


## Discourse markers in the speech of Italian tourist guides: a corpus-based study

Iolanda Alfano

Università di Salerno, Italy ✉ 

<https://dx.doi.org/10.5209/clac.87653>

Received: 17 March, 2023 • Accepted: 17 April, 2024

**Abstract:** The aim of this paper is to illustrate the incidence and functions of Discourse Markers (DMs) in the speech of tourist guides (TGs), and further investigate their polyfunctionality. Approximately 3 ½ hours of speech were examined (corresponding to 30 429 graphic words). These consisted of audiovisual recordings of Neapolitan expert guides conducting guided tours of the San Martino Charterhouse (Naples), extracted from the CHROME Corpus (Origlia *et al.*, 2018; Alfano *et al.*, 2023). A total of 1759 occurrences were analyzed. DM functions were studied by means of an onomasiological approach considering interactional, metatextual and cognitive dimensions, and then annotated using a multilevel annotation scheme. The examined features concern frequency, variability in forms, the possibility of co-occurrence and polyfunctionality, on both the syntagmatic and the paradigmatic axes. The results show that DMs are powerful indicators of the type of register: their functional distribution responds to specific communicative needs and the choice of preferred DMs depends very much on their polyfunctionality: the more polyfunctional they are, the more frequent they are likely to be. Finally, polyfunctional clusters of DMs do not present restrictions with respect to functions, whereas monofunctional clusters of DMs tend to appear with a planning or focusing function.

**Keywords:** Discourse analysis; Discourse markers; corpus analysis; Italian speech; tourist guides.

**Index:** 1. Introduction 2. Discourse-markers in this research 2.1 Exclusions 3. Methodology 3.1 Corpus 3.2 DMs functional annotation 3.2.1 Interactional macro-function 3.2.2 Metatextual macro-function 3.2.3 Cognitive macro-function 4. Results 4.1 Incidence 4.2 Functions 4.3 Forms and paradigmatic polyfunctionality 4.4 Co-occurrences and syntagmatic polyfunctionality 5. Discussion and conclusions. References.

**How to cite:** Alfano, I. (2025). Discourse markers in the speech of Italian tourist guides: a corpus-based study. *Círculo de Lingüística Aplicada a la Comunicación* 102 (2025), 133-153. <https://dx.doi.org/10.5209/clac.101428>

### 1. Introduction

The speech of tourist guides (TGs) can be framed within the complex and varied field of the Language of Tourism (LT). LT is a type of Language for Special Purposes (LSP) used in a specialist field of knowledge and characterized by highly specific lexical, morphosyntactic and textual features (Dann, 1996; Boyer & Viallon, 2000; Gotti, 2003, 2006; Calvi, 2005, 2006; Nigro, 2006; Maci, 2010; Calvi & Mapelli, 2011; Francesconi, 2014; Gotti *et al.*, 2017; Elia, 2018; Maci *et al.*, 2018; Maci, 2020; Denti, 2021; Yanoshevsky, 2021; Antelmi, 2022). The main lexical features identified in the literature for this LSP are a specific terminology characterized by monoreferentiality, clarity, conciseness and accuracy (a.o., Boyer & Viallon, 2000) and, more recently, the use of foreign words, especially Anglicisms due to English's status as a universal language on the Web (Giménez Folqués, 2015). As far as morphosyntax is concerned, the marked frequency of nominalization and passive forms stems from a general need to depersonalize and objectivize discourse in order to gain credibility. Moreover, the massive use of evaluative forms, above all positive adjectives and superlative forms, is linked to the desire to speak in brilliant terms about the services or goods being promoted (a.o., Calvi, 2006; Pierini, 2009). A general preference for the present tense is noted, as if to temporally untie what is asserted and make it eternal. The requirements of clarity and accuracy also have an impact in determining the textual organization, arranged according to compositional canons peculiar to the textual genre to which they belong (a.o., Nigro, 2006). However, tourism genres are the result of a process of interdiscursivity (Bhatia, 2008, 2010): they borrow from several domains (such as history, geography, art, architecture, etc.) and a mix of specialized languages with no clear-cut boundaries (Calvi, 2005, 2010), as they integrate several aspects of different genres and are combined with visual, text-external factors and with interdiscursive practices (Denti, 2021).

Tourist texts vary according to writer/reader or speaker/listener interaction, spatial and temporal relationships, and communicative goals, but also to the use of different semiotic resources (Santulli *et al.*, 2009; Denti, 2021; Antelmi, 2022). A crucial axis of variation depends on the participants' degree of specialization in the communication, which results in an abstract division into three macro-areas of interaction: experts-to-experts, experts-to-semi-experts, and experts-to-non-experts (Calvi, 2005). Communication between experts is characterized by a high density of concepts and specific terminology. Here, texts follow precise structures and are as objective and neutral as possible. The purpose of texts that fall into this category is to present the research results or compare different opinions; it will therefore be expository-argumentative in nature. In the second case, i.e. communication between experts and semi-experts, the specific content and terminology are slightly simplified, commented and explained in order to make the specialist text more accessible. The main purpose is didactic, in the sense that it aims to train future trainers. In the third ideal case, i.e. communication between experts and the general public, the specialized language is cleansed of excessive technicalities and is constantly explained and reformulated (Calvi, 2005). The transmission of this knowledge is achieved according to the general or more specific purposes and functions of the different texts, which can be defined as: 1) informative, of a practical nature (transport, timetables, etc.), as well as of a cultural, historical, artistic or other nature of a tourist destination; 2) persuasive, with the aim perhaps being to sell a product (travel, service, etc.) and/or promote an image and, more generally, a vision of the world; 3) directive, containing instructions to be followed by the tourists in the course of the trip (Calvi & Bonomi, 2008).

However, there are no pure genres: hybrid texts are often the norm and not an exception, due to rapid changes in the fruition of texts (Calvi, 2010) and the limits between genres becoming increasingly fuzzy (Calvi, 2017). Participants in the communication can present a highly heterogeneous and unpredictable real degree of specialization. Moreover, although TGs seek primarily to provide their listeners with information (such as historical facts), they recommend and express opinions through their very personal experience, entertain tourists, and aim to create a good impression of the tour and the place. Consequently, in order to investigate the speech characteristics of TGs, an accurate description of the specific situational context is absolutely essential. Far from being simply pre-formulated talks and monological lectures, guided tours are best viewed as being interactive and shaped around the moment's unfolding challenges and opportunities, in which the audience, regarded as individuals rather than a single unit, play an active role as learners (Best, 2012). Moreover, participants engage in actions, such as asking questions or giving explanations, in mobile vs. stationary specific interactional spaces within particular arrangements of the bodies in action while talking and walking. Questions and answers and, more generally, turn-taking management and the whole interaction are sensitive to the spatial ecology of this activity (Mondada, 2009, 2014a, 2017, 2018; Broth & Mondada, 2013).

While the linguistic literature on written guides has an established tradition (for an overview, see Nigro, 2006; Mapelli, 2013; Calvi, 2017; Elia, 2018; Antelmi, 2022), relatively less attention has been paid to human guides and their speech. In TG speech, the organization of turns and sequences, coherence and cohesion are achieved using various strategies and a combination of multimodal resources: lexical and morphosyntactic choices (Mauri 2014), phonetic and prosodic cues (Cataldo *et al.*, 2019; Alfano *et al.*, 2021; Schettino *et al.*, in press), gazes, gestures and body movements (Mondada, 2014b; Cutugno & Campisi, 2022) and also through the use of discourse markers (DMs) (Rosypalová, 2012).

In spite of the extraordinary attention that DMs have enjoyed for decades now, to the best of our knowledge, they have not been extensively studied in the speech of TGs. The interest in this research stems from the empirical observation of the special pervasiveness of these elements in this specific register. The hypothesis underlying this work is that the high frequency of DMs is closely related to their polyfunctionality (Degand *et al.*, 2022). Therefore, the aim of this paper is to investigate the incidence and functions of these elements by conducting a corpus-based analysis with an onomasiological approach (López & Borreguero, 2010; Borreguero, 2015; Sansò, 2020). Detailed studies on the functional range of a single DM abound, whereas a bird's-eye perspective seems, in our view, to provide a different way of looking at the data that is useful in further delineating the characteristics of this speech. Therefore, this study addresses these two main questions: (1) How are DMs used and distributed in the Italian speech of TGs and which DMs are the most frequent? (2) How is their frequency related to polyfunctionality?

This paper is organized as follows. In Section 2 an overview of the properties defining the DMs in this study is given. In Section 3 the methodology is presented: Section 3.1 is devoted to the corpus with details on the situational context, and Section 3.2 to the functional tag-set and the annotation process. Section 4 is devoted to the results of this investigation. Finally, Section 5 contains discussion, conclusions and future avenues for research.

## 2. Discourse markers in this research

We do not address the overabundance of terms and definitions for DMs, but we think it is useful to point out that the difference in terminology reflects the wide range of approaches used to identify them, as well as the multiple functions that they seem to fulfill (see, a.o., Schourup, 1999; Dér, 2010; Brinton, 2017). Since there is no agreement on what constitutes a DM, it is necessary to identify the traits held to be their defining properties in this study.

There is widespread agreement in the literature that DMs belong to different lexical categories, such as adverbs, verbs, nouns, etc. (a.o., see Schourup, 1999; Bazzanella, 2011) and that there is no formal criterion to identify them, since they form a highly heterogeneous functional class. This can be explained by the fact

that they are essentially non-referential indexes and do not contribute to the propositional content of the sentence, but rather convey pragmatic flavors connected to it or signal a semantic relation between the two sentences (among the classic studies, Schrifflin, 1987; Fraser, 1999). Despite a certain lack of agreement among linguists, the term “discourse” is apparently used to stress that the role of DMs must be considered at the level of discourse, rather than the sentence or the speech act. The term “marker” underlines the fact that DMs serve to signpost, mark, indicate and/or signal, rather than describe something (Blakemore, 2002; Borreguero, 2015). Depending on cotextual and contextual variables, DMs have a rich indexical meaning (Aijmer, 2002) related to their sequential position. Indeed, they are found in several positions within the clause, the utterance or other units, even though they seem to prefer the “peripheries” (Traugott, 2012; Beeching & Detges, 2014; Salameh *et al.*, 2018; Pons, 2018a, 2018b; Sansò, 2020). They can co-occur in the same unit and can point backwards and forwards in the context (Bazzanella, 2011; Fraser, 2015; Pons, 2018b; Haselow, 2019; Sansò, 2020; Crible & Degand, 2021; Pons & Fisher, 2021; Traugott, 2022). Finally, they can operate on several levels simultaneously, as pointed out by Bazzanella’s (2011) distinction between paradigmatic (*in absentia*) and syntagmatic (*in presentia*) polyfunctionality. The author defines “paradigmatic polyfunctionality” as the possibility for a DM to perform different functions depending on the linguistic and extralinguistic context, and “syntagmatic polyfunctionality” as the compresence of several functions conveyed by a single DM in a given occurrence.

And yet, DMs are not the only expressions which have multiple discourse functions and which operate as indicators at the level of discourse. According to recent analysis perspectives, the property distinguishing them from other discourse indicators lies in the relationship between conceptual and procedural meaning (Sperber & Wilson, 1986; Blakemore, 2002). Unlike other categories, DMs serve to guide the hearer to the interpretation of the utterance. We fully embrace this defining trait and adopt Bazzanella’s (2011) definition in this work: starting from their original meaning, DMs acquire further functions depending on the context, and serve to organize the discourse structure, to connect elements within the sentence and between sentences, to make the position of the utterance explicit in the interpersonal dimension and to highlight ongoing cognitive processes. This means that they lose part of their original meaning and gain a new meaning and new functions as a result of the grammaticalization, pragmaticalization or cooptation process (see Degand & Evers-Vermeul, 2015; Heine *et al.*, 2017; Brinton, 2017; Heine *et al.*, 2021; Traugott, 2022). This desemantization only concerns the conceptual meaning, which thus ceases to constitute the *core meaning*. On the contrary, the procedural meaning is acquired and becomes the core meaning (Borreguero, 2015: 158).

“As far as conceptual meaning itself is concerned, it can be misleading as a criterion in that even if the meaning of a discourse marker is (or seems to be) conceptual, its function may clearly be pragmatic and text-level [...] What kind of meaning a given DM has essentially depends on how far it has got in this grammaticalisation process” (Dér, 2010: 12). The different degree of desemantization implies that some DMs retain part of their original meaning, while others are completely “opaque” and acquire a procedural meaning that is often not directly derivable from their conventional meaning (at least from a synchronic point of view). Therefore, DMs considered in this research present different degrees of transparency/opacity in their original meaning. For example, some DMs considered as politeness markers – see the examples in Sansò (2020: 30) *scusa*, ‘sorry’, *cortesemente*, ‘courteously’, *per favore*, ‘please’, *gentilmente*, ‘kindly’ – maintain much of their conceptual meaning.

Here we adopt an inclusive approach, in the sense that we have considered all types of “functional markers” (Ghezzi & Molinelli, 2014): markers of functions belonging to the interactional domain, ensuring textual cohesion and indicating something about the speaker’s evaluation of the information status of their utterance (known as *modal particles* in other studies, see Degand *et al.*, 2013). As far as their realization is concerned, they may exhibit phonetic reduction (Schubotz *et al.*, 2015; Yaru *et al.*, 2021) and may be realized as a separate tone unit, although they may show different prosodic realizations (Raso & Vieira, 2016).

In line with the framework outlined above, this paper regards DMs as all those indexical elements with procedural value that have lost (a variable part of) their original meaning and have gained new discourse functions and values that may change rapidly over time with unforeseen outcomes. In our study, therefore, prototypical DMs:

- are pragmatic and textual cues with a procedural meaning,
- are loosely connected to the syntactic structure or are outside it,
- do not contribute to the truth-compositional semantics of the proposition,
- are multifunctional,
- are not identified on the basis of formal or positional features.

For an in-depth discussion of why these traits are peculiar to “prototypical” DMs and are not always true in the same way, see Dér (2010).

This set of properties rules out almost no elements a priori, but excludes many elements often regarded as DMs. For the sake of clarity, let us consider some exclusions in our work by giving some examples in the following section.

## 2.1. Exclusions

We did not consider proper interjections or filled pauses: they perform functions that are similar to those fulfilled by DMs and may have a procedural meaning, but they have no other non-discursive function from which the discursive function has developed (Borreguero, 2015). Moreover, we also excluded several frequent constructs of contemporary

Italian speech, such as *quello che*, lit. ‘that which is’, considered by Cella (2015: 160) to be an example of structural change currently underway. The structure is formed by the demonstrative pronoun *quello*, ‘the one’ – which either functions as the (direct or indirect) object of a preceding verb or is linked to a noun – leading to a relative clause introduced by *che*, ‘that’, followed by the verb *to be* (acting as the copula) and a noun. Let us consider the example in (1) from which we could remove *quello che* without losing anything from a semantic point of view:

(1) Consideriamo *quello che* è il gusto personale.

‘Let us consider personal taste’. (lit. ‘Let us consider *that which* is the personal taste’)

The construct shares several functions with those performed by DMs, such as gaining precious time to search for the most suitable word to say or mitigating (see D’Achille, 2016), but, once again, according to the criteria adopted in this work, it does not present defining properties of DMs because it does not gain new discourse functions or values.

Similarly, let us consider the following example (2) extracted from a webpage (Sansò, 2020: 48):

(2) IO CREDO IN TE. Ho scelto questa frase perché è meraviglioso sentirla dalle persone che amiamo, *non lo pensi anche tu...!??* Credere negli altri ed in noi stessi a volte non è facile, ma quando succede è fantastico...! Dentro al nostro cuore ci sentiamo più sereni e gioiosi

‘I BELIEVE IN YOU. I chose this sentence because it is wonderful to hear it from the people we love, *don’t you think...!??* Believing in others and in ourselves is sometimes not easy, but when it happens it is fantastic...! Inside our hearts we feel more serene and joyful’

Sansò explains that the entire proposition *non lo pensi anche tu...!??* (‘don’t you think...!??’) could be considered as a DM, because it performs the same discourse functions (a request for confirmation/agreement) as *no?*, ‘isn’t it?’ or *vero?* ‘right?’. In a scale of explicitness and transparency, it is situated in the maximum positive pole. Moreover, the author claims that it is preferable to adopt a broad view with the procedural meaning as the only discriminating element, not least because only in this way can the different diachronic paths of these elements be examined. Indeed, fully transparent procedural expressions often turn into opaque elements (as in the case of (1) *dunno*, coming from *I don’t know what else to say about this topic*, Sansò, 2020: 50).

Nevertheless, we do not agree with this view: from a synchronic perspective, the procedural meaning of this kind of expression (such as *non lo pensi anche tu...!??*) does not derive from a desemantization process but can be ascribed to the more general metalinguistic function whereby, when we speak, we continuously give instructions for interpretation to our interlocutors.

For the same reasons, if we compare two expressions with a topic-orientation function (resumption) such as *tornando a quanto dicevo prima* (lit. ‘going back to what I was saying before’, ‘back to my point’) and *in ogni caso* (lit. ‘in any case’, ‘anyway’), only the latter was considered to be a DM in the present study, assuming it does not have the conventional meaning of “in each of the cases considered” but serves to resume a previously mentioned topic. Both can perform the same function in discourse but, in our restrictive view, only *in ogni caso* (lit. ‘in any case’) can be considered a DM with multifunctionality and a primarily pragmatic meaning, while examples like the former rely on their conventional meaning and “are more appropriately thought of as conjunct adverbials (if phrasal, e.g. *back to my point*) or as clause chunks (if they have a clausal structure, e.g. *I totally forgot*)” as discussed by Traugott (2020: 19) regarding the marker *back to my point*.

### 3. Methodology

#### 3.1. Corpus

The dataset is derived from a corpus of Italian tourist guides’ speech, the CHROME Corpus, collected within the CHROME project (*Cultural Heritage Resources Orienting Multimodal Experience*, Origlia et al., 2018; Alfano et al., 2023). The CHROME corpus consists of audio-visual recordings of three female Neapolitan tourist guides, each leading four tours at the San Martino Charterhouse in Naples. Each tour is organized into six architectural and artistic Points of Interest (Pols):

- 1) *The Pronaos* (outside the doorstep of the church, where the tour introduction and a brief overview on the history of the Naples Charterhouse are recorded);
- 2) *The Great cloister* (a large outdoor space, where details of the monks’ everyday life are given);
- 3) *The Parlor* (the first internal area, where details of the Charterhouse rules are provided);
- 4) *The Chapter Halls* (where frescoes and other elements are described);
- 5) *The Choir* (behind the altar, where the history of the church decoration process is given);
- 6) *The Treasure Hall* (deeper inside the monastery, where the relationship between the monks and the governing parties in Naples is described and architectural details are illustrated).

In each Pol one camera was aimed at the guide while another was aimed at the audience, and speech was recorded by three microphones: one on a headset worn by the TG and two at about one meter equidistant from the TG and directed towards the tourists. Thus, even though the tour took place in both mobile and stationary interactional spaces, speech recordings were made in stationary spaces, corresponding to the six Pols. Although tourists could freely intervene by interrupting the TG’s explanations, they did not do so often.



Before moving on to the next Pol, TG asked if there were any questions. Hence, Q&A generally occurred at the end of the explanation of a Pol and technical considerations mean that speech in mobile interactional spaces does not form part of the recordings. This should be kept in mind when the study is placed within the overall framework of mobile vs. stationary interactional spaces of the interlocutors while talking and walking (Mondada, 2009, 2014a, 2017, 2018; Broth & Mondada, 2013). Our focus is exclusively on the TGs' speech rather than on the audience's, although the latter's presence is crucial to determine this complex expert-to-non-experts interaction (Calvi, 2005), with an inevitably varying degree of audience specialization. TGs have the important task of providing an adequate key to interpreting historical, artistic and natural heritage, while achieving mediation between specialized language and everyday speech, typical of communication between experts and the general public.

Table 1 shows the CHROME dataset analyzed in this study, corresponding to 7 Pols for each TG. The oral texts have an overall duration of 3 hours 21minutes 37 seconds.

Table 1. Dataset analyzed in this study

Speaker	Duration	Words
G01	01:01:25	8 220
G02	01:16:45	12 843
G03	01:03:27	9 366
Total	03:21:37	30 429

### 3.2. DMs functional annotation

DMs were identified by following the criteria laid out in Section 2. The annotation was performed three times by one coder (the author) over a period of two years and disagreements between the three versions helped to improve the annotation scheme and the consistency of annotations. Criteria that were not sufficiently operational or replicable were removed or reformulated, particularly where fuzzy boundaries between labels were concerned.

The identification procedure and the label assignment were performed by carefully assessing the context and evaluating the procedural meaning(s) of the DM. In a second step, the same labels were compared in order to check that they annotated the same function(s). Although we will focus on examples in micro-sequences, the analysis is based on the whole conversation they are extrapolated from.

Labelling was conducted using the *ELAN* software for multilevel linguistic annotations (Sloetjes & Wittenburg, 2008), where the category labels for each tier are assigned from a pre-specified vocabulary. Each tier in *ELAN* corresponds to a dimension or a macro-function (M-F) with its functions indicated by mutually exclusive labels entered in the *ELAN* vocabulary. In the type of speech considered in this study, two tags may co-occur only in the interactional dimension: a request for confirmation may also be turn-yielding, but in our corpus this is very rare. However, when it did occur, we did not consider the tag indicating turn management, assuming it to be implicit and labeled it as a request for confirmation. On the contrary, there are no a priori constraints between the different annotation dimensions. Therefore, every DM can receive several labels in the different dimensions so that the typical polyfunctionality of these elements can be respected.

On a separate level, speech was segmented into major prosodic units, identified by clear acoustic boundaries (Degand & Simon, 2009), in order to consider the position of the DM. In this work, we do not address the relation between functions and positions of DMs, for which we refer to Alfano & Schettino (2023). Major prosodic units were isolated by considering a number of phonetic boundary markers, namely the presence of a (potential) final pause;  $f_0$  declination of both  $f_0$  and energy; parametrical reset at the beginning of a new TU; prepausal lengthening. We segmented speech by following perceptive criteria and, when in doubt, we moved from *ELAN* to *Praat* (Boersma & Weenink, 2019) and manually followed the rules for splitting speech into major prosodic units identified by Degand & Simon (2009). These units, often delimited by real or potential pauses, are generally characterized by an initial  $f_0$  reset, followed by a declination, an overall fall in  $f_0$  and intensity, and a variety of final contours. In our examples, we place them between square brackets, using a transcription system based on Savy (2005). Table 2 shows the main symbols used for orthographic transcription.

Table 2. Main symbols used for orthographic transcription

Elements	Symbols
false starts	/
word fragments	+
brief silent pauses	<sp>
long silent pauses	<lp>
filled pauses with vocalization	<eeh>
filled pauses with nasalization	<ehm>
sound prolongations	<vv> and <cc>
backchannels	<eh>, <ah>, <mh>, <ahah>, <mhmh>
non-verbal vocalizations	<laugh>, <tongue-click>, <breath>, <cough>, <clear-throat>

As for the translation of the Italian DMs into English, we indicate a literal translation for the item under investigation, which corresponds to one of the first senses of the lemma in the Collins Unabridged Italian to English online dictionary. This choice aims at a simple reproduction of the form in order to give an idea of the conventional meaning associated with it. In many cases, the translation of a DM is not only particularly difficult (Aijmer *et al.*, 2006; Bazzanella, 2006a), but the best translation may simply be its omission in the target language or its replacement by another lexical item (Borreguero, 2011). If the examples contain other occurrences of DMs, in addition to the form under investigation, the proposed translation seeks to comply with the procedural meaning, values and functions played by the DM.

In addition, a separate ELAN tier was used to indicate clusters of DMs, identifying co-occurring DMs with the same function (*catene*, ‘chains’, Bazzanella, 2011):

(3) A: Mi parli dei neogrammatici!

B: Sì (-) *dunque allora* i neogrammatici *cioè* [silenzio] (Extracted from Bazzanella, 2011)

‘A: Tell me about Neogrammatics!

B: Yes (-) *dunque allora* (lit. ‘so then’) Neogrammatics *cioè* (lit. ‘that is’) [silence]’

or with different functions (*cumuli*, ‘clusters’, Bazzanella, 2011):

(4) *ma* (presa di turno) *guardi* (richiesta d’attenzione), io eh quello che posso dire è questo (Extracted from Bazzanella, 2011)

‘*ma* (lit. ‘but’) (turn-taking) *guardi* (lit. ‘look’) (request for attention) what I can say is this

However, unlike Bazzanella (2011), we classified as *catene* and *cumuli* only consecutive co-occurring DMs in the same major prosodic unit (not simply “juxtaposed”, see Cuenca & Crible, 2019), exemplified in (5) and (6) respectively. In the name of the audio file, G indicates the guide, V the tour and P the point of interest. Henceforth, we indicate it in brackets for each example:

(5) [abbiamo attraversato il chiostro dei procuratori] [che<ee> erano<oo> insomma delle figure abbastanza importanti] [<ehm> e funzionali alla vita alla alla vita della Certosa] <breath> [*cioè ovvero* erano coloro che si occupavano delle <ehm> come dire dell’amministrazione della vita quotidiana] [di tutto ciò che concerne la vita quotidiana e i rapporti con l’esterno dei certosini] (Extracted from G01V01P02)

‘[we went through the cloister of the procurators] [who were basically quite important figures] [<ehm> and functional to life at the Charterhouse] <breath> [*cioè ovvero* (lit. namely, i.e.) they were those who were in charge of the <ehm> how to say the administration of daily life] [of everything related to daily life and the Carthusians’ relations with the outside world]’

(6) [abbiamo addirittura una crocifissione del Cristo nella lunetta proprio della della parete di fondo della zona del coro] [*quindi assolutamente* <eh> *diciamo* classicheggiante] (G02V01P05)

‘[we even have a crucifixion of Christ in the lunette of the very back wall of the choir area] [*quindi assolutamente* <eh> *diciamo* (lit. therefore absolutely <eh> let’s say) classical]’

In (5) and (6) we find clusters of DMs, but in (5) both *cioè* and *ovvero* perform the same reformulation function with explicative purposes (Flores Acuña, 2009; Dal Negro & Fiorentini, 2014; Solsona, 2014; Sansò & Fiorentini, 2017): they are used to illustrate, expand or clarify the previous segment (“procurators were quite important figures and were functional to life at the Charterhouse”). On the contrary, in (6) *quindi*, *assolutamente* and *diciamo* perform different functions. *Quindi* is used to indicate a deduction of a logical consequence of something said or inferable: the crucifixion of Christ is presented as a quintessential classical element, which justifies the chosen epithet. *Assolutamente* functions as a boosting device, to increase intensity relating to the propositional content and illocutionary force of the deduction. Finally, *diciamo* fulfills the function of speech management, covering the programming times necessary for speech production (the cooccurrence with the filled pause <eh> reinforces this kind of interpretation, see Waltereit, 2006) while at the same time responding to a preparatory tempering mitigation strategy, and functions as a bush to reduce the precision of the propositional content (Caffi, 2007). The lexical choice *classicheggiante*, not a classic style but a style that imitates or tends toward the classical, has an approximation value in itself, reinforced by the DM *diciamo*, which functions as a further approximating device. Clearly, functions are assigned on a contextual basis: each occurrence of *diciamo* does not necessarily perform both functions and, obviously, may perform others (reformulation, for instance).

The tag-set follows the model devised by Bazzanella (1995, 2006a, 2006b, 2011) and other scholars (López & Borreguero, 2010; Molinelli, 2014; Borreguero, 2015; Sansò, 2020) and includes three macro-functions (M-Fs): interactional, metatextual and cognitive M-Fs.

The interactional M-F comprises the functions that serve to manage the interaction and social identity of speakers (for instance, those operations that pertain to turn management or phatic communication). The metatextual M-F consists of those functions related to the organization and the structuring of the discourse as well as the linguistic formulation of the text (e.g., the introduction, change, and closing of topics and the strategies to plan an utterance). Lastly, the cognitive M-F encompasses the functions related to cognitive processes, such as inference, speaker’s subjectivity and commitment, propositional content and illocutionary force.

Our tag-set is not intended to be exhaustive, but encompasses the functions identified in our corpus, summarized at the end of Section 3 in Table 3. Although the examples contain DMs that perform more than one function simultaneously, below we only illustrate the one for the label we wish to exemplify. Remember that square brackets indicate major prosodic units. Where there are exchanges between the TG and the audience, the latter's turns are quoted in a new line and indicated explicitly by "**audience:**". If there is no indication, it is the TG's speech.

The following sections will offer an example for each function and will explain the functions that appear to be less shared, in order to clarify which criteria were used in assigning the respective labels.

### 3.2.1. Interactional macro-function

As to the interactional macro-function (M-F), we do not dwell on labels used to manage turn-taking processes, of which we have no occurrences of turn-maintaining or yielding but only of turn-taking (T-T, see example 7).

(7) [*allora*] <lp> [ce lo stavamo anticipando prima] (G02V01P05)

'[*allora* (lit. 'then') <lp> [we were anticipating this before]'

Not surprisingly, this function is carried out in the vast majority of cases in this corpus by *allora* (Bazzanella *et al.*, 2007; Bazzanella & Borreguero, 2011, Ghezzi & Molinelli, 2020; De Cristofaro *et al.*, 2022). We also do not dwell on the politeness DMs (*Po*), which in our corpus correspond to formulaic phrases of greeting and parting or apologizing (see example 8), or on DMs used to request agreement and/or confirmation (*R-A-C*, see example 9),

(8) [buongiorno] <breath> [*allora benvenuti*] (G02V01P01)

'[good morning] <breath> [well *benvenuti* (lit. welcome)]'

(9) [e anche per esempio la parte di Gesù come pescatore di genti] [*no?*] [quindi il miracolo dei pesci] (G02V01P03)

'[and also for example the part of Jesus as the fisher of men] [*no?* (lit. no?)] [so the miracle of the fish]'

to convey agreement (*Agree*):

(10) [quanti ne contate?]

**audience:** [sette]

[*benissimo*] [sono sette] (G03V01P03)

'[how many do you count?]

**audience:** [seven]

[*benissimo* (lit. very good)] [there are seven]'

Instead, we wish to focus on what is meant by Marking of shared knowledge (*M-S-K*, see example 11) and by Request for attention (*R-At*, see example 12).

(11) [Quindi *vedete*] <breath> [la zona del coro la chiesa tutta hanno dietro uno studio e una eccellenza dei materiali e delle tecniche assolutamente straordinari] (G02V01P05)

'[So *vedete* (lit. you see)] <breath> [the choir area the whole church reveal an absolutely extraordinary study and excellence of materials and techniques]'

(12) [il parato ligneo] [però qui /] [*vedete* lì era meno informale] (G02V01P04)

'[the wooden wallpaper] [but here /] [*vedete* (lit. you see) there it was less informal]'

In (11), after explaining some types of techniques and clarifying the difference between mosaic and committed marble, the TG uses a DM, corresponding to a *fatismo* in Bazzanella (2011), to emphasize the social cohesion and highlight the shared knowledge between her and her listeners. *Vedete* helps the TG to say something like "now, on the basis of what we have seen, discussed and imagined, you know and understand very well that the church, and in particular the choir, were made with excellent materials and extraordinary techniques". The same form can obviously perform other functions in other contexts. This is the case of the example in (12). Here, the TG refers to the wooden wallpaper she had mentioned earlier in another Pol. *Vedete* not only has the conventional meaning of inviting the audience to take a look, but it also functions as a DM used to request the audience's attention because the wallpaper described by the TG at the moment of the utterance is not like the previous one. Hence, it is as if she were saying something like "pay attention in order to appreciate the difference between the two wallpapers".

### 3.2.2. Metatextual macro-function

As regards the metatextual M-F, we first distinguished topic management markers from generic markers, annotating the latter with the label *Demarcation* (*Dem*). In (13) the TG is talking about the great cloister, the significance of the place and the habits of the monks.

(13) [Per controllare che i certosini facessero] [seguissero <ehm> la regola] [passassero la maggior parte del tempo nelle loro celle] <sp> [<eh> non rompessero il silenzio all'interno<oo> appunto del chiostro grande] [*in realtà* il Priore aveva diretto accesso <eh> a questo luogo] [per fare come diciamo<oo> delle visite a sorpresa] [per vedere se appunto tutto era fatto a dovere] (G01V01P02)

'[To check that the Carthusians did] [follow <ehm> the rule] [spent most of the time in their cells] <sp> [<eh> did not break the silence inside the great cloister] [*in realtà* (lit. actually) the Prior had direct access <eh> to this place] [to make as we say surprise visits] [to see if everything was done properly]'

Here the DM *in realtà* (lit. 'actually') does not introduce something that goes against expectations: there is nothing in the context to suggest that the prior did not have direct access to the cloister. In this sense, *in realtà* does not have its conventional meaning, but rather serves to delimit textually and pragmatically meaningful portions of speech – which here are also syntactic and prosodic units –, recognizable to the addressee, performing a function similar to punctuation and spacing in writing (Voghera, 2017: 91).

Topic management marker labels follow Sansò (2020: 20), with further differentiations. In order to distinguish between cases where the DM serves to organize information relating to the same or another topic (Fraser 2009), we separated topic shift and topic continuation. Thus, in our coding scheme, we separated DMs that

i. introduce a new topic (*Intro*):

(14) [*allora*] <lp> [siamo arrivati nella chiesa] (G02V01P05)

'[*allora* (lit. then) <lp> [we have arrived in the church]'

In (14) *allora* serves to introduce a new place she is going to illustrate, the church;

ii. shift to another topic (*Shift*):

(15) [e *allora* altro elemento fondamentale] <sp> [quello di suddividere la propria giornata <breath> sull'impostazione di una regola di un ordine già esistente] [era la regola di san Benedetto] (G03V01P01)

'[and *allora* (lit. then) another fundamental element] <sp> [that of dividing one's day <breath> on the setting of a rule of an existing order] [it was the rule of St. Benedict]'

In (15) the TG is explaining the characteristics of the Carthusian order, and *allora* signals the shift to another feature, so as to introduce the division of the day into several predetermined times;

iii. signal continuation with the current topic (*Cont*):

(16) [si occupa soprattutto dei marmi colorati] [*quindi* lo incontreremo anche nella parte della chiesa] (G02V02P01)

'[He deals mainly with the colored marbles] [*quindi* (lit. therefore) we will also meet him in the church]'

In (16) the TG is talking about Cosimo Fanzago and *quindi* adds something more about the topic Fanzago that later will be found in another place;

iv. mark a closure of a topic (*Clo*):

(17) [Rispetto al quartiere<ee> commerciale<ee> e animato che è oggi] [provate a immaginare una collina dove *in pratica* non c'era nulla] (G01V01P01)

'[Compared to the commercial and bustling neighborhood that it is today] [try to imagine a hill where *in pratica* (lit. in practice) there was nothing]'

In (17) the TG is explaining the location of the Carthusian monastery, the Vomero hillside, which is now a lively neighborhood, but at the time was an isolated place. The DM *in pratica* (lit. 'in practice') is not opposed to "in theory", but it means that at that time there was absolutely nothing. After describing it as a place that was "completely uninhabited" at that time, *in pratica* helps to close the description of the isolated place;

v. signal the beginning of a digression (*Dig*):

(18) [io con le scuole *per esempio* faccio sempre un<nn> esempio molto semplice] [come i cellulari] <lp> [passano subito appunto di moda] (G03V02P01)

'[with schools *per esempio* (lit. for example) I always give a very simple example] [like cell phones] <lp> [they go out of fashion right away]'

In (18) the TG is explaining that the 14th-century castle was modified in the 16th century for defensive reasons. Just as technology quickly became obsolete, similarly the original form of the castle was no longer functional to the new ways of waging war. To explain the parallelism, the TG makes a digression, introduced by *per esempio*;

vi. return to a prior topic, resuming it after a digression (*Res*):

(19) [*quindi* negli anni venti incominciano i lavori] <lp> [<eeh> affidati tra gli altri anche a Tino di Camaino] <sp> [che qui lavora in veste di architetto] [...] [incominciano *quindi* i lavori] (G01V01P01)

'[so the 1920s saw the beginning of works] <lp> [<eeh> entrusted to Tino di Camaino, among others] <sp> [who worked here as an architect] [...] [works *quindi* (lit. therefore) began]'

In (19) the TG had introduced the topic "works" and makes a digression on Tino di Camaino to explain that he had worked as an architect in the Carthusian monastery, but he had also worked for the king as a sculptor and his works can be appreciated in various Neapolitan churches. Therefore, *quindi* lets the TG return to the already mentioned topic "works";



- vii. make a reference to a previously mentioned topic (*Ref*):

(20) [furono realizzati *appunto* da questo <ehm> pittore <eeh> spagnolo attivo attivo qui a Napoli] (G01V01P05)

'[were made *appunto* (lit. actually) by this <ehm> Spanish <eeh> painter active active here in Naples]'

In (20) the TG explains that the choir is decorated with some great masterpieces by painters active in Naples. Among them, she mentions Domenichino and the Spaniard Ribera and concludes by pointing out that the prophets depicted on the nave were painted *appunto* by this previously named Spanish artist active in Naples at that time.

The last two labels, *Res* (19) and *Ref* (20), are very close to each other, but in our tag-set a *Ref* DM is a type of metatextual instruction informing the listeners that this is an already discussed topic (it could be paraphrased with "as I told you before"), whereas a *Res* DM not only indicates a topic resumption but also that, after digressing, the speaker is going on with that topic.

Another metatextual function is the one used to manage the speech flow (*Plan*). Following Voghera (2017: 91), DMs can fulfill the function of covering programming times necessary for speech production. To avoid the occurrence of too many disfluencies, resulting inevitably in fragmentation and the consequent need for repair mechanisms, these DMs help to plan the utterance. Briz (1998: 211) speaks of a "delaying role" and compares these DMs to pauses, which serve to think about and plan what one will say. Even though they may co-occur with silent or filled pauses and with segmental prolongations, the verbal sequence is not really altered and no reconstruction intervention by the receiver is required. In (21), *insomma* helps the TG to go on with her explanation, reducing the duration of the silent pauses and/or avoiding the occurrence of other filled pauses.

(21) [qualcuno si era risentito <eeh> del<ll> <sp> *insomma* dell'operato <eeh> di questi pittori in città]

'[someone had resented <eeh> the <sp> *insomma* (lit. in conclusion) the work <eeh> of these painters in the city]'

Not all linguists include the focalizers (*Foc*) category among DMs because they are syntactically integrated into the clause in which they appear and do convey meaning (Borreguero, 2015). In this paper we have not considered all focalizing adverbs as *Foc* DMs, but only those that present a shift in their meaning, acquiring further discourse functions and procedural value in the discourse at the same time. They serve to emphasize a portion of text evoked by a series of implicit alternatives, as in the following example:

(22) [è una rappresentazione di quella che è la vocazione la conversione stessa del personaggio] [che poi diventa <eh> il protagonista *soprattutto anche* di questa Certosa] (G02V01P03)

'[it is a representation of what is the vocation, the very conversion of the character] [who then becomes <eh> the protagonist *soprattutto anche* (lit. above all also) of this Charterhouse]'

The TG is talking about St. Martin, depicted as a mercenary saint going to meet the Carthusians, with an iconographic reference to the horse episode: Martin went out on horseback covered with his cloak. Along the way he saw a poor man, half-naked and staggering from the cold. Martin, having no money with him, took pity on him and cut his cloak in half, donating part of it to the poor man. The TG points out that this is not an actual episode, but rather a depiction of the character, who becomes "the protagonist *soprattutto anche* of this Charterhouse". *Soprattutto* (lit. 'above all') and *anche* (lit. 'in addition') go beyond their conventional meaning. *Soprattutto* does not literally indicate "most importantly", "more than anything else" or "especially": there is in fact no direct comparison with other elements in the cotext. Nor does *anche* truly convey the meaning of "in addition": in fact, there is no other episode, mentioned or inferable, of which St. Martin is mentioned as the protagonist, other than being the protagonist of the Carthusian monastery. These DMs serve to stress that he alone and not someone else is the real protagonist of the Charterhouse, from which it takes its name. Although not explicitly mentioned in the cotext, there are other important figures of saints in the history of the Charterhouse, such as St. Bruno. These DMs presuppose the existence of a paradigm of alternatives to the element on which they semantically operate, often called the *focus* of the focusing adverb and give further prominence (in addition to the prosodic prominence on the focus "il protagonista" in example 22). In the light of previous research on focusing Italian adverbs, we assume that they depend on the interplay of prosodic, syntactic and informational factors and behave in a very complex and modular way, determining different focusing effects (in particular on *anche* and *soprattutto*, see De Cesare 2010, 2022).

With regard to the so-called reformulation DMs, the literature on this topic has traditionally been shaped around the division between paraphrastic and non-paraphrastic reformulation. However, reformulation is a complex category and the two terms *paraphrastic* and *non-paraphrastic* include very different discursive functions, such as identification, specification, clarification, correction, conclusion or consequence, recapitulation, reconsideration or the partial or total loss of relevance of what has been previously expressed (Garcés, 2008). Considering our corpora and the observations of Pons (2013), we have grouped all the possibilities into four categories: *Paraphrasing* (*Paraph*), *Correction* (*Cor*), *Exemplification* (*Ex*) and *Reformulation* (*Reform*).

We have considered paraphrasing (*Paraph*) as an explanatory reformulation with an identity relation between the reference segment and the reformulated one, including variation, expansion and reduction (Gülich & Kotschi, 1983: 328):

(23) [poiché siamo nel coro dei padri] [*quindi* nel coro dei monaci effettivi] (G02V01P05)

‘[since we are in the choir of the fathers] [*quindi* (lit. *therefore*) in the choir of the actual monks]’

In (23), if we consider the difference between the Carthusian monks, the fathers, priests or those destined to become priests that lived in the strictest solitude, and the others, the brothers (convers or donate), we can say that the DM *quindi* introduces an identity relation: the fathers are the “actual”, the “real” monks and are the legitimate inhabitants of the monastery.

The label *Cor* in our tag-set indicates a corrective reformulation with a substitution relation between the reference segment and the reformulated one:

(24) [nel senso che l’appezzamento di terreno gli è dato in concessione] [*ma proprio* in donazione sostanzialmente perché poi loro ne sono gli amministratori] (G02V02P01)

‘[in the sense that the plot of land is given to them in concession] [*ma proprio* (lit. *but just*) in donation basically] [because they then are the administrators]’

In (24), without going into the details of the different types of donations, it is clear from the context that the speaker intends to substitute the term *concessione*, ‘concession’, which seems to be a temporary gift, with the term *donazione*, ‘donation’, to indicate a gift that cannot be revoked. Therefore, she replaces the latter term with the former, which seems more appropriate, and uses *ma proprio* to correct herself.

We have used the label *Ex* to designate an explanatory reformulation through one or more possible examples. For instance, in example (25) *quindi* is a reformulation DM that clearly introduces some examples of the furnishings proper to the Eucharist, such as cups or bowls:

(25) [e quindi parliamo di tutti quelli che erano i paramenti utili all’ufficio della messa] [di tutto quello che poteva essere <breath> l’arredo proprio per l’eucarestia] [*quindi* coppe, boccali, tessuti estremamente preziosi dei paramenti sacerdotali] (G02V01P06)

‘[and so we are talking about those things that were the vestments useful for the office of the Mass] [of all that could be <breath> the proper furnishings for the Eucharist] [*quindi* (lit. *therefore*) cups, bowls, extremely precious fabrics of priestly vestments]’

Finally, we have labelled all the other cases as *Reform*, in that they were neither *Paraphrasing*, *Correction*, or *Exemplification*, for example cases of clarification or recapitulation:

(26) [...] nel corso del cantiere seicentesco] [*quindi* tra gli anni venti e gli anni cinquanta del Seicento] (G02V01P05)

‘[in the course of the seventeenth-century construction period] [*quindi* (lit. *therefore*) between the 1620s and the 1650s]’

In (26), if we consider the reference segment (in the course of the seventeenth-century construction period) and the reformulated one (between the 1620s and the 1650s), we can say that it is a specification of the right period. It is, in fact, not properly a paraphrase (the seventeenth-century is a longer period), nor a correction (the seventeenth-century is not wrong), nor an exemplification (between the 1620s and the 1650s is not an example).

### 3.2.3. Cognitive macro-function

Moving to the cognitive or modal dimension, we used three labels to categorize DMs with a cognitive M-F. *Inference marker (Inf)* DMs used to indicate a deduction of a logical consequence of something said or inferable (Sansò 2020: 26):

(27)[con la rifusione <breath> del bronzo del Pantheon addirittura utilizzato da Bernini] [*quindi* è chiaro che ci vuole essere un richiamo a dei grandissimi <breath> elementi <eh> diciamo di confronto con Roma] (G02V01P04)

‘[with the recasting <breath> of the bronze of the Pantheon even used by Bernini] [*quindi* (lit. *therefore*) it is clear that there is meant to be a reference to very large <breath> elements <eh> let’s say of comparison with Rome]’

In (27), the TG explains the presence of several classical elements in the Charterhouse, culminating in the most emblematic example, Bernini’s bronze: *quindi* makes it explicit that it is legitimate a direct comparison with Rome.

The second label in this dimension, *Mitigation*, is used in a broad sense to identify all kinds of marks on the epistemic and evidential axes expressed by DMs. They serve to decrease or hide obligations on the speaker, making the latter’s position explicit with respect to the low degree of certainty and/or commitment to the propositional content of the utterance, reducing their subscription to its illocutionary force and establishing a distance from what they are saying (bushes, hedges and shields respectively, Caffi, 2007). We do not know whether Caffi would place these markers in the cognitive dimension, but among other possible mitigation strategies, the author also mentions DMs (e.g., *per così dire*, ‘so to speak’ and *diciamo così*, ‘let’s say’ (Caffi, 2007: 115).

We can see in (28) that DMs help the TG to convey the isolation of these monks, albeit within a community, and help her not to assume too specific a stance on the correct expression of the spirituality of the Carthusian rule. As an art historian, she probably does not want to take responsibility for the religious orthodoxy of what she reports on Carthusian life:

(28) [Sono degli eremiti che vivono in comunità] [quindi è una regola ibrida *per certi per certi versi*] [quindi <eh> di monaci che scelgono una forma <ehm> di allontanamento dal mondo] [<ehm> che però non rinunciano *diciamo in qualche modo* a una vita di comunità] [quindi una sorta di formula a metà a metà strada *in un certo in un certo senso*] (G01V01P02)

‘[They are hermits who live in community] [so it is a hybrid rule *per certi per certi versi* (lit. *in some in some ways*)] [that is of monks who choose a form of detachment from the world] [but who do not renounce, *diciamo in qualche modo* (lit. *let’s say in some way*) a life of community] [so a sort of half-baked formula *in un certo in un certo senso* (lit. *in a certain in a certain way*)]’

*Per certi versi, diciamo, in qualche modo, in un certo senso* weaken the speaker’s degree of certainty about the accuracy of the propositional content and mark it as an approximation: they are hermits living in a sort of oxymoronic “isolated community” and adopting a sort of “hybrid rule”.

The last label in the cognitive M-F, *Intensification (Int)*, indicates DMs used as mechanisms to increase intensity relating to propositional content and illocutionary force and yield an intensifying interpretation: in other words, an increase in the quantity or quality of something gradable, an emphasizing interpretation underpinning a central/prototypical property, and a modal interpretation, i.e. a strengthening of the speaker’s attitude toward the truth of the expressed proposition and its illocutionary force (De Cesare 2000). In the following example:

(29) [e allora qui] <sp> [all’interno della chiesa] [vedremo *assolutamente* quello che è il linguaggio barocco] (G03V01P01)

‘[and so here] <sp> [inside the church] [we will *assolutamente* (lit. *absolutely*) see what the Baroque language is]’

The TG explains that little of the original Gothic layout remains in the present Carthusian monastery. In fact, renovation works starting in the late sixteenth century and throughout the seventeenth century ended up transforming it: the original Gothic appearance was changed and replaced with its current Baroque style. She goes on to explain that while there are still classical references on the outside, such as the white and gray color scheme, the interior of the church features a veritable explosion of colorful marbles, typical of the Baroque style. *Assolutamente* functions to stress the prototypical property of the church’s Baroque style and, at the same time, serves to increase the degree of commitment to the propositional content of the utterance and intensify the illocutionary force of her assertion.

All the functions identified in our corpus are summarized in Table 3.

Table 3: The functional tag-set

M-Fs		Functions	Labels
Interactional	to manage turn-taking processes	Turn-taking	<i>T-T</i>
	to make the utterance polite	Politeness	<i>Pol</i>
	to establish a social connection and to mark shared knowledge	Marking of shared knowledge	<i>M-S-K</i>
	to guarantee or request attention, comprehension, confirmation or agreement	Request for attention	<i>R-At</i>
		Request for agreement and/or confirmation	<i>R-A-C</i>
		Agreement	<i>Agree</i>
Metatextual	to organize information not related to topic entities in discourse	Demarcation	<i>Dem</i>
	to manage discourse topics	Introduction	<i>Intro</i>
		Shift	<i>Shift</i>
		Continuation	<i>Cont</i>
		Closure	<i>Clo</i>
		Digression	<i>Dig</i>
		Resumption	<i>Res</i>
		Reference	<i>Ref</i>
	to gain time for managing speech production	Planning	<i>Plan</i>
	to highlight an element (by evoking a set of alternatives)	Focusing	<i>Foc</i>
	to express the same semantic content	Paraphrasing	<i>Paraph</i>
	to correct	Correction	<i>Cor</i>
	to provide an example	Exemplification	<i>Ex</i>
	to (non-paraphrastically) reformulate	Reformulation	<i>Reform</i>

M-Fs		Functions	Labels
Cognitive	to express inference and logical argumentative connection	Inference	<i>Inf</i>
	to weaken speech act responsibility in different ways	Mitigation	<i>Mit</i>
	to intensify illocutionary force	Intensification	<i>Int</i>

## 4. Results

We illustrate the results obtained by first showing the incidence of DMs and the overall frequencies of occurrence for the three M-Fs (§ 4.1) and then point out the functions identified for each one (§ 4.2). Next, we present the frequency list of the corresponding forms and their paradigmatic polyfunctionality (§ 4.3). Finally, we focus on clusters of DMs and their syntagmatic polyfunctionality (§ 4.4).

### 4.1. Incidence

Our analysis focuses on 1759 occurrences of DMs. If we look at the incidence of DMs per 100 words, we obtain a general frequency close to 6 % (5.78 %, or, to employ a more usual normalized frequency in Corpus Linguistics, a frequency of 57.80 per thousand words (ptw)). In order to understand the actual incidence of these elements, we have considered orthographic words, excluding nonverbal vocal elements (indicated in hooked brackets in the transcription, such as filled pauses). Therefore, in a string made up of repetitions or truncated words, each one counted as a word: *sulla nel soff+ sul soffitto* (lit. ‘on the in the ceil+ on the ceiling’) counted as 5 words.

If we look at M-Fs and their combinations marked by a “+” (Figure 1), not surprisingly the prevailing dimension is the metatextual one, both in isolation and in combination with the cognitive M-F.

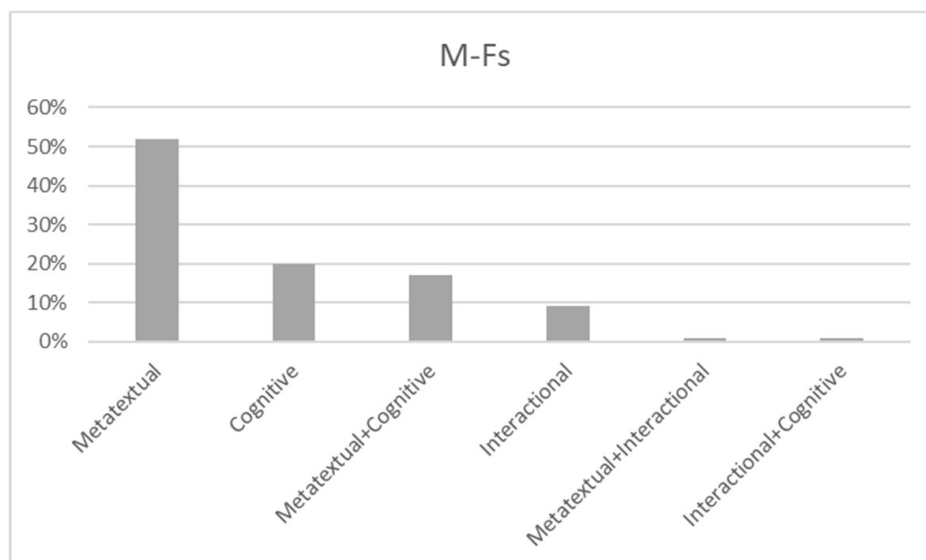


Figure 1. M-Fs (%) in our dataset of the Chrome corpus

### 4.2. Functions

The description of the M-Fs provides a general overview, but the analysis of the various functions within each macro-type offers a more accurate picture. In the interactional dimension (see Figure 2), we can see a distribution of various functions, but there is a clear prevalence of the function *Marking of shared knowledge* (M-S-K). DMs that function as turn-taking (T-T) and politeness (Pol) devices are obviously present. Finally, there are DMs that function as a clear attempt at direct audience involvement: *Request for attention* (R-At) and *Request for agreement and/or confirmation* (R-A-C).

Moving to the metatextual dimension (see Figure 3), we again observe a strong prevalence of one type: *Plan*. In almost half the cases, DMs serve to manage speech flow, gaining precious seconds for planning and production. The next most frequent functions are *Paraphrasing* and *Focusing*, with paraphrastic reformulation strategies (*Paraph*, *Ex*) outnumbering the non-paraphrastic ones (*Cor*, *Reform*). While *Cor* and *Reform* are more clearly non-paraphrastic, it is debatable whether or not *Ex* is. We emphasize this finding to note the prevalence of an explanatory type of reformulation, which highlights the TG's need to provide explanations. However, besides the semantic information, the reformulated segment may also convey a change in the argumentative orientation, in the presuppositions and implicatures and in the polyphonic structure (Pons, 2013).

As far as topic management is concerned, we see that DMs are used more to shift to another topic (*Shift*), to close it (*Clo*), to go on with the same topic (*Cont*) or to make a reference to a previously mentioned not surprisingly, (*Ref*) than to introduce (*Intro*) or resume (*Res*) it.



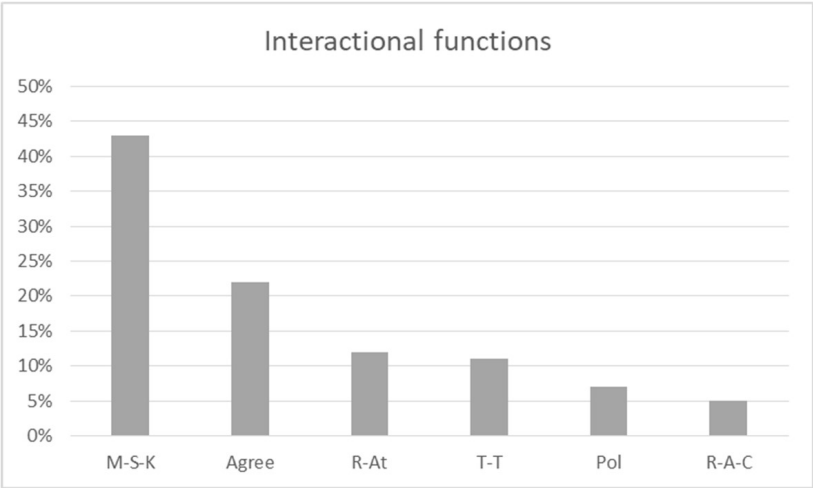


Figure 2. Interactional functions (%)

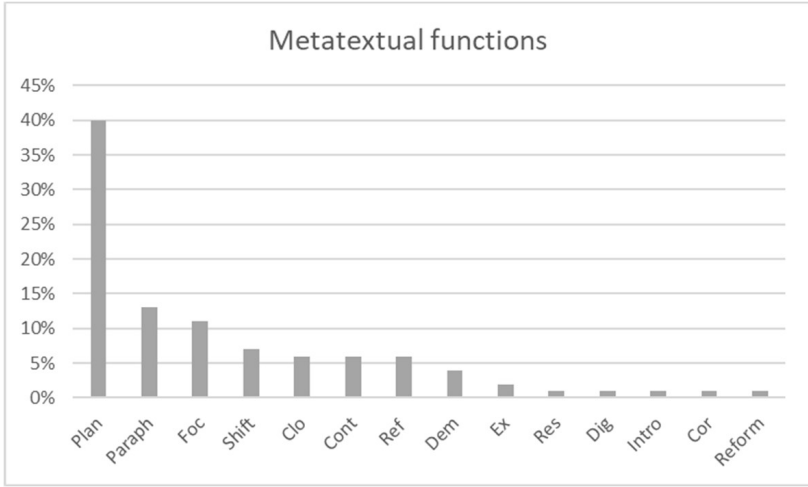


Figure 3. Metatextual functions (%)

In the cognitive dimension (see Figure 4), we observe the presence of all types of DMs considered, but *Inference (Inf)* and *Mitigation (Mit)* DMs prevail over *Intensification (Int)*.

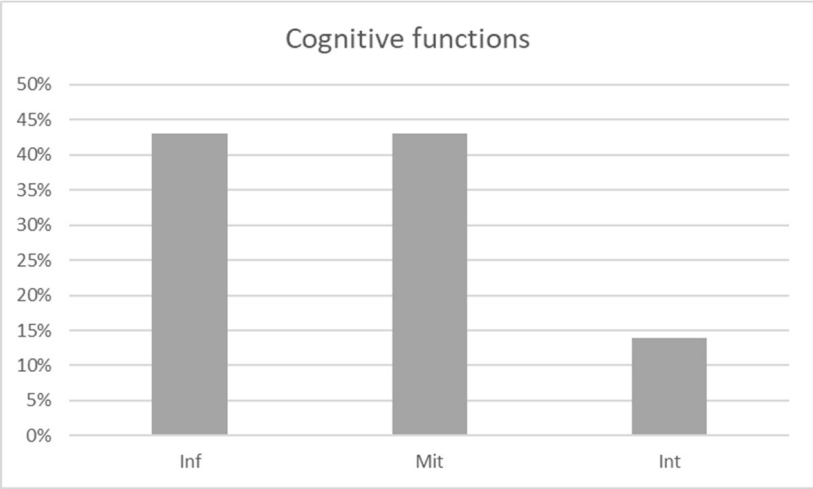


Figure 4. Cognitive functions (%)

4.3. Forms and paradigmatic polyfunctionality

Having provided an overview of the functions performed, we now focus on the forms they take. Table 4 shows the frequency list of the first 20 DMs found in the dataset (a total of 1759 tokens, 102 types) and the number of functions performed by each form. They account for 84 % of the total, while “other” refers to a total of 82 forms that occur less than 13 times. The first 20 DMs are not only very frequent, but also very polyfunctional: the number of functions performed by each form averages 5.65, indicating that relatively few forms may perform many different functions.

Table 4. Frequency list of the first 20 DMs in the oral corpus and their number of functions

	Rank	Number of occurrences	%	Number of functions
<i>quindi</i> (lit. 'therefore')	1	358	20 %	12
<i>appunto</i> (lit. 'actually')	2	304	17 %	12
<i>diciamo</i> (lit. 'let's say')	3	196	11 %	7
<i>sostanzialmente</i> (lit. 'basically')	4	120	7 %	6
<i>proprio</i> (lit. 'just')	5	75	4 %	4
<i>poi</i> (lit. 'then')	6	49	3 %	6
<i>insomma</i> (lit. 'in conclusion')	7	48	3 %	6
<i>assolutamente</i> (lit. 'absolutely')	8	48	3 %	3
<i>chiaramente</i> (lit. 'clearly')	9	43	2 %	5
<i>in realtà</i> (lit. 'actually')	10	42	2 %	10
<i>allora</i> (lit. 'then')	11	41	2 %	6
<i>cioè</i> (lit. 'namely')	12	25	1 %	4
<i>ovviamente</i> (lit. 'obviously')	13	21	1 %	3
<i>ecco</i> (lit. 'here, there, so, okay')	14	21	1 %	6
<i>praticamente</i> (lit. 'practically')	15	19	1 %	5
<i>anche</i> (lit. 'also')	16	19	1 %	4
<i>in un certo senso</i> (lit. 'in some way')	17	17	1 %	1
<i>come dire</i> (lit. 'how to say')	18	14	1 %	2
<i>comunque</i> (lit. 'anyway')	19	13	1 %	5
<i>per l'appunto</i> (lit. 'exactly')	20	13	1 %	5
other		299	16 %	

The distribution of these more frequent forms indicates that the first three DMs, *quindi*, *appunto* and *diciamo* account for 49 % of the total (see Table 4). On the other hand, if we look at less frequent DMs ("other" in Table 4), 90 % of these forms perform only one function.

Low semantic definition, which does not mean low information content or low interpretability, is described as the most functional strategy in face-to-face communication, since it allows for incremental progress in meaning construction (Voghera, 2017). This strategy is implemented in speech through the use of high-functional performance words or expressions with a wide range of meanings, i.e. those that show greater adaptability and flexibility in a greater number of contexts. Despite certain obvious differences, this preference also seems to manifest itself for DMs.

Their polyfunctionality makes them more employable for more purposes and in more contexts and thus encourages their greater frequency of use (see Degand *et al.*, 2022).

#### 4.4. Co-occurrences and syntagmatic polyfunctionality

As for syntagmatic polyfunctionality, that is to say how many DMs perform more than one function at the same time, we identify a value of 19 % in our dataset, with a marked prevalence of the association between metatextual and cognitive functions (17 %), while the other two combinations (metatextual M-F + interactional M-F and interactional M-F + cognitive M-F, only account for 1 % (see Figure 1). The most frequent functional associations are shown in Table 5.

Table 5. The most frequent combinations in syntagmatic polyfunctional DMs

Metatextual + Cognitive M-Fs	
<i>Plan + Mit</i>	53.1 %
<i>(Intro/Cont/Clo/Shift/Res) + Inf</i>	27.5 %
<i>Paraph + Inf</i>	10.5 %
other	8.9 %

*Plan + Mit* is the prevalent expression of syntagmatic polyfunctionality in our dataset: in this case the DM helps to manage the discourse flow by preventing occurrences of disfluencies and, at the same time, functions as an approximating device that reduces the precision and the adequacy of the lexical choice: for instance, *diciamo* in example (6) is simultaneously a speech management and an approximation device.

*Inf* DMs also function simultaneously as topic management. In (30), we see an instance of a *Clo* DM:

(30) [finge l'argento] <breath> [ma è cartapesta] <sp> [quindi anche in questo caso un materiale povero]

'[it fakes silver] <breath> [but it is papier mâché] <sp> [quindi (lit. 'therefore') again a poor material]'

Here the TG had presented the altar and emphasized the fact that, despite appearances, it is made of wood, namely a poor material. She then moves on to the angels on either side of the altar and concludes that, again, it is made of papier mâché, a poor material. In the context, the defining feature of the material used is whether or not it is a poor material. Hence, the deduction made explicit by the DM *quindi* also indicates a topic closure: we can conclude that all elements of the altar are made of poor materials.

Finally, *Inf* DMs also function simultaneously as *Paraph* DMs, as in the following example:

(31) [e questa poderosa struttura] [che viene ad essere costruita nello stesso tempo della prima certosia] <sp> [quindi sempre nel Milletrecento] (G03V02P01)

'[And this mighty structure] [which comes to be built at the same time as the first Carthusian monastery] <sp> [quindi (lit. 'therefore') still in the year 1300]'

Here a relationship of semantic identity is established between the reference segment and the reformulated segment, in the sense that the first Carthusian monastery had been built in the year 1300. At the same time, equivalence is also indicated by the DM as inferable information, made explicit by the reformulated segment.

As far as co-occurrences are concerned, consecutive co-occurring DMs account for 12 % of the total; 6 % of the DMs appear in *cumuli* (clusters of DMs, each one with different functions) and 6 % occur in *catene* (clusters of DMs, all performing the same function or functions).

*Cumuli* do not present restrictions with respect to the combination of functions. Indeed, their distribution does not show statistically significant trends ( $\chi^2 = 81.77$ ,  $df = 54$ ,  $p = 0.069$ ): each function may co-occur with other ones. However, combining both monofunctional and polyfunctional DMs, the functions that most often occur in *cumuli* are monofunctional DMs of *Plan* (12 %), *Inf* (10 %) and *M-S-K* (8 %), that is, they can be found in all three M-Fs. As regards the sequencing of monofunctional DMs, there are no rigid orders among the functions, but looking at the M-Fs, we can see that, when present, interactional DMs tend to appear first, then metatextual DMs, and then cognitive DMs. Figure 5 depicts functions in *cumuli*, without considering functions that have a 1 % occurrence rate (which collectively account for 9 % of the total).

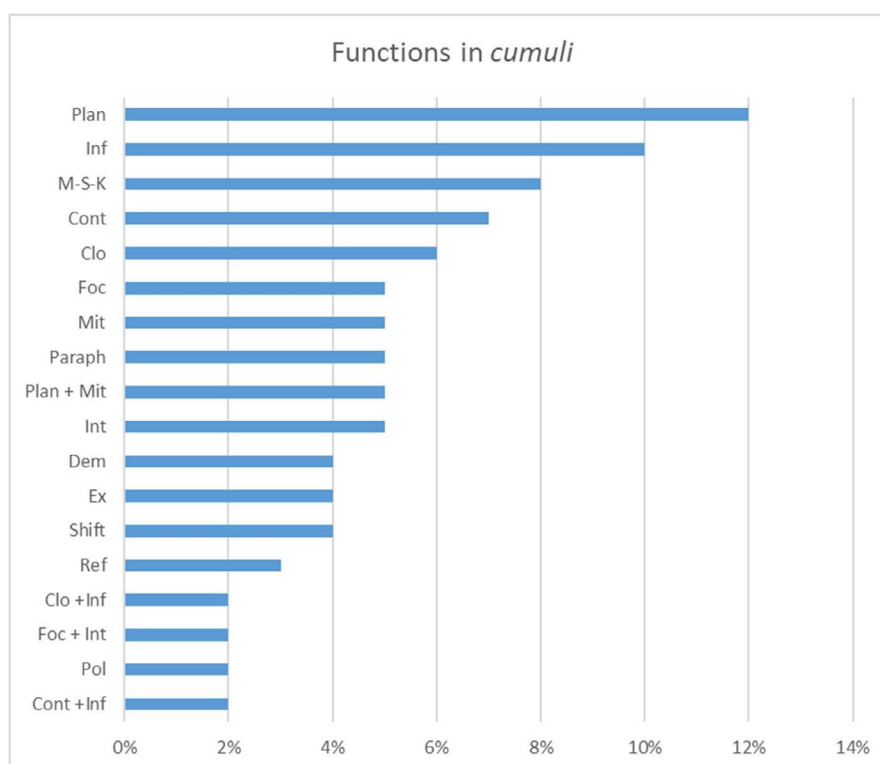


Figure 5: Most frequent functions of DMs in *cumuli*

Figure 5 points out that DMs occurring in *cumuli* are mainly monofunctional (85 %). Indeed, syntagmatic polyfunctionality in *cumuli* can take the form of several combinations, but none are very frequent.

*Catene* show a narrower range of functions than *cumuli* (see Figure 6) and present clearer trends: the functions most likely to appear in consecutive co-occurrences are *Focusing* and *Planning*, reaching statistical significance ( $\chi^2 = 206.74$ ,  $df = 54$ ,  $p = 0.0005$ ).

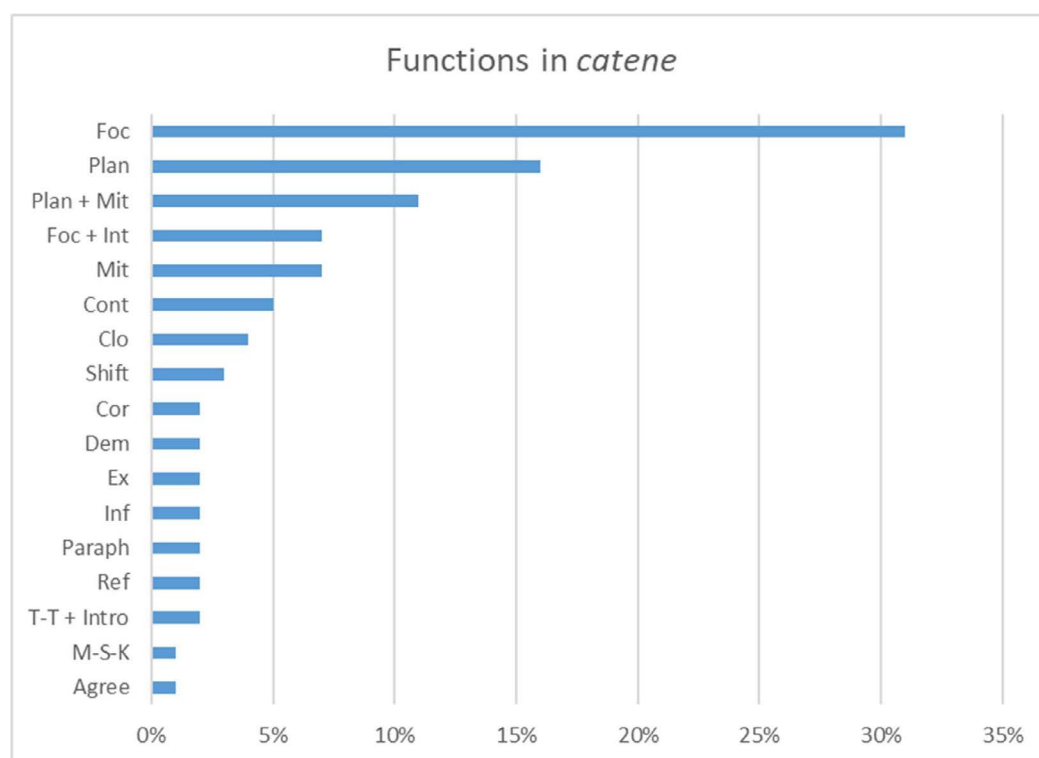


Figure 6. Functions of DMs in *catene*

Like *cumuli*, DMs occurring in *catene* are mainly monofunctional (80 %). In this kind of speech, the most frequent polyfunctional DMs in *catene* serve to plan and mitigate. Both *cumuli* and *catene* almost always consist of sequences of two elements and rarely of three DMs (see Alfano & Schettino, 2023).

## 5. Discussion and conclusions

Although the frequency and use of DMs certainly depend on each individual and their preferences, the conducted analysis highlights interesting and little examined aspects of TGs' speech and offers new data in the wide and varied world of discourse and pragmatic markers, which we have referred to as DMs.

The last thirty years of research into DMs from a variety of theoretical frameworks have produced a wealth of knowledge in many languages, including Italian. Yet, the large amount of research on this subject has not produced corpus-based studies on Italian with an onomasiological approach. Therefore, we cannot compare our results on functions and forms with those of other studies on Italian speech. The overall frequency indicates a pervasive use of DMs in our dataset (frequency: 57.80 ptw), comparable to that found for other languages in different communicative situations. Adopting a more inclusive definition of DMs than ours, Crible (2018: 82) reports about 60 DMs ptw in French and 49 in English, with an important variation across registers: the highest relative frequency in conversational genres (private conversations and phone calls, around 70 DMs ptw overall), closely followed by interviews (67 DMs ptw) and, to a lesser extent, classroom lessons and radio interviews (54 DMs ptw). Despite some differences between French and English, the incidence of DMs is related, albeit not deterministically, to the degree of preparation and interactivity: the lower the degree of preparation, the greater the attention for self-monitoring, resulting in an increase in interruptions, reformulations and speech-supporting devices; on the other hand, the more interactive the setting, the higher the number of DMs used.

Monofunctional and polyfunctional metatextual DMs are the most frequent in our dataset, indicating that the TG's primary need is to organize the discourse and express cognitive intentions regarding text management and structure in order to inform and entertain the audience. In communicative discourse, the TG needs to negotiate meaning with the audience so as to orient attention and strengthen a positive image of the site. Planning DMs prevail over the others. They play a vital role in ensuring the fluency of speech and, by avoiding too many disfluencies, contribute to the TG's self-confidence and the audience's perception of that confidence. It is obviously quite common for explanatory reformulation, focalization and for topic management to be introduced by DMs: in this interaction between experts and the general public, the specialized language, cleansed of excessive technicalities, needs to be constantly explained and reformulated (Calvi, 2005).

With regard to the interactional M-F, our data indicate that interactional sensitivity pervades the work of TGs, and that actions are crafted so as to seek and ensure the audience's empathy, engagement and attention by leveraging on (presumed) shared knowledge, fostering interest and encouraging orientation towards the element under consideration. However, in order to better grasp the role played by DMs in this kind of speech, we think it is essential to stress that the so-called cognitive DMs are used by the TG not only to express inference markers and logical argumentative connection, but also to mark subjectivity and commitment and to modulate propositional content and illocutionary force. Finally, intensifiers are mainly used to emphasize an appeal to the visitors' emotions (Rosypalová, 2012).



With regard to syntagmatic polyfunctionality, we have seen that almost one in 5 DMs performs more than one function simultaneously. More interestingly, looking at paradigmatic polyfunctionality, the average of the functions performed by each form is 5.65 and ranges widely from 1 or 2 more conventionalized functions of forms such as *in un certo senso* (lit. 'in some way') (*Mit*) or *come dire* (lit. 'how to say') (*Plan* and *Mit*) to 12 functions of the very polyfunctional *quindi* (lit. 'therefore') (*Inf*, *Cont*, *Ref*, *Clo*, *Shift*, *Res*, *Intro*, *Dig*, *Paraph*, *Ex*, *Ref*, *T-T*) and *appunto* (lit. 'actually') (*Plan*, *Dem*, *Foc*, *Clo*, *Shift*, *Res*, *Paraph*, *Ex*, *Ref*, *Inf*, *Int*), which cover an exceptionally wide functional space. As the total number of forms amounts to 102 and the first 5 alone (*quindi*, lit. 'therefore'; *appunto*, lit. 'actually'; *diciamo*, lit. 'let's say'; *sostanzialmente*, lit. 'basically'; *proprio*, lit. 'just') account for almost half of the occurrences, we have a measure of the extraordinary plasticity and extensibility of the functions. As we have seen, frequency is related to polyfunctionality. Similarly to what happens at the lexical level in the choice of high-functional performance words or expressions with a wide range of meanings, i.e. those that show greater adaptability and flexibility in a greater number of contexts, the choice of DMs in this speech also falls on the most polyfunctional forms. The TG often has to interrupt the talk to show something to the audience. In this talking and showing process, the TG needs a reduced set of DMs to constantly pick up the threads. Moreover, the most frequently used forms also seem to be characterized by a fairly good degree of formality on the diaphasic dimension, and are not diastatically or diatopically marked, in the sense that on the diaphasic and diastatic axes they would be located in the center of the linguistic repertoire (Berruto, 2012[1987]) and on the diatopic axis, they are not regionally connotated. There are few studies on Italian diatopically marked DMs, but we can certainly say that the forms found to be most frequent in this study do not coincide with any of those identified for Neapolitan Italian (D'Aguanno, 2019) and are not immediately traceable to any regionalism. In our opinion, this absence of markedness on the axes of variation favors their exploitability.

As far as co-occurrences of DMs are concerned, previous studies indicate that this phenomenon is relatively frequent. Analyzing categorical combinations of DMs in Catalan and Spanish oral narratives, Cuenca & Marín (2009) find around 17 % of all DMs occurrences. Combinations of DMs occur in our corpus in a similar proportion: they account for 12 % of the total. A significant number of *Focusing* and *Planning* functions were found for *catene* (as pointed out, a.o., by Bazzanella, 2011), while *cumuli* in our dataset do not present restrictions or correlations between possible or acceptable functions, nor with respect to the order in which they occur. However, looking at M-Fs, our data follow the expected trend of the "DM Sequencing Hypothesis" (Haselow, 2019): monofunctional *cumuli* in initial position (for the relation between functions and positions, see Alfano & Schettino, 2023) tend to appear in the sequence interactional>metatextual>cognitive DMs. Moreover, all types of co-occurrences appear in most cases as sequences of two elements and rarely of three DMs, as stated for different languages (Pons, 2018b; Cuenca & Crible, 2019) and between monofunctional metatextual DMs, but the result is in line with their general prevalence in the dataset.

Although the specific forms are likely to vary from speaker to speaker, we believe that the identified features of incidence, distribution, functional variability, plasticity in forms, possibility of co-occurrence in clusters, paradigmatic and syntagmatic polyfunctionality do characterize this specific register. More precisely, while results on incidence and co-occurrence are overall comparable to those obtained for other types of speech in different communicative situations, a characteristic pattern of DMs in this type of speech emerges clearly: discourse planning benefits from a high frequency of relatively few DMs endowed with wide and flexible functional variation. Whether and to what extent this narrow polyfunctional pattern is also found in other types of speech in similar communicative conditions remains to be further investigated in future research.

## References

- Aijmer, Karin (2002). *English discourse particles. Evidence from corpora*. John Benjamins. <https://doi.org/10.1075/scl.10>
- Aijmer, Karin, Foolen, Ad, Simon-Vandenberghe & Anne-Marie (2006). Pragmatic Markers in Translation: a Methodological Proposal. In Fischer, K. (Ed.): *Approaches to Discourse Particles*. Amsterdam: Elsevier, 101-114. [https://doi.org/10.1163/9780080461588\\_007](https://doi.org/10.1163/9780080461588_007)
- Alfano, Iolanda, Cataldo, Violetta, Orrico, Riccardo, & Schettino, Loredana (2021). Sentence topics in Italian: An analysis on the CHROME Corpus. *Loquens*, 8(1-2), e083. <https://doi.org/10.3989/loquens.2021.083>
- Alfano, Iolanda, Cataldo, Violetta, & Schettino, Loredana (2023). CHROME Corpus: Multi-level annotation of an audiovisual corpus. *CHIMERA: Revista De Corpus De Lenguas Romances Y Estudios Lingüísticos*, 10, 135-153. <https://revistas.uam.es/chimera/article/view/16013>
- Alfano, Iolanda, Schettino, Loredana (2023). Segnali discorsivi in italiano. Funzioni e posizioni. *Cuadernos de Filología italiana*, 30, 207-227. <https://doi.org/10.5209/cfit.84010>
- Antelmi, Donatella (2022). *Il turismo come discorso. Generi e testi dal racconto al Web*. Roma: Dino Audino.
- Bazzanella, Carla (1995). I segnali discorsivi. In L. Renzi, G. Salvi & A. Cardinaletti (Eds.), *Grande grammatica italiana di consultazione*. Il Mulino, 225-257.
- Bazzanella, Carla (2006a). Discourse markers in Italian: towards a "compositional" meaning. In Fischer, K. (Ed.): *Approaches to Discourse Particles*. Amsterdam: Elsevier, 449-464. [https://doi.org/10.1163/9780080461588\\_024](https://doi.org/10.1163/9780080461588_024)
- Bazzanella, Carla (2006b). Segnali discorsivi e sviluppi conversazionali. In F. Albano Leoni & R. Giordano (Eds.), *Italiano parlato. Analisi di un dialogo*. Liguori, 137-158.
- Bazzanella, Carla (2011). Segnali discorsivi. In *Enciclopedia dell'italiano, Istituto della Enciclopedia Italiana*. [https://www.treccani.it/enciclopedia/segnali-discorsivi\\_\(Enciclopedia-dell%27italiano\)](https://www.treccani.it/enciclopedia/segnali-discorsivi_(Enciclopedia-dell%27italiano))

- Bazzanella, Carla & Borreguero Zuloaga, Margarita (2011). 'Allora' e 'entonces': problemi teorici e dati empirici. In Elisaveta Khachatryan (Ed.), *Discourse markers in Romance languages*, *Oslo Studies in Language*, 3 (1), 7-45. <https://doi.org/10.5617/osla.181>
- Bazzanella, Carla, Bosco, Cristina, Garcea, Alessandro, Gili Fivela, Barbara, Miecznikowski, Johanna & Tini Brunozzi, Francesca (2007). Italian allora, French alors: functions, convergences and divergences. *Catalan Journal of Linguistics*, 6, 9-30. <https://doi.org/10.5565/rev/catjl.122>
- Beeching, Kate & Detges, Ulrich (Eds.) (2014). *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*. Brill. <https://doi.org/10.1163/9789004274822>
- Berruto, Gaetano 2012 [1987]. *Sociolinguistica dell'italiano contemporaneo*. 2nd edn. Roma: Carocci (1987: Roma: La Nuova Italia Scientifica).
- Best, Katie (2012). Making museum tours better: Understanding what a guided tour really is and what a tour guide really does. *Museum Management and Curatorship*, 27(1), 35-52. <https://doi.org/10.1080/09647775.2012.644695>
- Bhatia, Vijay K. (2008). Genre analysis, ESP and professional practice. *English for Specific Purposes*, 27(2), 161-174. <https://doi.org/10.1016/j.esp.2007.07.005>
- Bhatia, Vijay K. (2010). Interdiscursivity in professional communication. *Discourse & Communication* 21(1), 32-50. <https://doi.org/10.1177/1750481309351208>
- Blakemore, Diane (2002). *Relevance and Linguistic Meaning. The Semantics and Pragmatics of Discourse Markers*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511486456>
- Boersma, P. & Weenink. 2019. Praat: doing phonetics by computer [Computer program]. Version 6.1.08. <http://www.praat.org/> (accessed December 5, 2019)
- Borreguero Zuloaga, Margarita (2011). La traducción de los marcadores del discurso. Valores, funciones, posiciones y otros problemas. In D. Sáez Rivera et al. (Eds.) *Últimas tendencias en traducción e interpretación*. Frankfurt: Iberoamericana/Vervuert, 123-40. <https://doi.org/10.31819/9783865278777-009>
- Borreguero Zuloaga, Margarita (2015). A vueltas con los marcadores del discurso: de nuevo sobre su delimitación y sus funciones. In A. Ferrari & L. Lala (Eds.), *Testualità. Fondamenti, unità, relazioni*. Franco Cesati, 151-170.
- Boyer, Marc & Viallon, Philippe (2000). *La comunicazione turistica* [Italian translation by A. Pedon. Roma: Armando Editore].
- Brinton, Laurel J. (2017). *The Evolution of Pragmatic Markers in English: Pathways of Change*. Cambridge University Press. <https://doi.org/10.1017/9781316416013>
- Briz, Antonio (1998). *El español coloquial*. Ariel.
- Broth, Mathias & Mondada, Lorenza (2013) Walking away. The embodied achievement of activity closings in mobile interactions. *Journal of Pragmatics* 47: 41-58. <https://doi.org/10.1016/j.pragma.2012.11.016>
- Caffi, Claudia (2007). *Mitigation*. Amsterdam: Elsevier. <https://doi.org/10.1163/9780080466224>
- Calvi, Maria Vittoria (2005). Il linguaggio spagnolo del turismo. Viareggio: Baroni Editore.
- Calvi, Maria Vittoria (2006). *Lengua y comunicación en el español del turismo*. Madrid: Arco/Libros.
- Calvi, Maria Vittoria (2010). Los géneros discursivos en la lengua del turismo: una propuesta de clasificación. *Ibérica* 19, 9-32.
- Calvi, Maria Vittoria (2017). The Language of Tourism in New Travel Guides: Discursive Identities and Narratives. In M. Gotti, S. Maci & M. Sala (Eds.), *Ways of Seeing, Ways of Being: Representing the Voices of Tourism*. Bern: Peter Lang.
- Calvi, Maria Vittoria & Bonomi, Milin (2008). El lenguaje del turismo de los textos especializados a la comunidad del viajero. In C. Navarro, F. Dalle Pezze, R. Miotti & R.M. Rodríguez (Eds.), *La comunicación especializada: terminología y traducción*, 181-202. Bern: Peter Lang, 139-156.
- Calvi, Maria Vittoria & Mapelli, Giovanna (Eds.) (2011). *La lengua del turismo. Géneros discursivos y terminología*. Bern: Peter Lang.
- Cataldo, Violetta, Schettino, Loredana, Savy, Renata, Poggi, Isabella, Origlia, Antonio, Ansani, Alessandro, Sessa, Isora & Chiera, Alessandra (2019). Phonetic and functional features of pauses, and concurrent gestures, in tourist guides' speech. In D. Piccardi, F. Ardolino & S. Calamai (Eds.), *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale*, 205-231. Milano: Officinaventuno.
- Cella, Roberta (2015). *Storia dell'italiano*. il Mulino.
- Crible, Ludivine (2018). *Discourse Markers and (Dis)fluency. Forms and functions across languages and registers*. John Benjamins. <https://doi.org/10.1075/pbns.286>
- Crible, Ludivine & Degand, Liesbeth (2021). Co-occurrence and ordering of discourse markers in sequences: A multifactorial study in spoken French, *Journal of Pragmatics*, 177, 18-28. <https://doi.org/10.1016/j.pragma.2021.02.006>
- Cuenca Ordiñana, Maria Josep & Crible, Ludivine (2019). "Co-occurrence of discourse markers in English: from juxtaposition to composition", *Journal of Pragmatics*, 140:2, 171-184.
- Cuenca Ordiñana, Maria Josep & Marín, Maria Josep (2009). "Co-occurrence of discourse markers in Catalan and Spanish oral narratives", *Journal of Pragmatics*, 41:5, 899-914.
- Cutugno, Francesco & Campisi, Emanuela (2022). Linguistica applicata alle esperienze di cultura digitale. Analisi multimodale nel progetto CHROME. In J. Saturno & L. Spreafico (Eds.), *Fare linguistica applicata con le digital humanities*. *Studi AltLA* 14. Milano: Officinaventuno, 9-25.
- D'Aguanno, Daniele (2019). Su alcuni tratti pragmatici dell'italiano di Napoli. *Annali-sezione romanza LXI*, 1, 93-120.

- Dal Negro, Silvia & Fiorentini, Ilaria, (2014). Reformulation in bilingual speech: Italian *cioè* in German and Latin. *Journal of Pragmatics*, 74, 94-108. <https://doi.org/10.1016/j.pragma.2014.09.002>
- Dann, Graham M.S. (1996). *The Language of Tourism. A Sociolinguistic Perspective*. Oxon: CAB International. <https://doi.org/10.1079/9780851989990.0000>
- D'Achille, Paolo (2016). Rispondiamo a quelli che sono i quesiti più frequenti. <https://accademiadellacrusca.it/it/consulenza/rispondiamo-a-uno-di-quelli-che-sono-i-quesiti-pi%C3%B9-frequenti/1034>
- De Cesare, Anna Maria (2000). Sulla semantica di alcuni tipi di intensificazione in italiano: „Davvero, è proprio molto interessante!“ *Romanistisches Jahrbuch*, 51(1), 87-107. <https://doi.org/10.1515/roja-2000-0105>
- De Cesare, Anna Maria (2010). On the focusing function of focusing adverbs. A discussion based on Italian data. *Linguistik online*, 44, 4/10, 99-116. <https://doi.org/10.13092/lo.44.406>
- De Cesare, Anna Maria (2022). To be or not to be focus adverbs? A corpus-driven study of It. *anche* in spontaneous spoken Italian. In D. Garassino & D. Jacob (Eds.), *When Data Challenges Theory: Unexpected and paradoxical evidence in information structure*. John Benjamins, 203-238. <https://doi.org/10.1075/la.273.07de>
- De Cristofaro, Elisa, Crocco, Claudia, Badan, Linda & Plevoets, Koen (2022). Discourse markers and turn-planning at the pragmatics-prosody interface: The case of *allora* in spoken Italian. *Journal of Pragmatics*, 198, 125-141. <https://doi.org/10.1016/j.pragma.2022.06.012>
- Degand, Liesbeth, Broisson, Zoé, Crible, Ludivine & Grzech, Karolina (2022). Cross-linguistic variation in spoken discourse markers: distribution, functions, and domains. In *Discourse-pragmatic variation and change: theory, innovations, contact*, 83-104.
- Degand, Liesbeth, Cornillie, Bert & Pietrandrea, Paola (2013). *Discourse Markers and Modal Particles: Categorization and Description*. John Benjamins. <https://doi.org/10.1075/pbns.234>
- Degand, Liesbeth & Evers-Vermeul, Jacqueline (2015). Grammaticalization or pragmaticalization of discourse markers?: more than a terminological issue. *Journal of Historical Pragmatics* 16 (1), 59-85. <https://doi.org/10.1075/jhp.16.1.03deg>
- Degand, Liesbeth & Simon, Anne C. (2009). On identifying basic discourse units in speech: theoretical and empirical issues. *Discours* 4. <http://discours.revues.org/index5852.html> (accessed January 7, 2022). <https://doi.org/10.4000/discours.5852>
- Denti, Olga (2021). Building the tourist destination through multiple semiotic modes. *International Journal of English Linguistics* 11(4), 83-104.
- Dér, Csilla I. (2010). On the status of discourse markers. *Acta Linguistica Hungarica* 57 (1), 3-38. <https://doi.org/10.1556/aling.57.2010.1.1>
- Elia, Antonella (2018). Il linguaggio del Turismo: *WikiVoyage* e l'evoluzione delle guide turistiche online. *Trakya Üniversitesi Edebiyat Fakültesi Dergisi*, 8 (15), 119-155.
- Flores Acuña, Estefanía (2009). La reformulación del discurso en español en comparación con el italiano. Estudio contrastivo de los marcadores de reformulación. In Garcés Gómez, M.P. (Ed.), *La reformulación del discurso en español en comparación con otras lenguas*. Universidad Carlos III y Boletín Oficial de Estado, Madrid, 93-136.
- Francesconi, Sabrina (2014). *Reading Tourism Texts. A Multimodal Analysis*. Bristol: Channel View Publications.
- Fraser, Bruce (1999). What are discourse markers? *Journal of Pragmatics*, 31 (7), 931e952. [https://doi.org/10.1016/s0378-2166\(98\)00101-5](https://doi.org/10.1016/s0378-2166(98)00101-5)
- Fraser, Bruce (2009). Topic orientation markers. *Journal of Pragmatics*, 41, 892-898. <https://doi.org/10.1016/j.pragma.2008.08.006>
- Fraser, Bruce (2015). The combining of Discourse Markers – A beginning, *Journal of Pragmatics*, 86, 48-53. <https://doi.org/10.1016/j.pragma.2015.06.007>
- Garcés Gómez, María del Pilar (2008). *La organización del discurso: marcadores de ordenación y de reformulación*. Iberoamericana-Vervuert. <https://doi.org/10.31819/9783865278661>
- Ghezzi, Chiara & Molinelli, Piera (Eds.) (2014). *Discourse and Pragmatic Markers from Latin to the Romance Languages: New insights*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199681600.003.0001>
- Ghezzi, Chiara & Molinelli, Piera (2020). Connectives and cyclicity: from the Latin temporal phrase *illa hora* to the Italian discourse marker *allora*. *Journal of Historical Pragmatics*, 21 (2), 208-235. <https://doi.org/10.1075/jhp.00042.ghe>
- Giménez Folqués, David (2015). Anglicisms in Tourism Language Corpora 2.0. *Procedia – Social and Behavioral Sciences* 198, 149-156. <https://doi.org/10.1016/j.sbspro.2015.07.431>
- Gotti, Maurizio (2003). *Specialized Discourse: Linguistic Features and Changing Conventions*. Bern: Peter Lang.
- Gotti, Maurizio (2006). The Language of Tourism as Specialized Discourse. In O. Palusci & S. Francesconi (Eds.), *Translating Tourism: Linguistic/Cultural Representations*, 15-34. Trento: Università di Trento.
- Gotti, Maurizio, Maci Stefania & Sala, Michele (Eds.) (2017). *Ways of Seeing, Ways of Being: Representing the Voices of Tourism*. Bern: Peter Lang.
- Güllich, Elisabeth & Kotschi, Thomas (1983). Les marqueurs de reformulation paraphrastique. *Cahiers de Linguistique Française*, V, 305-351.
- Haselow, Alexander (2019). Discourse marker sequences: Insights into the serial order of communicative tasks in real-time production. *Journal of Pragmatics* 146, 1-18. <https://doi.org/10.1016/j.pragma.2019.04.003>
- Heine, Bernd, Kaltenböck, Gunther, Kuteva, Tania Long, Haiping. 2017. Cooptation as a discourse strategy. *Linguistics* 55: 1-43. <https://doi.org/10.1515/ling-2017-0012>



- Heine, Bernd, Kaltenböck, Gunther, Kuteva, Tania & Long, Haiping (2021). *The rise of discourse markers*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108982856>
- López Serena, Araceli & Borreguero Zuloaga, Margarita (2010). Los marcadores del discurso y la variación lengua hablada vs. Lengua escrita. In Ó. Loureda & E. Acín (Eds.), *Los estudios sobre marcadores del discurso en español, hoy*. Arco Libros, 415-495.
- Maci, Stefania (2010). *The Language of Tourism*. Bergamo: CELSB.
- Maci, Stefania (2020). *English Tourism Discourse. Insights into the professional, promotional and digital language of tourism*. Milano: Hoepli.
- Maci, Stefania, Sala, Michele & Godnič Vičič, Šarolta (Eds.) (2018). The Language of Tourism: An Introduction to the Topical Issue. *Scripta Manent* 12.
- Mapelli, Giovanna (2013). El léxico de las guías descriptivas. In L. Chierichetti & G. Garofalo (Eds.), *Discurso profesional y Lingüística de corpus. Perspectivas de investigación*. Cerlis 3. Bergamo: CELSB Libreria Universitaria, 125-138.
- Mauri, Elena (2014). L'italiano delle guide turistiche: Uno studio della lingua orale del turismo. *Italiano LinguaDue* 6 (1), 232-248. <https://doi.org/10.13130/2037-3597/4234>
- Molinelli, Piera (2014). Orientarsi nel discorso: segnali discorsivi e segnali pragmatici in italiano. In *Discorso e cultura nella lingua e nella letteratura italiana*. Franco Cesati.
- Mondada, Lorenza (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics* 41: 1977-1997. <https://doi.org/10.1016/j.pragma.2008.09.019>
- Mondada, Lorenza (2014a) Bodies in action: Multimodal analysis of walking and talking. *Language and Dialogue* 4(3): 357-403. <https://doi.org/10.1075/ld.4.3.02mon>
- Mondada, Lorenza (2014b) Pointing, talk and the bodies: Reference and joint attention as embodied interactional achievements. In Seyfeddinipur M., Gullberg M. (Eds.) *From Gesture in Conversation to Visible Utterance in Action*. Amsterdam: Benjamins, 95-124. <https://doi.org/10.1075/z.188.06mon>
- Mondada, Lorenza (2017). Walking and talking together: Questions/answers and mobile participation in guided visits. *Social Science Information* 2017, Vol. 56(2) 220-253. <https://doi.org/10.1177/0539018417694777>
- Mondada Lorenza (2018). Questions on the move. The ecology and temporality of question/answers in mobility settings. In A. Deppermann & J. Streeck (Eds.), *Time in Embodied Interaction. Synchronicity and sequentiality of multimodal resources*. Amsterdam: Benjamins, Pragmatics & Beyond New Series 293, 161-202. <https://doi.org/10.1075/pbns.293.05mon>
- Nigro, Maria Giovanna (2006). *Il linguaggio specialistico del turismo. Aspetti storici, teorici e traduttivi*. Roma: Aracne. <https://doi.org/10.24310/entreculturasertci.vi9.11299>
- Origlia, Antonio, Savy, Renata, Poggi, Isabella, Cutugno, Francesco, Alfano, Iolanda, D'Errico, Francesca, Vincze, Laura, & Cataldo, Violetta (2018). An Audiovisual Corpus of Guided Tours in Cultural Sites: Data Collection Protocols in the CHROME Project. In *Proceedings of the 2018 AVI-CH Workshop on Advanced Visual Interfaces for Cultural Heritage*, 1-4.
- Pierini, Patrizia (2009). Adjectives in Tourism English on the Web: A Corpus-based Study. *CLAC. Círculo de Lingüística Aplicada a la Comunicación* 40, 93-116.
- Pons Bordería, Salvador (2013). Un solo tipo de reformulación. *Cuadernos AISPI*, 2, 151-170.
- Pons Bordería, Salvador (2018a). Paths of grammaticalization: beyond the LP/RP debate. In S. Pons Bordería & Ó. Loureda (Eds.), *Beyond Grammaticalization and Discourse Markers: New Issues in the Study of Language Change*, Leiden/Boston, Brill, 334-383. [https://doi.org/10.1163/9789004375420\\_012](https://doi.org/10.1163/9789004375420_012)
- Pons Bordería, Salvador (2018b). The combination of discourse markers in spontaneous conversations. Keys to untying a Gordian knot. *Rev. Romane* 53 (1), 121-158. <https://doi.org/10.1075/rro.00008.pon>
- Pons Bordería, Salvador & Fischer, Kerstin (2021). Using discourse segmentation to account for the polyfunctionality of discourse markers: The case of *well*. *Journal of Pragmatics*, 173, 101-118. <https://doi.org/10.1016/j.pragma.2020.11.021>
- Raso, Tommaso & Vieira, Marcelo A., (2016). A description of Dialogic Units/Discourse Markers in spontaneous speech corpora based on phonetic parameters. *Chimera: Romance Corpora and Linguistic Studies* 3 (2), 221-249.
- Rosypalová, Jana (2012). *Language of Tourism: Corpus Analysis of Spoken Discourse of Tour Guides and Written Discourse of Tourist Guides*. LAP LAMBERT Academic Publishing.
- Salameh Jiménez, Shima, Estellés Arguedas, María & Pons Bordería, Salvador (2018). Beyond the notion of periphery: An account of polyfunctional discourse markers within the Val.Es.Co. model of discourse segmentation. In K. Beeching, P. Molinelli & C. Ghezzi (Eds.), *Positioning the self and others: Linguistic traces*, Amsterdam/Philadelphia, Benjamins. <https://doi.org/10.1075/pbns.292.05sal>
- Sansò, Andrea (2020). *I segnali discorsivi*. Carocci.
- Sansò, Andrea & Fiorentini, Ilaria (2017). Reformulation markers and their functions: Two case studies from Italian. *Journal of Pragmatics*, 120, 54-72. <https://doi.org/10.1016/j.pragma.2017.08.010>
- Santulli, Francesca, Antelmi, Donatella & Held, Gudrun (2009). *Pragmatica della comunicazione turistica*. Roma: Editori Riuniti.
- Savy, Renata (2005). Specifiche per la trascrizione ortografica annotata dei testi. *Italiano Parlato. Analisi di un dialogo*. Napoli: Liguori [CD-ROM].
- Schettino, Loredana, Betz, Simon, Cutugno, Francesco & Wagner, Petra (in press). Hesitations and individual variability in Italian tourist guides' speech. In *Atti del XVII Convegno Nazionale dell'Associazione Italiana di Scienze della Voce*.
- Schiffrin, Deborah (1987). *Discourse Markers*. Cambridge University Press, Cambridge.



- Schubotz, Louise, Oostdijk, Nelleke & Ernestus, Mirjam, (2015). Y'know vs. you know: what phonetic reduction can tell us about pragmatic function. In S. Lestrade, P.J.F. Swart & L. Hogeweg (Eds.), *Addenda: Artikelen voor Ad Foolen*. Radboud University, 361-380.
- Schourup, Lawrence (1999). Tutorial overview: Discourse markers. *Lingua* 107, 227-265. [https://doi.org/10.1016/S0024-3841\(96\)90026-1](https://doi.org/10.1016/S0024-3841(96)90026-1)
- Sloetjes, Han & Wittenburg, Peter (2008). Annotation by category-ELAN and ISO DCR. In 6th international Conference on Language Resources and Evaluation (LREC 2008). Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands. <https://tla.mpi.nl/tools/tla-tools/elan/>
- Solsona Martínez, Carmen (2014). Valores metatextuales de “cioè” en el marco de la gramática contrastiva italiano/español: de la reformulación a la conexión deductiva. *Philologia Hispalensis*, 28(2). <https://doi.org/10.12795/ph.2014.v28.i02.10>
- Sperber, Dan & Wilson, Deirdre (1986). *Relevance. Communication and cognition*. Blackwell.
- Traugott, Elisabeth (2012). Intersubjectification and clause periphery. *English Text Construction* 5, 1, 7-28. <https://doi.org/10.1075/etc.5.1.02trau>
- Traugott, Elisabeth (2020). Is 'Back to my Point' a Pragmatic Marker? An Inquiry into the Historical Development of some Metatextual Discourse Management Markers in English. *Catalan Journal of Linguistics*, 13-29. <https://doi.org/10.5565/rev/catjl.307>
- Traugott, Elisabeth (2022). *Discourse structuring markers in English: A historical constructionalist perspective on pragmatics* (Constructional Approaches to Language 33). Amsterdam, Philadelphia: John Benjamins. <https://doi.org/10.1017/s1360674322000351>
- Voghera, Miriam (2017). *Dal parlato alla grammatica*. Carocci.
- Waltereit, Richard (2006). The rise of discourse markers in Italian: A specific type of language change. In Fischer, K. (Ed.): *Approaches to Discourse Particles*. Amsterdam: Elsevier, 61-76. [https://doi.org/10.1163/9780080461588\\_005](https://doi.org/10.1163/9780080461588_005)
- Yanoshevsky, Galia (2021). Introduction: les discours du tourisme, un objet privilégié pour l'analyse du discours, *Argumentation et Analyse du Discours* [Online], 27. <https://doi.org/10.4000/aad.5425>
- Wu, Yaru, Hutin, Mathilde, Vasilescu, Ioana, Lamel, Lori, Adda-Decker, Martine & Degand, Liesbeth (2021). *Fine phonetic details for discourse marker disambiguation: a corpus-based investigation*. DiSS 2021, the 10th Workshop on Disfluency in Spontaneous Speech, Aug 2021, Saint-Denis, France, 23-28.