, DP '*nicht*' bears no resemblance any more to standard negation. That is, it is not the same item which by virtue of being in a different syntactic position than usual results in a different meaning than it would in a lower position. The result is two apparently accidentally homophonous items with completely different meanings. As Goodhue (2019: 1) puts it, several languages possess high negation and "this remarkable cross-linguistic fact merits explanation". However, it is a defining feature of German DPs that they virtually all have homophonous counterparts, usually other particles like conjunctions and adverbial elements to which they are historically related (e.g., Bayer & Obenauer 2011). To illustrate, *ja* has counterparts in a homophonous answer particle, a stressed version used only in imperatives and a tag marker. In this light, the fact that in German a proposed DP has a homophonous counterpart is not a problem but rather expected. The fact that propositional negation is universally available in the world's languages further demystifies the cross-linguistic appearance of this homophony and rather hints at a common route of grammaticalization. To conclude the argument, we treat DP '*nicht*' like DP '*etwa*' in the next subsection whose homophony to its truth-conditional counterpart meaning *circa* or *roughly* is not deemed to be in need of a synchronic explanation.

Let us now illustrate the merits of a lexical account of DP '*nicht*': on the one hand, it gets rid of the oxymoron of an uncancellable conversational implicature and explains why its presence always contributes the same meaning. This is a favourable outcome. On the other hand, it avoids several difficulties that arise from treating 'high negation' as negation. First, there is so far no mechanism or motivation why negation is moved to a higher than usual position. DPs, on the other hand, are assumed to modify speech acts in the line of research pursued here and, therefore, scope high by default. Secondly, there is also no reason why placing negation higher in the structure leads to the presence of an epistemic or verum operator (AnderBois' (2019) proposal does not suffer from this problem). Another question that is to our knowledge not discussed is why high negation can only appear in questions. Standard negation can appear in every sentence type. If high negation and standard negation are the same item, leading only to different effects by negating different complements, high negation should be able to appear in imperatives and assertions as well, an apparently unaccounted phenomenon. For DPs on the

other hand, idiosyncratic distribution across sentence types is again an expected and unsurprising property.

The only approach mentioned that avoids those difficulties is the one by Gieselman & Caponigro (2013) that differs strongly from the ones previously presented: they make use of Gunlogson's (2008) framework for discourse structure; treating them as DPs, they describe *nicht* (and its polar opposite *etwa*) as imposing discourse conditions regarding the evidence and beliefs of the speaker. Gunlogson (2008) accounts for rising declaratives by proposing contingent commitment, a weaker form of commitment that lets the speaker commit to the proposition only as long as it is supported by the authority of the addressee. This models the in-between character of rising declaratives as declaratives that require an answer like questions. The discourse effect of a question with *etwa* and *nicht* is similar in Gieselman & Caponigro's (2013) approach in that it adds a contingent commitment to the commitment set of the speaker, with the additional requirement that the speaker has some prior evidence for p ($\neg p$ in the case of *etwa*). Although our approach is similar by treating *nicht* as modifying discourse functions, we reject this proposal based on its reliance on notions like evidentiality.

We now turn to determine the core contribution of *nicht*. AnderBois (2019) provides a detailed and critical overview of the various contributions of high negation proposed in the previous literature which he groups into four categories: the first is a perceived emphasis on the truth of the proposition; we will not address this pragmatic connotation here. The second is contextually given negative evidence going against the speaker's bias/belief (AnderBois' Example 35).

(34) A: Okay, now that Stephan has come, we are all here! Let's go!
B: Isn't Jane coming?

As AnderBois discusses such a tension is not a necessary condition for the use of *nicht* (see e.g. (32) where no tension is present). Additionally, we are not concerned here with the situations in which *nicht* can be strategically used but its meaning. This leaves two major notions that AnderBois distinguishes (although the distinction is not necessarily always made overtly in the literature): one is the speaker attitude towards the proposition which runs under different names like positive prior belief, a positive epistemic/doxastic implicature or a bias towards the truth of the proposition. The other is a bias in the sense of an expectation/desire for a positive answer. We attempt to show that notions concerning both speaker's beliefs and desires are cancellable.

It is possible to use *nicht* in genuine information seeking questions:

(35) *Online question about how to certify knowledge of Microsoft Office for job applications without paying a lot of money*
Gibts *nicht* vielleicht einen Online-Test, der aber trotzdem seriös genug ist, um ihn in den Lebenslauf zu setzen?
'Isn't there maybe an online test that's still professional enough to put it in your CV?'[1]

Here the person who is asking does not know of any such possibilities (hence, their question), s/he merely rather wants the outcome to be positive. In the following example, the speaker attitude is not epistemic but bouletic in nature:

(36) Mir ist langweilig. Hättest du *nicht* Lust, klettern zu gehen?
'I'm bored. Wouldn't you like to go climbing?'

Here, knowledge or even belief about the truth of the proposition is not necessary (just as Gieselman & Caponigro's (2013) requirement of prior evidence), and mere desire for its truth suffices in the above examples. However, it is equally possible to find cases where a positive answer is the one that is known or believed but not desired:

(37) A: Are you up for a LAN-party tomorrow?
B: Müssen wir *nicht* fünf Essays für Morphologie bis Freitag schreiben?
B: 'Don't we have to write five essays for morphology until Friday?'

Nonetheless, it is still the case that in all these examples one question resolution receives a privileged status although it cannot invariably be based on a single notion of likelihood or desired outcome. What remains is the expectation of a positive answer to a question with *nicht*. The F&B model allows us to encode the privileged status of a specific answer while also circumventing the problem of defining the source for speaker expectation by using the *PS*. F&B (2010: 88) introduce the *PS* so that "the anticipatory nature of certain conversational moves is captured". The task of the *PS* is to identify privileged future states; e.g. while many reactions are possible to an assertion, confirmation has a special or default status. This is captured by the *PS* projecting confirmation as the privileged future state. This is invariant even where a contradiction by the interlocutor is expected. Default polar questions project both members of their denotation as possible answers since both options are equally privileged (as far as grammar is concerned).

1. Available on the following website: https://www.gutefrage.net/frage/fuer-den-lebenslauf-microsoft-office-kenntnisse-nachweisen (last access: 18 March 2021).

We propose that it is this abstract, grammatically relevant form of default resolution, independent of likelihood-based expectation, that *nicht* expresses. Just as confirmation is the default response to an assertion, the default reaction to a *nicht*-question is the answer that confirms the proposition expressed in its radical. Formally, *nicht* modifies moves to project only their positive sentence radical, turning a polar question from projecting two possible futures $\{p,\neg p\}$ into one where it only projects p.

(38) **nicht**(PQ(S[I], K_i)) = K_o such that
i. $T_o = push([S[I]; \{p,\neg p\}], T_i)$
ii. $PS_o = PS_i \cup \{p\}$

As for the previously proposed meanings for *nicht* it is plausible that they arise as common inferences from the underlying meaning. A prior belief that p is the case, or a desire for it to be so are plausible reasons to ask a question with a privileged outcome. That a question that is biased towards confirmation can appear in contexts where a prior belief is apparently challenged is a natural use. That the relevance of the truth of such an issue receives emphasis in such a situation is also expected.

4.2 *Etwa*

Finally, we turn to *etwa* which in Gieselman & Caponigro's (2013) conception is the opposite side of the coin. The proposed contributions of *etwa* are parallel to the ones for *nicht* except for the reversed polarity. Thurmair gives the following example for the use of *etwa* which in its discourse particle use can only occur in polar interrogatives (Thurmair 1989: 170, translation for all her examples by us):

(39) Uwe: Gestern war 'Dallas' vielleicht wieder spannend!
Uwe: "Dallas' was exciting again, yesterday!'
Mona: Was?! Schaust du dir den Quatsch *etwa* an?
Mona: 'What?! Do you watch that nonsense ETWA?'

She paraphrases the contribution of *etwa* as indicating that Uwe's utterance contradicts Mona's prior assumptions about his TV habits. For the use of *etwa*, the speaker needs some form of evidence that p could be the case. What causes the speaker to pose a question with *etwa* is the conflict between their prior expectation and the current evidence that they are faced with (this is the reverse to prior positive belief and tension with negative evidence in the case of *nicht*).

Another connotation that accompanies the use of *etwa* is that the speaker expects or hopes for a negative answer, or conversely, fears a positive answer. Further examples illustrate this (ibid.):

(40) *Max puts on his jacket:*
Bea: Willst du *etwa* jetzt noch weggehen? Es ist 1 Uhr!
Bea: 'Do you ETWA want to go out now? It's 1am!'

(41) Die Whisky-Flasche ist schon wieder leer. Trinkst du *etwa* heimlich?
'The whisky bottle is empty again. Are you ETWA drinking secretly?'

To summarize, Thurmair (1989) takes the contribution of *etwa* to be the following: the speaker has evidence that p (and is surprised about that), and the speaker expect(ed) $\neg p$ to hold and hopes that $\neg p$ be the case. We pursue the same strategy as with *nicht* by trying to show that those are cancellable contributions. What remains is the abstract notion of question bias which we argue is informed by but nonetheless independent of notions like expectations and wishes.

As already conceded by Thurmair (1989), there are examples where the negative attitude towards p is almost absent. The following example shows that the opposite can be true where the speaker wishes p to be the case:

(42) *Party guests arrive with an enormously large cake.*
Host: Ist der *etwa* für mich?
Host: 'Is that ETWA for me?'

While in the above example modesty requires projection of a *no*-answer, it shows that *etwa* can felicitously be used where the speaker hopes for a positive answer. Csipak & Zobel (2014) construct similar examples and argue that only a change in the likelihood assigned to contrasting beliefs is necessary. They propose that *etwa* signals that the speaker realized she mistakenly believed the negative answer to the question to be more likely than the positive one.

While Thurmair talks about the (prior) expectation of $\neg p$ and evidence for p, Gieselman & Caponigro (2013) argue that *etwa* requires (prior) evidence for $\neg p$, with no mention of a necessary conflict between expectation and evidence. Their example involves a situation in which famous German actresses visited a school, but the speaker has some evidence to believe a specific actress was absent and can felicitously ask (their Example 1b and 4):

(43) War Veronica Ferres *etwa* gestern auch an der Schule?
'Was Veronica Ferres also at school yesterday by any chance?'

Parallel to their treatment of *nicht*, they propose that a question with *etwa* adds a contingent speaker commitment with the additional requirement that the speaker have some prior evidence for $\neg p$. The literature, therefore, appears contradictory; at least it is unclear if what Gieselman & Caponigro (2013) call evidence is not better termed speaker expectation. In (42) e.g., it is difficult to find what the source of the negative evidence is. Take (44):

(44) *A couple is walking around in the MOMA. One of them notices a particularly unconventional piece*:
Soll das *etwa* Kunst sein?
'Is that ETWA supposed to be art?'

Here, all available evidence points towards a positive answer. With *etwa*, the speaker insinuates judgement by expressing preference for a negative answer. As with *nicht*, this indicates that the relevant attitude need not be epistemic but can border on a bouletic/deontic reading:

(45) Ist das *etwa* deine Vorstellung von Anstand?
'Is that ETWA your idea of decency?'

In conclusion, while a negative answer is projected in every case, its motivation can range from prior expectation (40), hope (41) or a bouletic preference bordering on rhetorical questions (44–45), yet the polar opposite of each attitude can appear if it is not the source of negative projection (positive evidence/expectation (44), hope/bouletic preference for p (42)).

As with *nicht*, we circumvent attempts to find an all-encompassing description of potential speaker states by reducing the meaning of *etwa* to the independent grammatical primitive of unmarked utterance resolution and by relegating questions of possible motivations for such an utterance to pragmatic considerations. The formal account of *etwa* is equivalent to the one of *nicht*, modulo the reversed polarity. A polar question modified by *etwa* projects only a default negative resolution $\{\neg p\}$ in its *PS*.

(46) **etwa**$(\mathrm{PQ}(S[\mathrm{I}], K_i)) = K_o$ such that
i. $T_o = push([S[\mathrm{I}]; \{p, \neg p\}], T_i)$
ii. $PS_o = PS_i \cup \{\neg p\}$

As with *nicht*, similar but polar opposite contexts are natural occurrences of *etwa* such as resolving tension, eliciting a specific answer or insinuating a bias. However, what informs the bias and what motivates a speaker to use a bias question should not be accounted for in a semantic account of its meaning.

To conclude this section, a comment on the evidence used throughout this paper. A reviewer raises the relevant question whether we indeed got rid of attitudes since they still play an influential role in the argumentation for the final definition. While this is true, all attitudes have only been used *ad absurdum*, not to argue for yet another attitude, showing that *no* attitude is uniquely predictive of a DP's use. The reviewer comments that the predictions of the F&B model, with the current approach to these DPs, are not fully understood, that it could not a priori predict e.g. if *etwa* can be used in questions where one hopes for one but expects the

opposite answer, or the other way around; or which responses are ungrammatical. The aim of this paper is to reduce DP effects to structural properties of discourse. The ban of *ja* in answers provides good negative evidence that its effect violates a fundamental principle of discourse grammar. Negative evidence in the case of *etwa*/*nicht* is admittedly harder to come by since the *PS* provides only a markedness ranking for resolutions without grammatically excluding some. However, the same situation obtains with the assertion operator. Assertion confirmation as unmarked resolution is a reasonable assumption, yet negative evidence is hard to come by. But the assertion operator is not expected to predict (meta-theoretic reasoning over issues such as) whether assertions convey a hope or likelihood of confirmation or fear of contradiction (since its confirmation projection stays constant in both confrontational debates and dialogues with like-minded people). If we accept a privileging of resolutions as part of our discourse grammar and a DP's effect is an altered privileged response (while resisting any attempt to base this ranking on some attitude), then we should not expect the theory to predict such speaker states for the use of those DPs, either. We illustrated above that none of the previously proposed attitudes are primitive notions of grammar. We, therefore, view it as a merit, not a problem, that the F&B model has nothing to say about these connotations.

5. Conclusion

We set out to reduce DPs to the primitives of the F&B model, driven by the motivation to replace descriptive attitude-based notions with formally defined discourse effects. For each of the three particles, their apparent indication of a speaker attitude can be cancelled in an appropriate context. We take the fact that such counterexamples can be found in the general case as indicator that it is a useful heuristic to treat such notions as epiphenomenal ('*Vulcan hypothesis*'), as well as validation for the criticism towards attitude-based approaches expressed in the first section.

Moreover, every particle was found to alter the discourse behaviour of its host utterance, true to their name. With the proposed discourse effects, it was still transparent how the respective attitudes arise as frequent inferences; that no single such notion has to be the necessary and sufficient condition for a particle's use, however, is an advantage of this discourse approach. We have taken the ban on *ja* in answers as indicator that the particle modifies its host utterance's interaction with the *QUD*, parallel to a proposal about non-restrictive relative clauses in the F&B model; a welcome reduction of superficially distinct phenomena if this parallelism turns out to be exact. Some form of epistemic justification for an imposed *CG* update is expected, but no single one is necessarily present in every utterance. The effect that *etwa* and *nicht* indicate a privileged answer to a polar question could be

modelled using the *PS*. A polar question that has an unmarked default resolution in the negative or positive supersedes a definition of bias in terms of epistemic or bouletic attitudes. Such stances are predestined to inform biases but again do not represent the core contribution of these particles. If the current approach to (a subset of) German DPs is on the right track, investigating the speaker attitudes of typical contexts for a particle can only be descriptive work since one would otherwise conflate secondary effects with the DP's actual contribution.

The modelling of DPs as speech act modifiers also makes it possible to propose an answer to the question what the allowed variation of the meaning of DPs is (the second motivation behind this paper). If DPs can only cause a single modification to the way speech acts affect the discourse components, their variation is significantly restricted due to the small number of these components, and their status as a small, closed class of elements follows naturally. This work focused on two components, the *Table* and *CG* that model the (non-)interaction with the *QUD* and the *PS* that captures default resolutions of speech acts. The particle *ja* encodes the marked version of the binary choice whether a move contributes a *QUD* (i.e. whether or not it interacts with the *Table*); the particles *etwa* and *nicht* each subtract one possible default resolution to an open polar question, realizing both options of the simplest form of variation in the *PS* for questions. Put differently, if one would have to formulate hypotheses about potential variation within these components for non-standard speech acts, the above effects would be natural examples.

Further research is required to determine which other particles can be described using the definitions proposed here, and which discourse components they interact with.

References

AnderBois, Scott. 2019. Negation, alternatives, and negative polar questions in American English. In *Questions in Discourse – Volume 1: Semantics* [Current Research in the Semantics/Pragmatics Interface 35], Klaus von Heusinger, Malte Zimmermann & Edgar Onea (eds), 118–171. Leiden/Boston: Brill.

AnderBois, Scott, Brasoveanu, Adrian & Henderson, Robert. 2015. At-issue proposals and appositive impositions in discourse. *Journal of Semantics* 32(1): 93–138. https://doi.org/10.1093/jos/fft014

Bayer, Josef & Obenauer, Hans-Georg. 2011. Discourse particles, clause structure, and question types. *The Linguistic Review* 28: 449–491. https://doi.org/10.1515/tlir.2011.013

Csipak, Eva & Zobel, Sarah. 2014. A condition on the distribution of discourse particles across types of questions. In *Proceedings of NELS* 44(1), Jyoti Iyer & Leland Kusmer (eds), 83–94. Amherst, MA: GLSA.

Döring, Sophia. 2016. Modal Particles, Discourse Structure and Common Ground Management. PhD dissertation, HU Berlin.

Egg, Markus & Mursell, Johannes. 2016. The syntax and semantics of discourse particles. In *Discourse Particles: Formal Approaches to their Syntax and Semantics* [Linguistische Arbeiten 564], Josef Bayer & Volker Struckmeier, 15–48. Berlin/Boston: de Gruyter. https://doi.org/10.1515/9783110497151-002

Farkas, Donka. 2007. The grammar of polarity particles in Romanian. Manuscript, <https://people.ucsc.edu/~farkas/papers/polarity_particles.pdf> (17 March 2021).

Farkas, Donka. & Bruce, Kim. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27: 81–118. https://doi.org/10.1093/jos/ffp010

Gieselman, Simone & Caponigro, Ivano. 2013. Two sides of the same pragmatic move: the German discourse particles *etwa* and *nicht*. In *Proceedings of NELS* 40(1), Seda Kan, Claire Moore-Cantwell & Robert Staubs (eds), 231–244. Amherst, MA: GLSA.

Goodhue, Daniel. 2019. High negation questions and epistemic bias. In *Proceedings of Sinn und Bedeutung* 23(1), M. Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally & Cristina Real-Puigdollers (eds), 469–486. Barcelona: Universitat Autònoma de Barcelona.

Grosz, Patrick. 2014. German *doch*: an element that triggers a contrast presupposition. In *Proceedings of the Forty-Sixth Annual Meeting of the Chicago Linguistic Society* 46(1), Rebekah Baglini, Timothy Grinsell, Jon Keane, Adam Roth Singerman & Julia Thomas (eds), 163–177. Chicago: Chicago Linguistic Society.

Gunlogson, Christine. 2008. A question of commitment. *Belgian Journal of Linguistics* 22(1): 101–136. https://doi.org/10.1075/bjl.22.06gun

Gutzmann, Daniel. 2016. Modal particles ≠ modal particles (= modal particles). In *Discourse Particles: Formal Approaches to their Syntax and Semantics* [Linguistische Arbeiten 564], Josef Bayer & Volker Struckmeier, 144–172. Berlin/Boston: de Gruyter. https://doi.org/10.1515/9783110497151-007

Jacobs, Joachim. 1991. On the semantics of modal particles. In *Discourse Particles: Descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German* [Pragmatics & Beyond New Series 12], Werner Abraham (ed), 141–162. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.12.06jac

Kaufmann, Magdalena & Kaufmann, Stefan. 2012. Epistemic particles and performativity. In *Proceedings of SALT* 22, Anca Chereches (ed), 208–225. Ithaca, NY: CLC Publications. https://doi.org/10.3765/salt.v22i0.2635

Kratzer, Angelika. 1999. Beyond ouch and oops: how descriptive and expressive meanings interact. Handout for a talk at the Cornell Conference on theories of context dependency, Cornell University: Ithaca, NY.

Kratzer, Angelika & Matthewson, Lisa. 2009. Anatomy of two discourse particles. Handout for a talk at SULA 5, Harvard.

Ladd, Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society* 17(1), Robert A. Hendrick, Carrie S. Masek & Mary Frances Miller (eds), 164–171. Chicago Linguistic Society.

Lindner, Katrin. 1991. 'Wir sind ja doch alte Bekannte' The use of German *ja* and *doch* as modal particles. In *Discourse Particles: Descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German* [Pragmatics & Beyond New Series 12], Werner Abraham (ed), 163–202. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.12.07lin

Romero, Maribel & Han, Chung-Hye. 2004. On negative yes/no questions. *Linguistics and Philosophy* 27: 609–658. https://doi.org/10.1023/B:LING.0000033850.15705.94

Rudin, Deniz. 2020. Intonational Commitments. Manuscript, USC, <https://denizrudin.github.io/papers/rudin_2020_intonational_commitments.pdf> (17 March 2021).
Stalnaker, Robert. 1978. Assertion. *Syntax and Semantics* 9: 315–332.
Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen*. Tübingen: Niemeyer. https://doi.org/10.1515/9783111354569
Viesel, Yvonne. 2015. Discourse structure and syntactic embedding: the German discourse particle *ja*. In *Proceedings of the 20th Amsterdam Colloquium* 20, Thomas Brochhagen, Floris Roelofsen & Nadine Theiler (eds), 418–427. Amsterdam: University of Amsterdam.

Language index

Subject index

Discourse particles have often been treated as a phenomenon restricted to Germanic languages (Abraham 2020) and they still raise questions about their nature as an independent category. This book reveals that this phenomenon exists in other languages as well, and provides evidence for their nature as a separate category. The volume brings together a collection of nine papers that focus on three research topics: (a) the diachronic development of discourse particles; (b) their syntactic analysis; and (c) the study of their semantic-pragmatics. Furthermore, it also discusses other issues less often dealt with in the literature but of great interest for linguistic theory, such as the acquisition of discourse particles by children or the analysis of elements not usually considered discourse particles but whose historical path or microvariation indicates otherwise. Additionally, the book offers a cross-linguistic perspective as it discusses various languages including Basque, Catalan, German, Italian, Laz, Mandarin Chinese, Old English, Portuguese, and Spanish.

ISBN 978 90 272 1107 1

John Benjamins Publishing Company