

Zoé Broisson and Liesbeth Degand

6 How egocentric is discourse marker use? Investigating the impact of speaker orientation and cognitive load on discourse marker production

Abstract: Language has often been described as produced ‘for the hearer’. A case in point is the study of discourse markers (DMs), often presented as linguistic expressions that ‘provide instructions to the hearer’ (Hansen, 2006: 25). But what about their role for the speaker? Psycholinguistic research showing that speakers demonstrate egocentric behavior when their cognitive load increases (e.g. Keysar et al. 1998, Roßnagel 2000, Vogels et al. 2015) suggests that the lack of consideration for the role played by constraints on the speaker leaves us with an incomplete account of language production. To assess whether speakers use DMs with a communicative intention, i.e. purposefully and intentionally, this contribution reports on two experiments investigating the impact of speaker orientation (allocentric vs. egocentric) and cognitive load on the use of DMs. Results show that speakers use DMs both in the allocentric and egocentric conditions, but with systematic differences in terms of frequency, diversity and DM types. Cognitive load, however, does not seem to have any of the hypothesized effects on DM use.

Keywords: Discourse markers, cognitive load, speaker orientation, speech elicitation, spoken interaction, French

1 Introduction

Language is an inherent part of human communication. Despite its ubiquity in our everyday lives, using language for comprehension or production remains a demanding task from a cognitive and motor perspective (Levelt 1989: 1). As Tanenhaus (2013: 1) explains, understanding is difficult because it involves making inferences about a speaker’s intentions on the basis of (sometimes vague) acoustic cues, and this under a limited time constraint. But equally taxing is the act of speaking, as it represents the engagement of our motor commands to allow the

Zoé Broisson, Université catholique de Louvain

Liesbeth Degand, Université catholique de Louvain, liesbeth.degand@uclouvain.be

<https://doi.org/10.1515/9783110790351-006>

translation of a complex conceptual piece of information into a series of sounds or hand gestures (MacDonald 2013: 2). From either perspective, it seems that attending to the competing demands of both the speaker and the hearer is crucial to achieve successful communication in any interaction.

While dividing the efforts involved in language use evenly between speaker and hearer would be optimal, it has been suggested that one of the two may well have the upper hand in the cognitive processing-driven organization of grammar. But whose processing considerations shape the way that we use language to express ourselves (see e.g. Pijpops et al. 2018)? Those of our interlocutors, or our own? Many theorists and researchers have attempted to answer this question, which continues to resonate in recent research (e.g. Debska and Raczaszek-Leonardi 2018, Galati et al. 2018, Hawkins 2004, Horton and Gerrig 2016, Jaeger and Tily 2011, Yoon and Brown-Schmidt 2018). Nevertheless, the scientific community shows little convergence on this issue. Classical theories of language assume that spoken conversation is only successful when it is collaborative (Grice 1975). For this reason, most scholars believe that it makes more sense to think that language is produced directly for the hearer: if speakers put effort into producing language that is tailored to their addressee, their own intentions will be understood better and more quickly, which in turn will reduce the overall effort required in communication (e.g. Achim, Achim and Fossard 2017, Yoon and Brown-Schmidt 2017, Brennan and Hanna 2009, Clark 1996). Next to the cooperation principle (Grice 1975), several theories and models of language production embrace this hearer-oriented or allocentric view, among them the audience design model (Blokpoel et al. 2012); common ground theory (Clark and Carlson 1981, 1982) or the perspective-taking model (Galinsky et al. 2008). Following Levelt's (1989) model for the speaker, these theories place audience design processes during the early stages of language production. However, certain researchers have challenged this assumption that audience design would be a systematic step in the planning process. In Barr and Keysar's (2006: 902) words, "the question is no longer whether speakers design their speech for their listeners, but how and when they do so." More precisely, the hearer-oriented speech planning would require far too many cognitive resources to be used efficiently in every day speech (e.g. Bard et al. 2000, Brown and Dell 1987, Horton and Keysar 1996). Instead, it has been suggested that speakers may well produce language as if it was intended for themselves, at first, and then, only if they have enough time or memory resources or strong enough cues (Vogels et al. 2020), adapt it to their addressee. Models and theories defending this speaker-oriented view comprise the generic- vs. partner-listener adaptation model (Dell and Brown 1991), the perspective monitoring and adjustment model (Horton and Keysar 1996, Keysar et al. 1998, 2000), and the dual-process model (Bard et al. 2000, Roßnagel 2000). These theories of language assume that the

incorporation of common ground during language processing (be it production or comprehension) is controlled, optional and subject to restrictions on memory capacity, and that it takes place during the later stages of processing such as monitoring, rather than in the initial utterance plan. Concretely, when speech production is difficult (e.g. low working memory, high cognitive load), hearer-oriented (utterance) planning is impaired and might be supplanted by egocentric strategies (MacDonald 2013). These two divergent lines of thought have governed the field of language processing research from its development in the late 1980s until today.

In this contribution, we propose to focus on discourse markers as a linguistic category having the potential to offer an innovative vantage point on the issues tied to language production in terms of speaker vs. hearer orientation.

2 Discourse markers as a lens

Discourse markers (henceforth, DMs) are words or expressions such as *well*, *you know*, *I mean*, *anyway* that are highly frequent and widespread in language in use, and therefore have the potential to inform us about inherent aspects of human communication. While the related literature shows little agreement on the limits of the DM category as a linguistic class (see Introduction to the volume), authors do agree on the fact that one specific DM can fulfil more than one function, either simultaneously (multifunctionality) or in different discourse contexts (polyfunctionality) given their high degree of polysemy (Bolden 2015, Crible 2017b: 107, Fischer 2006). A great deal of research supports the view of DMs as a multi-faceted phenomenon, with a general consensus that DMs operate as a pragmatic category in the metalinguistic domain, beyond the syntactic and semantic aspects of discourse (Martín Zorraquino and Portolés Lázaro 1999: 4057). Regarding their role and function in language processing, it seems that DM studies are marked by a strong hearer-oriented bias, both in Coherence Theory and in Relevance Theory, the two main theoretical perspectives regarding the function of DMs.

The Coherence Theory approach to DMs defines their role as one that signals and explicitly embodies the coherence relations that bind the elements of discourse together, therefore facilitating its overall comprehension (Schiffrin 1987: 49, Taboada 2006: 567–568). However, this model has been criticized for relying too heavily on the textual aspect of discourse focusing on connectivity and therefore failing to account for other aspects of meaning that might be signalled differently than through relationality, such as ones expressing the speaker's attitude towards their discourse or issues of speaker-hearer management, mostly present

in speech (see e.g. Maschler 2009, Crible and Cuenca 2017). The second leading theoretical approach to the functional definition of DMs is based on Sperber and Wilson's (1986) Relevance Theory. In terms of discourse analysis, this framework defines the role of DMs as one that facilitates the cognitive aspects of discourse processing. According to Blakemore (1987), DMs fulfil this function by signalling the relevance of an utterance within the discourse context, therefore making the processing efforts involved in the interpretative process of the speaker's intended meaning by the hearer less cognitively costly. In other words, this theoretical framework mostly focuses on the procedural meaning conveyed by DMs, which consists in providing metadiscursive instructions to ease the interpretation of information (Brinton 2008: 1, Hansen 2006: 25).

Thus, both of the "mainstream" theoretical frameworks describe DMs in terms of hearer benefit, their purpose being mostly understood as facilitating comprehension one way or another. Yet, a few scholars have suggested that DM use may also benefit speakers in some ways, such as by facilitating their own integration of common ground (Fuller 2003, Jucker and Smith 1998), or by 'adding some spatiality to the temporality of speech' (Crible 2018: 48) and allowing both speaker and hearer to project or backtrack their attention along the flux of speech. As far as processing is concerned, the study of DMs has shown that their use reflects the underlying processes of language planning, as they constitute major indicators of (dis)fluency besides other linguistic signals of fluency and disfluency ('fluen-cemes' in Götz' (2013) words), such as filled or unfilled pauses and repetitions (Crible, Degand and Gilquin 2017). However, these accounts remain scarce and the dominant paradigm in DM research continues to follow a bias towards the hearer, made evident by the majority of psycholinguistic studies investigating a restricted set of DMs (mainly connectives) and their impact on comprehension, and this mainly in non-interactive contexts (e.g. Brehm-Jurish 2005, Canestrelli, Mak and Sanders 2013, Sanders and Noordman 2000, Traxler, Bybee and Pickering 1997).

One might argue that this approach fails to account for one of the key aspects of language by not considering the role of DMs vis-à-vis production. Indeed, a closer look at results from corpus analyses reveals evidence suggestive of an 'easy-on-the-speaker' use (MacDonald 2013). For instance, DM polyfunctionality could be interpreted as benefitting the speaker by allowing him or her to (re) use the same forms in different contexts to express different functions, keeping them easily retrievable. In a similar logic, DM multifunctionality could allow the speaker to mitigate planning demands by using one form to perform several functions at once. Finally, a phenomenon known as underspecification (Spooren 1997), that is 'the mismatch between a discourse marker's semantics and its enriched interpretation in context (e.g. *and* in a contrastive relation)' (Crible et al.

2019: 142), could well be used as a mechanism to reduce processing costs if we consider that speakers allocate underspecified meanings to ambiguous DMs to avoid making further disambiguation steps, in that they do not develop a detailed mental representation of the relation expressed by the DM but rather stop at one that is ‘good enough’ to save resources (Crible et al. 2019, Ferreira, Bailey and Ferraro 2002, Frisson 2009, Irmer 2011). This type of corpus evidence, however indirect, highlights the potential for more related work in the area of language processing to investigate if and how the functional-pragmatic DM category could be used in ways that are not always communicative in nature but rather represent speaker strategies. This, in turn, would further inform us about how humans may mitigate the demands of utterance planning, and about the extent to which these impact speech production and speakers’ capacity to adapt to their addressee.

In this study, we will try to find out whether DMs are (primarily) used as a signal to facilitate the addressee’s comprehension and/or as a trace of the speaker’s production difficulties. This research question is addressed through one primary research question: What is the impact of allocentric v. egocentric speaker orientation on DM production? And a secondary research question: What is the impact of cognitive load on DM production, especially in the allocentric condition? More concretely, do (allocentric) speakers behave more egocentrically when put under augmented cognitive load?

3 Rationale of the study and hypotheses

To assess whether speakers use DMs with a deliberate communicative intention, this chapter proposes to investigate the use of discourse markers in a set of two experiments manipulating speaker orientation (egocentric vs. allocentric), in one of two cognitive load conditions (high vs. low). The expected outcome of this experiment is to identify potential speaker-oriented DMs and potential listener-oriented DMs. The cognitive load manipulation will aim at furthering the results obtained by assessing whether, in the allocentric experiment, high cognitive load elicits the use of more speaker-oriented DMs, while low or normal cognitive load elicits the use of more listener-oriented DMs. Thus, we expect an interaction between the two variables, such that under high cognitive load speakers become more egocentric, using less listener-oriented DMs in the allocentric condition. The strong assumption underlying this design is that advancing knowledge about how DMs work and why speakers use them will in turn shed more light on who between the speaker or hearer ultimately has the upper hand in the processing-driven shaping of grammar.

In this respect, the following two research questions will be addressed:

- RQ1. Is there a difference between the allocentric and the egocentric output, in terms of DM density, diversity and functional use?

Obviously, we expect to find more listener-oriented DMs in the allocentric task, and more speaker-oriented DMs in the egocentric task. Yet, what characterizes these two types of DMs is not clearly established. Based on the literature review, we expect to find higher density and lower diversity of DMs in the output of the allocentric task than in that of the egocentric task, due to the naturally demanding nature of the linguistic processing inherent to interactive communication (Levelt 1989) and the strong link between DMs and disfluency due to processing pressure (Crible 2018, Crible, Degand and Gilquin 2017, Götz, 2013), which is characterized by the increased use of DMs and the repetition of the same forms (Crible 2018: 87). In this (high-pressure) interactive context, speakers can be hearer-oriented, dedicating sufficient resources to the needs of the listeners, or may turn to more egocentric behaviour, if the cognitive demands of the communicative task are too high. Our assumption is that speakers will be able to dedicate sufficient resources to their hearers under “normal cognitive load” in the allocentric task. Therefore, we expect that speakers will use fewer polysemous (different meanings) DMs, because polysemy is known to be harder to process (Frisson 2009, Rice et al. 2019), and is therefore more taxing on the hearer. Also, given the obvious influence of the communicative situation on DM use, we expect to find more interpersonal (speaker-hearer management) DMs in the allocentric condition. In terms of meaning relations, it is also expected to find more markers of reformulation, specification and contrast of information in the allocentric task since these have been identified as representative of discourse functions geared towards the hearer’s needs (Ciabbari 2013, Cuenca 2003, Cuenca and Bach 2007). Spelling out the hypotheses for RQ1, these are:

- H1: There is a higher density of DMs in the allocentric task than in the egocentric task
- H2: There is a lower diversity of DMs in the allocentric task than in the egocentric task
- H3: Speakers use fewer polysemous DMs in the allocentric condition than in the egocentric task (under “normal cognitive load”)
- H4: There are more interpersonal DMs in the allocentric task than in the egocentric task.
- H5: Speakers use more DMs with a hearer-benefitting meaning (reformulation, specification, contrast) in the allocentric task than in the egocentric task

In addition to these specific hypotheses, we will sketch a distributional picture of DM use in the two speaking conditions in order to find out which DMs, if any, can be considered as indices of speaker v. listener orientation.

- RQ2. Is there a difference between the high and low cognitive load conditions, in terms of DM density, diversity and functional use?

The general hypothesis for the second research question is that increased cognitive load should make speakers in the allocentric condition more egocentric. Yet, given the relation between DMs and (dis)fluency (cf. *supra*), it is expected that a higher density and a lower diversity of DMs will be found in the high cognitive load condition, both in the egocentric and in the allocentric condition. In addition, for a similar reason, DMs performing ‘symptom’ rather than ‘signal’ uses (i.e. those with a monitoring or repairing function), in the words of Crible (2018: 53), are expected to be found in larger proportions in the high cognitive load condition. Finally, DMs that display a high degree of polysemy are expected to be found in larger proportions in the high cognitive load condition, in line with their previously hypothesized ‘easy-on-the-speaker’ use. The hypotheses for RQ2 are spelled out as follows:

- H6: There is a higher density of DMs in the high cognitive load condition than in the low cognitive load condition
- H7: There is a lower diversity of DMs in the high cognitive load condition than in the low cognitive load condition
- H8: Speakers use more polysemous DMs in the high cognitive load condition than in the low cognitive load condition.

To answer the two research questions presented above and test their related hypotheses, a usage-based, empirical approach to DMs in elicited speech will be employed, therefore requiring a set of experimental data as working material. These data along with the methods employed in their collection and analysis are detailed in the following section.

4 Experimental design

In this study, speech was elicited from participants performing two speaking tasks carried out consecutively as part of an experiment in which we manipulated two variables, the cognitive load of both participants as well as speaker orientation. In this design, the dependent variable is the DM output produced by participants.

In the first task, two participants had to jointly assemble a tangram puzzle. The dialogic and collaborative nature of this task determined its orientation as allocentric, that is centred on the addressee. One participant, the director, had to give blind verbal instructions to another participant, the matcher. The perspectives of the director and the matcher were dissociated by giving access to the puzzle silhouette to the director only. This was done in an attempt to trigger the use of potential hearer-oriented DMs as part of perspective-adaptation mechanisms. In this task, cognitive load was manipulated by having both participants conduct a (secondary) verbal memory task while assembling the puzzle. The use of dual tasks is an established method for eliciting cognitive load which researchers have been employing for a few decades. In this study, a verbal rather than visual dual task (e.g. Rosa and Arnold 2017) was selected to ensure that it would interfere with memorizing processes or attending to discourse information (tied to the phonological loop of working memory, cf. Baddeley et al. 2020) rather than with visually attending to the geometrical shapes and the silhouette of the tangram puzzle (tied to the visuo-spatial sketchpad of working memory). This design choice was inspired by Vogels et al. (2015), who followed recommendations from a study testing the effectiveness of various dual tasks on verbal, visual and spatial working memory conducted by Kellog, Olive and Piolat (2007).

In the second task, all participants had to retell themselves the instructions to the puzzle they had solved (or had attempted to solve) in the first task, using verbal cues only. The monologic nature of this task determined its orientation as egocentric, that is centred on the self. In this task, no perspective difference was therefore induced. However, cognitive load was manipulated by asking participants who were in the high cognitive load condition to recall these instructions from memory, a method previously employed in Roßnagel (2000). Participants in the low cognitive load condition were presented with the assembled puzzle while giving their instructions. Since the cognitive load manipulation for the egocentric task was very different from that for the allocentric task, the comparison between the effects of cognitive load in the allocentric and egocentric tasks may be hard to interpret.¹ Therefore, comparisons will be made only within each speaker orientation task. Our focus will be mainly on the allocentric task where we expect a clear impact of the cognitive load manipulation, with the cognitive load manipulation in the egocentric task serving mainly as a control.

¹ We would like to sincerely thank one of the reviewers for pointing out this potential methodological flaw in our data.

4.1 Participants

Data was collected from twenty-four participants (12 male, 12 female) recruited as acquaintances of the first author. As a result of this sampling method, all participants knew each other to some degree. This is a limitation to the present study, due to the established negative effect of the ‘closeness-communication bias’ on humans’ ability to take their addressee’s perspective into account (cf. Savitsky et al. 2011). However, since the degree of closeness was relatively stable across participants, all being either friends, partners or family members, we believe the bias was similar across all pairs. This method of participant selection was chosen due to its convenience. All participants reported being native speakers of Belgian French, and all were residents of the Hainaut province. Their age ranged from 19 to 55 years old ($\mu=28.9$, $s.d.=12.8$, $Md=24$), although younger speakers between the ages of 19 and 26 were preferred due to the known effect of age on the mitigation of demands created by dual tasks (cf. Kemper et al. 2009), and therefore formed the majority of the sample ($n=17$, 70.83%). Half of the participants acted as directors in the first speaking task, and the other half acted as matchers. Half of the speakers ($n=12$, 8 males, 4 females) performed the elicitation tasks under high cognitive load conditions, while the other half ($n=12$, 4 males, 8 females) performed the tasks under low/normal cognitive load conditions.

4.2 Materials

The experimental items consisted of a set of seven printed color-coded geometrical shapes commonly used to solve tangram puzzles, and two puzzle silhouettes alternatively given to participants, one representing a cat, as shown in Figure 1, and the other one representing a butterfly. Participants were also provided with a written copy of instructions to keep with them at all times, a bell which they could ring to call the experimenter into the room if a (substantial) problem arose, as well as a timer.

4.3 Procedure

The recordings took place in a quiet but non-acoustically treated room premised at a private home. Before taking part in the experiment, participants were asked to provide information about their age and first language, and to sign a consent.

At the beginning of the first speaking task, participants were randomly assigned to the role of director or matcher. They were told that the experiment was investi-



Figure 1: Cat tangram puzzle (silhouette on the left, coloured solution on the right).

gating the way pairs of friends, partners or family members gave each other verbal instructions in problem-solving settings. For this purpose, they would have to keep communicating as much as possible during the task and avoid blanks. The experimenter gave detailed instructions about how the two phases of the experiment would proceed, and gave each participant an individual written copy of those instructions in case it was needed for future reference. Then, the experimenter explained how tangram puzzles could be assembled. After participants had time to ask questions, they were seated at a table opposite each other. A blind attached to the table prevented participants from establishing visual contact and had participants communicate purely verbally. Finally, the experiment could begin and participants were given ten minutes to solve the tangram puzzle together. The experimenter exited the room and gave a signal to the participants to start solving the puzzle.

In the high cognitive load group, participants were each given a laptop at the same time as they were being seated at the table. Then, the experimenter gave them additional instructions and demonstrated how the secondary verbal memory task was to be performed. A questionnaire created on the www.typeform.com platform was loaded onto each laptop. As participants completed the questionnaire, words such as MOT and LOT (French for ‘word’ and ‘prize’, respectively), chosen for their close graphic and phonetic representations, were continually prompted onto the screen throughout the experiment. Every time a new word appeared, participants were asked whether it was the same as the first word presented to them at the start of the experiment and had to press a ‘Yes’ or ‘No’ button. Participants did not receive feedback on their answers. These dual-task stimuli were an adaptation of Vogels et al. (2015). It was designed in such a way to have participants keep a verbal item in their working memory at all times, thereby burdening it.

After having completed the first part of the experiment, participants were given a five minutes break. Then, they orally received detailed instructions about

the second speaking task from the experimenter, and once again were given a hardcopy. In this phase of the experiment, participants were told they would be recorded while verbally retelling themselves how to solve the puzzle they had just assembled with their partner, as if they would need to use these instructions (and these instructions exclusively) to assemble it again in three months' time. Participants were given a time limit of ten minutes to complete this task, and had access to the tangram puzzle and its assembled solution the whole time. After having left time for both participants to ask questions, the experimenter and one of the participants left the room, and gave the other participant a signal that he or she could start recording themselves and started the timer. Once the first person was finished, he or she rang the bell to call back the experimenter into the room (as previously instructed). Then, the other participant repeated the protocol by themselves.

In the high cognitive load condition, participants were given an additional five minutes to remember how to assemble the puzzle after being given the instructions about the task, and before beginning to record themselves. Again, they were given a time limit of ten minutes to complete the task, although this time they did not have access to the puzzle nor its solution while recording themselves, but rather had to recall the instructions entirely from memory.

4.4 Data set

The speech data collected through the elicitation tasks was compiled as a corpus, which we will onwards refer to using the acronym TACO, which stands for 'TAngram COrpus', in reference to the main topic represented in its contents. TACO can be defined as a synchronic, monolingual Belgian French corpus of elicited speech displaying a low degree of formality collected in the year 2018 in Belgium. It contains 36 oral texts, has a total duration of almost 3 hours (2:53:17s), and comprises 25,473 word tokens. TACO is (approximately) balanced between both conditions: the high cognitive load (HCL) component makes up 53% of the corpus with 13,406 word tokens, and the low cognitive load (LCL) component makes up 47% of the corpus with 12,067 word tokens.

Each text composing TACO was transcribed orthographically in standard French (retaining standard contracted forms only) and aligned to its corresponding audio file using the annotation software Praat (v. 6.1) (Boersma and Weenink 2019). The following technical guidelines were used for the transcription: individual tiers were created for each speaker and assigned an ID created using the following structure: [Speaker's initials (2 letters) + Sex (M, F) + Age], e.g. VEF55 [VE + F + 55]. Any instance of a person's name was anonymized and transcribed either as 'NAME' or

using the appropriate participant ID. Silent pauses were marked by an underscore ('_'), filled pauses were transcribed as *euh*, *euhm*, *hum*, *m/*, *mm/*, *h/*, *pf/*, *hm*, *tch/*, *ts/*, or FP if their articulation was ambiguous, and both types of pauses were transcribed within their own boundaries (i.e. separate from spoken language). Truncations were marked by a forward slash ('/'), and any unclear text was marked as 'UNCLEAR'. The textual output of each transcript was extracted from Praat's .textgrid format into a .txt file using a Praat script ('save_conversation_tiers_as_text_file.praat') written by Lennes (2019).

Table 1 provides an overview of the data set of the transcribed contents in TACO. All 24 participants performed both the allocentric and the egocentric tasks, either in the high or the low cognitive load condition, and either as director or matcher in the allocentric task.

Table 1: Overview of contents in TACO.

TASK	COGNITIVE LOAD	TEXTS	WORD TOKENS	DURATION
Dialogic (allocentric)	High	6	9511	59:26s
	Low	6	8169	53:43s
Monologic (egocentric)	High	12	3898	30:41s
	Low	12	3895	29:27s
Total		36	25473	2:53:17s

Examples (1) and (2) illustrate the kind of speech available in the data set, with identified DMs marked in bold.

Allocentric High Cognitive Load condition

- (1) SAF24: **bon alors** *euh le carré mauve tu le places sur sa pointe en face de toi*
'well then uhm the purple square you put it on its tip in front of you'
ASF24: **okay** *ouais*
'okay yeah'
SAF24: **okay**
okay
ASF24: *ouais*
'yeah'
SAF24: **ensuite** *attends je teste non merde c'est pas ça euh*
'then wait I test it out no shit that's not it uhm'
ASF24: **ben limite** *essaye de me décrire la forme*
'well maybe try to describe the form to me'
SAF24: *ouais c'est un chat*
yeah it's a cat

Egocentric High Cognitive Load condition

- (2) COM19: **donc** euh je commence par le parallélogramme qui se met euh en bas à droite **ensuite** je fais la tête du chat **donc** la tête du chat c'est euh le carré **mais** qui se met euh en biais **donc** de manière à former quasi un losange **puis** j'assemble les deux petits triangles sur le carré de manière à faire les deux euhm les deux oreilles **puis** je reviens au parallélogramme

'so uhm I start with the parallelogram that is put uhm at the bottom on the right then I make the head of the cat so the head of the cat that is uhm the square but that is put uhm across so in such a way to form nearly a diamond then I join the two small triangles on the square in order to make the two uhm the two ears then I come back to the parallelogram'

4.5 DM identification and annotation

The full annotation of the TACO data was carried out in two consecutive steps by a single annotator for the whole corpus. To ensure the validity of the annotation, Kappa statistics (Cohen, 1960) were computed as intra-annotator reliability measures for each level of disambiguation. In this section, their results and the procedure involved in each of the annotation steps are described in more detail.

The first step of the annotation process consisted in the chronological and manual identification of DM candidates in the TACO corpus. This was done by carefully reading and listening to the sound-aligned transcripts using the EXMAR-aLDA software (Schmidt and Wörner 2012). The identification of DMs was carried out using the following open-class definition of DMs taken from Crible (2018) and a set of additional restrictive criteria added by the same author in Crible and Zufferey (2015):

DMs are a grammatically heterogeneous, syntactically optional, polyfunctional type of pragmatic marker. Their specificity is to function on a metadiscursive level as procedural cues to constrain the interpretation of the host unit in a co-built representation of on-going discourse. They do so by either signaling a discourse relation between the host unit and its context, making the structural sequencing of discourse segments explicit, expressing the speaker's meta-comment on their phrasing, or contributing to the speaker-hearer relationship.

The restrictive criteria further applied to this definition in view of improving its replicability are the following, quoted directly from Crible and Zufferey (2015: 4):

- [strict] syntactic optionality: their removal does not alter the grammaticality of the utterance;

- scope over syntactically and semantically independent units: there must be a finite or implicit predicate, which excludes relative and non-finite clauses, and nominal phrases except when these are acting as a-verbal predicates;
- high degree of grammaticalization: fixed multi-word units, frequently used (not idiosyncratic) and semantically non-compositional;
- incompatibility with membership in the categories of fillers, interjections, response signals, epistemic parentheticals, general extenders, tag questions and editing terms.

The combined use of the definition and the criteria was tested on a recording (HCL_SAAS_ALLO) representing approximately 7% of the whole TACO corpus, at three weeks' interval. It allowed for the identification of 242 DMs with a Kappa value of 0.935, which is indicative of almost perfect agreement according to Landis and Koch (1977).

The second and final step of the annotation process consisted in the chronological and manual classification of the previously identified DMs using Crible and Degand's (2019) functional annotation model. Of particular interest in this model are the two independent dimensions of analysis that aim at accounting for the full functional spectrum of DMs. On the one hand, DMs are coded according to the discourse domain in which they work, corresponding more or less to the communicative intention of the speaker, what he or she aims at achieving through their contribution. Four domains are distinguished, depending on whether a given DM is used to link events in the world (ideational domain), to link mental states or attitudes (rhetorical domain), to structure the discourse (sequential domain) or to manage the speaker-hearer interaction (interpersonal domain). Independently from the domain annotation, DMs also receive a discourse function corresponding to the specific meaning of the marker in the given context (one of fifteen functions among which [consequence], [addition], [temporal], [specification], [contrast], etc.). Theoretically, any function can combine with any of the four domains giving rise to sixty (4x15) possible domain-function combinations visualized in Table 2. Cross-linguistic corpus studies have indeed demonstrated that many functions can combine with the four domains, but that languages differ in the ways DMs may actually express specific combinations (Degand, Broisson, Crible, and Grzech 2022).

Given the theoretical weight of the combinations made possible between domains and functions to encompass DM poly- and multifunctionality, the annotation was carried out on a DM-by-DM basis following the operationalisations for domains and functions in Crible and Degand (2019). No double tags were used as part of the disambiguation method. Once again, the reliability of the annotation was tested, this time by coding a sample of 250 DMs (10% of the whole DM

Table 2: Crible and Degand's (2019) functional annotation scheme.

DOMAINS	[Ideational] [Rhetorical] [Sequential] [Interpersonal]
FUNCTIONS	[Addition] [Alternative] [Cause] [Concession] [Condition] [Consequence] [Contrast] [Hedging] [Monitoring] [Temporal] [Specification] [Agreeing] [Disagreeing] [Topic] [Quoting]

sample). Kappa scores of 0.84 and 0.70 were obtained for domains and functions respectively, indicating almost perfect and substantial levels of agreement.

Following the completion of the annotation procedure, the data were processed using MS Excel and R Studio (RStudio Team, 2015) for subsequent analyses. The R packages 'vcd' and 'dplyr', were used to carry out chi-squared tests. Visual representations of the data were produced with the packages 'ggplot2'. Measures of intra-rater agreement were computed using Geertzen's (2012) online kappa calculator.

5 Analysis and discussion

5.1 Discourse markers frequency (density)

Table 3 presents an overview of the DMs frequency in the TACO data. A total number of 2700 DMs were identified.

Table 3: Raw and relative frequencies (ptw) of DMs by speaking task and cognitive load condition

	ALLOCENTRIC TASK			EGOCENTRIC TASK			TOTAL
	High	Low	High+Low	High	Low	High+Low	
<i>Corpus size</i> (words)	9,511	8,169	17,680	3,895	3,898	7,793	25,473
<i>DMs</i>	1,158	981	2,139	284	277	561	2,700
<i>ptw</i>	121.75	120.08	120.98	72.91	71.06	71.99	105.99

Zooming in on the figures reported in Table 3 for each subcorpus of TACO, two general observations can be made. The first concerns the much higher overall relative frequency of DMs in the allocentric task (120.98 ptw) than in the egocentric task (71.99 ptw). This difference was found to be statistically significant by using

a log-likelihood test ($LL=131.76$, $p<0.01$) (Brezina 2018: 84–85). On the basis of the reported frequencies and the significant log-likelihood test, we may thus confirm H1 (RQ1): speakers use more DMs in the allocentric task than in the egocentric task, both in raw and in relative frequencies.

The second observation to be made concerns the apparent stability of relative frequencies found for DMs across cognitive load conditions, in both the allocentric (high cognitive load: 121.75 ptw, low cognitive load: 120.08 ptw) and the egocentric (high cognitive load: 72.91 ptw, low cognitive load: 71.06 ptw) task. Indeed, the differences between the frequencies reported in each condition were not found to be statistically significant, whether it be in the allocentric task ($LL=0.1$, $p>0.05$) or its egocentric counterpart ($LL=0.09$, $p>0.05$). These results contradict H6 (RQ2), according to which a higher density of DMs would be found in the high cognitive load condition of each task, due to their established link to planning pressure as markers of disfluency.

5.2 Discourse markers diversity

Moving onto diversity, comparing the type-token ratios (TTR) in the different speaking tasks is a way to uncover to what extent the speaking situation influences the number of different DM types speakers use. It is known, however, that TTR is a simple measure of lexical density which is sensitive to text length, in that it decreases when a text becomes longer and the same words are used again (Brezina, 2018: 58). As a consequence, its reliability can be questioned when texts of different lengths are compared against each other, as it is the case here with the allocentric and the egocentric speaking tasks. As Figure 2 shows, texts produced in the allocentric task and those produced in the egocentric task possess different typical lengths, as characterized by the key values of their distributions. This may be due to a number of factors, including how fast pairs of participants solved the puzzle in the collaborative task (3 out of 12 pairs of speakers [25%] finished under their allocated time limit of ten minutes), the role held by speakers in the collaborative task (texts produced by directors have a mean length of 1,003.92 words, whereas that of matcher texts is 469.42 words), and potentially the fact that giving instructions to oneself simply requires a lot less time and effort than it does to give instructions to someone else, as the chance of misunderstanding is lifted (although given the same time limit to perform both tasks, most speakers completed the egocentric task in under three minutes, $M=2.5'$). In addition, since the egocentric speaking task systematically took place after the allocentric task, we cannot exclude that the repetition of two similar tasks impacted on the number of words produced as previous studies have shown that repetition of (conversational) tasks leads to lesser word use (e.g. Clark and Wilkes-Gibbs 1986).

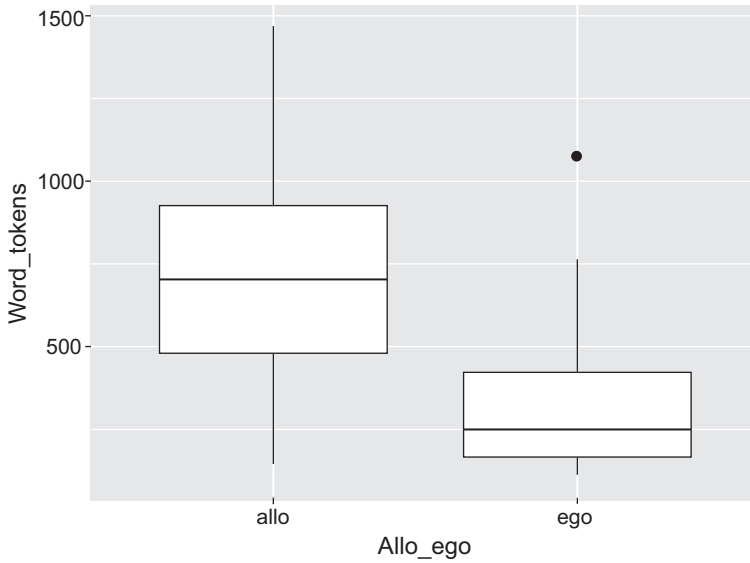


Figure 2: Dispersion of text length (in words), per speaker per speaking task.

Since texts produced by matchers in the allocentric task and texts produced in the egocentric task have relatively more comparable lengths ($M=469.42$ for the matcher group, $M=324.71$ for the egocentric task), we resampled the data set according to this criterion of individual text length. Table 4 displays the type-to-token ratios for the different speaking tasks on the basis of the restriction applied to the allocentric dataset.

Table 4: Type-token ratio of DMs per speaking task (resampled data set) and cognitive load condition.

	ALLOCENTRIC TASK			EGOCENTRIC TASK			TOTAL
	High	Low	High+Low	High	Low	High+Low	
<i>Types</i>	46	41	54	32	34	47	56
<i>TTR</i>	11.5	11.11	7.02	11.27	12.27	8.38	8.19

The figures in Table 4 show that as expected (H2, RQ1), DMs are used less diversely by speakers in the allocentric task ($TTR=7.02$) than in the egocentric task ($TTR=8.38$), although the difference between the two is very slight. This is interesting, as in previous work Crible (2018: 87) noted a sharp TTR gap between the more monologic and prepared speech and those involving a higher degree of interaction and a lower degree of preparation. The latter are assumed to entail a greater pressure

on working memory as speakers plan language – as is the allocentric task here – and thereby trigger the reuse of similar DMs. However, this is not reflected by the present findings. As for the cognitive load manipulation, we expected a lower diversity of DMs in the high cognitive load condition (H7, RQ2), which is not confirmed in the allocentric task (HCL: TTR=11.5; LCL: TTR=11.11), and only very slightly in the egocentric task (HCL: TTR=11.27, LCL: TTR=12.27). Given the similarities shared by the two speaking tasks in terms of content (instructions on a tangram puzzle), any differences (or similarities) in DM diversity are likely to be due to the manipulations of speaker orientation and cognitive load applied during the data collection process. It can therefore be concluded that, as shown by the results of the experiment so far, speakers appear to use DMs for their addressees more than they do for themselves, although both allocentric and egocentric motives seem to drive their use, given the high density of DMs overall in the egocentric output. While speakers use DMs significantly more frequently when taking the perspective of an addressee is required for communication, the increased burden this entails on the speakers does not seem to prevent them from varying their use of these discourse-structuring devices. Whether they use the same or different DM types, and for what functions, will be investigated throughout the following sections.

5.3 Discourse markers distribution: Descriptive account

Before turning to the hypotheses concerning the use of polysemous DMs in the different conditions (H3, H8), we want to give an overview of the types of DMs used, focussing on their frequency as a potential characteristic of speaker-oriented v. listener-oriented DMs.

As mentioned above, the design of the experiment carried out for the present purposes resulted in an imbalance of data between the allocentric and egocentric subcorpora of TACO, due to their corresponding speaking tasks producing different amounts of output (17,689 words vs. 7,793 words, respectively). As the present study focuses on a highly frequent and pervasive phenomenon, viz. DMs, this shortcoming arguably has limited impact on the normalized frequencies and diversity information presented thus far. However, in order to prevent this bias from influencing type and function distributions, and for comparability purposes, the further analyses will be carried out on a resampled dataset comprising a (near-)equal number of DMs in all conditions. This balanced dataset was obtained by matching the number of DMs in the lowest category ($n=561$), the egocentric task, with that of the highest category, the allocentric task. To do this, an equal number of DMs ($n=24$) were randomly selected from the allocentric output of each of the 24 speakers represented in the dataset ($n=576$), and compiled with

the existing egocentric output. The new dataset presented in Table 5 comprises 1137 DMs of 66 unique types.

Table 5: Distribution of DMs in the balanced sample, per speaking task and cognitive load condition.

	ALLOCENTRIC TASK	EGOCENTRIC TASK	Total
<i>High cognitive load</i>	288	284	572
<i>Low cognitive load</i>	288	277	561
Total	576	561	1,137

The following sections describe the peculiarities of this frequency distribution and its variations across experimental conditions.

5.3.1 Discourse markers types and frequency bands

To investigate a potential interaction between frequent vs. infrequent DMs and speaker orientation and/or cognitive load, all 66 DM types found in the dataset were sorted into four groups or ‘bands’ of frequency, identified as apparent gaps between clusters of DMs displaying similar frequency ranges in the data. The four groups classify DM frequency as ‘very high’ (>169 tokens, $n=2$ types, *donc* ‘so’, *et* ‘and’), ‘high’ (49–81 tokens, $n=5$ types, *mais* ‘but’, *ensuite* ‘then’, *okay*, *alors* ‘well, then’, *voilà* ‘that’s it’), ‘moderate’ (6–35 tokens, $n=22$ types, e.g. *ben* ‘well’, *en fait* ‘in fact’, *hein* ‘right’, *maintenant* ‘now’, *tu vois* ‘you see’), . . .), and ‘low’ (<6 tokens, $n=35$ types, e.g. *puisque* ‘since’, *par conséquent* ‘as a consequence’, *écoute* ‘listen’, *par contre* ‘in contrast’, . . .). The full list of DM types is provided in Table A1 in the appendix. This skewed DM frequency distribution with few DMs being used very frequently, and many DMs being used with a moderate or low frequency has been found to be typical of the class of DMs (e.g. Crible 2018, Martín Zorraquino and Portolés Lázaro 1999). The distribution of frequency bands across speaker orientation (allocentric vs. egocentric) is shown in Table 6 below.

Overall, DM tokens are distributed rather evenly across the top three categories, namely ‘very high’, ‘high’, and ‘moderate’, only the bottom category, ‘low’, counts approximately five times fewer tokens than the others despite counting the highest number of distinct types ($n=35$). A closer look at the distribution of more vs. less frequent DMs per speaking task (or ‘orientation’) reveals that more ‘very high’ frequency DMs seem to occur in the egocentric task than in the allocentric task, while more ‘high’, ‘moderate’ and even ‘low’ frequency DMs appear more

Table 6: Distribution of frequency bands across speaker orientation.

	VERY HIGH	HIGH	MODERATE	LOW	Total
Allocentric	95	201	234	46	576
Egocentric	271	149	120	21	561
Total	366	350	354	67	1137

often in the allocentric task. A chi-squared test of independence was performed to examine the relationship between speaker orientation and DM frequency bands. A significant association was found: $X^2(3)=138,23$, $p<0.0001$. The overall effect is moderate: Cramér’s $V=0.3549$. All cells but the ‘high’ frequency were found to contribute significantly to the chi-squared value at the level of $p<0.05$. ‘Moderate’ and ‘low’ frequency band DMs were found to be overrepresented in the allocentric task and underrepresented in the egocentric task, whereas ‘very high’ frequency band DMs were found to be strongly overrepresented in the egocentric task and underrepresented in the allocentric task. A possible explanation is that those ‘very high’ frequency DMs (*et, donc*) are frequent because they are underspecified and multifunctional (Crible et al. 2019), and as a consequence easier for speakers to retrieve from memory. Thus, their use is potentially triggered not by the intention to accommodate speech to a particular addressee’s perspective one way or another, but rather by production economy habits on the part of the speaker. Therefore, they are more likely to be found in the egocentric task where they are not in competition with other, more hearer-oriented DMs.

If highly frequent DMs are ‘easy-on-the-speaker’, then their use should increase together with the amount of cognitive load imposed on speakers. As Table 7 reveals, this holds true in the two speaking tasks, although the difference between the high and low cognitive load conditions appears rather slight.

Table 7: Distribution of frequency bands per cognitive load condition, per speaker orientation.

		VERY HIGH	HIGH	MODERATE	LOW	Total
<i>Allocentric</i>	High	50	101	114	23	288
	Low	45	100	120	23	288
<i>Egocentric</i>	High	144	71	58	11	284
	Low	127	78	62	10	277
Total		366	350	354	67	1137

Once again, chi-squared tests of independence were performed, this time to examine the relation between cognitive load and DM frequency bands within each speaking task. Significant associations were found neither in the allocentric task nor the egocentric task ($X^2(3)=1.49$, $p>0.05$), meaning that cognitive load had no effect on DM frequency bands, similarly to what we found for their frequency and diversity.

5.3.2 Discourse markers frequency and polysemy

To assess whether the high frequency of some DMs could indeed be linked to their polysemy, the degree of polysemy of the DM types found in the dataset was measured in each condition. It was operationalized as follows: DMs characterized by ‘monosemy’ were defined as those performing a single function (cf. Section 4.5, the functions level in the annotation); those characterized by ‘low polysemy’ as performing two to three functions; and those characterized by ‘high polysemy’ as performing five or more functions. DMs with a frequency of three or less were discarded as hapax legomena of which the polysemy could not be established. This leaves us with 1097 DM tokens and 36 DM types. The resulting annotations per DM frequency band are presented in Table 8.

Table 8: Distribution of polysemy levels per DM frequency band.

	HIGH POLYSEMY	LOW POLYSEMY	MONOSEMY	Total
Very high freq.	366	0	0	366
High freq.	181	169	0	350
Moderate freq.	76	181	97	354
Low freq.	0	0	27	27
Total	623	350	124	1097

These figures clearly suggest a relationship between the two variables: all of the DMs identified as ‘very high’ frequency also display a high degree of polysemy (100%), while all DMs identified as ‘low’ frequency (yet with a frequency of more than three tokens in the dataset) are monosemous (100%). The other two frequency bands, ‘high’ and ‘moderate’, span more polysemy categories, where ‘high’ frequency shows a slight tendency to attract ‘high’ polysemy DMs (in 51.71% of cases), and ‘moderate’ an affinity for low polysemy DMs (in 51.13% of cases), without excluding highly polysemous and monosemous DMs. By means of a chi-squared test, a significant association was found between frequency band and polysemy ($X^2(6)=760.85$, $p<0.0001$). The overall effect is large, Cramér’s $V=0.589$. The analysis of the stand-

ardized residuals indicates that all but one cell (high frequency * high polysemy) contributed significantly to the chi-squared value at the level of $p < 0.01$ with residual values above or below +2.58 and -2.58 respectively (Levshina 2015: 221).

Together with the previously established egocentric preference for ‘very high’ frequency DMs, and allocentric preference for ‘high’, ‘moderate’ and ‘low’ frequency DMs, these results suggest that speaker-oriented DMs uses are characterized by a higher degree of polysemy, and listener-oriented DMs are characterized by a lower degree of polysemy. In other words, speaker orientation would not as much impact the selection of high or low frequency DMs per se, but rather their degree of polysemy.

Indeed, as Table 9 shows, greater counts of highly polysemous DMs are found in the egocentric output than in its allocentric counterpart, while monosemous, although less frequent overall, seem to be typical of the allocentric orientation.

Table 9: Distribution of polysemy levels per speaker orientation.

	HIGH POLYSEMY	LOW POLYSEMY	MONOSEMY	Total
Allocentric	262 (47.46%)	201 (36.41%)	89 (16.12%)	552
Egocentric	361 (66.24%)	149 (27.34%)	35 (6.42%)	545
Total	623	354	122	1097

To assess the significance of this relationship, a chi-squared test of independence was performed. A significant association was found between speaker orientation and degree of DM polysemy ($X^2(2)=46.93$, $p < 0.0001$), with an overall small effect (Cramér’s $V=0.207$). There is a significant positive association at the level of $p < 0.01$ between monosemous DMs and allocentric speech, and another between highly polysemous DMs and egocentric speech. On the other hand, allocentric speech significantly disprefers highly polysemous DMs, while monosemous DMs are dispreferred in egocentric speech. These results support H3 (RQ1), that speakers use fewer polysemous DMs in the allocentric condition. On this basis, we may tentatively conclude that certain DMs and uses may well be indices of speaker orientation: listener-oriented DMs leaning towards monosemy, which is easier to process for the hearer, and speaker-oriented DMs being those of the highly polysemous kind.

The fact that highly polysemous DMs occurring in the ‘high’ and ‘very high’ frequency bands, make up approximately half (49,86%, see Table 8) of the total number of DMs found in the output of the (resampled) allocentric task point to a use of DMs that is less heavily centred on the hearer’s needs than previously thought (e.g. Hansen, 2006), but rather balanced between those of both speakers and hearers.

To find out whether a higher cognitive load triggers the use of more ‘speaker-oriented’ highly polysemous DMs within each task (H8, RQ2), similar tests were run contrasting the data in the high vs. low condition (Table 10).

Table 10: Distribution of polysemy levels per cognitive load condition, per speaker orientation.

		HIGH POLYSEMY	LOW POLYSEMY	MONOSEMY	Total
<i>Allocentric</i>	High	133	96	46	275
	Low	129	105	43	277
<i>Egocentric</i>	High	183	71	22	276
	Low	178	78	13	269
Total		623	350	124	1097

Unsurprisingly in light of the inconclusive results previously obtained regarding the distribution of frequency bands across cognitive load conditions, no significant associations were found between cognitive load and degree of DM polysemy, be it in the allocentric data ($X^2(2) = 0.56$, $p > 0.05$), or in the egocentric data ($X^2(2) = 2.62$, $p > 0.05$). At this point, we cannot state that polysemous DMs are “easy-on-the-speaker” (nor “hard-on-the-speaker”, for that matter), but rather that they are listener-unfriendly. In the following Section we will further deepen the impact of speaker orientation on DM use, matters of cognitive load will not be further analysed given the inconclusive results so far.

5.4 Discourse markers distribution: Domains and functions

The analysis so far has focused on providing a more descriptive account of the DM category and its formal behaviour in the experimental conditions under study. In this Section, we turn to the more functional potential of DMs to provide insight into the debated allocentricity of language. To this end, we will make use of the two-dimensional annotation in Domains and Functions, following Crible and Degand’s (2019) protocol (cf. Section 4.5).

5.4.1 Domains across conditions

Table 11 displays the number of DM tokens used in the different domains across the two speaker orientation tasks.

Table 11: Distribution of domain tokens per speaker orientation (balanced sample).

	IDEATIONAL	RHETORICAL	SEQUENTIAL	INTERPERSONAL	Total
Allocentric	29	105	174	268	576
Egocentric	86	91	368	16	561
Total	115	196	542	284	1,137

Overall, the sequential domain displays the highest frequency in the balanced TACO sample. This higher sequential use is in line with previous results across different languages (Crible 2018, Crible and Degand 2019, Degand et al. in press). Here, the results are skewed by the egocentric output, however, where the sequential domain dominates by far in comparison with its three alternatives. In the allocentric output, it is the interpersonal domain which seems to have the upper-hand, although DM tokens are distributed relatively more evenly in this condition, that is with a large number of DMs also found in the sequential and rhetorical domains. Interestingly, a small yet noticeable amount of interpersonal domain tokens can be found in the egocentric task. Upon closer analysis of those instances, it seems that their use stems from one of two possible explanations, or a combination of both. The first is related to the fact that participants were asked to give instructions to themselves ‘as if they had to use these instructions again in three months’ time to reassemble the puzzle’. A number of them did indeed just that, and addressed an impersonated version of their future self in the recording, such as by using the second person singular (*tu*). Their speech was thus punctuated with ‘interpersonal’ monitoring DMs including *hein*, *attends* and *quoi*, whenever they stalled due to having to recall information (3) and/or had to apply a repair (4,5), for instance, even though no addressee was there to whom they had to indicate their willingness to keep their speaking turn.

- (3) *et tu vas le placer m [filled pause] attends hein je ne sais plus le triangle vert*
‘and then you put it uh [filled pause] wait hang on I don’t know anymore the
green triangle’ HCL_CHF52_EGO
- (4) *à côté du des deux côtés euh m non attends du côté des deux côtés*
‘next to the the two sides uhm no wait on the same side as the two sides’
HCL_MAM19_EGO
- (5) *un centimètre au-dessus euh collé euh un elle doit elle doit dépasser d’un cen-
timètre sur le triangle jaune quoi*
‘one centimetre above um stuck on uh a it needs it needs to cover one
centimetre of the yellow triangle actually’ LCL_VEF55_EGO

The alternative explanation is that due to the observer's paradox (Labov, 1972), participants behaved in the egocentric task as they would have if the experimenter had been present, due to knowing they were being recorded. The feeling of being observed would have therefore triggered the use of such DMs. This suggests one of two things, either that the experimental condition was overall not constraining enough to prevent speakers from using these particular interpersonal DMs, or that they are not truly 'interpersonal' (i.e. produced 'for the hearer'). The latter option would imply that the use of such DMs is more automatic than controlled, as it does not appear to be affected by reduced working memory capacity, and more 'universal' at that, as suggested by Fuller (2003).

Returning to the distribution of domains in the two speaking tasks, their relative proportions are presented in Figure 3.

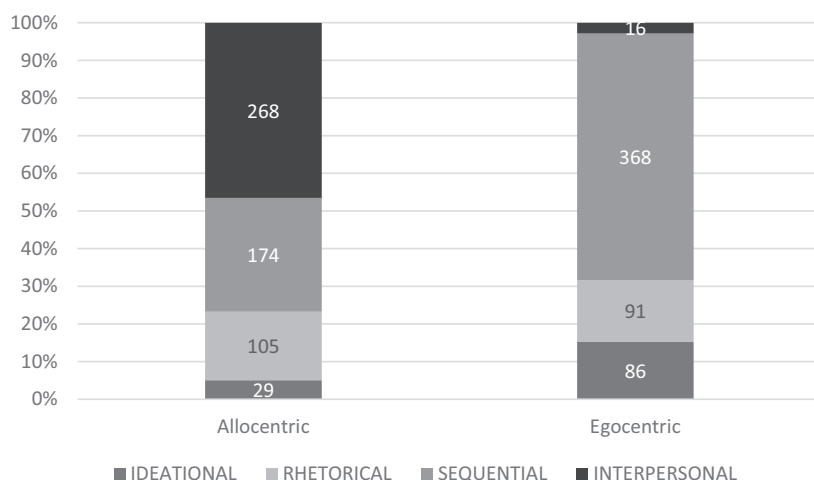


Figure 3: Proportions of DM domains per speaker orientation.

This visualization of the relative density of domains in the dataset clarifies the gaps distinguishing the speaking tasks. Overall, DMs produced in the allocentric task appear more functionally diverse than those produced in the egocentric task at this level of analysis. A chi-squared test revealed a significant association between domain distribution and speaker orientation: $X^2(3)=322.16$, $p<0.001$, with an overall large effect (Cramér's $V=0.532$). The results of the analysis revealed three significant positive associations, two between the ideational and the sequential domains and the egocentric speaker orientation, and one between the interpersonal domain and the allocentric orientation. Negative significant associations

were found between the ideational and the sequential domain and the allocentric orientation, and between the interpersonal domain and the egocentric orientation.

While the association between the interpersonal domain and the allocentric speaker orientation seems obvious, confirming H4 (RQ1), it is less straightforward to interpret the results indicating that DMs used to structure discourse (sequential) and to express (relations between) facts (ideational) are not likely to be used for communicative purposes, but are rather associated with the egocentric perspective of speakers. Apparently, the structuring function of DMs is more about organizing discourse from one's own perspective, than it is to facilitate the listener's comprehension. A closer look at the specific DM meanings (discourse functions) expressed in the sequential domain nuances this observation. A qualitative analysis of the functional DM distribution in the two speaker orientations focusing on the sequential domain only suggests a number of tendencies (see Table 12). The egocentric condition seems to attract three types of discourse relations. In order of decreasing frequency these are: monitoring, marked by DMs such as *alors* ('well, then'), *donc* ('so'), *okay*, *voilà* ('that's it'), temporal relations, marked mainly by *ensuite* ('then, after that'), *maintenant* ('now'), *puis* ('then') to express chronological ordering, and additive relations marked mainly by *et* ('and'). In the allocentric condition, there is only one type of discourse relation that stands out in terms of frequency, namely monitoring, with a larger variety of DMs to mark the relation. In addition to the DMs used in the egocentric task, *ben* ('well') and *ouais* ('yeah') seem to be more specific of the allocentric condition. Strikingly, a number of discourse relations occur (almost) only in the allocentric task, namely causal (marked by *parce que* 'because'), concessive (*mais*, 'but') and consequence (marked by *donc*, 'so') relations, when used sequentially. Such uses are not present in the egocentric dataset. While they fulfil primarily a sequential role in organizing the flow of speech they also directly address the listener, as in (6).

- (6) Receiver : **et donc** tu veux que je fasse l'inverse que je mette les pointes contre les pointes
 'and so you want me to do the opposite to put the tips against the tips'
 Director: ouais mais moi je sais pas le truc que tu dessines **hein** NAME
 'yeah but I don't know the thing that you are drawing right'

This preliminary description of which type of DMs are put to use to express different types of discourse relations in the sequential domain suggests that when speakers make the structure and organization of their discourse flow explicit, they do this in different ways depending on the speaker orientation they take.

Table 12: Distribution of DM meanings in the sequential domain across speaker orientation.

	ADD	ALT	CAU	CCS	CSQ	HDG	MNT	SPE	TMP	TOP	Total
ALLOCENTRIC	12	9	2	5	12	2	102	1	20	9	174
EGOCENTRIC	71	18	0	0	5	2	153	8	105	6	368
Total	83	27	2	5	17	4	255	9	125	15	542

ADD: additive, ALT: alternative, CAU: causal, CCS: Concession, HDG: hedging, MNT: monitoring, SPE: specification, TMP: temporal, TOP: topic

Coming now to the higher proportion of ideational DM uses in the egocentric domain, the (relational) meaning distribution across the two speaker orientations displayed in Table 13 shows that ideational uses in the egocentric task are on the account of mainly two discourse relations, i.e. addition (marked exclusively by *et*) and consequence (marked by *donc* ('so'), *et donc* ('and so') *du coup* ('consequently')), accounting together for nearly 77% of the cases. It does seem that when speakers are giving instructions to themselves, they mainly describe the events in a list like way, or, more rarely, as following from one another.

Table 13: Distribution of DM meanings in the ideational domain across speaker orientation.

	ADD	ALT	CAU	CCS	CSQ	CDN	CTR	SPE	TMP	Total
ALLO	6	6	3	3	2	1	2	1	5	29
EGO	55	0	4	4	11	0	2	4	6	86
Total	61	6	7	7	13	1	4	5	11	115

ADD: additive, ALT: alternative, CAU: causal, CCS: concession, CSQ: consequence, CDN: condition, CTR: contrast, SPE: specification, TMP: temporal

5.4.2 Speaker- and hearer-oriented functions of discourse markers

At this level of analysis, it was hypothesized (H5, RQ1) that DMs functioning as markers of reformulation, specification and contrast of information were expected to be found in larger proportions in the allocentric output than in the egocentric output, as previous literature had identified them as representative of discourse functions geared towards the hearer's needs (Ciabbarri 2013, Cuenca 2003, Cuenca and Bach 2007). In Crible and Degand's (2019) taxonomy, specification (SPE) and contrast (CTR) do have their own label, but reformulation (paraphrastic or non-paraphrastic) is expressed via the combination of the rhetorical domain and the 'alternative' function (ALT) for the reformulation of two full units, while

repairs due to a change in phrasing, or the reformulation of incomplete units is marked by the combination of the sequential domain and the alternative function. An additional function in the taxonomy that may be relevant for the present discussion on hearer-orientation is concession (CCS), when “the marker signals that the segment it connects denies one or several expectations related to the other segment” (Crible and Degand, 2019). Table 14 zooms in on the distribution of these alleged hearer-oriented discourse relations across the two speaker orientations in the balanced sample. For ease of presentation, reformulation here groups sequential and rhetorical alternative.

Table 14: Distribution of “hearer-oriented” functions per speaker orientation (balanced sample).

	ADD	ALT	CAU	CCS	CSQ	CDN	CTR	SPE	TMP	Total
ALLO	6	6	3	3	2	1	2	1	5	29
EGO	55	0	4	4	11	0	2	4	6	86
Total	61	6	7	7	13	1	4	5	11	115

At first glance, the data in Table 14 show that there is no association between speaker orientation and alleged hearer-oriented DM functions. With the exception of concession, none of the other functions of contrast, specification and reformulation associate with the expected speaker orientation. Yet, when we look at the type of DMs used, we see a different distribution pattern across the conditions. Strikingly, the polysemous DM *donc* is used overwhelmingly in the egocentric task (24 out of the 41 egocentric reformulation markers) to express reformulation, while speakers resort to more variation to reformulate in the allocentric condition (*enfin* (‘I mean’), *donc* (‘so’), *ou alors* (‘or else’), *alors* (‘then’)). The same pattern arises for the specification relation, where *donc* again takes the majority of the egocentric uses on its account (15 out of the 25 DMs used in the egocentric task), thus confirming the selection of highly frequent, polysemous DMs when speakers take an egocentric perspective, regardless of the type of discourse domain or discourse relation. In contrast, the allocentric speaker orientation associates with concessive relations, overwhelmingly expressed by *mais* (‘but’). So, while speakers may reformulate, specify and even contrast for themselves or for others, establishing an explicit concessive relation is most often directed towards the hearer.

Figure 4 displays the full range of DM functions per speaker orientation in the balanced sample. The most frequent function across the two is monitoring, corresponding to speakers’ need to manage the discourse flow. In the allocentric task, top DM uses also include expressing agreement with an interlocutor

and contradicting or nuancing (conceding) information. In the egocentric task, speakers mostly use DMs to add information and define the temporality of the events they talk about.

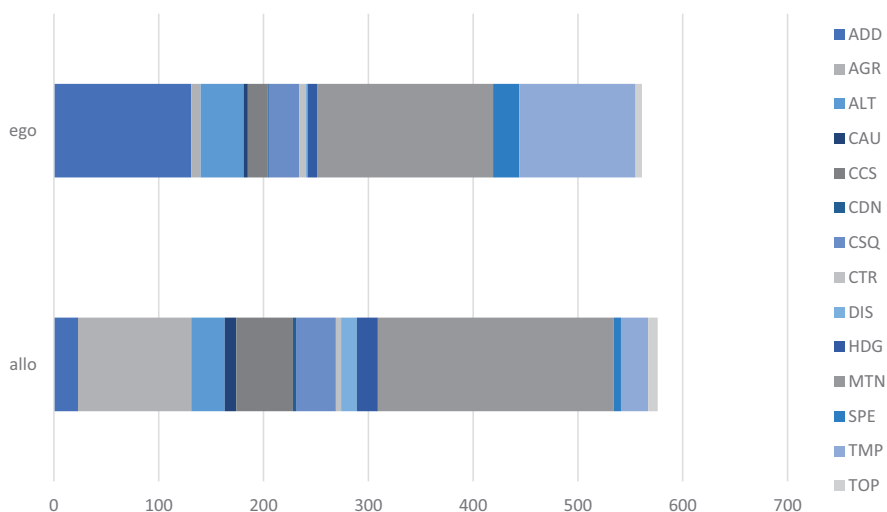


Figure 4: Raw frequency of DM functions per speaker orientation (balanced sample).

Legend: ADD: additive, AGR: agreement, ALT: alternative, CAU: causal, CCS: concession, CDN: condition, CSQ: consequential; CTR: contrastive, DIS: disagreement, HDG: hedging, MTN: monitoring, SPE: specification, TMP: temporal, TOP: topic.

To assess whether the trends of attraction detailed above were statistically valid, a chi-squared test of independence was computed. To do so, three functions had to be regrouped as ‘other’ due to their low number of occurrences: cause, condition and disagreement. A significant association was found between speaker orientation and DM function ($X^2(11) = 267.89$, $p < 0.001$). An overall moderate effect was found (Cramér’s $V = 0.485$). A significant positive association was found between the allocentric orientation and agreement, concession, hedging, monitoring and other, and between the egocentric orientation and the addition, specification and temporal functions.

From these results, it seems that more hearer-oriented uses of DMs include mostly showing agreement (AGR) to confirm (7) or check (MTN) (8) the correspondence of information between speaker and hearer information states. Indicating a shift in perspective orientation on the part of the hearer may be necessary through approximation (HDG) (9) or concession (CCS) (10). Arguably, most are uses that would indeed be more in line with an addressee-oriented definition of discourse markers such as Hansen’s (2006), and could be defined as more ‘con-

scious', controlled uses linked to the integration and/or negotiation of common ground between interlocutors. Exceptions could perhaps include some 'interpersonal' monitoring DMs linked to floorholding which apparently also occur in egocentric situations and could potentially perform more universal functions, as discussed in the previous section.

- (7) *d'accord et euh est-ce que ça doit longer euh le triangle euh*
'alright and does it need to go along uh the triangle
um' HCL_ASF24_ALLO
- (8) *comme ça tu as les deux oreilles tu vois (pause) bon*
'that way you have both ears, see (pause) alright' LCL_ANM55_ALLO
- (9) *tu le fais rentrer de genre deux trois millimètres*
'you make it go inwards by like two or three millimetres' LCL_PAF25_ALLO
- (10) *il faut qu'on prenne le bleu et le vert mais je sais pas comment les mettre*
'we need to use the blue one and the green one but I don't know how to
place them' LCL_CAF39_ALLO

DMs identified as associated to the egocentric orientation, on the other hand, tend to be used by speakers to signal the introduction of discourse-new information or express the continuity linking two consecutive ideas or segments (ADD) (11); elaborate on a previous segment by providing a detail which can be 'subjectively appreciated by the speaker' (Crible and Degand, 2019) (SPE) (12), and structure the temporality (or order) of the events or arguments presented (TMP) (13). These are uses that could be seen as 'framing' and enhancing the way information is presented (Jucker and Smith, 1998).

- (11) *le parallélogramme qui euh va venir faire la la deuxième pointe et qui va toucher à la fois le carré*
'the parallelogram which uh is going to complete the the second peak and
at the same time come into contact with the square' LCL_VIM19_EGO
- (12) *l'un de ses petits côtés à savoir celui qui est le plus haut dans ta forme*
'one of its smaller sides that is the one placed higher inside your shape'
HCL_SAF24_EGO

- (13) *l'hypoténuse étant ainsi diagonale ensuite on vient prendre le grand triangle*
 'the hypotenuse therefore being diagonal then we shall take the big triangle'
 LCL_CLF21_EGO

6 Conclusions

The aim of this contribution was to pursue the following two research objectives: (1) uncover potential speaker-oriented and hearer-oriented uses of DMs based on their frequency distribution, linguistic features and functional behaviour in an experiment manipulating speaker orientation and cognitive load; and (2) to situate evidence from DMs within the wider discussion concerning the egocentric or allocentric basis of language production.

Our main findings are that DMs occur with a higher overall relative frequency in the allocentric task than in the egocentric task. This frequency effect does however not extend to the cognitive load manipulations. Beyond mere frequency, DMs were found to be used more diversely in the egocentric experiment than in its allocentric counterpart, although the overall difference was not great. Similarly to frequency, no major differences were found between DMs identified in either cognitive load condition, be it in one task or the other. This pattern of findings is repeated throughout the analysis, as the results tend to show a greater impact of speaker orientation over cognitive load on the production of DMs. Overall, these results confirm the hypotheses formulated for RQ1 in terms of DM frequency and diversity, but not those of RQ2 which predicted that DMs would be produced more frequently but less diversely under a higher mental load.

A more fine-grained approach to DM frequency distribution and their linguistic behavior revealed that speakers in the egocentric context used a more restricted set of highly frequent DM types characterized by their high degree of polysemy (more than five different senses). On the other hand, speakers in the interactive context preferred a wider variety of less polysemous types distributed more evenly across frequency bands, including a number of monosemous DMs. These results were taken as evidence that speakers, when they behave egocentrically, show an affinity for forms that are more economical in terms of processing, such as underspecified DMs that can perform several functions. On the other hand, the use of fewer polysemous and more monosemous DMs in the allocentric condition leads us to conclude that speakers do indeed take their hearers' processing cost into account using more specific DMs that are easier to interpret in context.

Turning to the functional analysis of DMs, the objective was to try identifying potential speaker- or hearer-oriented uses within the limitations of the present off-line experimental design to elicit such information, and the reliability of its accompanying approach to DM annotation. Regarding domain distribution corresponding more or less to the speaker's communicative intention, the data revealed that hearer-oriented DM use performs mostly in the interpersonal domain, as well as in the rhetorical domain to a lesser extent. More speaker-oriented DMs, on the other hand, operate in the sequential and ideational domains of discourse and function as 'frames' that structure the temporality of discourse, fillers, or serve as markers of self-monitored adaptations or repair of speech content.

From a theoretical point of view, the starting point of the present study lies in the widespread claim that discourse phenomena, and language more generally, take place 'for the hearer'. A case in point are the numerous contributions focusing on a restricted number or types of DMs and the role they play in guiding comprehension processes. Such studies have led authors to extrapolate that DMs are indeed hearer-oriented, despite a wealth of psycholinguistic research claiming that speakers' processing considerations may be more central to language production and grammar shaping than previously thought. This claim is based mostly on evidence that speakers display egocentric behavior when their working memory is impaired, although this is hotly debated.

The present research showed that the way speakers used DMs varied depending on whether they gave instructions to themselves or to an addressee, suggesting that specific DM types or uses may be geared towards the hearer, and others to the self. The fact that DMs identified as speaker-oriented were found in large proportions in both the egocentric and allocentric tasks goes to show that, while DMs appear to be hearer-oriented to a certain extent, their speaker-oriented and hearer-oriented functions seem to co-exist and competitively weigh in the decisions that determine use in context to ultimately benefit both interlocutors when it comes to processing cost.

Importantly, the results stemming from the cognitive load manipulation applied in the experiments were mostly inconclusive, indicating minor to no differences across conditions. The first explanation coming to mind is that the cognitive load manipulation imposed on speakers during the experiment simply was not efficient. Arguably, however, this is unlikely for two reasons: first, high cognitive load was elicited differently across the two speaking tasks (verbal memory task in the allocentric task and recalling instructions from memory in the egocentric task), thus not relying on a single cognitive task. Second, the ways in which high cognitive load was elicited here were chosen for their established effectiveness in previous linguistic studies (Vogels et al. 2015, Roßnagel 2000). Therefore, the fact that stability was found across the treatment conditions in both tasks

is interesting, and suggests that higher memory demands do not substantially affect the production of DMs. One could argue that this claim is conditional to the assumption that speech produced in the low cognitive load condition was in fact produced under low cognitive load. This being said, the notion of ‘low’ cognitive load here (and in similar studies) needs to be taken with a grain of salt, as it is widely acknowledged that language production in itself is demanding processing-wise (Levelt 1989, Tanenhaus 2013), and therefore prone to requiring a certain level of mental effort from speakers regardless of any additional (artificial) burdens. In sum, no matter how high or low the cognitive load of the participants in this study was to begin with, the methodology employed in the treatment conditions should in any case increase it. Our general conclusion therefore goes against the hypothesis of an egocentric bias in language production put forward by Keysar and colleagues (Keysar et al. 1998, Keysar et al. 2000) bringing new perspectives to the table that carry potential weight as they concern a phenomenon highly pervasive in language-in-use, namely DMs.

Appendix

Table A1: Overview of DM types per frequency band (N= 1137).

Very high frequency >169 tokens	High frequency 49–81 tokens	Moderate frequency 6–35 tokens	Low frequency <6 tokens
donc (so), et (and)	mais (but), ensuite (then), okay, alors (so/ well/then), voilà (that’s it)	allez (come on), après (after), attends (wait), ben (well), bon (well), c’est ça (that’s it), d’abord (first), du coup (consequently), en fait (actually), en gros (approximately), enfin (finally/well), et alors (and then), et donc (and so), hein (you know), là (there), maintenant (now), non (no), oui (yes), parce que (because), puis (then), quoi (hé), tu vois (you see)	à savoir (that is), bah, bien (good), ça va (right), d’accord (okay), de toute façon (in any case), déjà (already), écoute (listen), eh, eh ben (well), en tout cas (in any case), en vrai (for real), enfin bon (well then), et après (and after), et ensuite (and then), et puis (and then), finalement (finally), genre (like), je vois (I see), limite (limit), mec (dude), ou (or), ou alors (or else), par conséquent (as a consequence), par contre (in contrast), premièrement (first), puisque (since), quand même (anyway), si (if), si tu veux (if you want), t’inquiète (no worries), une fois (once), vas-y (go), viens (come)

Note: Translations of DMs are very approximate

References

- Achim, Amélie M., André Achim & Marion Fossard. 2017. Knowledge likely held by others affects speakers' choices of referential expressions at different stages of discourse. *Language, Cognition and Neuroscience* 32(1). 21–36. doi:10.1080/23273798.2016.1234059.
- Baddeley, Alan, Michael W. Eysenck & Michael C. Anderson. 2020. *Memory*. 3rd edn. London: Psychology Press.
- Bard, Ellen Gurman, Anne H. Anderson, Catherine Sotillo, Matthew Aylett, Gwyneth Doherty-Sneddon & Alison Newlands. 2000. Controlling the Intelligibility of Referring Expressions in Dialogue. *Journal of Memory and Language* 42(1). 1–22. doi:10.1006/jmla.1999.2667.
- Barr, D. J. & Boaz Keysar. (2006). Perspective Taking and the Coordination of Meaning in Language Use. In M. J. Traxler & M. A. Gernsbacher (eds.), *Handbook of Psycholinguistics*. 2nd edn, 901–938. London: Academic Press.
- Blakemore, Diane Lesley. 1987. *Semantic constraints on relevance*. Oxford: Basil Blackwell.
- Blakemore, Diane. 1993. The relevance of reformulations. *Language and Literature* 2(2). 101–120. doi:10.1177/096394709300200202.
- Blokpoel, Mark, Marlieke van Kesteren, Arjen Stolk, Pim Haselager, Ivan Toni & Iris van Rooij. 2012. Recipient design in human communication: Simple heuristics or perspective taking? *Frontiers in Human Neuroscience* 6. Article 253. doi:10.3389/fnhum.2012.00253.
- Boersma, Paul & Daniel Weenink. 2019. *Praat: doing phonetics by computer* (Version 6.1). from <http://www.praat.org/>.
- Bolden, Galina B. 2015. Discourse Markers. *The International Encyclopedia of Language and Social Interaction*. Hoboken, NJ: John Wiley and Sons, Inc. doi:10.1002/9781118611463.wbielsi031.
- Brehm-Jurish, Eva Ute. 2005. *Connective ties in discourse : three ERP-studies on causal, temporal and concessive connective ties and their influence on language processing*. Potsdam: University of Potsdam Doctoral Dissertation.
- Brennan, Susan E. & Joy E. Hanna. 2009. Partner-Specific Adaptation in Dialog. *Topics in Cognitive Science* 1(2). 274–291. doi:10.1111/j.1756-8765.2009.01019.x.
- Brezina, Vaclav. 2018. *Statistics in Corpus Linguistics: A Practical Guide*. Cambridge: Cambridge University Press. doi:10.1017/9781316410899.
- Brinton, Laurel J. 2008. *The Comment Clause in English: Syntactic Origins and Pragmatic Development*. Cambridge: Cambridge University Press.
- Broisson, Zoé. 2019. *How egocentric is language production? Evidence from discourse marker use under cognitive load*. Louvain-la-Neuve: Université Catholique de Louvain. Master thesis.
- Brown, Paula M. & Gary S. Dell. 1987. Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology* 19(4). 441–472. doi:10.1016/0010-0285(87)90015-6.
- Canestrelli, Anneloes R., Willem M. Mak & Ted J. M. Sanders. 2013. Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes* 28(9). 1394–1413. doi:10.1080/01690965.2012.685885.
- Ciabarri, Federica. (2013). Italian reformulation markers: a study on spoken and written language. In Catherine Bolly & Liesbeth Degand (eds.), *Across the Line of Speech and Writing Variation*, 113–128. Louvain-la-Neuve: Presses universitaires de Louvain.

- Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.
- Clark, Herbert H. & Thomas B. Carlson. 1981. Context for comprehension. In John Long & Alan Baddeley (eds.), *Attention and performance IX*, 313–330. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Clark, Herbert H. & Thomas B. Carlson. 1982. Hearers and Speech Acts. *Language* 58(2). 332–373. doi:10.2307/414102.
- Clark, Herbert H. & Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition* 22 (1). 1–39.
- Crible, Ludivine. 2017. Towards an operational category of discourse markers: A definition and its model. In Chiara Fedriani & Andrea Sanso (eds.), *Discourse markers, Pragmatics Markers and Modal Particles: New Perspectives*, 101–126. Amsterdam and Philadelphia: John Benjamins.
- Crible, Ludivine. 2018. *Discourse Markers and (Dis)fluency. Forms and functions across languages and registers*. Amsterdam and Philadelphia: John Benjamins. doi:10.1075/pbns.286.
- Crible, Ludivine, Ágnes Abuczki, Nijolė Burkšaitienė, Péter Furkó, Anna Nedoluzhko, Sigita Rackevičienė, Giedrė Valūnaitė Oleškevičienė & Šárka Zikánová. 2019. Functions and translations of discourse markers in TED Talks: A parallel corpus study of underspecification in five languages. *Journal of Pragmatics* 142. 139–155. doi:10.1016/j.pragma.2019.01.012.
- Crible, Ludivine & Maria-Josep Cuenca. 2017. Discourse Markers in Speech: Distinctive Features and Corpus Annotation. *Dialogue and Discourse* 8(2). 149–166.
- Crible, Ludivine & Liesbeth Degand. 2019. Domains and Functions: A Two-Dimensional Account of Discourse Markers. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics* (24). doi:10.4000/discours.9997. <http://journals.openedition.org/discours/9997> (8 November, 2019).
- Crible, Ludivine, Liesbeth Degand & Gaëtanelle Gilquin. 2017. The clustering of discourse markers and filled pauses: a corpus-based French-English study of (dis)fluency. *Languages in Contrast* 17(1). 69–95.
- Crible, Ludivine & Sandrine Zufferey. 2015. Using a unified taxonomy to annotate discourse markers in speech and writing. In Harry Bunt (ed.), *Proceedings of the 11th Joint ACL – ISO Workshop on Interoperable Semantic Annotation (isa-11)*. 14–22. London, UK: ACL.
- Cuenca, Maria-Josep. 2003. Two ways to reformulate: a contrastive analysis of reformulation markers. *Journal of Pragmatics* 35(7). 1069–1093. doi:10.1016/S0378-2166(03)00004-3.
- Cuenca, Maria Josep & Carme Bach. 2007. Contrasting the form and use of reformulation markers. *Discourse Studies* 9(2). 149–175.
- Dębska, Agnieszka & Joanna Rączaszek-Leonardi. 2018. What Makes Us More Egocentric in Communication? The Role of Referent Features and Individual Differences. *Discourse Processes* 55(1). 1–11. doi:10.1080/0163853X.2016.1198137.
- Degand, Liesbeth, Zoé Broisson, Ludivine Crible & Karolina Grzech. 2022. Cross-linguistic variation in spoken discourse markers: Distribution, functions, and domains. In Elizabeth Peterson, Turo Hiltunen & Joseph Kern (eds.), *Discourse-Pragmatic Variation and Change: Theory, Innovations, Contact*, 83–103. Cambridge: Cambridge University Press.
- Dell, Garry S. & Paula M. Brown. 1991. Mechanisms for listener-adaptation in language production Limiting the role of the “model of the listener.”. In Donna Jo Napoli & Judy Anne Kegl (eds.), *Bridges between psychology and linguistics*. 105–129. New York: Academic Press.

- Ferreira, Fernanda, Karl G. D. Bailey & Vittoria Ferraro. 2002. Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science* 11(1). 1–15. doi:10.1111/1467-8721.00158.
- Frisson, Steven. 2009. Semantic Underspecification in Language Processing. *Language and Linguistics Compass* 3(1). 111–127. doi:10.1111/j.1749-818X.2008.00104.x.
- Fischer, Kerstin (ed.). 2006. *Approaches to discourse particles*. (Studies in Pragmatics 1). Bingley, UK: Emerald Group Publishing.
- Fuller, J.M. 2003. Discourse marker use across speech contexts: A comparison of native and non-native speaker performance. *Multilingua* 22(2). 185–208.
- Galati, Alexia, Anthi Diavastou & Marios N. Avraamides. 2018. Signatures of cognitive difficulty in perspective-taking: is the egocentric perspective always the easiest to adopt? *Language, Cognition and Neuroscience* 33(4). 467–493. doi:10.1080/23273798.2017.1384029.
- Galinsky, Adam D., William W. Maddux, Debra Gilin & Judith B. White. 2008. Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science* 19(4). 378–384. doi:10.1111/j.1467-9280.2008.02096.x.
- Geertzen, Jan 2012. *Inter-Rater Agreement with multiple raters and variables*. Retrieved March 23, 2019, from <https://nlp-ml.io/jg/software/ira/>
- Götz, Sandra. 2013. *Fluency in native and nonnative English speech*. Amsterdam and Philadelphia: John Benjamins. doi:10.1075/scl.53.
- Grice, Herbert Paul. 1975. *Logic and conversation*. *Speech Acts*. 41–58. New York: Academic Press.
- Hansen, Maj-Britt Mosegaard. 2006. A dynamic polysemy approach to the lexical semantics of discourse markers (with an exemplary analysis of French 'toujours'). In Kerstin Fischer (ed.), *Approaches to discourse particles*, 21–41. Oxford: Elsevier.
- Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Horton, William S. & Richard J. Gerrig. 2016. Revisiting the memory-based processing approach to common ground. *Topics in Cognitive Science* 8(4). 780–795. doi:https://doi.org/10.1111/tops.12216.
- Horton, William S. & Boaz Keysar. 1996. When do speakers take into account common ground? *Cognition* 59(1). 91–117. doi:10.1016/0010-0277(96)81418-1.
- Irmer, Matthias. 2011. Bridging inferences. Constraining and resolving underspecification in discourse interpretation. Berlin and Boston: De Gruyter Mouton.
- Jaeger, T. Florian & Harry Tily. 2011. On language 'utility': Processing complexity and communicative efficiency. *Cognitive Science* 2(3). 323–335. doi:10.1002/wcs.126.
- Jucker, Andreas H. & Sara Smith. 1998. And people just you know like 'wow'. Discourse markers as negotiating strategies. In Andreas H. Jucker & Yael Ziv (eds.), *Discourse Markers. Descriptions and theory*, 171–201. (Pragmatics and Beyond New Series 57). Amsterdam and Philadelphia: John Benjamins.
- Kellogg, Ronald T., Thierry Olive & Annie Piolat. 2007. Verbal, visual, and spatial working memory in written language production. *Acta Psychologica* 124(3). 382–397. doi:10.1016/j.actpsy.2006.02.005.
- Keysar, Boaz, Dale J. Barr & William S. Horton. 1998. The egocentric basis of language use: Insights from a processing approach. *Current Directions in Psychological Science* 7(2). 46–49. doi:10.1111/1467-8721.ep13175613.

- Keysar, Boaz, Dale J. Barr, Jennifer A. Balin & Jason S. Brauner. 2000. Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science* 11(1). 32–38. doi:10.1111/1467-9280.00211.
- Landis, J. Richard & Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159–174. doi:10.2307/2529310.
- Lennes, Mietta. 2019. SpeCT – *The speech corpus toolkit for Praat*. Retrieved May 13, 2019, from <https://lennes.github.io/spect/>
- Levelt, William J. M. 1989. *Speaking: From intention to articulation*. Cambridge, Mass.: The MIT Press.
- Levshina, Natalia. 2015. *How to do Linguistics with R: Data exploration and statistical analysis*. Amsterdam and Philadelphia: John Benjamins. doi:10.1075/z.195.
- MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4. 226. doi:10.3389/fpsyg.2013.00226.
- Maschler, Yael. 2009. *Metalanguage in Interaction: Hebrew Discourse Markers*. Amsterdam and Philadelphia: John Benjamins.
- Martin Zorraquino, Maria Antonia & José Portoles Lazaro. 1999. Los marcadores del discurso. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, 4051–4213. Madrid: Espasa-Calpe.
- Pijpops, Dirk, Dirk Speelman, Stefan Grondelaers & Freek Van De Velde. 2018. Comparing explanations for the Complexity Principle: Evidence from argument realization. *Language and Cognition* 10(3). 514–543. doi:10.1017/langcog.2018.13.
- Rice, Caitlin A., Natasha Tokowicz, Scott H. Fraundorf & Teljer L. Liburd. 2019. The complex interactions of context availability, polysemy, word frequency, and orthographic variables during lexical processing. *Memory and Cognition* 47(7). 1297–1313. doi:10.3758/s13421-019-00934-4.
- Rosa, Elise C. & Jennifer E. Arnold. 2017. Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language* 94. 43–60. doi:10.1016/j.jml.2016.07.007.
- Roßnagel, Christian. 2000. Cognitive load and perspective-taking: applying the automatic-controlled distinction to verbal communication. *European Journal of Social Psychology* 30(3). 429–445. doi:10.1002/(SICI)1099-0992(200005/06)30:3<429::AID-EJSP3>3.0.CO;2-V.
- Sanders, Ted J. M. & Leo G. M. Noordman. 2000. The role of coherence relations and their linguistic markers in text processing. *Discourse Processes* 29(1). 37–60.
- Savitsky, Kenneth, Boaz Keysar, Nicholas Epley, Travis Carter & Ashley Swanson. 2011. The closeness-communication bias: Increased egocentrism among friends versus strangers. *Journal of Experimental Social Psychology* 47(1). 269–273. doi:10.1016/j.jesp.2010.09.005.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Schmidt, T. & K. Wörner. (2012). EXMARaLDA. In Jacques Durand, Ulrike Gut & Gjert Kristoffersen (eds.), *The Oxford handbook on corpus phonology*, 402–410. Oxford: Oxford University Press.
- Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Cambridge MA: Harvard University Press.
- Spooren, Wilbert. 1997. The Processing of Underspecified Coherence Relations. *Discourse Processes* 24(1). 149–168.
- Taboada, Maite. 2006. Discourse markers as signals (or not) of rhetorical relations. *Journal of Pragmatics* 38(4). 567–592. doi:10.1016/j.pragma.2005.09.010.

- Tanenhaus, Michael K. 2013. All P's or mixed vegetables? *Frontiers in Psychology* 4. doi:10.3389/fpsyg.2013.00234.
- Traxler, Mathew J., Michael D. Bybee & Martin J. Pickering. 1997. Influence of connectives on language comprehension: eye tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 50(3). 481–497.
- Vogels, Jorrig, Emiel Krahmer & Alfons Maes. 2015. How cognitive load influences speakers' choice of referring expressions. *Cognitive Science* 39(6). 1396–1418. doi:10.1111/cogs.12205.
- Vogels, Jorrig, David M. Howcroft, Elli Tourtouri & Vera Demberg. 2020. How speakers adapt object descriptions to listeners under load. *Language, Cognition and Neuroscience* 35(1). 78–92. doi:10.1080/23273798.2019.1648839.
- Yoon, Si On & Sarah Brown-Schmidt. 2018. *Aim low: Mechanisms of audience design in multiparty conversation*. *Discourse Processes* 55(7). 566–592. doi:10.1080/0163853X.2017.1286225