



Discours

Revue de linguistique, psycholinguistique et informatique

35 | 2024
Varia

Discourse Markers as Cues for Topic Shift Interpretation: Experimental Evidence from Spanish

Marqueurs discursifs comme indices pour l'interprétation du changement de topique: preuves expérimentales en espagnol

Mercedes Villalobos Cardozo, Ludivine Crible and Liesbeth Degand



Electronic version

URL: <https://journals.openedition.org/discours/13307>

DOI: 10.4000/1356z

ISSN: 1963-1723

Publisher:

Laboratoire LATTICE, Presses universitaires de Caen

Electronic reference

Mercedes Villalobos Cardozo, Ludivine Crible and Liesbeth Degand, "Discourse Markers as Cues for Topic Shift Interpretation: Experimental Evidence from Spanish", *Discours* [Online], 35 | 2024, Online since 23 December 2024, connection on 23 January 2025. URL: <http://journals.openedition.org/discours/13307> ; DOI: <https://doi.org/10.4000/1356z>



The text only may be used under licence CC BY 4.0. All other elements (illustrations, imported files) are "All rights reserved", unless otherwise stated.

Discourse Markers as Cues for Topic Shift Interpretation: Experimental Evidence from Spanish

Mercedes Villalobos Cardozo

University of Louvain, Institute for Language and Communication (ILC)

Ludivine Crible

Ghent University, Linguistics Department

Liesbeth Degand

University of Louvain, Institute for Language and Communication (ILC)

.....
Mercedes Villalobos Cardozo, Ludivine Crible, Liesbeth Degand, «Discourse Markers as Cues for Topic Shift Interpretation: Experimental Evidence from Spanish», *Discours* [En ligne], 35 | 2024, mis en ligne le 23 décembre 2024.

.....
URL : <http://journals.openedition.org/discours/13307>

.....
Titre du numéro : *Varia*

Coordination : Anne Le Draoulec & Josette Rebeyrolle

Date de réception de l'article : 09/04/2024

Date d'acceptation de l'article : 10/07/2024

Discourse Markers as Cues for Topic Shift Interpretation: Experimental Evidence from Spanish

Mercedes Villalobos Cardozo

University of Louvain, Institute for Language and Communication (ILC)

Ludivine Crible

Ghent University, Linguistics Department

Liesbeth Degand

University of Louvain, Institute for Language and Communication (ILC)

.....

Despite the complex terminological panorama regarding discourse markers (hereafter DMs) and the elusiveness of the notion of discourse topic, the relationship between these linguistic phenomena is a long-standing one. Descriptions linking the two are present in early work on DMs and are still the subject of ongoing discussion. Often, however, these characterisations, which present certain DMs as cues for topic management in interaction, are made from a corpus-based perspective. Here, as part of a larger mixed-methods project investigating DMs from production to prediction, we explore participants' interpretations of discourse topic in conversation when introduced by three DM configurations in Spanish: *y* ("and"), perceived lengthened *y* ("and"), and the combination *bueno y* ("well and"). This is achieved through an experimental forced-choice task and the qualitative analysis of the material presented, based on authentic data in Spanish. By doing so, we examine the intricate relationship between DMs and their context to disentangle their actual contribution to the interpretation of discourse topic. Results show that these DMs do not override the semantic content of the utterances and, therefore, only moderately contribute to recognising topic continuity or topic discontinuity. An analysis of the content of the utterances presented in our stimuli allowed us to create a set of strong and weak strategies for topic (dis)continuity that seem to help speakers with the topic interpretation task.

Keywords: Spanish, discourse markers, discourse topic, experimental, interaction

.....

Malgré la complexité terminologique entourant les marqueurs discursifs (MD) et la nature évasive du concept de topique discursif, la relation entre ces phénomènes linguistiques est étudiée depuis longtemps. Des descriptions les reliant apparaissent dès les premiers travaux sur les MD et continuent de faire l'objet de discussions actuelles. Cependant, la plupart de ces études, concluant que certains MD fonctionnent comme des indices de gestion du topique en interaction, sont exclusivement basées sur corpus. Ici, dans le cadre d'un projet plus vaste utilisant des méthodes mixtes pour étudier les MD de la production à la prédiction, nous explorons les interprétations des participants concernant le topique discursif dans une conversation, lorsqu'il est introduit par trois configurations de MD en espagnol: y («et»), y («et») perçu comme prolongé, et la combinaison bueno y («bien et»). Pour ce faire, nous utilisons une tâche expérimentale de choix forcé et l'analyse qualitative du matériel présenté, basée sur des données authentiques en espagnol. Ce faisant, nous examinons la relation complexe entre les MD et leur contexte pour démêler leur contribution à l'interprétation du topique discursif. Les résultats indiquent que le poids de ces MD ne dépasse pas celui du contenu sémantique de l'énoncé et qu'ils n'ont donc

pas un impact significatif sur l'interprétation du topique. L'analyse du contenu des énoncés de nos stimuli a permis d'identifier des stratégies, tant fortes que faibles, pour la (dis)continuité du topique, qui semblent aider les locuteurs dans la tâche d'interprétation du topique.

Mots clés : *espagnol, marqueurs du discours, topique discursif, expérimental, interaction*

.....
Funded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 859588. We would like to thank the two anonymous reviewers for their detailed and constructive comments. All remaining errors are ours.

1. Introduction

1 Discourse markers (henceforth DMs) are metalanguaging expressions that provide information about the use of language itself and refer to relations between speakers and/or cognitive processes (e.g., Maschler, 2009). Although the study of DMs is a prolific and well-developed field, and several models have been proposed over the years (Maschler & Schiffrin, 2015; Fischer, 2006), part of its complexity stems from the lack of unanimity on the functions of DMs (Crible, 2017).

2 The mention of topic management in relation to DMs is present from early studies on. In her seminal work, Schiffrin (1987: 26 [our emphasis]) states that “part of an idea structure is its organisation of topics and subtopics – *what is being talked about*. Unfortunately, *I have no solution to propose as to how to find topics and subtopics*, although it often *seems intuitively very clear*, especially when topics shift”. Within the debate around discourse topic, this excerpt already refers to two important points: the lack of a unanimous definition for the concept of (discursive) topic and the difficulty in identifying topic boundaries.

3 Since then, many studies have focused on the role of DMs in managing topicality (e.g., Charolles, 2020; Fagard & Charolles, 2020; Ionescu, 2020; Traugott, 2020). However, most of the claims have been based on corpus work only. Although this method provides robust results on naturally produced data, corpus-based observations remain insufficient to explain the causality of these patterns (see Zufferey, 2020). Furthermore, corpus studies can account for the production of language, but not for the interpretation of what is being said from the listener's perspective.

4 To complement our knowledge of the interplay between DMs and context from an experimental perspective, we designed and applied an acceptability judgment task where participants decided whether two utterances were about the same topic or different ones. The goals of this study are twofold: to quantify the extent to which DMs influence topic interpretation, and to describe the resources present in the stimuli that help maintain topic continuity or mark its discontinuity. In this framework, we hypothesise that some DMs are better predictors of topic discontinuity than others, but also that the context of occurrence will be influential. Thus, the

overarching goal is to explore the interaction between DMs and their context to assess their actual contribution to topic interpretation, based on a combination of quantitative and qualitative analysis of experimental data.

- 5 This paper is organised as follows: first, we introduce the notions of discourse topic and topic continuity in discourse. Then, we discuss the role of DMs as topic management devices and follow with a description of the DMs of interest. Afterwards, we present our methods and report the results of our analysis, both quantitative and qualitative. Finally, we discuss our findings in the context of existing research and conclude by summarising our main results and suggesting potential directions for future study.

2. Approaches to discourse topic and discourse markers

2.1. About discourse topic

- 6 Defining discourse topic has proven to not be a simple task. A first distinction should be made between the concept of sentence topic – or theme – and discourse topic. Although related, the former is typically studied within the framework of information structure and corresponds to given information in a sentence or what the sentence is about (Lambrecht, 1994). This binary opposition does not apply to the latter notion (Riou, 2015), which is not restricted to the sentence, but studied at the discourse level and will be the focus of our paper.

- 7 Contrary to the ease with which speakers manage topicality in conversation, the effort of arriving at a theoretically unified notion has uncovered discrepancies and overlaps between multiple different approaches to discourse topic. In fact, even if the importance of this concept has been demonstrated for human communication (Palomares et al., 2006), discourse topic has been considered “the most frequently used, unexplained, term in the analysis of discourse” (Brown & Yule, 1983: 70) as well as “the most recalcitrant to systematic analysis” (Atkinson & Heritage, 1985: 165). This complexity of discourse topic suggests that a comprehensive understanding requires a holistic approach, integrating linguistic, cognitive, and interactional dimensions¹.

- 8 In line with the aims of our study, our approach to understanding discourse topic will centre on a key distinction between content-focused and structure-focused approaches – essentially, the “what” versus the “how” (see Goutsos, 1997).

- 9 Content-based approaches are primarily interested in what the discourse is about (Zellers, 2013; Riou, 2015) and understand topic as “the content, subject matter, issues, foci, or ideas on which individuals converse” (Palomares et al., 2006: 60). A major limitation of these frameworks is that, although they seem to have

1. See Givón (1983) for functional approaches, Van Dijk (1981) for cognitive perspectives, and Asher (2004) and Asher and Lascarides (2003) for formal semantic frameworks.

great intuitive resonance, definitions of discourse topic remain vague and difficult to apply to empirical data. Nevertheless, we agree that the “‘what’ of interpersonal communication [...] may be what most people mean by ‘topic’ in the everyday lexicon” (Palomares et al., 2006: 63). It follows that topic is recognisable to speakers and that they rely on aboutness to differentiate between a known topic or a new one.

10 Turning to structure-based approaches, discourse topic is defined as “something procedural that speakers achieve via orderly practices” (Couper-Kuhlen & Selting, 2017: 520). This proposal is defended by conversation analysts, among others, who focus on the devices, patterns, and strategies that participants in an interaction use to achieve topicality management.

11 Within the conversation-analytic approach to discourse topic, a further distinction is made between two types of topic change: stepwise and disjunctive. The first type, stepwise topic shift, is gradual and almost imperceptible: speakers add new information to what has already been said until this becomes the new topic of the interaction (Zellers, 2013). In this case, boundaries between topics are fuzzy and complicated to establish, as they are virtually imperceptible and conversation just seems to follow its “natural” flow (Couper-Kuhlen & Selting, 2017), as in [1]².

- [1] A: ¿A usted le gusta la cocina?
 B: Comerla.
 A: O sea, no sabe cocinar nada.
 B: No, no, no. Nada, nada.
 A: Y <alargamiento> si tuviera que preparar u organizar una fiesta, un cumpleaños, una comunión, un bautizo, ¿cómo lo haría?
 B: Muy bien, diciéndoselo a mi mujer, que es la que se entera ella.
 A: Do you enjoy cooking?
 B: Eating.
 A: So, you don’t know how to cook at all.
 B: No, no. Not at all.
 A: Y <lengthening> if you had to prepare or organise a celebration, a birthday, a communion, a christening, how would you do it?
 B: Ok, I would tell my wife. She is the one who knows about those things.
 (PRESEEA, VALE_H31_050)

12 In example [1], the interaction gradually progresses from a brief discussion initiated by the interviewer (speaker A) about the interviewee’s (speaker B) cooking skills to the hypothetical scenario of speaker B having to organise a celebration. By building on the topic of food, the interviewer wonders how the interviewee would handle the task of planning a celebration, given his acknowledgment of not knowing how to cook.

2. All examples are authentic data from the PRESEEA corpus described in Section 3.1.

- 13 In contrast, in the second type of shift discussed here – disjunctive topic shift – “participants work expressly to mark their turn as not fitted to what has gone before” (Couper-Kuhlen & Selting, 2017: 522). These moves are characterised by disengaging from previous talk in an abrupt manner, which could potentially result in the conversation being perceived as more fragmented, as in [2].

[2] B: [...] me encanta todo el campo y las flores y las plantas, que paso todo el día fuera trabajando el jardín.
 A: A mí también me gusta mucho el jardín y vivo en un chalé y también me dedico los sábados y domingos a eso.
 B: Luego me entenderá.
 A: Sí, sí, sí. Bueno y, una cosa que me interesa: ¿cómo era la relación paterno-filial, es decir, tuya con tus padres y tuya con tus hijos? ¿Has notado algún...?
 B: Mucha diferencia.
 B: [...] I love the countryside, the flowers, and the plants, so I spend all day outside working in the garden.
 A: I also really like gardening, and I live in a house with a garden, so I spend my Saturdays and Sundays doing that too.
 B: Then you'll understand me.
 A: Yes, yes, yes. Well, there's something I'm interested in: how was the parent-child relationship? I mean, yours with your parents and yours with your children. Have you noticed any...?
 B: A big difference.
 (PRESEEA, VALE_M22_035)

- 14 In [2], the interviewee (speaker B) is in the middle of describing her hobbies, specifically gardening, before the interviewer (speaker A) transitions to a considerably different topic: the relationship the interviewee had with her parents compared to the one she maintains currently with her children.

- 15 Regardless of how discourse topics are defined, topic continuity is key to coherence and cohesion. In the next section, we will explore the framework proposed by Givón (1983), which lies at the basis of our own approach, and the scale of resources that support topic continuity, ensuring a coherent flow in discourse.

2.2. Topic continuity in discourse

- 16 Topic continuity in interaction is negotiated by both speaker and hearer, who cooperate to attain conversational coherence (Schiffrin, 1985), which plays a central role in the understanding of discourse and the way it is processed and stored (Das & Egg, 2023). According to Givón (2017: 4):

Human discourse is typically multi-propositional. That is, we string together verbal event-or-state clauses in coherent sequences, ones that maintain a high degree of continuity. The sub-elements – strands – of discourse coherence tend to persist

from one clause to the next across stretches of discourse or clause chains. The overall thematic coherence of human discourse is then the tapestry-like product.

17 Unlike several “micro traditions” that split sentences into “focus” (new information) and “topic” (old information), Givón (1983: 7) views topicality as a “non-discrete entity” within a “macro tradition”. This broader approach offers greater functional and psychological significance and contributes to the notions of discourse continuity and predictability. From this standpoint, the sentence (and not the word) is the “basic information processing unit in human discourse”, which combines with other sentences to create the so-called thematic paragraph (TP henceforth): “the most immediately relevant level of discourse within which one can begin to discuss the complex process of continuity in discourse” (Givón, 1983: 7).

18 Givón’s approach identifies three types of continuity: thematic, action, and participant continuity. These are mediated through the TP and expressed using structural, grammatical, and syntactic elements in sentences, bridging the gap between the “macro” and “micro” levels of language organisation.

19 Thematic continuity is the “overall matrix” for all continuities. It often overlaps with participant and action continuity within a TP, which focuses on the same theme. Thematic continuity is weakly expressed in grammar, typically through conjunctions or clause markers, depending on the language. In turn, action continuity refers to temporal sequentiality in the TP, which is coded via the tense-aspect-modality subsystem; action can shift without changing the theme. Lastly, participant continuity refers to the most crucially involved participant in the action sequence and the most likely to be coded as the grammatical subject of the TP. This continuity is coded through grammatical structures that identify, maintain, and track the topic throughout the discourse. While participants may change, the overall theme or action doesn’t necessarily have to.

20 From this perspective, as shown in Figure 1, devices can be ranked according to their degree of referential continuity. This scale is useful for analysing how speakers choose different referential forms to maintain topic continuity in discourse based on the accessibility of the referent. Since these forms must be studied in a language-specific manner, our qualitative analysis (see Section 4.2) will focus on those strategies present in the stimuli of our task in Spanish.

21 Givón (1983) also elaborates on the several factors that influence how available a discourse topic is, amongst them the distance from the last mention (the shorter the gap of absence, the easier to identify), the presence of other topics (the more topics are present, the more difficult to correctly identify a topic) and the availability of semantic information (the more redundant information within the clause, the more probable that a topic could participate in the clause in the specific semantic/grammatical information).

22 Although this literature does not explicitly mention DMs, we set the present study in this line of device-based approaches and will now turn to exploring the relationship between DMs and discourse topic.

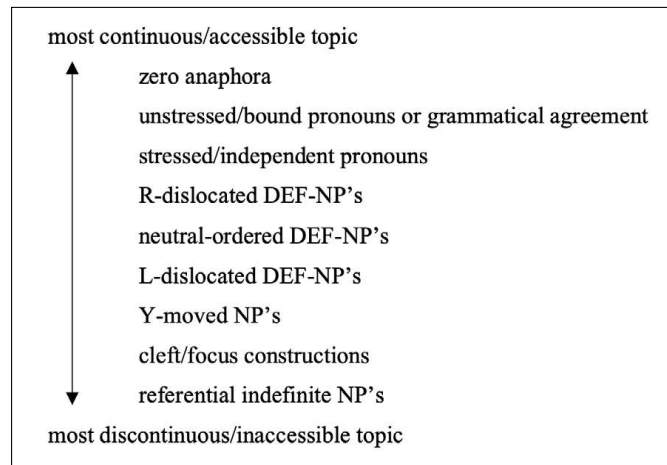


Figure 1 – Scale of referential continuity (based on Givón, 1983: 17)

2.3. Discourse markers as signals (or not) of topic change

- 23 In the realm of DMs studies, topic and DMs interact via the so-called markers of topic shift – also known as topic shifters –, a much less studied category in the field, compared to markers of rhetorical relations (see Fagard & Charolles, 2020). Fraser (2009) calls them “topic orientation markers” and defines them as a type of discourse management marker that signal a meta-comment on the structure of the emerging discourse. These include items such as *back to the point*, *as I was saying*, *speaking of* and *on a different note*. “Topic orientation markers” often co-occur with “attention markers”, such as *alright*, *ok*, *now*, and *well*, which indicate “there is going to be a change, but they do not typically indicate the nature of the change” (Fraser, 2009: 896). Topic orientation markers and attention markers together signal the orientation of discourse in four ways: return to a prior topic; add to, or continue with the present topic; digress from the present topic (e.g., Pons Bordería & Estellés, 2009); and introduce a new topic.
- 24 Besides these specific expressions, DMs in general have been persistently described as strategic devices for topic management, solidifying the idea that they are intentionally used by speakers to signal upcoming topical discontinuity (e.g., Estellés & Pons Bordería, 2014) and interpreted by hearers as such. This confirms the idea of a signal-inference dynamic between speakers during interaction (Couper-Kuhlen, 2004). This inference presumably relies on the semantics of the DM, although this is an issue under discussion, where two main positions are opposed: DMs either create or display meaning. We will review these positions in the following.
- 25 The first approach proposes that DMs create coherence relations and add meaning to discourse. DMs do this by encoding procedural meaning and, to some extent, conceptual meaning (depending on the core meaning of the marker): while the former constrains the possible interpretations of the relations between utterances from the

presence of the DM (Escandell Vidal & Leonetti Jungl, 2004; Escandell Vidal et al., 2011; Zufferey & Degand, 2024), the latter codifies concepts (Murillo Ornat, 2010). This feature of procedurality has been persistently included in definitions of DMs (Fraser, 1999; Bolly et al., 2017; Portolés Lázaro, 2008). Couper-Kuhlen's (2004) view of a signal-inference dynamic is thus supported by this conception of DMs as carriers of procedural meaning. With respect to topic management, DMs are thus seen as carrying a specific instruction that exerts pressure on the surrounding context, which is then reinterpreted as expressing a new topic.

26 In the second approach to the semantics of DMs, they are seen as merely displaying a coherence relation that is already accessible in the context. In this view, the presence or absence of a DM “does not alter the potential discourse relationship between the message which follows and the foregoing discourse” (Fraser, 1990: 390). In other words, the interpretation of utterances is already constrained by the context and therefore already inferable from it. This is in line with Taboada (2006), who argues that DMs select but do not create coherence relations, so that the meaning of the marker must be compatible with the meanings of the surrounding discourse.

27 One of our goals is to contribute to this debate by showing whether DMs can actively constrain context interpretation or simply interact with it. The next section will review the specific markers under study that will be used as a test-bed on this issue.

2.4. The discourse markers in this study

28 This study examines three DM configurations: *y* (“and”), perceived lengthened *y* (“and”), and the combination *bueno y* (“well and”). These DMs were selected based on existing literature and a previous corpus analysis carried out within the framework of the same research project (Villalobos Cardozo et al., forthcoming a). Firstly, while *y* has a core meaning of addition (Gómez Txurruka, 2003), its high frequency and underspecification allow it to perform multiple functions in discourse. Next, for the case of the perceived lengthened version of *y*, corpus analysis suggests it may serve a distinct discourse function, potentially aiding in managing topic shifts (see Section 3.1). Finally, *bueno* (“well”) often functions as a marker of topic shift in Spanish and is typically paired with other conjunctions and connectives (see Fuentes Rodríguez, 2009). In this context, combining *bueno* with *y* strategically allows for the exploration of a continuum of *y* variations. By focusing on these three DM configurations, we can assess their role in topicality management, ranging from underspecification to specialisation. With this foundation in place, we now turn to the literature review to contextualise and support these choices.

29 First, *y* is usually classified as a connective and as a DM that functions to express a coordinating relationship in discourse (see Gómez Txurruka, 2003). Specifically in the realm of topic management, García García (2020) found three functions of *y* in speakers of Spanish as a second language: i) topic proffering – to initiate talk –, ii) a “masked” topic shift – to change topics in an unmarked way – and iii) topic continuation, following the “normal expectation” of the marker.

30

Previous work on *and* – its correlate in English – has shown that its meaning in discourse is mostly due to its use in context and marginally due to fixed semantic information (Turk, 2004: 224). In this line, Crible and Cuenca (2017) found up to 11 different types of discourse functions in spoken data for *and*, such as topic shift, topic resuming and opening boundary. Lastly, Schiffrin (1987: 146) argues that *and* can mark discourse topic at both local and global levels: the first by concatenating events of a same topic, as in [3]; the second by differentiating discourse topics on a global level, like in [4]. This second structure can also be conceived as sequences of actions such as different questions in question-answer pairs, in which case the new question is a specification of the first one or a new question.

[3] TOPIC 1 EVENT
and EVENT
and EVENT

A: Mmm, ¿molesta que alguna vez te traten de “tú” cuando tú estás tratando de “usted” o al revés? Es decir, ¿molesta...?

B: Bueno, si yo estoy tratando de “usted” y me tratan de “tú”...

A: Sí.

B: Mmm, si pongo algún reparo, ¿eh?

A: Y, ¿qué sientes? ¿Qué...? ¿Cuál es ese reparo que...?

B: Pues, que me parece que se toma un atrevimiento que no le corresponde.

A: Mmm, o sea, que hay que mantener las distancias, ¿no?

B: Eso es lo que yo creo.

A: Mmm... Does it bother you if someone addresses you as “tú” when you’re addressing them as “usted” or the other way around? I mean, does it bother you...?

B: Well, if I’m addressing someone as “usted” and they address me as “tú”...

A: Yes.

B: Mmm, I do take issue with that, you know?

A: AND what do you feel? What...? What is it that bothers you...?

B: Well, I feel like they’re taking a liberty that doesn’t belong to them.

A: Mmm, so you think it’s important to maintain that distance, right?

B: That’s what I believe.

(PRESEEA, MADR_H33_049)

[4] TOPIC 1
EVENT
EVENT...
and TOPIC 2
EVENT
EVENT...

B: Sí, y la verdad que esta gente está para sacarse el sombrero. O sea, uno justifica plenamente lo que uno ve como expresión de cariño hacia los policías, porque realmente están al servicio de la comunidad.

A: Wow, eso sí que es susto.

B: Eso fue susto.

A: Y, plancha vergüenza. ¿Cuál ha sido la plancha más grande que te haya tocado vivir en la vida?

B: Puf, ¿plancha? Sabes que no me recuerdo mucho de las planchas, no sé si es porque las bloqueo o no sé si principalmente porque, en general, soy una persona que tiene poco miedo al ridículo.

B: Yes, and honestly, these people deserve our utmost respect. I mean, you completely understand the expressions of affection toward the police because they are truly at the service of the community.

A: Wow, that was really scary.

B: That was scary.

A: **AND** embarrassing moments. What's the most embarrassing moment you've ever experienced?

B: Phew, embarrassment? You know, I don't really remember many embarrassing moments. I'm not sure if it's because I block them out or mainly because, in general, I'm a person who isn't afraid of being ridiculed.

(PRESEEA, SCHI_H23_085)

31 Next, to describe our second DM under study – lengthened *y* – we must briefly address prosody. Work on prosody and topic shift is vast and mostly focused on pitch and contour as cues for change (see Couper-Kuhlen, 2004). Nonetheless, within the field of DMs, interest in the prosody of these devices is slowly flourishing (for an overview, see Maschler & Schiffrin, 2015). Three studies focusing on prosodic features are particularly enriching to our work. First, and most importantly, Didirková et al. (2019), whose method we replicated in this paper, examined the impact of prosody on the interpretation of discourse relations expressed by *et* and *alors* via three perception experiments. Their findings suggest prosody helps speakers perceive and disambiguate discourse relations, although in the case of *et* prosody did not strongly facilitate the interpretation of topic shift specifically. Second, Matzen (2004) analysed the prosody of the DM *so* in relation to various discourse-marking functions. Findings show the lengthened version of the DM served as a distinctive factor only for the “closing-of-turn” category (Matzen, 2004: 84). Along with duration, pitch contour, position and sound were also considered in her analysis. Finally, Hansson (1999) investigated the differences between the prosodic profiles of sentential and discourse uses of Swedish *men* (“but”) and *så* (“so”). When used either to signal topic shift or return to previous topic, these DMs showed longer duration than sentential markers, which they argue relates to the need of planning time. It follows that we hypothesise that in Spanish interaction too, lengthened and non-lengthened *y* will fulfil different functions.

32 Finally, *bueno* – roughly translated as “well”, “okay”, “alright” in English – has often been associated with functions related to topic management. In her analysis of spoken data from Colombia, Travis (2005: 273) found this DM to be used by

speakers to signal reorientation in a conversation, amongst them introduction, return or closing of topic, and “other kinds of breaks in the flow of conversation”. In the same line, Candón Sánchez (2000: 152) described *bueno* as one of the most common markers used to “regulate the initiation, progression and closing of conversations”. In turn-initial position – she argues – it provides the speaker with time to plan the upcoming intervention. Finally, Maldonado and Palacios (2015) analysed spoken data from Mexican Spanish and concluded that *bueno* is moving away from its use as a device of acceptance of prior discourse to primarily function as a topic shifter and “window opener”.

- 33 Based on the literature reviewed above, we hypothesise that speakers will prefer
 a topic continuity interpretation in utterances introduced by *y*. In contrast, we
 anticipate that sentences introduced by both lengthened *y* and *bueno y* will be more
 likely to elicit the interpretation of discontinuity or new topic.

3. Methods

- 34 In this section we present the materials, participants, procedure, and data analysis
 of our task.

3.1. Materials

- 35 The stimuli in our task are based on prior corpus work from the same research
 project (Villalobos Cardozo et al., forthcoming a) on a sample from the “Proyecto
 para el estudio sociolingüístico del español de España y de América” (PRESEEA)
 (University of Alcalá, 2014).

- 36 PRESEEA is an open corpus of spoken Spanish developed by the University of
 Alcalá and compiled by several teams of researchers from various countries. It was
 initially designed for sociolinguistic objectives (see Moreno Fernández, 1996) and
 includes data from 1992 to 2019 from both Peninsular and Latin American Spanish.

- 37 This corpus consists of semi-structured interviews – each lasting at least
 45 minutes – based on thematic modules (e.g., weather, family and friends,
 traditions, anecdotes) and gathered through semi-elicitation in both formal and
 informal settings (Moreno Fernández, 2021a). As of December 2020, it contained
 approximately 396 hours of recorded material.

- 38 Orthographic transcriptions in the SGML4 system, following TEI standards
 (Moreno Fernández, 2021b), are readily available for researchers. These transcriptions,
 along with 10-minute samples in .mp3 format, can be downloaded from the site. All
 transcription files adhere to a standardised format and include metadata containing
 essential interview details such as date, location, transcriber and verifier information,
 as well as speaker demographics including age, education level, profession, and
 origin. Personal data within the material have been anonymised. Additionally,
 PRESEEA transcriptions are enhanced with phonetic and lexical labelling (see
 Moreno Fernández, 2021b), with lengthening being the most pertinent for our study.

39 In a previous corpus work in the same framework as this study (Villalobos Cardozo et al., forthcoming a), authors analysed a random sample of 400 instances from PRESEEA of two out of the three studied DMs: default *y* (“and”) and the lengthened *y* (“and”). Turn-initial occurrences produced by the interviewer were manually coded for topic continuity or discontinuity by two expert annotators. No specific instructions or criteria were provided beforehand. An interrater agreement (IRA henceforth) of 82.9%, with $\kappa = 0.631$ was achieved, indicating substantial agreement (Cohen, 1960), a reliable score for a discourse-level phenomenon (see Spooren & Degand, 2010). Given the high IRA, the extracted cases of *bueno y* were subsequently coded by a single annotator.

40 Stimuli for the task of the present study consisted of a total of 54 answer-question pairs, always in interviewee-interviewer order, selected from the PRESEEA material previously rated in the IRA by the expert annotators in the mentioned corpus work.

41 The second intervention in each of these sets – always a question – originally included a turn-initial occurrence of one of the three DMs under study. Based on this, we then designed a new set of alternate questions, followed by either topically continuous or discontinuous content.

42 To keep the utterances consistent in structure and number of words across the items, some modifications were made to the original material, which were also applied to the newly created continuations. Regional expressions were deleted, since speakers of both varieties (Peninsular and Latin American Spanish) would potentially participate. Filled pauses were accepted in the interviewee intervention, but not in the critical sentence uttered by the interviewer. The addressing form was also made homogeneous to *tú*. The average length of the continuations was 8.3 words. Table 1 shows an example of how the excerpts were modified.

43 The material to be included in the task was once again examined by two expert annotators in terms of topic. The utterances were then recorded in a soundproof booth by two trained native Spanish-speaking actors of Latin American descent. Recording and editing were performed using Pro Tools. In this task, the total duration for [y] was approximately 260 ms, lengthened [y:] was set to around 500 ms, and [bueno y] close to 670 ms. A 150 ms pause followed each DM. To avoid differences between stimuli, a single version of the three DMs was used in all recordings.

3.2. Participants

44 A total of 60 participants, native speakers of Spanish (80% from a Latin American variety) with an average age of 26.05 (SD = 6.12), took part in this study. From the sample, 37 individuals identified as female and 22 as male; one person did not provide an answer. Subjects were recruited via Prolific, a high-quality online data collection tool (Palan & Schitter, 2018; Peer et al., 2017), and received a £2 remuneration for participating, in alignment with Marie Skłodowska-Curie actions’ ethical payment conditions. To be eligible, participants had to meet the requirement of being native speakers and were limited to participating only once.

Status	Intervention	Translation
Original excerpt (LHAB_M33_103)	A: [...] también me gusta preparar cositas ricas. Siempre estoy buscando la manera de preparar cositas ricas y de darle a la comida la mejor <ruido = “I inhala” />, eh, belleza posible ehm, ehm. B: [Y] usted dice que ha hecho reformas en su casa. ¿Cómo era antes?	A: [...] I also like to prepare tasty little things. I’m always looking for ways to make delicious things and to give food the best <noise = “inhales” />, um, possible beauty, um, um. B: [And] you say you’ve made renovations to your house. What was it like before?
Modified interventions from the interviewee and interviewer	A: Siempre estoy buscando la manera de preparar cositas ricas y de darle a la comida la mayor belleza posible. B: [Y/Y:/Bueno y] ¿le has hecho reformas a tu casa recientemente?	A: I am always looking for ways to prepare tasty little things and make food look as good as possible. B: [Y/Y:/Bueno y] have you made any renovations to your house recently?
Modified intervention from the interviewee and constructed continuation from the interviewer (same topic)	A: Siempre estoy buscando la manera de preparar cositas ricas y de darle a la comida la mayor belleza posible. B: [Y/Y:/Bueno y] ¿sigues recetas o te inventas los platillos en el momento?	A: I am always looking for ways to prepare tasty little things and make food look as good as possible. B: [Y/Y:/Bueno y] do you follow recipes, or do you come up with the dishes on the spot?

Table 1 – Comparison of original PRESEEA excerpt vs final stimuli

3.3. Procedure

45 Our task was programmed under PsychoPy (v2021.2.3) (Peirce et al., 2019), an open-source software package designed to create and run behavioural science experiments. The task took approximately 15 minutes to carry out.

46 Introductory screens were shown to the participants, welcoming them to the experiment and collecting their informed consent. Afterwards, they were presented with the instructions for the task: they would listen to excerpts based on real interviews (consisting of a turn by the interviewee and then a question from the interviewer), which sometimes remained on the same topic and sometimes progressed onto a new topic. They were asked to categorise each case by making use of a five-point scale, with the options “same”, “similar”, “different” or “very different”. A non-labelled option appeared in the middle of the scale. Participants were asked to make use of it only in those cases in which it was impossible to provide a judgment.

Previous context (Interviewee)	Condition	Continuation (Interviewer)
I am always looking for ways to prepare tasty little things and make food look as good as possible.	[Y] same topic	[Y] do you follow recipes, or do you come up with the dishes on the spot?
	[Y:] same topic	[Y:] do you follow recipes, or do you come up with the dishes on the spot?
	[Bueno y] same topic	[Bueno y] do you follow recipes, or do you come up with the dishes on the spot?
	[Y] new topic	[Y] have you made any renovations to your house recently?
	[Y:] new topic	[Y:] have you made any renovations to your house recently?
	[Bueno y] new topic	[Bueno y] have you made any renovations to your house recently?

Table 2 – Example of stimuli

47 In each trial, images of conversation bubbles appeared as the recordings were automatically played with the intent of making it easier for the participants to follow the dynamic between the interviewee and the interviewer. After the interviewer's turn had finished playing, the question and the scale both appeared in the middle of the screen. Once participants had made their selection, the next trial began immediately.

48 A total of 324 trials were distributed in 6 counterbalanced lists of 54 items each. Each participant listened to 9 utterances per condition, which were presented in a randomised order. Linguistic judgment of the stimuli was collected via the five-point rating scale. Table 2 shows the different versions of the critical utterances.

3.4. Data analysis

49 The experiment logs were stored in a database on the Pavlovia server, from which they could be downloaded in Excel format. The data included the participant's choice on the scale, along with other details such as date, sound files played, participant ID (according to Prolific), operating system used, reaction times, and order of stimuli.

50 For data analysis and regression modelling, we transformed the Likert scale and treated responses as a binary variable. Same topic and similar topic were considered as "same topic", while slightly different and different topic were coded as "new topic". The "one cannot know" responses were removed to run the regression.

51 Two mixed-effects logistic regression models were run using the *glmer* function in R Statistical Software (R Core Team, 2024). The first model aimed to measure the effect of both the marker and the continuation on speakers' correct or incorrect identification of topic. For this, we used as reference the label given by the expert annotators (same topic or new topic) to each of the conversation excerpts. The second model measured how much the DM contributed to the participants' response (same or new topic), without considering their (dis)agreement with the expert annotators. This allowed us to explore whether they had specific preferences for topic (dis)continuity depending on the turn-initial DM that introduced the utterances.

4. Results

4.1. Quantitative analysis

52 From the total of 3,240 observations recorded, 192 (5.92%) consisted of a "One cannot know" answer (see Section 4.2.3). These trials were, therefore, removed from the analysis. A mixed-effects logistic regression model was fitted onto the data, including the interaction between DM and continuation and random effects by participant and by item. Results showed a significant effect of the continuation ($\beta = 117.7$, $SE = 1$, $p < 0.001$). The effect of the marker ($\beta = 0.99$, $SE = 2$, $p = 0.60$) and the interaction were not significant ($\beta = 1.58$, $SE = 2$, $p = 0.45$). Figure 2 presents the distribution of answers per continuation and DM. Responses are "correct" if they match the experts' topic annotation, and "incorrect" if they differ.

53 As shown in Figure 2, overall, scores are high, and participants tend to agree with the expert annotators when interpreting topic. Our task showed participants can recognise the topic relation between two utterances more easily when they relate to the same topic: 95.79% of correct answers versus 84.25% when the topic is different. However, the type of DM did not impact the interpretation, as 90% of the cases were correctly recognised regardless of the marker.

54 To explore participants' answers, independently of the experts' ratings, a second mixed-effects logistic regression model was fitted onto the data, including the interaction between DM and continuation and random effects by participant and by item. In this case, results show a significant effect of the continuation ($\beta = 790.1$, $SE = 1$, $p < 0.001$) but not of the marker ($\beta = 1.54$, $SE = 2$, $p = 0.46$). The interaction remained not significant ($\beta = 0.73$, $SE = 2$, $p = 0.69$). Figure 3 presents the distribution of answers (same and new topic) per DM.

55 As illustrated in Figure 3, the DM that introduced the utterances did not create preferences for topic (dis)continuity, as we had originally expected. Thus, our hypothesis that the DMs under study are predictive of (dis)continuity is not borne out. More precisely, it is not the case that DM [y] favours the interpretation of topic continuity, and it is not the case either that [y:] or [bueno y] would elicit topic discontinuity.

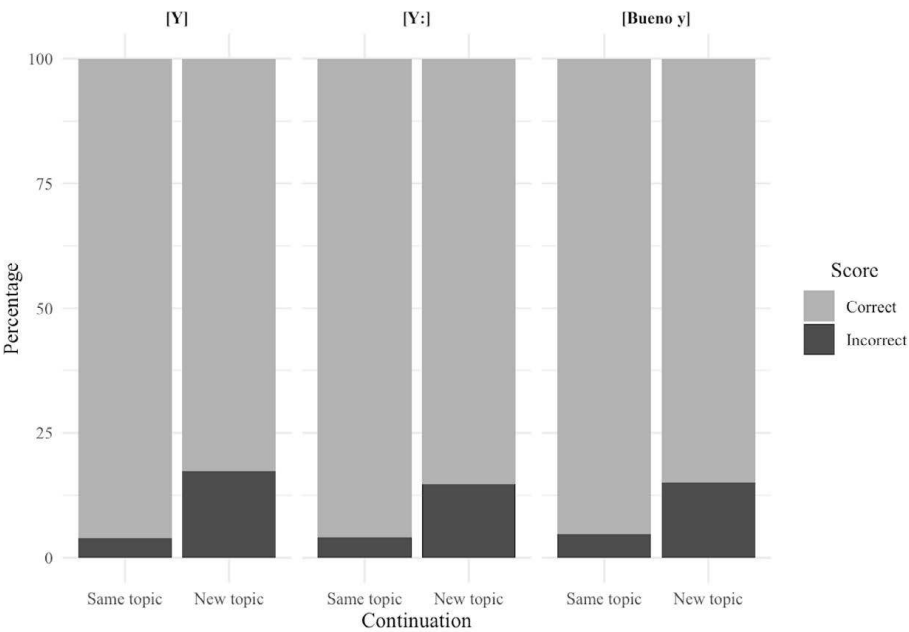


Figure 2 – Correct and incorrect answers per continuation and discourse marker

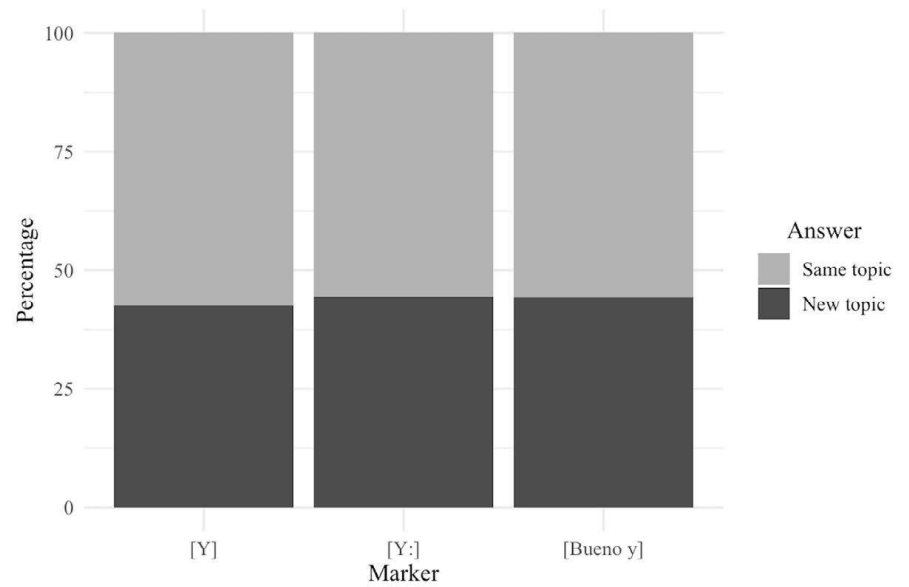


Figure 3 – Distribution of answers per discourse marker

4.2. Qualitative analysis: description of items and strategies for topic (dis)continuity

56 As demonstrated in the preceding section, speakers can reliably and consensually discern when adjacent utterances in a conversational setting stay on topic or change topic by using the content of the utterances rather than the DM that introduced them. For this reason, we now turn our focus to exploring the resources that contribute to topic continuity at the discourse level.

57 From the resources used by speakers in same-topic utterances and presented in the stimuli of our task, we identified strong and weak strategies of topic continuity. To establish a tiered system, we categorised them into three groups: “very successful” for those with a general agreement score of 85% or higher, “moderately successful” for scores between 70% and 85%, and “not successful” for those below 70%.

4.2.1. Strategies of topic continuity

58 Maximally continuous cases (85% of agreement and above) presented the most semantic information to the participant, like [5] and [6], via lexical repetition:

[5] B: Nosotros vivíamos aquí. Entonces... **Mis abuelos** vivían todos lejos; la única que venía de vez en cuando era mi **abuela**.

A: *Bueno* y, ¿era tu **abuela** materna o paterna?

B: We lived here. So... My grandparents all lived far away; the only one who came from time to time was my grandmother.

A: *Bueno* y, was it your maternal or paternal grandmother?

[6] B: Vivo aquí en el municipio Arroyo Naranjo. En la **Avenida de Los Apóstoles**, en Los Pinos.

A: *Bueno* y, ¿cómo es esa **avenida** en la que tú vives?

B: I live here in the Arroyo Naranjo municipality, on Avenida de Los Apóstoles, in Los Pinos.

A: *Bueno* y, what is that avenue like where you live?

59 In [5] the interviewee (speaker B) introduces *mis abuelos* (“my grandparents”) with a neutral-ordered definite noun phrase and then focuses on *mi abuela* (“my grandmother”). The interviewer (speaker A) directly refers to this referent, *tu abuela* (“your grandmother”), when asking for clarification. This lexical repetition across both turns reinforces discourse continuity and aligns with Givón’s idea that repetition of familiar entities strengthens continuity by reducing cognitive effort for the listener. Likewise, in the case [6] above, speaker B mentions specific locations and provides a stable introduction to the conversation’s main topic: her place of residence, “Avenida de Los Apóstoles”. By referencing *esa avenida*, speaker A maintains the conversation on the same discourse topic, referring to the “Avenida de Los Apóstoles” previously introduced by speaker B.

60 Another strategy used to maintain discourse continuity, following Givón's scale, is grammatical agreement or bound pronouns, frequent in Spanish as a pro-drop language. In our task, strategies like the ones presented in stimuli of the type [7] and [8] play a crucial role in ensuring a same-topic interpretation.

[7] B: **Mis papás** viven ahí y es imposible **convencerlos** para que vengan. Ya te imaginarás.
A: *Bueno y, ¿hace cuánto que están allá?*

B: My parents live there, and it's impossible to convince them to come here. You can imagine.

A: *Bueno y, how long have they been there?*

[8] B: Salgo sobre todo con **personas universitarias**. Igual... igual tengo mi grupo, pero me gusta más estar con la gente de la facultad porque tenemos horarios parecidos.

A: *Bueno y, ¿son todas de tu misma carrera?*

B: I mostly hang out with university people. I mean... I do have my own group, but I prefer being with people from the faculty because we have similar schedules.

A: *Bueno y, are they all from your same major?*

61 In [7] speaker B introduces *mis papás* ("my parents") using a neutral-ordered definite noun phrase to refer to a specific entity (her parents), which will be the base of the discourse topic. As shown in *es imposible convencerlos* ("it's impossible to convince them"), the bound pronoun *los* ("them") refers to "my parents". This follows Givón's scale, where bound pronouns are used for highly accessible topics. Lastly, speaker A's question of *¿hace cuánto que están allá?* ("how long have they been there?") reinforces this link, by using *están* – in third-person plural – to request more details about her parents. Likewise, in [8], topicality regarding speaker B's social interactions with peers from the university is maintained thanks to speaker A, who builds upon *personas universitarias* and *gente de la facultad*, by referring to the same group with *son*, ensuring that the referent remains clear.

62 Discourse topic continuity is maintained also via the use of words that belong to the same lexical field, with most items achieving an agreement score of 85% or higher in our task. Since semantically related words are highly predictable in discourse, they are easier for speakers to process. For example, in [9], *preparar cositas ricas* ("prepare/cook tasty little things"), *comida* ("food"), *recetas* ("recipes") and *platillos* ("dishes"), all contribute to discourse continuity.

[9] B: Siempre estoy buscando la manera de **preparar cositas ricas** y de darle a la **comida** la mejor, eh, belleza posible.

A: *Bueno y, ¿sigues recetas o te inventas los platillos en el momento?*

B: I am always looking for ways to prepare tasty little things and make food look as good as possible.

A: *Bueno y, do you follow recipes, or do you make up the dishes on the spot?*

63 There are, however, exceptions to the success of this last strategy of the lexical field. Out of the complete set of stimuli, two cases were rated as moderately continuous (70% to 85%) by participants, despite clear lexical or semantic connections.

[10] B: Pues, hasta ahora, con estos **cambios climáticos**, se han venido viendo muchos **incendios forestales** y han afectado a muchas personas, muchos animales.

A: *Bueno* y, ¿también se han dado **inundaciones** cuando llueve?

B: Well, so far, with these climate changes, there have been many wildfires, and they have affected a lot of people and many animals.

A: *Bueno* y, have there also been floods when it rains?

64 For instance, in example [10], experts initially considered the snippet to be a case of same topic, viewing wildfires and floods as part of a broad environmental discourse topic. This perspective aligns with Givón's notion that topics can expand in scope without losing continuity. However, because the link between climate change and specific events like wildfires and floods may not be immediately clear, participants might have interpreted the interviewer's question about floods as introducing a different, though related, subtopic, leading to moderate disruption in continuity.

4.2.2. Strategies of topic discontinuity

65 Results from the task indicate that the highest agreement scores for topic discontinuity (85% and above) use disjunctive topic shift. In these cases, the introduction of a new referent into the conversation, in an abrupt manner and without any link to the previous discourse, makes the new topic easier to identify, as in [11] and [12] below.

[11] B: **Mis papás** viven ahí y es imposible convencerlos para que vengan. Ya te imaginarás.

A: Y, ¿a **qué te dedicas** actualmente?

B: My parents live there, and it's impossible to convince them to come here. You can imagine.

A: Y, what do you do for a living?

[12] B: ¡Sí, sí, sí! **Eso** desapareció y hace poco pusieron un restaurante igual al final del boulevard, pero no es como el otro.

A: Y, ¿**algún recuerdo favorito de tu niñez**?

B: Yes, yes, yes! That disappeared, and recently they opened a similar restaurant at the end of the boulevard, but it's not like the other one.

A: Y, do you have any favourite childhood memory?

66 In [11] above (the new-topic version of the conversational snippet presented in [7]), speaker B talks about her parents and their reluctance to move. Immediately, with the question *¿a qué te dedicas?* ("what do you do for a living?"), the interviewer introduces a new, indefinite topic related to the interviewee's

profession, which has not been part of the conversation up to this point. In this sense, this question signals the start of a completely unrelated discourse topic, leading to a sense of high discontinuity. Likewise, in [12], speaker B talks about the disappearance of an old restaurant known to both participants in the conversation, and the establishment of a new restaurant at the end of the boulevard. Discontinuity in discourse topic occurs with the abrupt topic shift carried out by the interviewer, who inquires about the interviewee's childhood memories by asking *¿algún recuerdo favorito de tu niñez?* ("any favourite childhood memory?"). Here, speaker A uses an indefinite noun phrase, which typically introduces new topics or less accessible information.

67 In contrast to those cases that presented disjunctive topic shifts, items with stepwise topic shifts reached lower agreement scores. Interactional accounts suggest that gradual topic shifting is a frequent and preferred pattern in conversation. Although this adds to the fluidity of interaction, it does not necessarily aid in identifying topic boundaries. For instance, while experts coded case [13] as a new topic, participants were more sensitive to the similarities between topics.

[13] B: Antes, pues, terminábamos del cole y hacíamos los deberes. Nos íbamos a la calle todas las amigas; allí jugábamos.
A: Y, ¿cómo era tu barrio?

B: Before, we would finish school and do our homework. We would all go out to the street and play.
A: Y, what was your neighbourhood like?

68 In the example [13] above, the interviewee is talking about her childhood activities with her friends: *terminábamos del cole y hacíamos los deberes* ("we would finish school and do our homework") and *nos íbamos a la calle todas las amigas* ("we would all go to the street and play"). In this sense, the discourse topic proposed revolves around the daily routines of childhood and playtime with friends. Participants in our task may have perceived this snippet as the same topic because "neighbourhood" and "street" are lexically close. For that reason, *¿cómo era tu barrio?* ("what was your neighbourhood like?") might have been considered a natural and fluid progression of the conversation.

69 Finally, [14] showcases an interesting case where participants interpreted the exchange as a new topic, despite the lexical repetition and similarity in structure that led experts to code it as part of the same conversational project.

[14] B: Aunque sea más joven que yo o de mi misma edad, a **mi médico** lo **trato** de usted; una profesión que considero mucho mucho.
A: Y, ¿a tus **profesores** cómo los **tratas**?

B: Even if they are younger than me or the same age, I address my doctor as "usted"; it's a profession I hold in very high regard.
A: Y, how do you address your teachers?

70 In example [14] the underlying structure of the conversation remains the same, and even the verb *tratar* is repeated in both turns, which was expected to contribute to the continuity. However, it seems the change in professions (*médico* to *profesores*) might have made a difference to participants, who saw more discontinuity in the discourse topic than experts did.

4.2.3. “One cannot know” answers

71 To address potential reliability concerns and improve data quality, the Likert scale in the task included the option “One cannot know,” resulting in 5.92% answers. Participants were instructed to use this option strictly when they found it difficult to determine whether the utterance maintained the same topic or introduced a new one. Figure 4 focuses on the subset of these responses, showing their distribution across the DMs, compared to the judgment of expert annotators.

72 As depicted above, uncertainty was low: 192 of the total responses. Participants were generally able to provide clear interpretations for the tested stimuli. A closer analysis reveals that the use of the “one cannot know” option was more frequent in utterances classified by expert annotators as new topics. Examples [15] and [16] are the cases with most “one cannot know” responses from all the stimuli, for new topic and same topic instances, respectively.

[15] B: Yo trabajo por mi cuenta. Antes tenía una tienda de productos naturales, pero la vendí hace mucho tiempo ya.

A: *Bueno* y, ¿has disfrutado de los festejos?

B: I am self-employed. I used to own a natural products store, but I sold it a long time ago.

A: *Bueno* y, have you enjoyed the celebrations?

[16] B: Siempre he vivido en este pueblo. Lo más que he hecho es viajar un par de meses, pero vivir – lo que es vivir – solo aquí.

A: Y, ¿siempre ha sido igual desde que eras pequeña?

B: I’ve always lived in this town. The most I’ve done is travel for a couple of months, but as for living – truly living – only here.

A: Y, has it always been the same since you were little?

73 Example [15] presents the case with the highest number of “one cannot know” responses for the new-topic condition, with a total of 6 such answers. The remaining 24 responses categorised the conversational snippet as a new topic, with 7 participants selecting “different” and 17 choosing “very different”. Interestingly, this case presents cues of topic discontinuity, as it lacks an explicit topic continuity connection between speaker B’s work history and speaker A’s question about festivities.

74 Item [16] illustrates a same-topic instance with more “one cannot know” answers. Here, responses were more varied: 4 participants selected “one cannot know”, while 23 classified it as the same topic (17 as “same” and 6 as “similar”), and

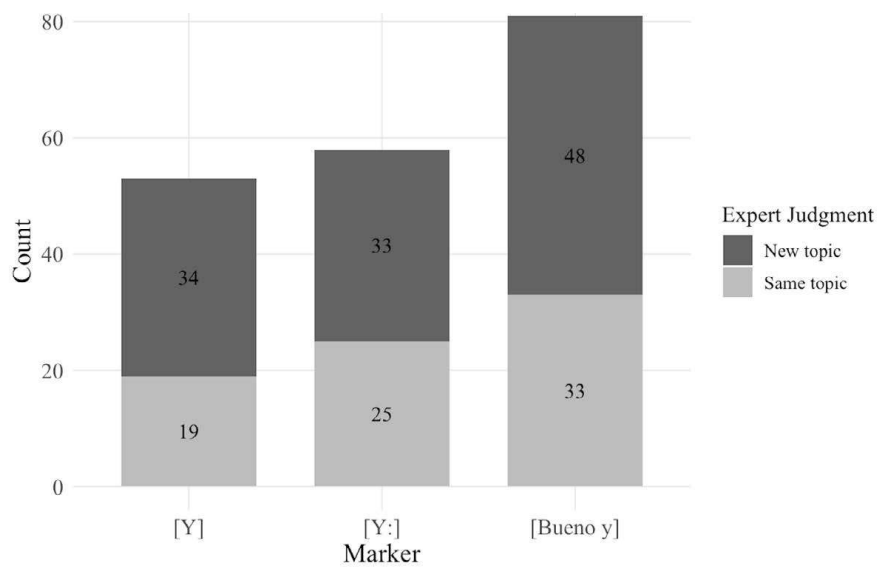


Figure 4 – Distribution of “One cannot know” responses per marker compared to expert annotation

only 3 considered it a new topic (2 as “different” and 1 as “very different”). In this case, the interviewee reflects on her lifelong stay in her hometown, except for some brief travelling. In response, the interviewer asks: *¿Siempre ha sido igual desde que eras pequeña?* (“Has it always been the same since you were little?”). Here, the use of the “one cannot know” answer seems more justifiable, since the verb lacks a clear referent, which could make participants doubt whether *ha sido* refers to the town or something else.

75 In sum, examining the “one cannot know” responses reveals that certain conversational exchanges were more challenging for participants to categorise due to ambiguity, subtle shifts, and lack of explicit connections. This uncertainty was predominantly observed in new-topic cases, supporting the notion that topic shifts increase uncertainty, whereas related content in same-topic utterances is generally easier to process, as it is more expected.

5. Discussion

76 Our findings suggest that the presence of either [y], [y:] or [bueno y] introducing an utterance does not significantly influence speakers’ interpretation of the aboutness of upcoming discourse nor their preference for the interpretation of the utterance as either a continuation of the same topic – in the case of [y] – or a shift to a different topic, as expected with [y:] and [bueno y]. These observations challenge

theoretical claims and most functional descriptions of DMs, which often attribute a central role in topic management to these devices.

77 Instead, our results are more consistent with recent research arguing for a more nuanced understanding of how discourse relations are signalled. For instance, Das and Taboada (2018) in their study on coherence relations, found over 80% of them were indicated by a wide variety of signals beyond DMs, such as referential, lexical, semantic and syntactic features, while only around 10% were exclusively signalled by DMs. Likewise, the study of Péry-Woodley et al. (2017), with a focus on enumerative structures, shows that the configurations of cues signalling these discourse patterns can occur at several levels of discourse organisation and are not limited to a specific linguistic form.

78 Considering this, we propose that the lack of significant effect observed for the DMs under study may be due to the broader range of cues presented in our stimuli, making the DMs less informative in guiding interpretation. These other linguistic signals, such as lexical repetition and anaphora, may have overshadowed the role of DMs as signals. This idea is consistent with ongoing research which indicates that when these same DMs are presented on their own in an auditory forced-choice task, participants tend to predict the upcoming utterance to be topically continuous when listening to [y] and discontinuous when hearing [y:] and [bueno y] (Villalobos Cardozo et al., forthcoming b). It is therefore reasonable to speculate that the extent to which DMs are recruited to make sense of discourse relations may depend on the presence or absence of accompanying cues.

79 Given these observations, we argue that the evidence remains inconclusive regarding whether the studied DMs entirely lack a signalling function for topic continuity or discontinuity in interaction. Instead, we propose that our findings may be influenced by methodological constraints inherent in the experimental design. The offline nature of the task, where participants have access to entire conversational snippets before making their decisions, is fundamentally different from natural conversation, where DMs and their timing may play a more significant role in signalling upcoming content and managing topic transitions. Furthermore, although functional descriptions of these markers often place them on a continuum from underspecification to specialisation, it is crucial to note that these combinations do occur naturally, and thus the violation of expectations may not be as strong as it would be with more conceptually explicit meaning.

80 Another explanation for our results could be the distinction between speaker-oriented and hearer-oriented functions of DMs, or what Broisson and Degand (2022) describe as egocentric vs. allocentric uses. This distinction proposes that DMs may serve not only as communicative signals for the listener but also as traces of the speaker's ongoing cognitive processing. This idea that DMs reflect the speaker's cognitive processing could suggest that the DMs in our study may not be significantly involved in topic interpretation. In this sense, the DMs

might help the speaker in the planning and production of a topic shift (egocentric function), but they may not necessarily be perceived and used as such by the listener (allocentric function).

81 Finally, our findings challenge the assumption of procedural rigidity often associated with DMs. The results of our task suggest that these DMs do not rigidly constrain context interpretation but instead interact with it. The procedural information proposed to be ingrained in these DMs appears not to dictate how upcoming content is understood, but rather it contributes to the interpretive process in a more flexible way.

6. Conclusion

82 In this study, we set out to test whether DMs are reliable signals for topic shifts, as suggested by many theoretical accounts. We designed a rating task where participants were asked to decide if two utterances presented the same topic or different topics. In this task we compared a supposedly strong topic shifter [bueno y], with a more continuous version [y], and [y:], based on a previous corpus study that showed an association between the choice of the markers and the topic (dis)continuity in interviews in Spanish.

83 Despite the theoretical challenges in defining discourse topic and establishing clear criteria to distinguish topic boundaries within a conversation, our task showed that speakers can intuitively navigate these distinctions reliably and without difficulty. However, results showed no effect of the DM in the process of topic interpretation. It seems that participants can discern between new-topic utterances and same-topic utterances independently from the DM that introduces them. We also did not find specific associations between the DMs under study and topic (dis)continuity. Therefore, our prediction regarding DM as a possible signal for topic shift and topic continuity was not confirmed.

84 Following these results, we conducted a qualitative study of the items, to explore the resources and strategies that could explain why the instruction of the marker did not prevail. We showed topic continuity was mostly expressed via lexical repetition, subject-verb agreement, and lexical field, although the latter is not always sufficient for the interpretation. On the other hand, topic shift is easier to identify when the second utterance introduces a new and distant referent, as opposed to more stepwise changes between somewhat related topics.

85 We can conclude that although DMs can potentially contribute to signalling topic boundaries, sentence-internal strategies seem to have more impact on the interpretation. This suggests that experiments and corpora sometimes diverge and that production patterns are not always reflected in comprehension tasks. Future research using diverse methodologies, including online and offline tasks, would significantly advance our understanding of the interplay between DMs and topic management in discourse.

86 Moreover, our study showed that topic continuity is easier to identify than topic shift, aligning with the notion that discourse is processed as continuous by default (Sacks, 1995; Murray, 1997). As mentioned by Givón (1983: 12),

[w]hat is continuing is more predictable; what is predictable is easier to process or conversely, what is discontinuous or disruptive is less predictable and what is less predictable hence surprising is harder to process.

87 Our findings support this idea, as the interpretation of adjacent utterances as topically discontinuous seems to require a strong signal to overrule this prediction of continuity. It appears that the DMs in our study do not qualify as such strong signals.

References

- ASHER, N. 2004. Discourse Topic. *Theoretical Linguistics* 30 (2-3): 163-201.
- ASHER, N. & LASCARIDES, A. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.
- ATKINSON, J.M. & HERITAGE, J. (eds.) 1985. *Structures of Social Action*. Cambridge: Cambridge University Press.
- BLAKEMORE, D. 1987. *Semantic Constraints on Relevance*. Oxford – New York: Blackwell.
- BOLLY, C.T., CRIBLE, L., DEGAND, L. & UYGUR-DISTEXHE, D. 2017. Towards a Model for Discourse Marker Annotation: From Potential to Feature-Based Discourse Markers. In C. FEDRIANI & A. SANSÒ (eds.), *Pragmatic Markers, Discourse Markers and Modal Particles: New Perspectives*. Amsterdam – Philadelphia: J. Benjamins: 71-98.
- BRIZ GÓMEZ, A., PONS BORDERÍA, S. & PORTOLÉS LÁZARO, J. (eds.) 2008. *Diccionario de partículas discursivas del español*. Madrid: Ministerio de Ciencia y Tecnología. Available online: <http://www.dpde.es/>.
- BROISSON, Z. & DEGAND, L. 2022. How Egocentric Is Discourse Marker Use? Investigating the Impact of Speaker Orientation and Cognitive Load on Discourse Marker Production. In M.-J. CUENCA & L. DEGAND (eds.), *Discourse Markers in Interaction: From Production to Comprehension*. Berlin – Boston: De Gruyter: 121-158.
- BROWN, G. & YULE, G. 1983. *Discourse Analysis*. Cambridge – London: Cambridge University Press [1st ed.].
- CANDÓN SÁNCHEZ, M.T. 2000. Los conectores pragmáticos en la conversación coloquial. Aplicación a la enseñanza del español como segunda lengua. In M. FRANCO FIGUEROA, C. SOLER CANTOS, F.J. de COS RUIZ, M. RIVAS ZANCARRÓN & F. RUIZ FERNÁNDEZ (eds.), *Nuevas perspectivas en la enseñanza del español como lengua extranjera: actas del X Congreso Internacional de ASELE (Cádiz, 22-25 de septiembre de 1999)*. Cádiz: Universidad de Cádiz. Vol. 1: 149-156.
- CHAROLLES, M. 2020. Discourse Topics and Digressive Markers. *Journal of Pragmatics* 161: 57-77.
- COHEN, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20 (1): 37-46.

- COUPER-KUHLEN, E. 2004. Prosody and Sequence Organization in English Conversation: The Case of New Beginnings. In E. COUPER-KUHLEN & C.E. FORD (eds.), *Sound Patterns in Interaction: Cross-Linguistic Studies from Conversation*. Amsterdam – Philadelphia: J. Benjamins: 335-376.
- COUPER-KUHLEN, E. & SELTING, M. 2017. *Interactional Linguistics: Studying Language in Social Interaction*. Cambridge – New York – Melbourne: Cambridge University Press.
- CRIBLE, L. 2017. Towards an Operational Category of Discourse Markers: A Definition and Its Model. In C. FEDRIANI & A. SANSÒ (eds.), *Pragmatic Markers, Discourse Markers and Modal Particles: New Perspectives*. Amsterdam – Philadelphia: J. Benjamins: 99-124.
- CRIBLE, L. & CUENCA, M.-J. 2017. Discourse Markers in Speech: Distinctive Features and Corpus Annotation. *Dialogue and Discourse* 8 (2): 149-166.
- DAS, D. & EGG, M. 2023. Continuity in Discourse Relations. *Functions of Language* 30 (1): 41-66.
- DAS, D. & TABOADA, M. 2018. Signalling of Coherence Relations in Discourse, beyond Discourse Markers. *Discourse Processes* 55 (8): 743-770.
- DIDIRKOVÁ, I., CRIBLE, L. & SIMON, A.C. 2019. Impact of Prosody on the Perception and Interpretation of Discourse Relations: Studies on “Et” and “Alors” in Spoken French. *Discourse Processes* 56 (8): 619-642.
- ESCANDELL VIDAL, M.V. & LEONETTI JUNGL, M. 2004. Semántica conceptual/semántica procedimental. In M. VILLAYANDRE LLAMAZARES (ed.), *Actas del V Congreso de Lingüística General (León, 5-8 de marzo de 2002)*. Madrid: Arco Libros. Vol. 2: 1727-1738.
- ESCANDELL VIDAL, M.V., LEONETTI JUNGL, M. & AHERN, A. (eds.) 2011. *Procedural Meaning: Problems and Perspectives*. Bingley: Emerald.
- ESTELLÉS, M. & PONS BORDERÍA, S. 2014. Absolute Initial Position. In S. PONS BORDERÍA (ed.), *Discourse Segmentation in Romance Languages*. Amsterdam – Philadelphia: J. Benjamins: 121-155.
- FAGARD, B. & CHAROLLES, M. 2020. Topic Shifters in a Contrastive Perspective. *Journal of Pragmatics* 165: 45-48.
- FISCHER, K. (ed.) 2006. *Approaches to Discourse Particles*. Amsterdam – London – Paris: Elsevier.
- FRASER, B. 1990. An Approach to Discourse Markers. *Journal of Pragmatics* 14 (3): 383-398.
- FRASER, B. 1999. What Are Discourse Markers? *Journal of Pragmatics* 31 (7): 931-952.
- FRASER, B. 2009. Topic Orientation Markers. *Journal of Pragmatics* 41 (5): 892-898.
- FUENTES RODRÍGUEZ, C. 2009. *Diccionario de conectores y operadores del español*. Madrid: Arco Libros.
- GARCÍA GARCÍA, M. 2020. Turn-Initial Discourse Markers in L2 Spanish Conversations: Insights from Conversation Analysis. *Corpus Pragmatics* 5 (1): 37-61.
- GIVÓN, T. 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam – Philadelphia: J. Benjamins. Topic continuity in discourse: An introduction: 1-42.
- GIVÓN, T. 2017. *The Story of Zero*. Amsterdam – Philadelphia: J. Benjamins.
- GÓMEZ TXURRUKA, I. 2003. El significado de la conjunción “y”. *Teorema: Revista internacional de filosofía* 22 (1-2): 55-84.

- GOUTSOS, D. 1997. *Modeling Discourse Topic: Sequential Relations and Strategies in Expository Text*. Norwood: Ablex.
- HANSSON, P. 1999. Prosodic and Lexical Correlates of Swedish Discourse Markers in Spontaneous Dialogue. In J.J. OHALA, Y. HASEGAWA, M. OHALA, D. GRANVILLE & A.C. BAILEY (eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences (ICPhS 99)*. Berkeley: University of California: 1533-1536. Available online: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14_1533.pdf.
- HASELOW, A. 2017. *Spontaneous Spoken English: An Integrated Approach to the Emergent Grammar of Speech*. Cambridge: Cambridge University Press.
- IONESCU, A. 2020. Topic Shifters in Romanian: A Contrastive Analysis. *Journal of Pragmatics* 156: 110-120.
- KELLERMANN, K. & PALOMARES, N.A. 2004. Topical Profiling: Emergent, Co-Occurring, and Relationally Defining Topics in Talk. *Journal of Language and Social Psychology* 23 (3): 308-337.
- LAMBRECHT, K. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge – New York – Melbourne: Cambridge University Press.
- MALDONADO, R. & PALACIOS, P. 2015. “Bueno”, a Window Opener. In J. DAEMS, E. ZENNER, K. HEYLEN, D. SPEELMAN & H. CUYCKENS (eds.), *Change of Paradigms – New Paradoxes: Recontextualizing Language and Linguistics*. Berlin – Munich – Boston: De Gruyter: 97-108.
- MASCHLER, Y. 1997. Discourse Markers at Frame Shifts in Israeli Hebrew Talk-in-Interaction. *Pragmatics* 7 (2): 183-211.
- MASCHLER, Y. 2009. *Metalanguage in Interaction: Hebrew Discourse Markers*. Amsterdam – Philadelphia: J. Benjamins.
- MASCHLER, Y. & SCHIFFRIN, D. 2015. Discourse Markers: Language, Meaning, and Context. In D. TANNEN, H.E. HAMILTON & D. SCHIFFRIN (eds.), *The Handbook of Discourse Analysis*. Chichester: Wiley: 189-221 [2nd ed.].
- MATZEN, L.E. 2004. Discourse Markers and Prosody: A Case Study of “So”. *LACUS Forum* 30: 73-94.
- MORENO FERNÁNDEZ, F. 1996. Metodología del “Proyecto para el estudio sociolingüístico del español de España y de América (PRESEEA)”. *Lingüística* 8: 257-287.
- MORENO FERNÁNDEZ, F. 2021a. *Metodología del “Proyecto para el estudio sociolingüístico del español de España y de América (PRESEEA)”*. Alcalá de Henares: Editorial Universidad de Alcalá: 1-31. Available online: [https://presea.uah.es/sites/default/files/2022-02/Metodología del Proyecto para el estudio sociolingüístico del español de España y de América \(PRESEEA\)_Moreno Fernández \(2021\).pdf](https://presea.uah.es/sites/default/files/2022-02/Metodología del Proyecto para el estudio sociolingüístico del español de España y de América (PRESEEA)_Moreno Fernández (2021).pdf).
- MORENO FERNÁNDEZ, F. 2021b. *Marcas y etiquetas mínimas obligatorias para materiales de PRESEEA*. Alcalá de Henares: Editorial Universidad de Alcalá: 1-29. Available online: [https://presea.uah.es/sites/default/files/2022-02/Marcas y etiquetas mínimas obligatorias para materiales de PRESEEA_Moreno Fernández \(2021\).pdf](https://presea.uah.es/sites/default/files/2022-02/Marcas y etiquetas mínimas obligatorias para materiales de PRESEEA_Moreno Fernández (2021).pdf).
- MURILLO ORNAT, S. 2010. Los marcadores del discurso y su semántica. In Ó. LOUREDA LAMAS & E. ACÍN VILLA (eds.), *Los estudios sobre marcadores del discurso en español, hoy*. Madrid: Arco Libros: 241-280.

- MURRAY, J.D. 1997. Connectives and Narrative Text: The Role of Continuity. *Memory and Cognition* 25 (2): 227-236.
- PALAN, S. & SCHITTER, C. 2018. Prolific.ac – A Subject Pool for Online Experiments. *Journal of Behavioral and Experimental Finance* 17: 22-27.
- PALOMARES, N.A., BRADAC, J.J. & KELLERMANN, K. 2006. Conversational Topic along a Continuum of Perspectives: Conceptual Issues. *Annals of the International Communication Association* 30 (1): 45-97.
- PEER, E., BRANDIMARTE, L., SAMAT, S. & ACQUISTI, A. 2017. Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research. *Journal of Experimental Social Psychology* 70: 153-163.
- PEIRCE, J., GRAY, J.R., SIMPSON, S., MACASKILL, M., HÖCHENBERGER, R., SOGO, H., KASTMAN, E. & LINDELØV, J.K. 2019. PsychoPy2: Experiments in Behavior Made Easy. *Behavior Research Methods* 51: 195-203.
- PÉRY-WOODLEY, M.-P., HO-DAC, L.-M., REBEYROLLE, J., TANGUY, L. & FABRE, C. 2017. A Corpus-Driven Approach to Discourse Organisation: From Cues to Complex Markers. *Dialogue and Discourse* 8 (1): 66-105.
- PONS BORDERÍA, S. & ESTELLÉS, M. 2009. Expressing Digression Linguistically: Do Digressive Markers Exist? *Journal of Pragmatics* 41 (5): 921-936.
- PORTOLÉS LÁZARO, J. 2008. Las definiciones de las partículas discursivas en el diccionario. In M.P. GARCÉS GÓMEZ (ed.), *Diccionario histórico: nuevas perspectivas lingüísticas*. Madrid – Frankfurt: Iberoamericana – Vervuert: 179-202.
- PRESEEA 2014. *Corpus del Proyecto para el estudio sociolingüístico del español de España y de América*. Alcalá de Henares: Universidad de Alcalá. Available online: <http://preseca.uah.es/>.
- R CORE TEAM 2024. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: <https://www.R-project.org/>.
- RIOU, M. 2015. A Methodology for the Identification of Topic Transitions in Interaction. *Discours* 16: 1-28. Available online: <https://journals.openedition.org/discours/8997>.
- SACKS, H. 1995. *Lectures on Conversation*. G. JEFFERSON (ed.). Oxford – Cambridge: Blackwell.
- SCHEGLOFF, E.A. 1990. On the Organization of Sequences as a Source of “Coherence” in Talk-in-Interaction. In B. DORVAL (ed.), *Conversational Organization and Its Development*. Norwood: Ablex: 51-77.
- SCHIFFRIN, D. 1985. Conversational Coherence: The Role of “Well”. *Language* 61 (3): 640-667.
- SCHIFFRIN, D. 1987. *Discourse Markers*. Cambridge – New York: Cambridge University Press [1st ed.].
- SPOOREN, W. & DEGAND, L. 2010. Coding Coherence Relations: Reliability and Validity. *Corpus Linguistics and Linguistic Theory* 6 (2): 241-266.
- TABOADA, M. 2006. Discourse Markers as Signals (or Not) of Rhetorical Relations. *Journal of Pragmatics* 38 (4): 567-592.
- TRAUGOTT, E. 2020. The Development of “Digressive” Discourse-Topic Shift Markers in English. *Journal of Pragmatics* 156: 121-135.

- TRAVIS, C.E. 2005. *Discourse Markers in Colombian Spanish: A Study in Polysemy*. Berlin – New York: De Gruyter.
- TURK, M.J. 2004. Using “And” in Conversational Interaction. *Research on Language and Social Interaction* 37 (2): 219-250.
- VAN DIJK, T.A. 1981. *Studies in the Pragmatics of Discourse*. The Hague – Paris – Amsterdam: De Gruyter.
- VILLALOBOS CARDOZO, M., CRIBLE, L. & DEGAND, L. forthcoming a. Analysing Production Patterns of Turn-Initial Discourse Markers in Spanish: A Corpus Study on Topic Management [journal article in preparation].
- VILLALOBOS CARDOZO, M., CRIBLE, L. & DEGAND, L. forthcoming b. Discourse Marker Preferences at Turn Beginnings: Predicting Topic (Dis)continuity in Spanish [journal article in preparation].
- ZELLERS, M. 2013. Prosodic Variation for Topic Shift and Other Functions in Local Contrasts in Conversation. *Phonetica* 69 (4): 231-253.
- ZUFFEREY, S. 2020. *Introduction to Corpus Linguistics*. London – Hoboken: ISTE – Wiley [1st ed.].
- ZUFFEREY, S. & DEGAND, L. 2024. *Connectives and Discourse Relations*. Cambridge: Cambridge University Press.