

Corinne Rossari and Claudia Ricci

3 The semantics and pragmatics of discourse markers

Abstract: This chapter deals with the semantic-pragmatic characterization of discourse markers. This long-debated question is first addressed by an overview of their classification within linguistic theories, focusing on the delimitation of the class and on the nature of their contribution, especially the articulation between their semantic and pragmatic meaning. We then present our own account proposing an analysis of a subclass of DMs – connectors – in French and Italian. To this aim, we adopt a textometric approach, combining qualitative observation with statistical calculations determining the preferred environment of connectors, i.e. the forms associated with them in a statistically significant way. This method reveals for instance similarities between connectors expressing different links, which can be interpreted as resulting from a common semantic trait, the involvement of the speaker in her utterance. These would remain undetectable within a qualitative approach.

Keywords: discourse markers, connectors, meaning, textometric approach, speaker's involvement

1 Introduction

As a category whose members have as their main defining commonality their dependence on the situation of utterance, discourse markers (DMs) have been a privileged object of pragmatic studies. That is not to say that all the other levels of linguistic analysis do not come into play in their characterization, as it has indeed been shown in the literature: prosody (for DMs' property of delimiting illocutionary units, see among others Moneglia/Cresti 2006; Raso 2014; Raso/Vieira 2016; Cresti 2018; Cuenca 2020; for a corpus-based account of how prosody notably allows one to distinguish different meanings conveyed by DMs, ↗5 Discourse markers and prosody); prosody counterpart in written texts, i.e. punctuation (cf. Ferrari et al. 2008; Ferrari/Pecorari 2018); syntax-semantics (see, e.g., Bernard/Danlos 2019; Rossari/Javez 2000; Ferrari et al. 2008; De Santis et al. 2014; Ferrari/Pecorari 2018 for a distributional analysis); semantics-pragmatics, particularly within the theory of "Language-Inherent Argumentation" (Anscombre/Ducrot 1976; Ducrot et al. 1980). This brief and non-exhaustive overview shows that, although some DMs appear to function exclusively on an interactional level,

Acknowledgements: This article is a contribution to the research project *The argumentative patterns of modal forms. A corpus-based study comparing French and Italian* (n. 100001F_192247) funded by SNF.

they are subject to constraints applying at all levels of linguistic analysis. This, among other things, explains the difficulty for learners in grasping their use (for an account of DMs and second language acquisition, see Discourse markers and second language acquisition). DMs must therefore be viewed as linguistic objects with form and content forming part of the system of a language, even if they are often analyzed in combination with non-verbal manifestations of language such as gestures.

It is, however, undeniable that, while the classification of most other linguistic items can in theory be done on the basis of form, DMs hardly allow such a categorization, since this class is intrinsically constituted by elements of heterogeneous nature (verbs, adverbs, adjectives, etc., see Abeillé/Godard 2021, 1983–1993, for an overview of connectors). The feature for which DMs are distinguishable and recognizable as a class distinct from other classes of linguistic items is their function as markers of discourse management, whether on a textual level (pragmatic and argumentative) or on an interactional level (turn-taking management, backchannelling, etc.). As such, a semantic-pragmatic look at DMs is a fundamental prerequisite to any attempt at characterization, and all other levels of description (syntactic, morpho-lexical, prosodic) appear to serve to account for semantic-pragmatic features that would otherwise be difficult to delimit precisely.

Two main issues are at stake concerning the semantic-pragmatic characterization of DMs within linguistic theories: i. the delimitation of the class; ii. the nature of their contribution (that is, the articulation between their semantic and pragmatic meaning). This chapter will be organized around these key questions: first, we will review the terminology used to label the different forms relative to this class (section 2), before tackling the way in which linguistic theories address the two above-mentioned issues, with particular focus on Romance languages (section 3). We will then proceed to present our own account proposing an analysis of a subclass of DMs – connectors – within a textometric perspective, showing what such an approach – hardly used, so far, for characterizing DMs – brings to their description (section 4).

2 Labels and theoretical frameworks

Terminology in the field of discourse markers is, to say the least, not systematic, and it is characterized by a great deal of vacillation. *Discourse marker* is the most frequently used label in the English-speaking literature, although it is by far not the only one, and its description varies according to the different researchers. As appropriately pointed out by Fraser (1999, 932):

During the past ten years, the study of DMs has turned into a growth industry in linguistics, with dozens of articles appearing yearly. Unfortunately, the term has different meanings for different groups of researchers, and we find work on DMs done under a variety of labels including, but not limited to cue phrases [. . .], discourse connectors [. . .], discourse operators [. . .], discourse

particles [. . .], discourse signalling devices [. . .], phatic connectors [. . .], pragmatic connectors [. . .], pragmatic expressions [. . .], pragmatic formatives [. . .], pragmatic markers [. . .], pragmatic operators [. . .], pragmatic particles [. . .], semantic conjuncts [. . .], sentence connectors [. . .].

Being often only defined on the basis of their function, in relation either i. to their interactive properties or ii. to their relational properties in discourse, DMs are labelled within two different traditions: one distinguishing between DMs as “discourse signals” on the one hand and connectors on the other hand, on the basis of the above-mentioned interactive properties.; ii. one that does not make this distinction. According to the different theories and perspectives involved, the label *discourse marker* can thus be used in a narrow sense or in a broad sense. As for the narrow sense of DMs, as shown by the overview provided by Romero-Trillo (2006), traditionally, a part of the English-speaking literature on the subject assigns the label of DM exclusively, or for the most part, to particles having a conversational or ‘social interaction’ function, like phatics (for an overview, see Cuenca/Degand 2022), i.e. to items facilitating turn-taking and interaction. This narrow definition also typically concerns forms carrying the label “pragmatic markers” (Fraser 2009; Ghezzi/Molinelli 2014) or “(modal) particles” (Degand/Cornillie/Pietrandrea 2013) and rules out items having a textual or discursive function, usually labeled as “(discourse) connectors” or “logical-argumentative connectors”. Within a discourse perspective, Fraser (2009) describes DMs as fundamentally working on previous discourse, and thus on a (con)textual, not situational or interactional level. His definition therefore excludes DMs typical of conversation by only including items which correspond to the category generally labelled as connectors. This restrictive view on DMs leads Fraser to use two distinct labels: DMs for the items usually labeled as connectors and pragmatic markers for the items usually labelled as DMs within the above-mentioned conversational perspective. In this chapter, we will follow the leading tradition distinguishing DMs (Fraser’s pragmatic markers) from connectors (Fraser’s DMs), as this tradition is largely dominant both in Romance linguistics and in the literature on English (see for instance Hall 2007, who chooses the label of ‘discourse connectives’ for items like *but*). In this respect, our labelling is less comprehensive than the one by Remberger (2021), according to whom,

“[d]iscourse and pragmatic markers are functional units [. . .] that deictically relate text fragments, propositions, utterances, and discourse chunks to the context of speech. They manage the interaction of the discourse participants in the speech situation and facilitate successful communication. This group of functional units includes elements as diverse as discourse and pragmatic markers in the broad sense, illocutionary markers, sentence particles, modal particles, and connectors”.

The complexity of the labelling issue is also visible in the fact that there are linking items to which the label of connector is applied only in some instances of their use and not in others, depending on specific criteria. For instance, in Groupe Lambda-I (1975), *parce que* ‘because’ is analyzed either as an operator or as a connector according to its syntactic scope (working within one proposition or linking two propositions). The semantic scope is also a criterion for assigning an item the label of connector. For instance,

Hansen (1998), applying a distinction introduced by Sweetser (1990) between propositional level, epistemic level, and speech act level, considers the same conjunction (*because*) as being a connector only when it works on an epistemic and speech act level:

John married Sue *because* she was rich

Cecily couldn't have killed Algemon, *because* she was out of town on the night it happened

Have you got a minute? – *Because* there's something I wanted to discuss with you (Hansen 1998, 238)

A similar definitory criterion is proposed in Ferrari et al. (2008) for Italian connectors – also illustrating it by means of the causal conjunction, *perché* –, as well as in Debaisieux (2016) for *parce que*.

An additional difficulty lies in the fact that labels used in languages other than English, notably in French, Italian or Spanish, do not always necessarily cover the same categories as those referred to by the labels used in the English-speaking literature.

The label *mots du discours*, 'discourse words' introduced in the eponymous work by Ducrot et al. (1980), erases the distinction between DMs (in the narrow sense mentioned above) and connectors, putting on the same level of analysis elements which would definitely be considered as pertaining to two different categories within most frameworks: *mais* 'but' and *eh bien!* 'well!/'so!' are both considered by Ducrot et al. (1980) as *mots du discours* and analyzed within the same framework, that of Language-Inherent Argumentation; according to the traditional dichotomy, the former would be labelled as a connector and the latter as a DM in a narrow sense (or as a pragmatic marker according to Fraser's 2009 terminology). Other labels more or less cover the above-mentioned English labels: *marqueurs de discours* is usually used in the same sense as "discourse markers", sometimes including connectors (see, e.g., Fagard/Charolles 2018 on *d'ailleurs*, *par ailleurs* 'besides!/'incidentally', 'moreover'); as for the narrow sense of DMs, multiple labeling is at play: *marqueurs discursifs* 'discourse markers' (Dostie/Lefevre 2017), *signaux discursifs* 'discourse signals' (Swiggers 1995), *opérateurs discursifs* 'discourse operators' (Caron 1984), *particules du discours* 'particles of discourse' (Vincent 1991), *particules discursives* 'discourse particles' (Dargnat 2020), *particules énonciatives* 'enunciative particles' (Fernandez 1994; Paillard 2015 for a subclass of DMs), *signaux back-channel* 'back-channel signals' (Laforest 1992); *marqueurs illocutoires* 'illocutionary markers' (Roulet et al. 1987), *marqueurs d'interactivité* 'interactivity markers' (Spengler 1980; Rubattel 1981). Such diversity is the consequence of the theoretical frameworks in which these items are analyzed: for instance, an interactional perspective in Laforest's (1992) and Vincent's (1991) work; here, the focus is on back-channeling phenomena and a whole range of manifestations, including gestures, are labeled as back-channel signals. On the other hand, what is central in Roulet et al.'s (1987) perspective is the organization of discourse; as a result, a marker used in a dialogical way (for instance, to react, as in *d'accord* 'all right', *ok*. . .) is labeled *marqueur de fonction illocutoire* 'illocutionary function marker'; a marker used in a monological way (e.g. to

reformulate), takes on the label of *connecteur interactif* ('interactive connector'). Terminology for Italian usually distinguishes between *segnali discorsivi* ('discourse signals'), which, drawing on Bazzanella (1995; 2006), correspond to DMs in their dialogical/interactive function, and *connettivi (testuali)* '(textual) connectors', referring mostly to logic-argumentative connectors. An interesting example of the relationship between the semantics and pragmatics of connectors and their labelling can be found in the distinction between the label *congiunzione testuale* 'textual conjunction' as opposed to *congiunzione (frasale)* 'phrasal conjunction', which has gained prominence in the Italian terminology through its dissemination as a label of dictionary entries in the *Sabatini Coletti* Italian dictionary (Sabatini/Coletti 2008). This label corresponds to a precise usage differentiation for the same conjunction, reflecting distinctions which are also syntactic (e.g. the mood of the sentence in the scope of the conjunction). For example, a complex conjunction such as *per quanto* 'however much' is labeled a (phrasal) conjunction in the first example below and a textual conjunction in the second:

Per quanto gridasse, nessuno la senti
'however much she shouted, no one heard her'

Rinuncio a questa nomina, per quanto ne avrei avuto tutti i diritti
'I renounce this appointment, though I would have had every right thereto'.

This terminological vagueness reflects the vagueness related to class delineation. Indeed, the main characteristic of most theoretical descriptions of discourse markers is their broadness, since DMs are difficult to delimit on the basis of linguistic properties, be they grammatical, lexical or semantic. The description of their specific contribution to discourse is thus the most challenging aspect of their characterization.

Before addressing this issue, we will briefly remark on two points that systematically emerge from research on DMs in Romance.

The study of DMs in Romance languages abounds with contrastive, cross-linguistic analyses of different DMs or classes of DMs. Such a contrastive perspective is not only interesting *per se*, but also in that it goes hand in hand with the diachronic dimension in the analysis of this functional category, which is fundamental insofar as different semantic and pragmatic values of DMs have developed diachronically from a semantic core through processes of language change. The debate concerns mainly the nature of such change: whether it is the result of a grammaticalization process (cf. for instance Giacalone Ramat 2017, Visconti 2009 for Italian; González 2015 for Spanish and Catalan), or of a pragmatization process, as documented in Dostie (2004) and Hansen (2008) for French, or both (see, e.g., the analysis of French *comme*, Italian *come*, Portuguese *como*, and Spanish *como* by Mihatsch 2009, exploring their grammaticalization paths but also highlighting an incipient process of pragmatization for these markers). Waltereit (2006) defends a more radical view, excluding the idea that the rise of a DM is a case of grammaticalization, insofar as many of the original constructions giving rise to DMs are not dependent on sentence grammar (e.g. they do not need complements), and

thus produce non-obligatory items (a characteristic typical of DMs), differently from grammaticalization processes.

The intertwining of the diachronic study and the cross-linguistic characterization of DMs in different Romance languages is particularly challenging in that the evolution of a common source marker is not necessarily the same in the different languages in which it is used. For instance *alors* in French and *allora* in Italian have developed both an interactive function and a logical (consequence) connector function from the original temporal value and referential function, maintaining the latter in today's uses; however, not all discursive functions overlap when taking into account the two languages considered; the same goes for different uses of *donc* and *dunque* highlighted by Ferrari/Rossari (1994). The descriptions of Italian *già* in comparison with French *déjà* by Hansen/Strudsholm (2008) and subsequently by Squartini (2014), in addition to Mihatsch's (2009) analysis mentioned above, are all studies in which the intersection between the diachronic evolution and the cross-linguistic dimension is at play. The cross-linguistic study of DMs is covered in this volume by 8 Discourse markers in a contrastive perspective.

3 The contribution of DMs: Semantic-pragmatic descriptions in the literature on Romance languages (with a focus on French, Italian and Spanish)

3.1 General theoretical and methodological issues

As indicated above, the second issue in the description of DMs concerns the type of contribution that DMs make to the discourse. This question is related to the well-known and controversial theoretical issue of the link between encoding and context, i.e. the place of the code in relation to the place left to the context.

The dichotomy 'code/context' covers the distinction between semantics (considered as conventional indications) and pragmatics (discourse and contextual meaning) (for a comprehensive overview of this distinction, see Bach 1997). This debate has been conducted within different theoretical frameworks: it was described by Halliday/Hasan (1976) in terms of semantic relations vs rhetorical / pragmatic relations, and taken up within the Relevance Theory framework (Blakemore 1987; 1992) with the distinction between 'conceptual' and 'procedural' elements.

Generally, DMs have been defined as belonging to the pragmatic-interactional layer (see, e.g., Bazzanella 1995), while connectors have been analyzed more at the level of their semantic layer. Within the English-speaking literature, the debate around the contribution of connectors revolves around the issue of whether it is essentially conceptual, i.e. mapping onto concepts, or if it is essentially procedural, i.e. giving a set of instruc-

tions on how the recipient should use the context to construct the meanings intended by the speaker in the different usages, the conceptual part being reduced to minimal indications such as logical relations, if any (on this topic, see, among others, Blakemore 1987; 1992; 2002; Hall 2007, in dialog with Bach 1999; Hansen 1998; Wilson/Sperber 2012).

The code-context debate is intertwined with the debate on the monosemic or polysemic nature of DMs. Monosemic accounts of DMs assume the existence of one and the same core meaning in all interpretations of a marker, which must be contextually enriched (cf. Schiffrin 1987 or Fraser 2009); polysemic descriptions assume the co-existence of several meanings for the same marker on a semantic level and address the question of how these different meanings are related. Different mechanisms are evoked, such as metaphorical and/or metonymic extension (see Hansen 1998; 2008 for such a conception).

As shown by some instances within the catalog above, analogously to the English-speaking literature, both a narrow and a broad sense of the notion of DM can be found in studies on Romance languages: Ducrot et al. 1980 include for instance both textual connectors and interactional markers in their analysis, regardless of the label adopted. However, in general, studies on Romance appear to dissociate context-related forms more strongly from text-related forms, which are considered as conveying coded information. The latter have been often described within specific frameworks constituting theoretical models of discourse/text structure, from the pioneering work on the linking of discourse by Roulet et al. (1987) and the macro-syntax framework by Berendonner (1990; 2003) for French, to the ‘Basel model for paragraph representation’ (Ferrari et al. 2008) for Italian, to the Val.Es.Co. model for Spanish (Grupo Val.Es.Co. 2014). Thus, less permeability between classes can be observed: on the one hand, we find studies linked to interactional linguistics, analyzing spontaneous conversation; in such a perspective DMs are pragmatic functional units helping in the interpretation process. On the other hand, a significant part of the research on DMs in Romance is framed within a largely established text/discourse-oriented research perspective, which is mainly interested in relations between discourse/enunciative units (in addition to the above mentioned Ducrot et al. 1980, see for instance Culioli 1990 with the concept of enunciative operations, as well as Rossari 1994; Rossari et al. 2004; 2007 on connectors and discourse relations), as well as in the monological/dialogical nature of such relations. In this perspective, the contribution of DMs is at least partly coded by semantic traits.

From a methodological point of view, theoretical accounts of both DMs and connectors often focus on specific forms used as an illustration of the working mechanism of a class, or better still, a communicative function, both in the English-speaking literature and in the French-speaking tradition. Whether or not they are conducted within a specific theoretical framework, studies on particular markers or categories of markers are abundant in the literature on Romance: to give but a very few examples, see for French the work on reformulation markers by Authier-Revuz (1995), Rossari (1994), Steuckardt (2018), and the research on *donc* by Carel/Ducrot (1999), Fernandez (1994), Ferrari/Rossari (1994). For Italian connectors, see for instance Manzotti (1998) for exem-

plification markers and Mazzoleni (2015) on concessive constructions. DMs of reformulation (*quer dizer, ou seja, isto é*) as well as of contrast (*ao invés, pelo contrário*) in contemporary Portuguese have been extensively studied respectively by Macário Lopes (2004) and Macário Lopes/Sousa (2014), while Romanian epistemic markers (*desigur, poate*) are analyzed, among others, by Zafiu (2018). In all of these studies on Romance languages, attempts to distinguish different forms pertaining to the same class have been systematically made following the work of Anscombe/Ducrot (1976; 1977) as for the distinction between *mais* and *pourtant* (see, e.g., Rossari 1994 for reformulation markers). In other cases, connectors are described within more general studies on a specific logic-argumentative relation (this is the case for the description of *ad esempio* within a broader study on the relation of exemplification by Manzotti 1998 mentioned above, or for the study by Mazzoleni 2015 on ‘pre-concessive’ relations).

In all, the question of coded or contextual meaning depends on the ‘lens’ through which DMs are considered, namely, the specific theoretical framework within which they are analyzed. Indeed, the literature on Romance has mobilized theoretical frameworks where semantics and pragmatics are intertwined, as in the theory of Language-Inherent Argumentation (Anscombe/Ducrot 1976) or of linguistic polyphony (Nølke/Fløttum/Norén 2004). All these perspectives are based on the conditions of use of each connector, in terms of truth-conditional semantics or ‘argumentative semantics’.¹ In addition, literature on French and other Romance languages specifically adopts an enunciative perspective, which considers all layers of discourse (propositional content, speech act level, enunciation level) in the analysis of connectors (see Rossari et al. 2004 and 2007).

To better show the impact of the theoretical framework on the way in which the contribution of DMs is understood, we present in more detail some of the main ways in which they have been analyzed.

3.2 Descriptions of DMs within a semantic-pragmatic framework

Among the first descriptions of DMs, that of Halliday/Hasan (1976) is primarily textual, in that it considers conjunction among the relations providing text cohesion. We can interpret the type of contribution they attribute to conjunctive items such as *and* or *but* as partly ‘chameleonic’: they keep their encoded meaning within any context compatible with it (e.g. *and* is additive, *but* is adversative), but can also take on semantic features of the utterances they link and can thus express different types of conjunctive relations.

¹ La sémantique argumentative considère que tout énoncé – qu’il prétende être objectif, qu’il prenne ouvertement parti ou transmette des sentiments – construit des points de vue formulables par des argumentations ‘Argumentative semantics considers that every statement – whether it claims to be objective, openly takes sides or conveys feelings – constructs points of view that can be formulated by argumentations’ (<https://semanticar.hypotheses.org/>), ‘carnet de recherche’ on argumentative semantics on OpenEdition.org. Editorial manager: Zoé Camus).

From a discourse-analytic perspective based on coherence, Schiffrin (1987) adheres to the idea that DMs can reflect an indication that is already present in discourse when there is compatibility between the co-text and the coded meaning of the marker. However, in the absence of such compatibility, the DM keeps its encoded meaning, and the discourse relation it conveys has to be understood as pertaining to another level. This is for example the case of *because*: in a context that does not express a causal relation between states of affairs, it must be analyzed as indicating a causal relation pertaining to the non-propositional level (see Sweetser 1990; Dancygier/Sweetser 2009, who show that the same causal connector can convey a relation that bears on states of affairs, beliefs or speech acts).

An in-between perspective is the one developed by Fraser (2009), according to which DMs are conceived as conveying a stable encoded (core) meaning, of a conceptual as well as a procedural nature; both these components are involved in the instructions given by DMs.

A more radical instructional view is developed within the framework of Relevance Theory: Blakemore (1987; 2002) for English DMs and Moeschler (1989) for French DMs see connectors as guiding the interpretation of the utterance in which they appear by encoding instructions about contextual effects. Moeschler's description is based on the idea of an attempt to maximize the understanding of the interlocutor:

Connections are not between discourse units, but between linguistic and non-linguistic knowledge, represented in propositional forms. [...] This means that pragmatic connectives are not specifically introduced by a speaker in order to increase clarity in his talk (which would be an explicit indication that the Gricean maxim of manner is satisfied, cf. Grice 1975), but rather to increase access to the context, more likely to maximise the relevance of his utterance. (Moeschler 1989, 333).

An instructional description of DMs is also provided by Hansen (1998, 236), according to whom “semantically, markers are best seen as processing instructions intended to aid the hearer in integrating the unit hosting the marker into a coherent mental representation of the unfolding discourse”. Drawing on examples from French and English, Hansen points out that different uses of a polyfunctional morpheme have different consequences for the ensuing discourse, arguing that a polysemic analysis of markers, assigning it more than one conventional meaning, offers a better account of its impact on the following discourse than a minimalist analysis describing the particle as having an abstract core meaning to be further specified by the context.

3.3 Descriptions of DMs within an integrated pragmatics framework

In the framework of Anscombe/Ducrot's Integrated Pragmatics (Ducrot 2010), which constitutes a milestone in the description of DMs in Romance languages, the role of the code is stable and of an instructional nature, but much richer. A *mot du discours* ‘word

of discourse' is an expression allowing the speaker to show her presence, marking her attitude towards what she is saying or the direction she wants to give to her speech. The connector is thus conceived as constraining the argumentative orientation of the utterances it connects. This view results from the authors' conception of the meaning of the utterance itself: an utterance cannot be understood independently of its argumentative orientation. A DM like *mais* ('but') is conceived as conveying an anti-argumentative orientation, presenting the orientation of the introduced utterance as stronger, i.e. involving a stronger commitment than that of the utterance preceding *mais*. For instance, in an utterance like *la maison est belle mais elle est chère* 'the house is beautiful but it is expensive', the orientation towards the conclusion 'I'm not buying it' conveyed by the sentence following *mais* prevails on that 'I'm buying it' conveyed by the sentence preceding it in terms of strength of commitment. This is a particularly 'heavy' instruction for a single DM. It shows in a salient way how this conception reverses the perspective on DMs, integrating in the coded meaning indications usually referred to as pertaining to the pragmatic level of analysis. Along the same lines, according to Rossari (1994), connectors can be described as imposing constraints – not only in terms of argumentative orientation, but also in terms of polyphony or illocution – on the portion of discourse to which they refer.

Another framework involving an explicitly declared discourse semantic view of DMs is that provided by Paillard (2015). Analyzing DMs *tout de même* and *quand même* 'all the same', Paillard argues that DMs can be described as forming a class that can be defined according to semantic and syntactic (i.e. combinatorial) criteria, just like other word classes. However, differently from the semantics of other word classes such as nouns or verbs, the semantic description employed by Paillard to define DMs erases the distinction between the act of saying and what is said, in a perspective in which the utterance more or less felicitously expresses the representation of a state of affairs by the speaker. On this view, the characterization of DMs involves the notion of 'aboutness': the function of DMs is "to give a particular discourse (i.e. enunciative) status to the sequence over which the DM has its scope" (Paillard 2015, 167, our translation). Within this framework, Paillard distinguishes four classes of DMs:

- i. *mots du discours* (not equivalent to the traditional Ducrot's label), specifying that the proposition on which the marker has its scope reflects the personal perception of the subject. This category includes, for example, adverbial expressions such as *malheureusement* 'unfortunately', or *en réalité*, 'in reality'.
- ii. *enunciative particles*, which characterize the proposition in their scope as taking into account the actualization, in the left context, of what Paillard calls a 'pole of otherness', i.e. a content with respect to which the content introduced by the marker constitutes an alternative. This class includes markers such as *d'ailleurs* 'on the other hand', *cependant* 'however', *tout de même/quand même* 'all the same', but also more interactional DMs like *bien*, 'well', as well as adverbs like *déjà* 'already' (Paillard 2015).

- iii. *modalizers*, which express an indetermination between the proposition that conveys a state of affairs and the actual state of affairs. These are expressions such as *en quelque sorte*, *genre*, *quoi* ‘sort of’, ‘kind of’, or ‘what’, often placed at the end of the sentence;
- iv. finally, the category of *mots du dire* (lit. ‘words of saying’), which imply a ‘disengagement of the subject’ in Paillard’s terms, putting distance between a “meaning to say”, which may be the speaker’s or another instance, and what is being said. This last category includes expressions often containing the verb *dire*, such as *pour ainsi dire* ‘so to speak’, *à vrai dire* ‘to tell the truth’, *c’est-à-dire* ‘that is (to say)’, ‘namely’.

It must be added that Paillard’s description of DMs is also syntactic and prosodic, in that the interpretation of the markers is also linked to their distribution (at the beginning of the utterance, at the end of the utterance or in the middle position), to their scope (use on the proposition, on an element of the proposition and absolute use, where the DM is used alone), as well as to their prosodic profile (detachment or non-detachment, marked by punctuation in writing).

3.4 A description of DMs within a pragmatic-dialogic framework

A typology illustrating the challenge of delimiting the class within a purely pragmatic framework is provided by Bazzanella (1995) for the description of DMs in Romance languages. The fundamental properties of discourse markers are polyfunctionality (meaning that their function varies according to the context in which they are used) and sensitivity to the context of appearance. Within this conception, DMs are classified according to the following categories – cf. Bazzanella (1995):

1. Interactive function carried out either by the locutor or interlocutor (phatics, modulation markers, feedback markers);
2. Metatextual functions (demarcatives, focalizers (like *proprio* ‘precisely’) and connectors expressing a reformulation function, as well as an illustration function);
3. Cognitive functions, further distinguished as procedural (i.e. involving inference processes) vs epistemic (involving the judgment of the speaker).

This type of classification has the advantage of showing the complexity in defining the boundaries between categories. As already mentioned, a number of markers actually straddle more than one category, for instance dialogic-interactive uses and textual uses for a single marker, such as French *alors*. Also, markers grouped in the same class according to one criterion differ strongly with respect to many other aspects of their functioning, be it syntactic or semantic. For instance, among markers having a metatextual function in many classifications, we find focalizers, which differ from other DMs

in that their scope is intra-propositional, i.e. usually on a lexical element (adjective or adverb). Also, DMs which are standardly assumed to be ‘meta-textual’, i.e. which work at the level of the organization of the text and its structuring, can serve argumentative functions, such as illustration or reformulation (see for instance Rossari’s 1994 analysis of *comunque* ‘anyway’).

3.5 Description of DMs within a syntactic-semantic framework

Bernard/Danlos (2019) provide a syntactic-based classification of conjunctions in written French. Based on Haegeman’s (2004) syntactic framework, they classify conjunctions as being either ‘central’ or ‘peripheral’, showing a systematic correlation between their syntactic status and the meaning of the conjunctions (propositional vs textual). For instance, *quand* as a temporal conjunction (‘when’) has a ‘central’ status, whereas when it serves the role of an opposition connector (‘whereas’) it is syntactically peripheral. In a formal semantic framework, Rossari/Jayez (2000) have highlighted the semantic constraints that characterize the uses of some French connectors.

Although this chapter is mostly devoted to the semantics and pragmatics of DMs in the descriptions of Romance languages, syntax, and consequently the distribution of DMs, must be mentioned as part of the defining parameters in comprehensive descriptions such that of Paillard (2015) described above, but also as a parameter in relation to information structure, within models providing a discourse segmentation framework for the description of DMs. These models have been very profitable for French and later for Italian and Spanish, particularly for the description of items that elude traditional classification but rather require description in informational and textual terms, such as DMs. Below is a brief overview of the main models based on discourse segmentation.

3.6 Descriptions of DMs within ‘macro-syntactic’ models of discourse segmentation

In a research tradition that has been developing since the ‘90s, models of discourse segmentation related to pragmatics and syntax, highlighting the complex nature of text organization, have analyzed DMs as clues to determine text segmentation and hierarchy. The first to propose a discourse segmentation model for French based on syntactic and informational parameters were Blanche-Benveniste et al. (1990), with the G.A.R.S. model, and Berrendonner (1990; 2003). They notably introduced the notion of ‘macro-syntax’ to characterize the organization of syntactic configurations difficult to analyze within the framework of traditional grammar, particularly in spoken language, and particularly for what concerns ‘detached’ elements.

Based on Speech Act Theory, but in a broader framework, that of discourse organization, Roulet et al. (2001) provide a hierarchical model which integrates hierarchical, referential, interactional, relational, operational, periodic, topical, compositional, polyphonic, and strategic dimensions in its description. In a similar vein, taking into account a multidimensional analysis (specifically, prosodic and pragmatic), the Val. Es.Co. (Valencia Español Coloquial) group studies markers in informal conversation (Cabedo Nebot 2014).

Subsequent models for text segmentation, providing a framework for the analysis of connectors, focus on information/topical organization of spoken or written discourse: we refer here to Moneglia/Cresti (2006), then Raso (2014) for spoken discourse, as well as to Ferrari et al. (2008; 2014), with the Basel model of paragraph segmentation, for written text.

As for spoken discourse, the model of Cresti (2018) links prosody to the illocutionary and informational function of discourse units. According to this model, the pragmatic foundations of the utterance and its information structure are closely connected to the role of prosody in their identification. The reference unit of speech is the utterance, which is the counterpart of a speech act. Each unit is assigned an information function, corresponding to (and thus being identifiable via) prosodic patterns. Thus, rather than being characterized in semantic terms, the different information functions are identified and classified (within a corpus-driven approach) taking into account their prosodic realization.

Although a distinction is made between the notion of DMs and that of discourse connectors, this distinction does not cover the full range of DMs generally analyzed in the literature, particularly connectors performing logic relation functions. The function of discourse connectors falling under this description is that of linking devices connecting discursive sequences with the purpose of creating and maintaining the “weaving” of the text. An analysis of the semantic contribution of a specific connector is not the purpose of this model.

On the other hand, a notion of DM which involves logical connectors is at work for written text in the Basel Model for Paragraph Segmentation, conceived by Ferrari et al. (2008). This kind of model focuses on the “semantico-pragmatic connections that weave the content of the discourse into a whole” (Ferrari 2014, 24), thus allowing a fine-grained linguistic analysis of the semantic contribution of the class of connectors, often from a contrastive cross-language perspective.

The information structure of the utterance, considered as the crossing zone between the linguistic system (lexicon, grammar, prosody) and the text, has also a central role in this model. Connectors are seen as a particularly felicitous example of the working mechanism of such an interface: the different uses and functions of a connector, as well as its scope within the portion of text, is linked to its position within the Utterance and specifically within the type of Information Unit within which it is situated (Ferrari et al. 2008, 376). For instance, Mandelli (2008) shows that a connector like *in effetti* in initial

position conveys a richer relation from a semantic-pragmatic point of view (in this case, motivation), as well as having a broader textual scope, whereas the same connector within an Appendix Unit, rather than having a fully argumentative function, seems to be more related to the management of discourse (e.g. reformulating previous content), as well as to have a more limited scope, therefore participating more ‘locally’ in the constitution of the textual architecture. Such a model allows to understand the textual behavior of polyfunctional markers. It has been used from a contrastive perspective in the analysis of Spanish connectors as compared to their counterparts in Italian, for instance in Esposito (2020), who describes the Spanish additive particle *también* and its equivalents in Italian so as to detect the most relevant levels of cross-linguistic contrast within a functional perspective (see also Borreguero Zuloaga 2014; De Cesare/Borreguero Zuloaga 2014 for other examples of cross-linguistic analysis). However, the description of DMs in such a model is not concerned by the specific semantic contribution of the DM.

For a comprehensive view of the analysis of DMs in relation to discourse units, we refer to *6 Discourse markers and discourse/conversational units* as well as to Pons Bordería (2014).

Overall, descriptions analyzing DMs within an oral or written discourse segmentation framework have two advantages: on the one hand, they are functional descriptions, independent of the lexical or grammatical nature of the linguistic form under consideration, which leads to the possibility of including forms that may or may not have the function of DMs depending on their use and the information unit in which they are inserted; on the other hand, this description brings out fully the complex relationship between the pragmatic function of an utterance and the forms it hosts: the function of connectors reveals the presence and the different functions of informational units in the textual fabric, and conversely, the description of the function of such units makes it possible, in some cases, to account for the diversity of uses or nuances of meaning of connectors. On the other hand, such descriptions do not seem to allow one to define the basic semantic contribution of a specific form, which seems to be already pre-determined, i.e. outside the defining possibilities of the framework.

In summary, a main issue arising from any perspective is how to determine the semantic and pragmatic part in the analysis of DMs. The contribution of a DM is hardly detachable from its discursive environment. Because of its chameleonic nature, it is difficult to extract what is due to the DM, i.e. what makes the marker compatible with a specific environment, as opposed to what is due to the discourse in which it occurs. Consequently, the determination of what is stable (that is, conventional and therefore semantic), as opposed to what depends on pragmatic constraints, appears to be difficult to disentangle.

Our contribution will not solve these issues, but the textometric approach that we adopt and present below can contribute to identify some stable traits pertaining to the specific subclass of DMs we investigate: connectors.

4 A textometric approach applied to connectors

The textometric approach, which has as its precursor lexicometry (Lebart/Salem 1994; Pincemin 2011; Rastier 2008) consists in applying statistical calculations to the study of a digitalized corpus of texts. Such an approach allows one to link quantitative and qualitative analyses by querying the privileged contexts of appearance of forms in a chosen corpus. It therefore enables one to see whether regularities emerge with respect to the combinatory properties² of a given form with other forms of the lexicon. As for connectors, we can exploit this approach to identify which markers they attract (i.e. associate with in a statistically significant way) and, in doing so, to see the emergence of similarities between connectors, similarities that can be interpreted as resulting from a common semantic trait.

The hypothesis that guides us is that connectors, being privileged instruments of argumentation, are sensitive to indications that reveal the involvement of the speaker in her utterance, and this, regardless of the relation they convey. More precisely, we suggest that they differ in relation to the extent to which they make apparent the speaker's involvement (SI) within the relation displayed: the same relation can be displayed showing more SI with one connector than with another. Depending on the strength with which the connector brings out this involvement, it will have a greater or lesser tendency to associate with forms related to different manifestations of involvement. We label *speaker's involvement form* (SiF) a form which gives indications about the speaker's attitude, her 'saying', or her 'said', as well as indications about the managing of 'voices' emerging within the discourse. As for the first class, we retain for the 'said' the following categories: modal verbs and adverbs (*pouvoir* 'can' / 'may', *vouloir* 'want'; *probablement*, 'probably' *certainement* 'certainly', . . .), modal adjectives, i.e. adjectives including in their semantics the idea of possibility – e.g. *mangeable* 'edible'), verbs of thought (*croire* 'believe', *penser* 'think', . . .), and for the 'saying', enunciation adverbs (*franchement* 'frankly', *clairement* 'clearly', . . .). As for the second class, we retain *dicendi* verbs (*dire* 'to say', *refuter* 'to refute', *admettre* 'to admit', . . .). Some of the above forms can be polysemous: in these cases, as well as in cases of homonymy, we have ensured that the majority of their uses were relevant for our classification.

The methodology we built within this approach aims to allow us to test some connectors pertaining to different semantic classes across corpora representing different discourse genres in French and Italian. For each semantic class of connectors, two, three or four items have been selected, according to an assumption concerning their capacity to reveal the involvement of the speaker to a greater or lesser extent. For some of these connectors, theoretical frameworks have already pointed out properties which can be related to speaker's involvement; for others, it is possible to reveal these prop-

² We use here the expression *combinatory properties* taking as a base the notion of *profil combinatoire* ('combinatory profile') introduced by Blumenthal (2002) and defined as "the schematic structure of the syntactic and semantic neighborhood of a pivot word as it manifests itself in a large corpus" (Blumenthal 2002, 115, our translation).

erties through example manipulation; for instance, within the consequence class, we tested the capacity of an item to introduce judgements as well as facts (*donc* ‘therefore’) or predominantly facts (*de ce fait* ‘hence’), as in:

L'examen était difficile, donc elle est intelligente/?de ce fait elle est intelligente.

‘The exam was hard, so she’s smart / hence she’s smart’³

Our goal is to differentiate connectors according to their ability to attract forms belonging to the SiF class. To this end, we first identified, for each connector, all linguistic items specifically associated with it; among these items, we then sorted and retrieved the forms mentioned above, considering them as part of the macro-class of SiFs. Finally, we calculated a Log-likelihood score for each connector with respect to the set of forms belonging to the SiF class.⁴

In Table 1, we present the selected connectors for each of the three classes taken into account. Table 2 shows the SiFs which have been retrieved with the specificity score.

Table 1: Selected connectors in French and Italian.

	Consequence	Opposition	Cause
French	DONC ‘so’, ‘therefore’	MAIS ‘but’	CAR ‘because’
	C’EST POURQUOI ‘that is why’	POURTANT ‘yet’	PARCE QUE ‘because’
	PAR CONSÉQUENT ‘consequently’	CEPENDANT ‘however’	PUISQUE ‘since’
	DE CE FAIT ‘hence’, ‘as a result’	TANDIS QUE ‘whereas’	
Italian	DUNQUE ‘so’, ‘therefore’	MA ‘but’	POICHÉ ‘since’, ‘because’
	QUINDI ‘so’, ‘therefore’	PERÒ ‘but’, ‘though’	
	PERCIÒ ‘so’, ‘therefore’	EPPURE ‘yet’	DATO CHE ‘since’, ‘given that’
	DI CONSEGUENZA ‘consequently’, ‘as a result’	MENTRE ‘whereas’	

Connectors such as *perché* or *visto che* have been excluded either because of polysemy (*perché* corresponds to *why*, *because*, *in order to*) or because of homonymy (*visto* is used both as a functional item being part of the connector *visto che* or as a lexical item as a

³ Our English translations never aim at retrieving the acceptability or not of French or Italian original examples, nor are they meant to give the exact English counterpart of French or Italian connectors. They are exclusively intended to render the meaning of the original.

⁴ The identification of all items being specifically associated to each connector was made possible by the TXM software (Heiden 2010), which calculates a *specificity index*. This measure makes it possible to identify, for a given item (the pivot), the items associated with it in a statistically significant manner within a given span. As for measuring a statistically significant association with respect to the class of SiFs in its whole, this required a calculation made independently through the software R and resulting in a *Log-likelihood* score, which is also a specificity score. With this measure, a score of 10.83 or above ensures that the probability for the association not to be due to chance is 99.9%. The log-likelihood score as well as the specificity index are not impacted by the overall frequency of the forms considered.

past participle of *vedere* ('to see') in perfect tenses). Even though the connectors selected can be considered as possible translations of each other, our selection was not based on the highest possible similarity between French and Italian, but on the fact that, in both languages, they reveal a difference in relation to the speaker's involvement.

Table 2: French and Italian Speaker's involvement Forms' (SiFs) retrieved with the specificity index.

	French	Italian
Modal adverbs	assurément certainement certes curieusement forcément éventuellement évidemment forcément franchement incontestablement nécessairement peut-être potentiellement probablement sûrement visiblement	apparentemente assolutamente certamente evidentemente forse necessariamente probabilmente presumibilmente sicuramente
Modal verbs	devoir falloir nécessiter paraître pouvoir sembler vouloir	bisognare dovere occorrere parere potere sembrare volere
Verbs of thought	croire décider douter penser raisonner supposer	capire considerare credere decidere dedurre dubitare immaginare intuire pensare ragionare ricredere rielaborare riflettere ritenere supporre

Table 2 (continued)

	French	Italian
<i>Dicendi verbs</i>	admettre	acclamare
	appeler	ammettere
	concéder	ammonire
	confirmer	chiacchierare
	contester	chiedere
	critiquer	contestare
	déclarer	denunciare
	déconseiller	dichiarare
	discuter	dire
	demander	discutere
	dire	gridare
	expliquer	parlare
	nier	precisare
	parler	proporre
	préciser	ribadire
	prétendre	ribattere
	proposer	riferire
	recommander	ringraziare
	répondre	riparlare
	surnommer	rispondere
		soprannominare
		sussurrare
Adjectives ending in: -ible/-able (French), -ibile/-abile (Italian)	compréhensible	accettabile
	contestable	affidabile
	critiquable	ammissibile
	décelable	applicabile
	discutable	apprezzabile
	disponible	attaccabile
	distinguable	calcolabile
	divisible	comprensibile
	entreposable	condivisibile
	exploitable	definibile
	faisable	eleggibile
	fiabile	equivocabile
	impossible	giudicabile
	imprescriptible	godibile
	incapable	impossibile
	increvable	impraticabile
	indispensable	imprendibile
	indéniable	incommensurabile
	inéluçtable	incurabile
	inexpulsable	indagabile
	intelligible	indispensabile
	invérifiable	inevitabile
	invisible	inflessibile

Table 2 (continued)

French	Italian
irréfutable	innegabile
irréversible	inoppugnabile
louable	invincibile
négligeable	invisible
plausible	irrealizzabile
possible	irremovibile
préférable	irrinunciabile
prévisible	leggibile
probable	misurabile
redoutable	numerabile
regrettable	passibile
remarquable	plausibile
rentable	possibile
réparable	prevedibile
répréhensible	probabile
respectable	rispettabile
souhaitable	sovrapponibile
susceptible	utilizzabile
valable	verificabile
variable	verosimile
vérifiable	visibile
visible	vivibile
vraisemblable	

Concerning the genre, we exploit discourse corpora which *a priori* lack an argumentative dimension (encyclopedic discourse), or which, on the contrary, do have one (press and political discourse). For the former, our corpus consists of a random selection of French and Italian Wikipedia entries; for the latter: a whole year of newspaper *Le Monde* and *L'Adige*, as well as French and Italian presidential speeches.⁵

4.1 Results

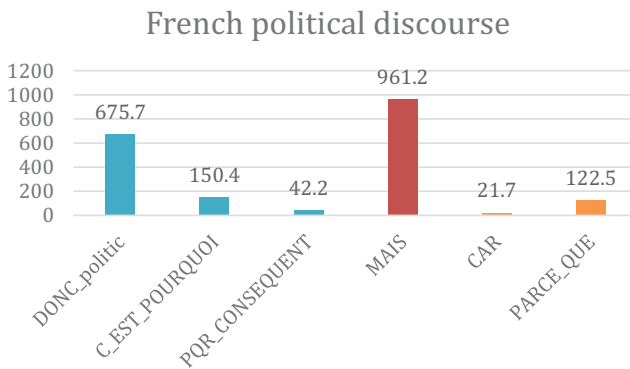
Data retrieved from our corpora show different, but very stable tendencies emerging across the three discourse genres with respect to the connectors we have considered, in both languages. For each class of connectors, it is possible to establish a ranking accord-

⁵ French newspaper *Le Monde* (17,895,009 tokens, covering the year 2010), Italian newspaper *L'Adige* (20,343,148 tokens, 38,496 articles covering the period 1996–2006); French Wikipedia (18,715,455 tokens, 10,571 articles, from 2019, without bibliographic sections); Italian Wikipedia (18,394,987 tokens, 15,905 articles from 2019, without bibliographic sections), Italian presidential speeches (2,639,007 tokens, covering the period 1955–2021), French presidential speeches (2,843,577 tokens, covering the period 1958–2020).

ing to the degree of specificity of their association with SiFs.⁶ Such stability in connectors of different semantic classes, in corpora representing discourse genres that strongly diverge with respect to their argumentative dimension, is an indication that the feature ‘speaker’s involvement’ can be assigned to the connector as a conventional trait.

Below, we present the results concerning the three corpora considered in French and in Italian.⁷

4.1.1 French corpora



AU: Please mention Figures 1–7 in the text.

Figure 1: Specific associations with SiFs for French connectors in political discourse (presidential speeches).

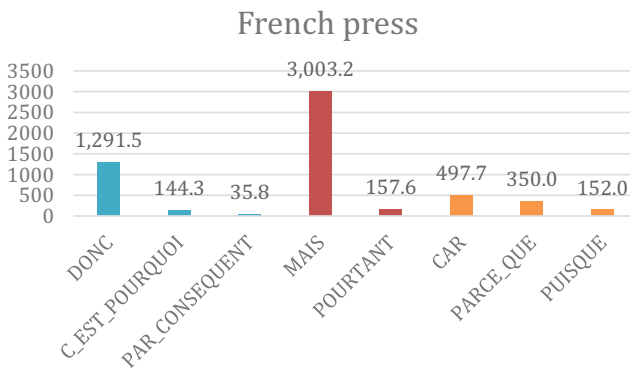


Figure 2: Specific associations with SiFs for French connectors in the press (newspaper *Le Monde*).

⁶ All scores equal to or above 10.83 can be linguistically interpreted as an attraction between a connector and SiFs. When the differences between the association scores is particularly marked, it can be linguistically interpreted as indicating stronger attraction between the connector and the form.

⁷ We thank Cyrielle Montrichard for the discussion and the calculations to obtain the log-likelihood scores for the associations of connectors with the class of SiFs.

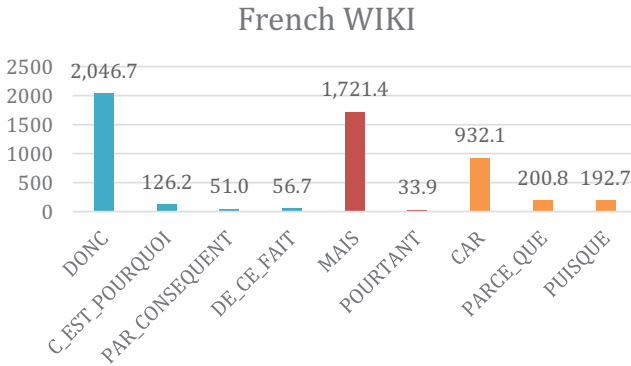


Figure 3: Specific associations with SiFs for French connectors in encyclopedic discourse (Wikipedia).

Relation of consequence (*donc, de ce fait, par conséquent, c'est pourquoi*)

In the three corpora analyzed, a ranking among connectors emerges, with *donc* clearly standing out compared to the other connectors, especially in the press and in Wikipedia. Also, we observe a difference in the association scores between *c'est pourquoi* (which appears as being closest to *donc*) and, on the other hand, *par conséquent* and *de ce fait*. This ranking is in line with the difference mentioned above, according to which *donc* is more prone to introducing a judgment than *de ce fait* and *par conséquent*. The intermediate position of *c'est pourquoi* can be accounted for by its focusing structure: focalizing a piece of information is in itself a clue of the speaker's involvement in her utterance. This is what distances *c'est pourquoi* from *par conséquent* and *de ce fait*. The retrieval of a sample of occurrences of the four consequence connectors in the press corpus through the keyword-in-context function shows results consistent with the above observation: out of 40 occurrences, 30–40% of *donc* and of *c'est pourquoi* connect judgments while no occurrences of this kind were found for *par conséquent* and *de ce fait*. That is not to say that a connector such as *de ce fait* is incompatible with a verb of thought; in this case the connector introduces the thought as a matter of fact. It remains that there is no statistically significant association between *de ce fait* and the expression of speaker's involvement in two corpora out of three, the score being extremely low in the third corpus.

Below, an example of a *donc* connecting judgments, in which the use of *de ce fait* would sound odd:

Les politiques macroéconomiques [. . .] ne se soucient nullement de la nécessité du développement et de la protection du marché national. La seule préoccupation de ces politiques est de produire à bas coût pour l'exportation vers le marché mondial." C'est donc/ ??de ce fait scandaleux d'entendre le FMI dire qu'il "se tient prêt à jouer son rôle avec le soutien approprié dans ses domaines de compétence (Le Monde).

'The macroeconomic policies [. . .] have no concern for the need for development and protection of the domestic market. The only concern of these policies is to produce at low cost for export to the

world market.” It is therefore outrageous to hear the IMF say that it “stands ready to play its role with appropriate support in its areas of competence”.

Conversely, in a context where the link concerns states of affairs, *de ce fait* sounds perfectly adequate:

En présence d'eau et de quelques sels, le dioxyde de soufre forme des sulfates d'acide sulfurique, aérosols – ou particules en suspension – de quelques microns qui réfléchissent une part du rayonnement solaire. De ce fait, ils ont un effet refroidissant sur la troposphère, couche inférieure de l'atmosphère (Le Monde).

‘In the presence of water and some salts, sulphur dioxide forms sulphuric acid sulphates, aerosols – or suspended particles – a few microns in size that reflect some of the sun’s radiation. As a result, they have a cooling effect on the troposphere, the lower layer of the atmosphere’.

Relation of opposition (*mais, pourtant, tandis que*)

As for opposition connectors, *mais* stands out in all corpora, followed by *pourtant* (though with much lower scores) in two corpora. Such a score is in line with the analysis concerning both connectors in the literature. In Anscombe/Ducrot (1977) *mais* is characterized by the hierarchical relationship it establishes in terms of *commitment* (a stronger commitment to the content to the right of the connector than to the content preceding it), thus imposing an orientation on the continuation of the discourse. This property is easily linkable to *mais*’s speaker involvement feature. Conversely, the research on *pourtant* (cf. Roulet et al. 1987; Anscombe 2002) considers it as a ‘symmetrical’ connector, i.e. as a connector not differentiating the constituents it links in terms of commitment. This property could be consistent with a weaker manifestation of the trait ‘speaker’s involvement’. *Cependant* does not appear in this ranking thus showing no affinity with SiFs. This may be due to the fact that, even though, similarly to *mais*, *cependant* hierarchizes the contents it links, the temporal relation it originally conveys (‘during this time’) plays a role in its semantic contribution, making the relationship more factual than subjective.

Relation of cause (*car, parce que, puisque*)

As regards the causal relation, the three connectors show attraction with SiFs, with the highest score for *car* in two corpora and the lowest (though not with a huge gap) for *puisque* in all the corpora. Such results are in line with the analysis made in the polyphony framework by Ducrot (1984) or Nølke/Fløttum/Norén (2004). These analyses highlight *puisque*’s property of introducing already acknowledged content. As a result, the speaker’s involvement may be less at stake, since she does not endorse the content that follows the connector. In that, *puisque* contrasts with *car*, which indicates a complete endorsement of the content by the speaker. *Parce que* shows a similar behavior to *car*, which is in line with its property of being a polyfunctional connector, allowing all kinds of connections (states of affairs, judgments, or illocutionary acts), as observed for *because* by Dancygier/Sweetser (2009) or by Debaisieux (2016) for *parce que*.

4.1.2 Italian corpora

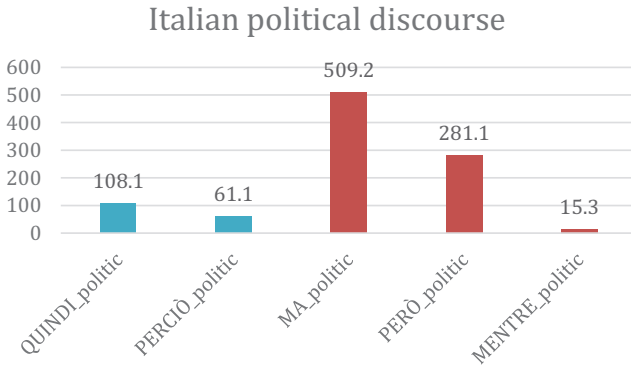


Figure 4: Specific associations with SiFs for Italian connectors in political discourse (presidential speeches).

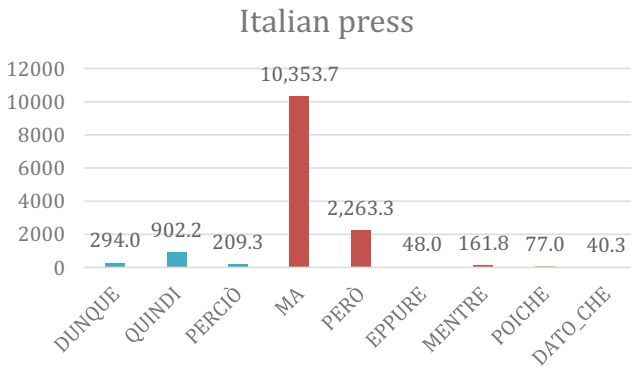


Figure 5: Specific associations with SiFs for Italian connectors in the press (newspaper *L'Adige*).

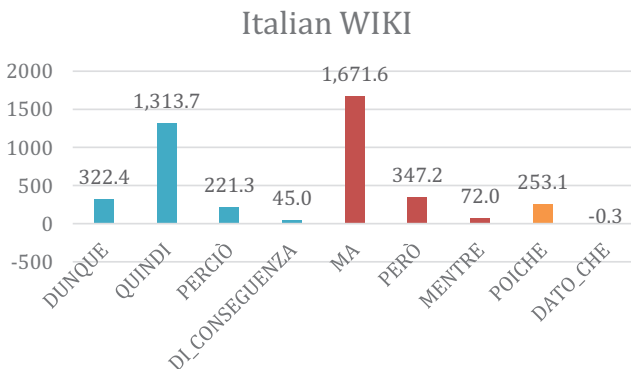


Figure 6: Specific associations with SiFs for Italian connectors in Wikipedia.

In all three corpora considered, we observe very clear-cut contrasts between – respectively – *ma* and *quindi* and the other connectors belonging to their semantic class. *Ma* is the connector which attracts SiFs, and its description in the literature (cf. Marconi/Bertinetto 1984) highlights a hierarchy in favor of the content introduced by the connector, although through different mechanisms compared to Anscombe/Ducrot's (1977) analysis of *mais*. The establishment of a hierarchy between the two contents related by *ma* is also relevant for the functioning of *però*. Conversely, *eppure*, which shows only one (much lower) specificity score, does not seem to establish a hierarchy between the contents it relates, but it rather seems to signal incompatibility, making it the topic of the discourse. This can be shown by the fact that *ma/però* can relate any contents (compatible or incompatible), because their hierarchizing property identifies the content on which the following discourse will continue, whereas *eppure*, when introducing compatible contents sounds odd, since it does not signal on which content the discourse will continue:

ci va a piedi, ma/però la strada è lunga
'she walks there, but it's a long road'

ci va a piedi, ma/però la strada non è lunga
'she walks there, but it's not a long road'

ci va a piedi, eppure la strada è lunga
'she walks there, yet she it's a long road'

?*ci va a piedi, eppure la strada non è lunga*
'she walks there, yet it's not a long road'

These different constraints in the use of *ma/però* vs *eppure* are in line with their SiFs scores – *eppure* being the only opposition connector showing very low (in one corpus only) or no attraction with SiFs.

As for *quindi*, we observe that, similarly to *donc*, it can easily relate judgments (at least 40% of the occurrences on a sample of 100 occurrences): its high association with SiFs is consistent with such a property. *Perciò*, similarly to *c'est pourquoi*, occupies an intermediate position between *quindi* and *di conseguenza*. Such a position is in line with the fact that *perciò* too can be analyzed as the result of a grammaticalization of a focalization (*per ciò*).

Typical uses where *quindi* and *perciò* connect judgments are illustrated by the following excerpts from our corpora:

quei trentini che fino ad oggi hanno beneficiato e goduto dei favori e dei contributi pubblici, che venivano dati in forma personale e non pubblica, hanno riconfermato metà degli uscenti, aggiungendo fra i nuovi anche gente che non avrebbe mai dovuto ricevere un voto. Quindi la colpa fondamentale è di quegli elettori che non hanno voluto cambiare, che hanno votato in forma clientelare, che non si sono mai interessati di come andavano le istituzioni, se non quella volta che andavano con il cappello in mano a cercare un privilegio o un favore (L'Adige).

‘those Trentino people who up to now have benefited and enjoyed public favors and contributions, which were given in a personal and non-public form, have reappointed half of the outgoing ones, adding among the new ones even people who should never have received a vote. So the fundamental fault lies with those voters who did not want to change, who voted in a patronage form, who never cared about how the institutions were going, except that time when they went with hat in hand to seek a privilege or favor’.

don Dante narra la sua esperienza a fianco dei diseredati e degli emarginati. [. . .] La scelta di vivere accanto a coloro che non sono in grado di stare al passo della vita frenetica dei nostri giorni. “La mia strada” è perciò un libro quasi sussurrato, dal linguaggio semplice ma ricco di contenuti (L’Adige). ‘Don Dante narrates his experience alongside the dispossessed and marginalized. [. . .] A choice to live alongside those who are unable to keep up with the hectic life of our days. “My pathway” is therefore an almost whispered story, whose language is simple, but rich in contents’.

In those excerpts, *quindi* as well as *perciò* introduce an axiological evaluation expressed respectively by *la colpa fondamentale* and *quasi sussurrato*. As for *di conseguenza*, while not incompatible with judgments, it is less frequent and much more attested connecting facts, as in the following example:

Invece anche prima di disporre di una maggioranza ha voluto imporre un presidente del consiglio provinciale e, pur dovendo prendere atto che non sarebbe stato votato, lo ha imposto una seconda ed una terza volta sempre con risultato ovviamente negativo, bloccando di conseguenza il funzionamento del consiglio (L’Adige).

‘Instead, even before he had a majority he wanted to impose a provincial council president and, while having to acknowledge that he would not be voted in, he imposed him a second and a third time always with obviously negative results, consequently blocking the functioning of the council’.

The Italian results for causal connectors also make one connector appear as more prone to attract SiFs than the others. In the two corpora where the connectors show attraction with SiFs, *poiché* is more strongly attracted to SiFs than *dato che*. This difference can be accounted for by the fact that *dato che* introduces already acknowledged contents (for a general semantic description of *dato che* (cf. Visconti 2000, 104), similarly to *puisque*.

4.2 Overall view on French and Italian connectors

As previously mentioned, the primary aim of our research is not to provide a comparative analysis of French and Italian, neither in terms of proximity between Italian and French connectors, nor in terms of their respective specific associations; our goal is to observe trends emerging in the two languages independently from each other. However, our research makes interesting trends emerge concerning both languages:

1. In French as well as in Italian, in each semantic category, there is an important discrepancy concerning at least one connector, which stands out with respect to the others pertaining to the same class: in French these are respectively *car* or *parce que*, *mais*, and *donc*, in Italian, *poiché*, *ma*, and *quindi*.

2. Concerning the relation of consequence, we observe parallelisms between connectors in the two languages: on the one hand, *quindi* and *donc* stand out with a higher score concerning their association with SiFs, on the other hand, the homologous connectors *par conséquent* and *di conseguenza* are both characterized by very low scores or by their absence in the graphs (which means that they have no attraction with SiFs).
3. The same parallelism is noticeable in the specific associations of the connectors conveying a relation of opposition: both French *mais* and Italian *ma* show the strongest affinity with SiFs, while *pourtant* and *eppure* stand on the exact opposite side, with significantly lower scores (or no presence at all depending on the corpus).

Such parallel trends in both languages strongly suggest that speaker's involvement is an intrinsic trait in the core meaning of connectors.

4.3 Overall view on genre variation

The graph below shows regularities in relation to corpus genre in both languages.

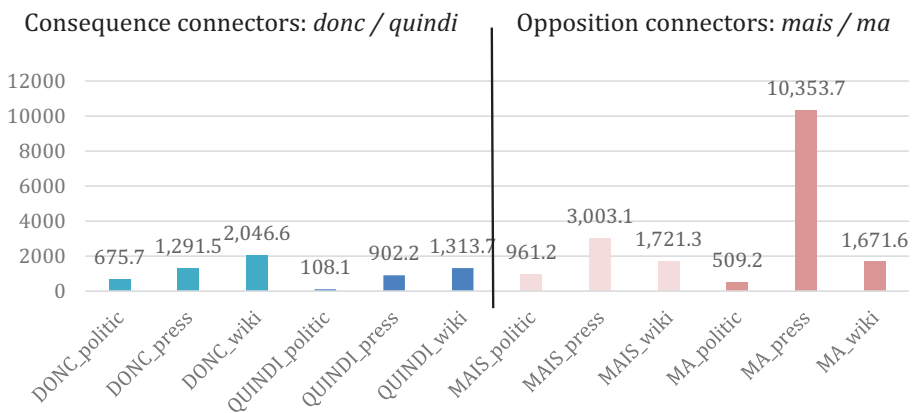


Figure 7: Behavior of connectors standing out for their association with SiFs within their class, in French and Italian.

In both languages, the two connectors that stand out within their category (*donc*, *quindi*, *mais*, *ma*) show parallel behaviors in their variation according to discourse genre: in French as well as in Italian, the connectors signaling the relation of consequence and of opposition both have lower association scores with SiFs in political speech. This difference between political discourse and the two other corpora seems at first sight puzzling, since the argumentative dimension of such a discourse is expected to be very pronounced. However, it can be explained in relation to the statistical method used: in

a discourse that globally conveys a very high degree of argumentation, we can expect the speaker to be more naturally inclined to use SiFs. If that is the case, the combination of connectors with SiFs is flattened, in comparison with less argumentative discourse, where the association of such forms, being less expected, emerges more strongly. Another strong parallelism between French and Italian concerns, on the one hand, encyclopedic discourse, and on the other hand the press. A higher score for connectors of consequence is found in Wikipedia, while a higher score for connectors of opposition is found in the press corpora. How can this variation be accounted for? For the reason just outlined above (the greater visibility of rare associations within a certain discourse genre), our assumption is that the high specificity scores of *donc* and *quindi* with SiFs in encyclopedic discourse are due to the 'neutral' nature of this type of discourse, making associations with SiFs noticeable. Conversely, the higher score of the opposition connector *ma/mais* in the press vs encyclopedic/political discourse can be put in relation with the dialogical value due to the concession relation that *ma* or *mais* establishes, which is particularly relevant in the press.

The overview of our three discourse genres allows us to see two manifestations of SI taking shape. One, which is relevant especially in the press, characterizes opposition connectors. These are related to the concession value, which involves a stronger commitment by the speaker of the content introduced by *mais* or *ma*. The other is relevant particularly to encyclopedic discourse, and is characteristic of consequence connectors. It is related to the ability of *donc* or *quindi* to put in relation not only states of affairs but also judgments, which necessarily involve a stronger commitment by the speaker. Since these differences between corpora and connectors concern, in a parallel manner, both languages, they can be considered as revealing very stable features of *donc*, *quindi*, and *mais*, *ma* – all conveying a strong SI feature due to hierarchization property for the latter and the judgment linking property for the former.

5 Conclusions and further research perspectives

In this chapter, we have proposed a review of the characterization of DMs at a semantic-pragmatic level of analysis, in articulation with the other levels of linguistic analysis. We have shown, for the subclass of connectors, how a textometric approach, based on the notion of statistically significant associations, can help us delineate a new outlook on the semantics and pragmatics of these forms, which goes beyond their relational properties: in this perspective, the connector can be a privileged conveyor of the involvement of the speaker. This characteristic, defining the very semantics of the connector regardless of the logical relation it conveys, results in an affinity with forms that also manifest, in different manners, the speaker's involvement. The analyses and statistical calculations carried out allowed us to delineate stable trends in the behavior of some connectors, with striking parallelisms between the degrees of specificities with

SiFs within the same genre in both Romance languages. This parallel behavior could not have been highlighted within a qualitative approach.

Such trends reveal:

- i. that measuring the extent to which the SI feature of connectives transpires through co-occurrence with other forms conveying it allows us to delineate groupings of connectives which are different from those based on the semantic relation they convey;
- ii. analogies and differences concerning discourse genres can be accounted for through the connector's particular ways of conveying SI: for instance, the behavior of *mais/ma*, *donc/quindi* is different in the press vs in encyclopedic discourse, as each pair attracts SiFs to a higher or lower degree.

The present research would benefit from an extension in the following aspects:

1. Deepening and specifying the SiFs paradigm, so as to delineate a constellation of forms defining more precisely the behavior of connectors through specific associations.
2. Broadening the research on the associations between DMs and SiFs not only to the other Romance languages but also to non-genetically related languages, while also maintaining an eye on genre differences.
3. Venturing into an application of this analysis to the category of DMs in spoken language, examining specific associations in speech corpora, to see if SI can be the defining link between different types of items, brought together under the label of *discourse markers*.

References

- Abeillé, Anne/Godard, Danièle (dir.) (2021), *La Grande Grammaire du français*, Paris, Actes Sud Imprimerie nationale Éditions, online edition: <https://www.grandegrammairedufrançais.com>.
- Anscombe, Jean-Claude (2002), “*Mais*”/“*pourtant*” dans la contre-argumentation directe. *Raisonnement, généricité, et lexicque*, *Linx* 46, 115–131.
- Anscombe, Jean-Claude/Ducrot, Oswald (1976), *L'argumentation dans la langue*, *Langages* 42, 5–27.
- Anscombe, Jean-Claude/Ducrot, Oswald (1977), *Deux mais en français?*, *Lingua* 43, 23–40.
- Authier-Revuz, Jacqueline (1995), *Ces mots qui ne vont pas de soi. Boucles réflexives et non-coïncidences du dire*, Paris, Larousse.
- Bach, Kent (1997), *The semantics-pragmatics distinction. What it is and why it matters*, *Linguistische Berichte* 8, 33–50.
- Bach, Kent (1999), *The myth of conventional implicature*, *Linguistics and philosophy* 22, 327–366.
- Bazzanella, Carla (1995), *I segnali discorsivi*, in: Lorenzo Renzi/Giampaolo Salvi/Anna Cardinaletti (edd.), *Grande grammatica italiana di consultazione*, vol. 3, Bologna, il Mulino, 225–257.
- Bazzanella, Carla (2006), *Discourse markers in Italian. Towards a “compositional” meaning*, in: Kerstin Fischer (ed.), *Approaches to discourse particles*, Amsterdam, Elsevier, 449–464.
- Bernard, Timothée/Danlos, Laurence (2019), *Connecteurs de discours centraux et périphériques*, *Linguisticae investigationes* 42:2, 262–297.

- Berrendonner, Alain (1990), *Pour une macro-syntaxe*, Travaux de linguistique 21, 25–36.
- Berrendonner, Alain (2003), *Éléments pour une macro-syntaxe. Actions communicatives, types de clauses, structures périodiques*, in: Antonietta Scarano (ed.) *Macro-syntaxe et pragmatique. L'analyse linguistique de l'oral. Actes du colloque international de Florence, avril 1999*, Roma, Bulzoni, 93–110.
- Blakemore, Diane (1987), *Semantic constraints on relevance*, Oxford, Blackwell.
- Blakemore, Diane (1992), *Understanding utterances. An introduction to pragmatics*, Oxford, Blackwell.
- Blakemore, Diane (2002), *Relevance and linguistic meaning. The semantics and pragmatics of discourse markers*, Cambridge, Cambridge University Press.
- Blanche-Benveniste, Claire/Bilger, Mireille/Rouget, Christine/Van den Eynde, Karel (1990), *Le français parlé. Études grammaticales*, Paris, CNRS.
- Blumenthal, Peter (2002), *Profil combinatoire des noms. Synonymie distinctive et analyse contrastive*, Zeitschrift für französische Sprache und Literatur 112, 115–138.
- Borreguero Zuloaga, Margarita (2014), *Focalizadores aditivos escalares y posición enunciativa. Un estudio contrastivo español / italiano*, Philologia hispálica 28/3–4, 13–57.
- Cabedo Nebot, Adrian (2014), *On the delimitation of discursive units in colloquial Spanish. Val.Es.Co. application model*, in: Salvador Pons Bordería (ed.), *Discourse segmentation in Romance languages*, Amsterdam/Philadelphia, John Benjamins, 157–184.
- Carel, Marion/Ducrot, Oswald (1999), *Le problème du paradoxe dans une sémantique argumentative*, Langue française 123, 6–26.
- Caron, Jean (1984), *Les opérateurs discursifs comme instructions de traitement*, Verbum 7, n° spécial: L'interlocution, 149–164.
- Cresti, Emanuela (2018), *The illocution-prosody relationship and the information pattern in spontaneous speech according to the Language into Act Theory*, Linguistik Online 88:1, 33–62.
- Cuenca, Maria Josep (2020), *Defective connective constructions. Some cases in Catalan and Spanish*, Corpus pragmatics 4, 423–448.
- Cuenca, Maria-Josep/Degand, Liesbeth (2022), *Discourse markers in interaction. From production to comprehension*, Berlin/Boston, De Gruyter Mouton.
- Culioli, Antoine (1990), *Pour une linguistique de l'énonciation. Opérations et représentations*, vol. 1, Paris, Ophrys.
- Dancygier, Barbara/Sweetser, Eve (2009), *Constructions with “if”, “since”, and “because”. Causality, epistemic stance, and clause order*, in: Elizabeth Couper-Kuhlen/Bernd Kortmann (edd.), *Cause – condition – concession – contrast. Cognitive and discourse perspectives*, Berlin/New York, De Gruyter Mouton, 111–142.
- Dargnat, Mathilde (2020), *Subjectivité et projection. Le cas des particules discursives*, SHS Web of Conferences, 78.
- De Cesare, Anna-Maria/Borreguero Zuloaga, Margarita (2014), *The contribution of the Basel Model to the description of polyfunctional discourse markers. The case of It. “anche”, Fr. “aussi”, and Sp. “también”*, in: Salvador Pons Bordería (ed.), *Discourse segmentation in Romance languages*, Amsterdam/Philadelphia, John Benjamins, 55–90.
- De Santis, Cristiana/Ferrari, Angela/Frenguelli, Gianluca/Gatta, Francesca/Lala, Letizia/Mazzoleni, Marco/Prandi, Michele (2014), *Le relazioni logico-sintattiche. Teoria, sincronia, diacronia*, Roma, Aracne.
- Debaisieux, Jeanne-Marie (2016), *Toward a global approach to discourse uses of conjunctions in spoken French*, Language sciences 58, 79–94.
- Degand, Liesbeth/Cornillie, Bert/Pietrandrea, Paola (2013), *Discourse markers and modal particles. Categorization and description*, Amsterdam/Philadelphia, John Benjamins.
- Dostie, Gaétane (2004), *Pragmaticalisation et marqueurs discursifs. Analyse sémantique et traitement lexicographique*, Bruxelles, De Boeck/Duculot.
- Dostie, Gaétane/Lefeuve, Florence (dir.) (2017), *Lexique, grammaire, discours. Les marqueurs discursifs*, Paris, Champion.

- Ducrot, Oswald (1984), *Le dire et le dit*, Paris, Les éditions de Minuit.
- Ducrot, Oswald (2010), *La pragmatique intégrée*, in: Martine Bracops (ed.), *Introduction à la pragmatique*, Bruxelles, De Boeck/Duculot, 161–208.
- Ducrot, Oswald, et al. (1980), *Les mots du discours*, Paris, Les éditions de Minuit.
- Esposito, Giorgia (2020), *Il modello di Basilea per l'analisi di "también". Uno studio contrastivo con l'italiano*, Cuadernos de filología italiana 27, 77–107.
- Fagard, Benjamin/Charolles, Michel (2018), *Ailleurs, d'ailleurs, par ailleurs. De l'espace à l'humain, de l'humain au discours*, Journal of French language studies, 1–25.
- Fernandez, Jocelyne (1994), *Les particules énonciatives dans la construction du discours*, Paris, PUF.
- Ferrari, Angela (2014), *The Basel Model for paragraph segmentation: the construction units, their relationships and linguistic indications*, in: Salvador Pons Bordería (ed.), *Discourse segmentation in Romance languages*, Amsterdam/Philadelphia, John Benjamins, 23–54.
- Ferrari, Angela/Cignetti, Luca/De Cesare, Anna-Maria/Lala, Letizia/Mandelli, Magda/Ricci, Claudia/Roggia, Enrico (2008), *L'interfaccia lingua-testo. Natura e funzioni dell'articolazione informativa dell'enunciato*, Alessandria, Edizioni dell'Orso.
- Ferrari, Angela/Pecorari, Filippo (2018), *Sintassi, punteggiatura e interpretazione dei connettivi. Il caso di "dunque" e di "tuttavia"*, *Lingua e stile* 53:2, 219–245.
- Ferrari, Angela/Rossari, Corinne (1994), *De "donc" à "dunque" et "quindi". Les connexions par raisonnement inférentiel*, *Cahiers de linguistique française* 15, 7–49.
- Fraser, Bruce (1999), *What are discourse markers?*, *Journal of pragmatics* 31:7, 931–952.
- Fraser, Bruce (2009), *An account of discourse markers*, *International review of pragmatics* 1:2, 293–320.
- Ghezzi, Chiara/Molinelli, Piera (edd.) (2014), *Discourse and pragmatic markers from Latin to the Romance languages*, Oxford, Oxford University Press.
- Giacalone Ramat, Anna (2017), *The diachronic development of the Italian focus particle "almeno"*, *Archivio glottologico italiano* 102, 205–224.
- González, Montserrat (2015), *From truth-attesting to intensification. The grammaticalization of Spanish "la verdad" and Catalan "la veritat"*, *Discourse studies* 17:2, 162–181.
- Groupe Lambda-I (1975), *Car, parce que, puisque*, *Revue romane* 10, 248–280.
- Grupo VAL.ES.CO. (2014), *Las unidades del discurso oral. La propuesta Val.ES.Co. de segmentación de la conversación (coloquial)*, *Estudios de lingüística del español* 35, 13–73.
- Haegeman, Liliane (2004), *The syntax of adverbial clauses and its consequences for topicalization*, in: Martine Coene/Greet De Cuyper/Yves D'Hulst (edd.), *Current studies in comparative Romance linguistics*, 107, Antwerp, Antwerp University, 61–90.
- Hall, Alison (2007), *Do discourse connectives encode concepts or procedures?*, *Lingua* 117:1, 149–174.
- Halliday, Michael A.K./Hasan, Ruqaiya (1976), *Cohesion in English*, London, Longman.
- Hansen, Maj-Britt Mosegaard (1998), *The semantic status of discourse markers*, *Lingua* 104, 235–260.
- Hansen, Maj-Britt Mosegaard (2008), *Particles at the semantics/pragmatics interface. Synchronic and diachronic issues. A study with special reference to the French phasal adverbs*, Amsterdam, Elsevier.
- Hansen, Maj-Britt Mosegaard/Strudsholm, Erling (2008), *The semantics of particles. Advantages of a contrastive and panchronic approach. A study of the polysemy of French "déjà" and Italian "già"*, *Linguistics* 46:3, 471–505.
- Heiden, Serge (2010), *The TXM platform. Building open-source textual analysis software compatible with the TEI encoding scheme*, 24th Pacific Asia Conference on Language, Information and Computation, Nov 2010, Sendai, Japan, 389–398.
- Laforest, Marty (1992), *L'influence de la loquacité de l'informateur sur la production de signaux backchannel par l'intervieweur en situation d'entrevue sociolinguistique*, *Language variation and change* 4:2, 163–177.
- Lebart, Ludovic/Salem, André (1994), *Statistique textuelle*, Paris, Dunod.
- Lyons, John (1977), *Semantics*, Cambridge, Cambridge University Press.

- Macário Lopes, Cristina (2004), *Contributo para o estudo sincrónico dos marcadores discursivos “quer dizer”, “ou seja” e “isto è” no português europeu contemporâneo*, *Diacrítica* 28, 33–50.
- Macário Lopes, Cristina/Sousa, Sara (2014), *The discourse connectives “ao invés” and “pelo contrário” in contemporary European Portuguese*, *Journal of Portuguese linguistics* 3:1, 3–27.
- Manzotti, Emilio (1998), *L'esempio. Natura, definizioni, problemi*, *Cuadernos de filología italiana* 5, 99–123.
- Marconi, Diego/Bertinetto, Pier Marco (1984), *Analisi di “ma” (parte prima: Semantica e pragmatica)*, *Lingua e stile* 19:2, 223–258.
- Mandelli, Magda (2008), *“In effetti” nel testo*, in: Emanuela Cresti (ed.), *Prospettive nello studio del lessico italiano. Atti del IX Congresso SILFI*, Firenze, FUP, 439–444.
- Mazzoleni, Marco (2015), *Connettori, grammatica e testi. “Ma” e “(ben)si” tra costrutti avversativi, sostitutivi e preconcensivi*, in: Angela Ferrari/Letizia Lala/Roska Stojmenova (edd.), *Testualità. Fondamenti, unità, relazioni*, Firenze, Cesati, 171–188.
- Mihatsch, Wiltrud (2009), *The approximators French “comme”, Italian “come”, Portuguese “como” and Spanish “como” from a grammaticalization perspective*, in: Corinne Rossari/Claudia Ricci/Adriana Spiridon (edd.), *Grammaticalization and pragmatics. Facts, approaches, theoretical issues*, (Studies in pragmatics 5), Bingley, Emerald, 65–91.
- Moeschler, Jacques (1989), *Pragmatic connectives, argumentative coherence and relevance*, *Argumentation* 3, 321–339.
- Moneglia, Massimo/Cresti, Emanuela (2006), *C-ORAL-ROM. Prosodic boundaries for spontaneous speech analysis*, in: Yuji Kawaguchi, Susumu Zaima, Toshiro Takagaki (edd.), *Spoken language corpus and linguistics informatics*, Amsterdam/Philadelphia, Benjamins, 89–114.
- Nølke, Henning/Fløttum, Kjersti/Norén, Coco (edd.) (2004), *ScaPoLine. La théorie scandinave de la polyphonie linguistique*, Paris, Kimé.
- Paillard, Denis (2015), *La notion de particules énonciatives en français. Analyse comparative de “tout de même” et de “quand même”*, in: Margarita Borreguero Zuloaga/Sonia Gómez-Jordana Ferary (edd.), *Les marqueurs du discours dans les langues romanes. Une approche contrastive*, Limoges, Lambert-Lucas, 165–178.
- Pincemin, Bénédicte (2011), *Sémantique interprétative et textométrie – version abrégée*, *Corpus* 10, 259–269.
- Pons Bordería, Salvador (ed.) (2014), *Discourse segmentation in Romance languages*, Amsterdam/Philadelphia, John Benjamins.
- Raso, Tommaso (2014), *Prosodic constraints on discourse markers*, in: Tommaso Raso/Heliana Mello (edd.), *Spoken corpora and linguistic studies*, Amsterdam/Philadelphia, John Benjamins, 411–467.
- Raso, Tommaso/Vieira, Marcelo Augusto (2016), *A description of dialogic units/discourse markers in spontaneous speech corpora based on phonetic parameters*, *CHIMERA. Romance corpora and linguistic studies* 1 (2016), 221–249.
- Rastier, François (2008), *Que cachent les “données textuelles”?*, in: Serge Heiden/Bénédicte Pincemin (edd.), *Actes des 9^{es} Journées internationales d'Analyse statistique des Données Textuelles*, vol. 1, Lyon, Presses Universitaires de Lyon, 13–26.
- Remberger, Eva-Maria (2021), *Discourse and pragmatic markers in the Romance languages*, in: *The Oxford research encyclopedia of linguistics*, Oxford, Oxford University Press (online), available at: <https://doi.org/10.1093/acrefore/9780199384655.013.675>.
- Romero-Trillo, Jesús (2006), *Discourse markers*, in: Keith Brown (ed.), *Encyclopedia of language and linguistics*, Oxford, Elsevier, 2–639.
- Rossari, Corinne (1994), *Les opérations de reformulation. Analyse du processus et des marques dans une perspective contrastive français-italien*, Bern, Lang.
- Rossari, Corinne/Beaulieu-Masson, Anne/Cojocariu, Corina/Razgouliaeva, Anna (2004), *Autour des connecteurs. Réflexions sur l'énonciation et la portée*, Bern, Lang.
- Rossari, Corinne/Beaulieu-Masson, Anne/Cojocariu, Corina/Ricci, Claudia (2007), *Les moyens détournés d'assurer son dire*, Paris, Presses de l'Université de Paris-Sorbonne.

- Rossari, Corinne/Jayez, Jacques (2000), *The semantics of pragmatic connectives in tree adjoining grammar. The French “donc” example*, in: Anne Abeillé/Owen Rambow (edd.), *Tree adjoining grammars*, Stanford, CSLI, 249–269.
- Roulet, Eddy/Auchlin, Antoine/Moeschler, Jacques/Rubattel, Christian/Schelling, Marianne (1987), *L’articulation du discours en français contemporain*, Bern, Lang.
- Roulet, Eddy/Fillietaz, Laurent/Grobet, Anne, avec la collaboration de Marcel Burger (2001), *Un modèle et un instrument de l’organisation du discours*, Bern, Lang.
- Rubattel, Christian (1981), *Remarques sur les performatifs fonctionnant comme marqueurs d’interactivité*, Cahiers de linguistique française 2, 89–92.
- Sabatini, Francesco/Coletti, Vittorio (2008), *Il Sabatini-Coletti. Dizionario della lingua italiana, diretto da Francesco Sabatini, Vittorio Coletti*, Milano, Rizzoli Larousse.
- Schiffrin, Deborah (1987), *Discourse markers*, Cambridge, Cambridge University Press.
- Spengler, Nina de (1980), *Première approche des marqueurs d’interactivité*, Cahiers de linguistique française 1, 128–148.
- Squartini, Mario (2014), *The pragmaticalization of “already” in Romance. From discourse grammar to illocution*, in: Chiara Ghezzi/Piera Molinelli (edd.), *Discourse and pragmatic markers from Latin to the Romance languages*, Oxford, Oxford University Press, 190–210.
- Steuckardt, Agnès (2018), *Les marqueurs de reformulation formés sur “dire”. Exploration outillée*, Langages 212, 17–33.
- Sweetser, Eve (1990), *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*, Cambridge, Cambridge University Press.
- Swiggers, Pierre (1995), *Synchronie, diachronie et pragmatique. D’une dichotomie de la langue à l’interaction dans la langue*, Langue française 107, 7–24.
- Vincent, Diane (1991), *Quelques études sociolinguistiques de particules du discours*, Revue québécoise de linguistique théorique et appliquée 10:3, 41–60.
- Visconti, Jacqueline (2000), *I connettivi condizionali complessi in italiano e in inglese. Uno studio contrastivo*, Alessandria, Edizioni Dell’Orso.
- Visconti, Jacqueline (2009), *From “textual” to “interpersonal”. On the diachrony of the Italian particle “mica”*, Journal of pragmatics, 41:5, 937–950.
- Waltereit, Richard (2006), *The rise of discourse markers in Italian. A specific type of language change*, in: Kerstin Fischer (ed.), *Approaches to discourse markers*, Amsterdam, Elsevier, 61–76.
- Wilson, Deirdre/Sperber, Dan (2012), *Meaning and relevance*, Cambridge, Cambridge University Press.
- Zafiu, Rodica (2018), *Epistemic and evidential markers in the rhetorical context of concession*, Journal of pragmatics 128, 116–127.