Linguistic Typology at the Crossroads 5-1 (2025): 181-234 ISSN 2785-0943

Beyond the topic-marking discourse function of Left Dislocations.

Evidence from Italian, Spanish, and English spoken corpora

DORIANA CIMMINO[1], VALENTINA SACCONE[2]

[1]UNIVERSITY OF VENICE

[2]UNIVERSITY OF FLORENCE

Submitted: 24/04/2025 Revised version: 3/07/2025

Accepted: 20/08/2025 Published: 13/11/2025

Abstract

Left Dislocations (LDs) are generally regarded as carrying a Topic/Comment information partition and performing an overall topic-marking discourse function. However, when confronted with actual texts, the description of both information structure and discourse functions fails to account for all the occurrences. In this paper, we present a cross-linguistic model for the description of LDs discourse functions, checked against Italian, English, and Spanish spoken corpora. Our results support Prince's view that discourse functions cannot be simply derived by form, demonstrating that prosody, syntax, information structure and discourse functions do not always align. Crucially, we show that LDs are not only topic-marking and that the function performed are shaped by context. Assuming a broader, non-utterance centered perspective, we define LDs as prominence cues used by speakers to signal a disruption in the ongoing discourse, the nature of this discourse prominence being dynamic and evolving as the text unfolds.

Keywords: left dislocation; topicality; prosody; discourse functions; cross-linguistic; spontaneous speech.

1. Introduction: On the topic-marking function of Left Dislocations

Left Dislocations (henceforth LDs) are generally regarded as syntactic structures in which a usually postverbal constituent is found in first position. Consider the following fictitious examples of LDs in Italian (ita; Indo-european, Romance), Spanish (spa; Indo-european, Romance), and English (eng; Indo-european, Germanic) – the languages of interest in this study:

(1)	(It.)	***La***	***torta,***	*Maria*	***la***	*mangi-a.*
		ART.DEF	cake	Maria	3SG.F.OBJ	eat-3SG.PRS
(2)	(Sp.)	***El***	***pastel,***	*Maria*	***lo***	*com-e.*
		ART.DEF	cake	Maria	3SG.M.OBJ	eat-3SG.PRS

(3) (En.) **The cake**, Maria eats **it**.

These structures are defined by Lambrecht (2001: 1050) based on four criteria: "(i) extra-clausal position of a constituent, (ii) possible alternative intra-clausal position, (iii) pronominal coindexation, (iv) special prosody". These criteria are claimed to apply cross-linguistically in prototypical instances, although only the first is considered necessary (though not sufficient). In Lambrecht's framework, the extra-clausal position of the left-dislocated constituent (criterion i) is directly related to its relevance within the structure and its prosodic prominence.

Given the extra-clausal position of the left-dislocated constituents 'La torta/ The cake/ El pastel' (in 1 to 3 above), Lambrecht argues that their relationship to the rest of the structure is a "kind of relevance relation between the TOP (Topic) entity and the proposition" (Lambrecht 2001: 1058). This relation involves both semantic traits, namely aboutness, identifiability, and definiteness, and contextual constraints, such as givenness and saliency. Thus, the left-dislocated constituent functions as a topic referent "matter of standing current interest and concern" (Lambrecht 1994: 121).

Other scholars contend that, in English and cross-linguistically, traits like topicality or relevance are not always associated with left-marked constituents; in fact, they argue that LDs are only generally discourse relevant (Birner & Ward 1998; Prince 1998). Based on corpus investigations, they claim that the association between syntactic and pragmatic features is "arbitrary" and "language-specific" (Prince 1998: 299).

Similarly, in some strands of Italian linguistic literature, the salience of the left-dislocated constituent is not necessarily tied to aboutness topicality or givenness. Berruto (1985: 227), among others (see also Sornicola 1981), defines it as the "center of interest or empathic focus," grounded in the speaker's "psychological reality" as reflected in the text. In the description of discourse functions, Ferrari (2003) and Cimmino (2024) emphasize the importance of both left and right co-texts and textual dimensions such as thematic progression, logic relations, and textual salience, besides the information structure of the LDs.

In classical Spanish literature as well (Gili Gaya 1961), LDs have been claimed to highlight the affective condition of the speaker or the psychological subject of the utterance (also in later works, such as Fernández Ramírez 1986). According to this view, LDs do not merely anticipate a thematic referent but serve to highlight a particular constituent in the discourse (Contreras 1978). However, in more recent works, the left-dislocated constituents are described as topic elements à la Lambrecht (Casielles Suárez 2003; Sedano 2012).

In Lambrecht's view, the syntactic independence of the left-dislocated constituent is also linked to its phonological realization (criterion iv). These constituents are realized as "independent phonological units," carrying a "degree of prosodic prominence" that depends on the "speech situation" (Lambrecht 2001: 1071), that is on discourse factors, such as the accessibility of the left-dislocated constituent. In a more radical approach, Frascarelli and Hinterhölzl (2007) and Frascarelli (2007) claim that the prosodic contour of LDs varies according to syntactic and functional features of the left-dislocated constituent. They identify three subtypes in which the left-dislocated constituent functions as a Topic performing different roles: aboutness-shift (A-Topic), contrastive (C-Topic), and familiar/given (G-Topic). Cross-linguistically, each Topic subtype is said to have a specific prosodic realization and a distinct function. This is crucial because, in this approach, the discourse functions of LDs are directly derived from formal traits such as prosody and syntactic position.

In both syntactically and functionally oriented strands of the classical literature on LDs, the equation between their prosodic and syntactic form and information function is generally accepted. LDs are typically described as carrying a Topic/Comment information structure and, as a result of this, fulfilling an overall topic-marking discourse function (for Italian: Benincà et al. 1988; Berretta 2002; Ferrari et al. 2008; for English: Reinhart 1981; Lambrecht 1994, 2001; for Spanish: Sedano 2012).

In the following classic example from Lambrecht (2001), the LD “Veal, it is worse” is used to introduce a new referent and to comment on it:

(4) (adapted from Lambrecht 2001: 1074)
Husband and wife at dinner table; H. looks at food on his plate:
‘H: It has no taste, this chicken.’
W: **VEAL**, **it** is worse.’

However, studies based on actual texts (e.g., for Italian: Sornicola 1981; Berruto 1985; Cimmino & Panunzi 2017; Garassino & Jacob 2018; for English: Prince 1998; Tizón-Couto 2012; for Spanish: Silva-Corvalán 1984; Downing 1997) have shown that this traditional account of LDs fails to capture the full range of their discourse functions. Indeed, LDs have been shown to operate not only in the thematic dimension but also at the logical level (Cimmino 2024) and in the interactional dimension of texts (Hidalgo Downing 2006). The main criticism to a form-function equation view is that the notion of an aboutness Topic used in classical literature is not always relevant in the analysis of real examples. Consider the following extract from Prince (1998), in which the speaker, a rose dealer, comments on the impossibility of cutting certain kinds of roses due to commercial restrictions:

(5) Sure, on the ones we can find locally, we can start cutting ourselves, and we’ll probably DO that with Lilac Charm. But **those others**, … **they**’ve got a lock on them; AND they are vicious people! Not many know this ... but they tried like hell to get S.A. put out of business. (Prince 1998: 287)

The left-dislocated constituent “those others”, referring to the restricted cuttings, does not serve the function of introducing a new referent. As Prince herself notes, the referent is not recalled in the subsequent discourse. Instead, the discourse shifts to focus on the “vicious people” preventing the cuttings; the dislocated referent is effectively abandoned.

Based on Frascarelli and Hinterhölz (2007: 87) and Büring (2016: 65) definitions of contrastive topics, Prince's example could be interpreted as such. In the example, the left-dislocated constituent “those others” is implied to be an alternative to “Lilac Charm”. Ultimately, looking at the relation between the co-text and LD’s information structure is necessary to account for their discourse function, avoiding context-free generalizations.

Following this last line of investigation, in this contribution we disentangle the functional account of LDs from the notion of topicality, showing that, although significant, the topic management functions of LDs are not the only ones at play in actual examples. We show that in spontaneous speech in Italian, European Spanish, and American English, the assumed prosodic and pragmatic relevance of the left constituent (Lambrecht's criterion iv) is not systematically associated with the syntactic traits of extra-clausal position (criterion i), alternative intra-clausal position (ii), and pronominal coindexation (iii). Furthermore, even when all four traits are simultaneously present, there is no consistent correlation between the pragmatic relevance of the left constituent and the assumed topic-marking discourse function of LDs.

Since, in real data, the prominence of the left constituent can serve a variety of purposes – not limited to establishing topicality – we argue that LDs' discourse functions are better understood through the broader notion of discourse prominence (von Heusinger & Schumacher 2019). In this contribution, we therefore describe LDs as prominence cues used by speakers to signal to the interlocutor a disruption in the ongoing discourse, not only at a topic level but also at a logic and interpersonal level.

In the remainder of the paper, we first present our data (§2.1) and the theoretical approach adopted (§2.2). We show that the use of spontaneous speech corpora, enriched with prosodic annotation, allows for the collection and analysis of authentic and naturalistic data in context, within a theoretical framework that integrates textual understanding. Subsequently, we present the results of the analyses at the interface between syntactic and prosodic traits (§3.1), and between prosodic traits and discourse functions (§3.2). We demonstrate that correlations between syntax, prosody, and information structure in LDs are language-dependent, and that discourse functions are ultimately determined by the context of use. We therefore provide further evidence for Prince's argument that discourse functions cannot be derived from form. Finally, we discuss the findings in light of von Heusinger and Schumacher's (2019) concept of discourse prominence (§4).

2. Materials and Methods

2.1. Corpora for interlinguistic comparison and data collection

The cross-linguistic analysis relates to three mini-corpora of spontaneous speech; the languages under consideration are Italian, European Spanish, and American English.

The mini-corpora were selected for their suitability to cross-linguistic comparison of information structure, comparability, and shared design principles. They are part of the C-ORAL project, firstly begun with the C-ORAL-ROM (Cresti & Moneglia 2005; Cresti et al. 2005), which gathers the four main Romance languages, i.e., French, Italian, Portuguese, and Spanish, then enlarged with non-Romance languages such as American English.

Each recording is transcribed following the CHAT-LABLITA format (Moneglia & Cresti 1997; MacWhinney 2000; see §2.2.1 below), an orthographic transcription enriched with prosodic-pragmatic annotations and the text-sound alignment. The process includes segmenting the speech flow into prosodically terminated sequences (TS) and prosodic-information units according to the Language into Act Theory[1] (L-AcT, Cresti 2000; Moneglia & Raso 2014). Within this framework, the TS is considered the basic unit of speech (Izre'el et al. 2020), defined as the smallest stretch of speech that is both pragmatically and prosodically interpretable in isolation. Pragmatic interpretability is ensured by the presence of at least one illocutionary unit, while the prosodic interpretability is marked by a boundary that conveys a sense of conclusion.

More in detail, the Italian mini-corpus (IT) comprises 20 texts (37355 words; 5663 terminated sequences; nearly 4h 37m) (Panunzi & Mittmann 2014). The Spanish mini-corpus (SP) comprises 37 texts (40586 words; 6445 terminated sequences; nearly 4h 7m) (Nicolás & Lombán 2018). The American English mini-corpus (AE) comprises 20 texts (26470 words; 3452 terminated sequences; nearly 2h 27m) (Ramos 2015; Cavalcante 2015; Cavalcante & Ramos 2016). The latter corpus was created by selecting texts of the Santa Barbara Corpus (Du Bois et al. 2000–2005) following the guidelines used for constructing the C-ORAL-ROM mini-corpora to ensure mutual comparability across resources. The AE mini-corpus is balanced in terms of communicative situations and information structure of TS[2]. Though it is smaller in

[1] Examples in the following sections include the original corpus ID, structured as follows: language (i = Italian, e = Spanish, a = American English); communicative context (fam = family, pub = public); event type (mn = monologue, dl = dialogue, cv = conversation); two digit recording number; and TS number, preceded by an underscore. For example, ifamcv01_3 refers to an Italian (i) family context (fam), conversation (cv) with a registration number of 01, and the third TS.

[2] Regarding information structure, the AE corpus is balanced to maintain a consistent ratio between simple and compound utterances. This allows for a relative quantitative comparison of the mini-corpora, despite differences in the absolute number of TS, in terms of the internal information articulation within the TSs.

overall size of Terminated Sequences compared to the other datasets, the ratio of words to TSs is similar across the three resources.

See Table 1 for a summary of the three mini-corpora.

Mini-corpus	Texts	Words	Terminated Sequences	Duration
Italian (IT)	20	37355	5663	4h 37m
Spanish (SP)	37	40586	6445	4h 7m
American English (AE)	20	26470	3452	2h 27m

Table 1: Size of the three mini-corpora.[3]

From the selected mini-corpora, we manually collected all occurrences of LDs based on a linear syntactic definition.[4] We defined LDs as syntactic structures in which the front constituent is extra-clausal and is reduplicated by a co-referential pronoun. This working definition remains valid cross-linguistically, despite the differences in word order and pronoun system between the languages object of scrutiny. As is known, Italian and Spanish are pro-drop languages, in which the subject can occupy both a pre-verbal and post-verbal position; moreover, the pronoun system includes both free and clitic instances (Benincà et al. 1988, for Italian; Fernández Ramírez 1986, for Spanish). On the contrary, English only shows free pronouns and the expression of the subject is obligatory (Biber et al. 1999). To avoid an unbalanced collection of data, we considered both clitic and free pronoun in all languages; hence, we succeeded in potentially including the whole range of dislocated constituents, from subject constituents (resumed by free pronouns in all languages) to direct and indirect objects (resumed by clitic free pronouns only in English).

The fact that the three corpora consist of spontaneous speech events enabled us to analyze authentic and naturalistic data, rather than relying on artificial examples. The prosodic tagging, which marks both terminal and non-terminal boundaries, allows us

[3] The Italian and the Spanish mini-corpora are freely available through the online database IPIC (www.lablita.it/app/dbipic, Panunzi & Gregori 2012).

[4] Part of the Italian data had previously been extracted and used in Cimmino & Panunzi (2017); Cimmino (2023) e (2024); part of the English data had been used in Cimmino (2023). However, in the present study, the dataset has been enlarged for both languages and the respective corpora and data have been checked for their prosodic features, as described further below.

to select those occurrences of LDs in which the dislocated constituent and the co-referential pronoun appear within the same TS – that is, they are not separated by a terminal prosodic boundary (cf. unplanned instances of LDs in Sornicola 1981: 136; Ozerov, this volume).

In particular, we selected 137 occurrences of LD across the three languages. Table 2 shows the number of LD occurrences identified in each mini-corpus, along with their relative frequency calculated both as a percentage of the total number of words and as a percentage of the total number of terminated prosodic sequences (TS).

Mini-corpus	LD occurrences	% of total words	% of total TS
IT	56	0.15%	1.0%
SP	73	0.18%	1.1%
AE	8	0.03%	0.2%

Table 2: Occurrences of LD in the three mini-corpora.

Occurrences are not evenly distributed across the languages. As expected, American English shows a low number of occurrences, compared to Spanish and Italian. Corpora studies based on British English (see a.o. Biber et al. 1999: 957) have indeed highlighted that the average rate in conversation is 200 occurrences per million words (0.02%). This means that in a corpus of our size, one would be expected to find around 5 or 6 occurrences of LD; consistently, we found 8 occurrences.

2.2. A Functional and Textual Theoretical Approach

Occurrences of LDs in Italian, Spanish, and American English have been analyzed using a functional and textual approach. This approach was initially developed for analyzing LDs and Preposings in Italian and English journalistic texts and was later refined for spoken data as well.[5] The main tenet of this approach is that the functions of LDs must be investigated by considering their interaction with the text in which

[5] The approach has been firstly proposed in Cimmino (2017), preceded by the preparatory works Cimmino (2014, 2016) on written Italian and English LMSs. Then it has been fine-tuned in studies on both written and spoken Italian and English (Cimmino & Panunzi 2017; Cimmino 2020, 2023, 2024).

they occur, in particular, with their thematic, logic and dialogic dimensions and the prominence of the text (as we will show in details below). Studies on spoken Italian LDs (see, in particular, Duranti & Ochs 1979; Ferrari 2003) have shown that the topicality of the left constituent can be exploited to emphasize the salience of a referent in the thematic progression of the text. More radically, in the approach adopted here, we assume that LDs operate beyond the level of the utterance, regardless of whether the structure has a Topic/Comment articulation or not. We thus move away from an utterance-centered perspective in favor of a pragma-discourse approach, which considers discourse segments independently of the informational properties of the LDs under analysis (Kecskes 2012: 294).

Building on this first tenet is the second guiding principle of the present approach: the functions of LDs are described by examining the interactions and preferential associations among their syntactic, prosodic-information, and textual features in context. However, each feature is investigated separately. This perspective aligns with the previously noted lack of direct correlation between form and function in English LDs (Prince 1989). The analyses presented here are thus conducted in a way that disentangles the syntactic and prosodic-information features from the functional description, with functions conceived as discourse-level phenomena.

In the following sections, we elaborate on the theoretical assumptions and tools used to describe prosodic-information and discourse-functional aspects, in §2.2.1 and §2.2.2, respectively.

2.2.1. The prosodic-information interface

The prosodic-information features are analyzed based on the framework of Language into Act Theory (L-AcT), developed by Cresti (2000) (see also Moneglia & Raso 2014; Cresti & Moneglia 2018a). L-AcT is a pragmatic approach used to describe and analyze data collected from spoken corpora in various languages including Romance and non-Romance languages. It arises from an interest in spontaneous language and has its roots in the central role played by the illocutionary activity conceived by the speaker during the communicative events, and prosody, by means of which the speaker produces, and the hearer interprets an utterance.

According to L-AcT, the utterance is the primary reference unit for spoken language. It is considered autonomous and conveys an illocutionary act. The

perception of an utterance in the speech flow is based on its prosodic profile,[6] and enclosed between two terminal prosodic boundaries[7] (Izre'el et al. 2020). Utterances can also be further segmented into smaller units (prosodic-information units), which are identified by non-terminal prosodic boundaries.[8] Although prosodic boundaries are primarily identified through tonal movements, they are in fact signaled by a complex configuration of prosodic and segmental cues, such as final lengthening, pitch reset, pause, pre-boundary creaky voice, intensity changes, and speech rate variations (Barbosa, 2008; Dilley et al., 1996; Mo & Cole, 2010; Mittmann & Barbosa, 2016). These cues often co-occur in complex patterns and may vary according to language-specific preferences and speech style (Barth-Weingarten, 2016; Hirst & Di Cristo, 1998). Importantly, the perception of boundaries – particularly at the utterance level – tends to be highly salient and naturally accessible to native speakers, as shown by high inter-rater agreement in segmentation tasks (Amir et al., 2004; Raso & Mittmann, 2009; Moneglia et al., 2010; Mello et al., 2012), even in spontaneous speech, and does not require any functional interpretation of the segmented unit. While no single cue is either necessary or sufficient on its own, the combination of multiple prosodic features provides robust perceptual signals for boundary detection.

L-AcT proposes an almost perfect isomorphy between the prosodic pattern (the sequence of prosodic units that make up utterances) and the information pattern (the way in which information is encoded in speech) (Cresti & Moneglia 2010). This isomorphy is intended to hold cross-linguistically and has been empirically validated in several languages.[9] The resulting information structure of an utterance is independent of the syntactic form of the locutive content.

Within an utterance, only one unit is perceived as autonomous and carries the illocution, and it is referred to as the Comment (COM), necessary and sufficient to

---

[6] Following the IPO methodology for perceptual intonation analysis, integrated by L-AcT ('t Hart et al. 1990), prosodic units are defined as intonation profiles – combination of one or more perceptually relevant pitch movements, shaped by f0 contour, its alignment with syllables, and the maximum and minimum points of f0.

[7] Marked with '//' in the transcriptions.

[8] Marked with '/' in the transcriptions.

[9] L-AcT was applied to the Italian LABLITA corpus (Cresti et al. 2018) and extensively tested in the annotation of Romance corpora, see C-ORAL-ROM (Cresti & Moneglia 2005), C-ORAL-BRASIL (Raso & Mello 2012), and Cor-DiAL (Nicolas Martinez 2012), Brazilian Portuguese (Panunzi & Mittmann 2014), Spanish (Nicolas Martinez & Lombán, 2018), and American English (Du Bois et al. 2000; Cavalcante & Ramos 2016).

form an utterance. Its prosodic contour can be described as a *root* unit, and varies widely along with its illocutionary value and the attitude with which it is conveyed (Mello & Raso 2011; Raso & Rocha 2015).

An utterance can be simple, consisting of only one information and prosodic unit (6), or complex composed of multiple units (7), where the COM is supported by other units to complete the utterance, each one with a recognizable prosodic contour (more or less fixed). Figures 1 and 2 show the prosodic contour and intonation pattern of the two examples.

(6) *FRA: *me* *lo* *ricordo*//$^{\text{COM}}$ (ifamdl12_55)
1SG.DAT 3SG.ACC.M remember:1SG.PRS
'*FRA: I remember it//$^{\text{COM}}$'

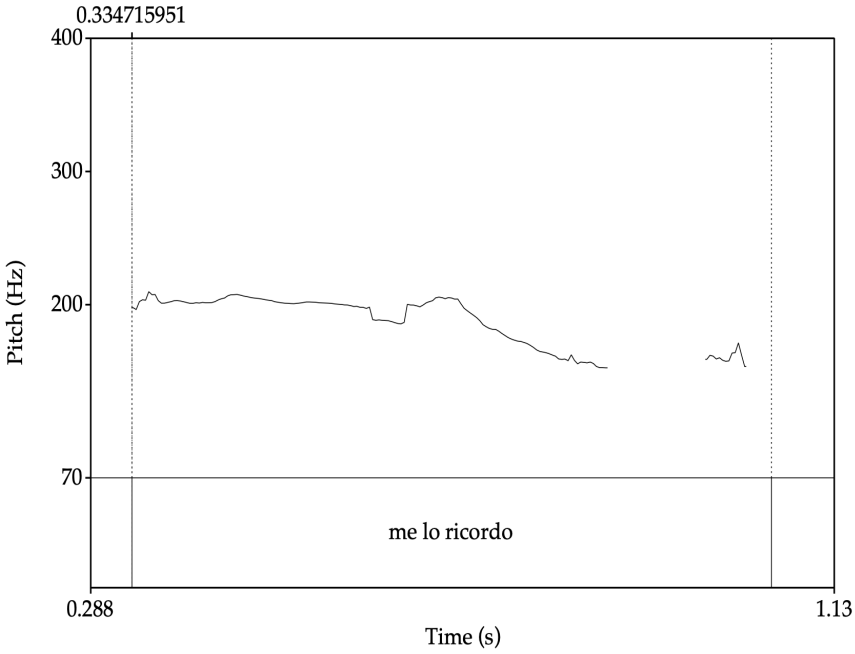

Figure 1: Prosodic contour of example (6) with transcription (female speaker).

(7) *VAL: *io*/$^{\text{TOP}}$ *ero* *parecchio* *nervosa*//$^{\text{COM}}$ (ifammn08_65)
I be:PST.1SG rather nervous:F.SG
'*VAL: I/$^{\text{TOP}}$ was quite nervous//$^{\text{COM}}$'

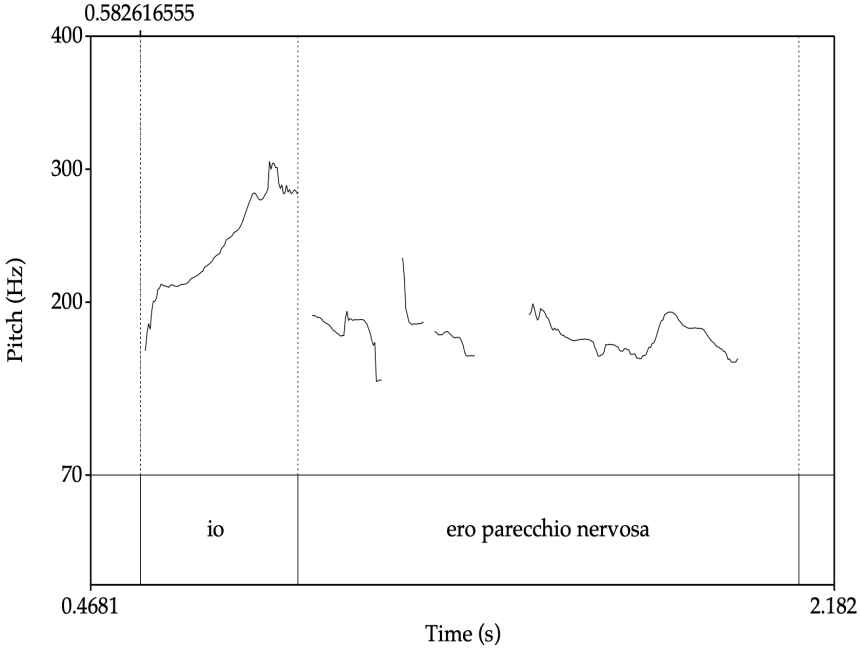

Figure 2: Prosodic contour of example (7) with transcription (female speaker).

The f0 contour of (6) is flat and continuous, falling at the end, with only one relevant pitch movement, whereas the contour in (7) is composed by more than one perceptually relevant movement. Specifically, Figure 2 illustrates an utterance comprising the necessary COM "ero parecchio nervosa" and a previous one that serves as the Topic "io", with a distinct rising pitch movement separated by a pitch reset from the subsequent unit.[10]

The COM unit may be preceded by a Topic unit (TOP). Following Moneglia & Raso (2014), we assume that the TOP provides the field of application for the illocutionary force of the Comment; it selects a domain of pragmatic relevance for the illocution. That is, utterances without a TOP necessarily refer to a contextually given domain (Hockett 1958; Cresti 2000).

The Topic unit is the main means of structuring information with respect to the Comment[11] (Cresti & Moneglia 2018 a, b). Regarding its distribution, it always precedes the COM and has a *prefix* prosodic contour ('t Hart et al. 1990).

[10] Pitch modulation marks boundaries between adjacent units, while perceived duration and volume also aid speech segmentation.

[11] For a quantitative resume of the incidence of utterances with Topics in multilingual corpora, see Table 1 in Cresti & Moneglia (2018b: 36). In short: Italian, 15.35%; Brazilian Portuguese, 9%; Spanish, 13.9%; and American English, 10.1%.

Both *prefix* and *root* prosodic contours can consist of a preparation and a nucleus. The nucleus corresponds to the minimal prosodic contour required to perform the information unit; its contour can be a simple movement (rising/falling/holding), or it can involve several movements aligned to the syllables participating in the contour (Cresti & Moneglia 2023). As a result, it is possible to identify a prosodically prominent part in both the Topic and Comment, relevance of which is connected to their functional value (Saccone et al. 2023). Research carried out on Brazilian Portuguese, European Portuguese and Italian (Mittmann 2012; Rocha 2012; Cavalcante 2015; Firenzuoli & Signorini 2003) identifies different types of *prefix* units on the base of the pitch movement in the syllable(s) of the prosodic unit who carries the prosodic prominence. As fully described in Cavalcante (2015), "the process is based on the assumption that only those f0 variations that result from voluntary production by the speaker show perceptual – and hence informational – relevance ('t Hart et al. 1990; Firenzuoli 2003)". The Topic profiles were then classified in the three types sketched in Figure 3.

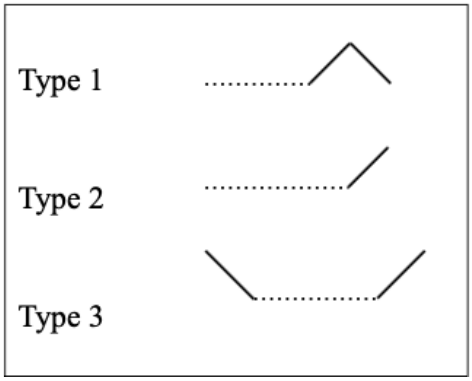

Figure 3: The 3 types of Topic prosodic forms (Cavalcante, 2015).

Summarizing Cavalcante (2015) and Firenzuoli and Signorini (2003), these types can be described as follows, based on empirical data from American English, Italian, Spanish, Portuguese (both European and Brazilian).[12]

- Type 1: rising-falling nucleus, aligned with the final stressed and any post-stressed syllables, with lengthening of the nuclear portion.
- Type 2: rising nucleus, aligned with the final stressed and any post-stressed syllables, also showing syllable lengthening.

[12] For a Kappa agreement test for TOP recognition, see Cavalcante et al. (2020).

- Type 3: two semi-nuclei (first and last syllables), possibly separated by a link; the first showing a falling and the second a (steeply) rising contour, both lengthened.

Apart from these types, some TOPs have been found exhibiting an overall flat f0 contour, considered a Type 3 subgroup (Cavalcante 2015), since they share characteristics with the first semi-nucleus, usually on a level that is higher than that of the final portion even if in a flatter form – hereinafter referred to as Flat.

Despite the distinction in f0 contour, studies carried out to date show no functional difference between the *prefix* prosodic forms, which are informally referred to as alloforms (Cavalcante 2020; Raso et al. 2017; Cavalcante et al. 2023). Our research questions call for a confirmation of this disunity.

2.2.2. Toward a Textual Description of the Discourse Functions of LDs

The discourse functions of LDs are described through the analysis of three parameters:

i) The degree of activation of the left-dislocated constituents,
ii) The relationship of the left-marked constituent with topic dynamism, and
iii) The relationship between the LDs' Comment and the logical dimension of the text.

Point (i), which corresponds to the analysis of givenness, is a classic parameter in descriptions of LD functions (Chafe 1987, 1994), whereas topic dynamism and the logic dimension require further clarification. We considered left constituents to be given, inferable or new based on the left co-text only, and on a closed set of relations inspired to Baumann and Riester 2012 (for further details we refer the reader to Cimmino 2017). We now detail the trickier concepts of topic dynamism and logic dimension which require further clarification.

The term topic dynamism refers to the organization of the Topics in the text – specifically, to the evolution of the field of application for the illocutionary force of the utterances of the text (Cimmino & Panunzi 2017: 141). When LDs display a topic partition, they contribute to structuring the text by providing contextual reference for the utterance. By observing the stability or shift in this reference, one can assess whether and how an LD signals a sudden change in the contextual coordinates

established by the text.[13] For instance, consider the following excerpt from a conversation between a goldsmith and a client, in which a third party asks whether the goldsmith can draw the requested design using a marker:

(8) *MAN: *ma* ***il*** ***pennarello***/TOP *non ce* ***l'*** *hai?*COM (ifamcv28_57)
but the marker NEG there it have:2SG
'*MAN: **the marker**/ TOP you don't have **it**? COM'

Here, the LD introduces a new referent into the discourse as a new Topic and, crucially, shifts the contextual coordinates for future utterances. From this point onward, the discourse centers around the marker and its role in the goldsmith's work.

With the term *logic dimension*, we refer to the logic relations present within the text and their coherent organization.[14] This includes logic relations between events (e.g., "If it rains, we stay home"), and between communicative units at both textual and illocutionary levels (Ferrari 2014) – that is, logic relations concerning the organization of ideas (e.g., "I read, thus I think") and between speech acts (e.g., "A: Do you like babies? B: Not at all"). In our analysis of LDs, the relevant logic relations are those between the Comment units of LDs and those of adjacent utterances, therefore we see such relations, including contrast, as discourse notions, which interacts with information-structure notions and not as information structurally related tout court (Brunetti 2024). Indeed, as noted in the literature (Huddleston & Pullum 2002 for English; Cimmino & Panunzi 2017 for Italian; Garassino & Jacob 2018 for Italian, French, and Spanish), LDs can highlight contrastive relations in the focal part of the structure. For example, the following LD occurrence highlights a contrast relation on the polarity of the verb to indicate that the speaker can indeed find a solution to a difficult problem:

[13] Since we conceive Topic as a pragmatic concept, we are not dealing with thematic progression, that is the unfolding of relevant referents in the text (Daneš 1974; Ferrari & De Cesare 2009). For further clarification see Cimmino (2023: 348).

[14] We here refer to the classification in Prandi (2006) and Ferrari (2014), but the identification of discourse relations has a long tradition, see at least Mann and Thompson (1988) and Asher and Lascarides (2003).

(9) A: *no, niente, eh, dobbiamo trovare una soluzione*.
B: *ah va be', **la soluzione** glie=**la** troviamo*
ah well ART.DEF solution[F] for.him=3SG.OBJ.F find:1PL
'A: nah, nothing, uh, we gotta find a solution.
B: well, **the solution** we'll find **it** for him'
(LIP corpus, in Garassino & Jacob 2018: 228)

Such occurrences alter the logic dimension of the text without engaging with its topical dimension.

By analyzing the parameters of givenness, topic dynamism, and logic relations independently and in interaction, it is possible to identify four main discourse functions. In previous analyses of Italian and English LDs in spontaneous spoken speech (Cimmino & Panunzi 2017; Cimmino 2023, 2024), we found that the givenness of left-dislocated constituents interacts primarily with the topic dynamism of the text. This interaction gives rise to two widely recognized discourse functions:

(I) The topicalization of new or non-active referents.
(II) The topicalization of old referents.

These functions are well-documented in the literature and represent classic uses of LDs in natural discourse (e.g., Benincà et al. 1998, for Italian; Lambrecht 1994, for English; Sedano 2012, for Spanish).

Less investigated is the possibility for LDs to interact with the logic dimension of texts. As seen above, LDs can highlight contrastive relations on the focal part of the structure, and in particular on polarity focus. Also, LDs can highlight corrective contrastive relations on the entire predicate; in the fictitious example below, speaker A asserts that s/he lies to Paolo to protect him, while speaker B asserts the opposite producing a corrective contrastive relations uttered via a LD:

(10) A: *Spesso ho mentito a Paolo per proteggerlo*.
B: ***A Paolo, gli** dico sempre quello che penso!*
to Paolo to.him say:1SG always that REL think:1SG
'A: I've often lied to Paolo to protect him.
B: **Paolo**, I always tell **him** what I think!'

Although the nature and scope of the contrastive relation may vary (see Cimmino 2024: 50 for a detailed discussion), this function can be generally described as:

(III) The highlighting of a contrastive relation in the discourse.

In performing this function, LDs do not explicitly encode the logical relationship but rather organize the text in a way that accentuates it.

Finally, prior analyses have identified instances where LDs influence neither topic dynamism nor the logic dimension. These LDs contribute to dialogic management by serving to:

(IV) Highlight the speaker's personal stance.

In Italian, this is often achieved through fixed expressions such as "a me mi" ('as for me I'), as seen in the example below:

(11) *MAR: [201] *qui posso girare//*
*MAX: [202] *così è pericoloso//*
*MAR: [203] ***a me*** *'un* ***mi*** *pare-a* *pericoloso//*[COM]
to me NEG to.me seem-PST.IPFV.3SG dangerous
[204] *tu se' te tu sei agitato/ come un coso//* [205] *io son tranquillissima//*
(ifamdl19_201-205)

'*MAR: [201] here I can turn//
*MAX: [202] this is dangerous//
*MAR: [203] as **for me I** don't think it is dangerous//[COM] [204] you are the one who is nervous/ like a thing// [205] I am very calm//'

This function was first observed in Italian by Duranti & Ochs (1979), who described it as a "floor-seeking" interactional move by the speaker. In a slightly different way, more in line with our observation here, Hidalgo Downing (2006) describes this in Spanish as the introduction of a new perspective or voice into the discourse. As will be shown in the analysis section, the expression of personal stance via LDs can occur when the speaker already holds the turn as well.

In the dataset selected for this study, we investigated the presence of the four discourse functions outlined above, as well as additional potential discourse roles played by LDs. In the following, we will thus be concerned not only with Italian and English, as in previous studies, but also with Spanish.

3. Italian, Spanish, and English LDs discourse functions in spontaneous speech

The description of LDs discourse functions occurring in our dataset follows the approach outlined in §2, therefore it consists of two separate steps, which have been the object of two parallel and separate investigations. First, we performed a prosodic and informative analysis to assess the presence of the information partition and the specific intonation features of each occurrence (§3.1.). Second, we conducted a functional analysis to examine the discourse role of LDs (§3.2). The results and details of the analyses are reported in what follows.

3.1. (Mis)matches between LDs' syntax, and prosodic-information features

The occurrences of LDs have been classified according to their prosodic-information structure, to understand whether the LD constituents were prosodically separated from the following material or not.

We first revised the L-AcT information tags previously present in the mini-corpora, which were the result of an inter-annotator agreement process. A new blind annotation was then carried out, independently of the existing tags. In cases of discrepancy between the two annotations, the relevant instances were re-examined in detail through acoustic analysis of the prosodic parameters previously described. Subsequently, we used the Praat software (Boersma & Weenink 2024) to observe the prosodic profile, obtain acoustic measurements of duration, f0, and intensity, and draw the intonation contours.

Examining the three mini-corpora, no direct correspondence emerged between syntactic form and information structure in LD occurrences. Specifically, LD syntactic partition does not consistently correspond to a specific information partition, nor does it necessarily entail a prosodic boundary between the dislocated element and the co-referential pronoun (§3.1.1). Therefore, the presence or absence of an information partition is influenced by factors beyond syntactic form or the syntactic weight of the

locutive content (§3.1.2). Moreover, no consistent acoustic-prosodic features systematically mark LDs (§3.1.3).

3.1.1 Information structure of LDs

Our classification distinguishes LDs where the dislocated constituent occurs in the same unit as its related pronoun, with no detectable prosodic boundary between them (*non-partitioned*), from LDs where the two occupy specific units, with the dislocated constituent followed by a boundary tone (*partitioned*). The results are shown in Table 3.

Mini-corpus	**LD non-partitioned**	**LD partitioned**	**Total**
IT	16 (29%)	40 (71%)	56
SP	45 (62%)	28 (38%)	73
AE	-	8 (100%)	8

Table 3: Information structure of LD occurrences.

We have observed both information structures in IT and SP. The IT mini-corpus shows a clear prevalence of the information partition (29% non-p. vs. 71% p.), while SP shows the opposite trend, with a less marked difference (62% non-p. vs. 38% p.). The AE mini-corpus distinctively presents LDs solely with an information partition.
Accordingly, we now examine the specific information traits associated with partitioned and non-partitioned LDs.

In the partitioned cases, we distinguish LD occurrences based on the information tag of their prosodic units, following L-AcT. The typical structure observed is the Topic/Comment pattern, in which each unit carries its own nuclear prosodic prominence: the dislocated constituent in the Topic, and the pronoun in the Comment, as illustrated in (12), (13), and (14).

(12) *MIC: ***il titolo***/$^{\text{TOP}}$ *non me* ***lo*** *ricordo*//$^{\text{COM}}$ (ifamdl01_38)
the title NEG REFL 3SG.OBJ.M remember:1SG
'*MIC: **the title**/$^{\text{TOP}}$ I do not remember **it**//$^{\text{COM}}$'

(13) *LUC: ***la*** ***nube***/TOP ***la*** *he* *hecho* *yo*//COM (efamcv02c_38)
the cloud[F] 3SG.OBJ.F have:1SG made I
'*LUC: **the cloud**/TOP I made **it**//COM'

(14) *COR: it was like/ **this one guy**/TOP **he** gets this master//COM (afammn05_20)

An examination of each mini-corpus individually allows us to highlight specific features of our dataset.

In SP, when the partition occurs, it always corresponds to the Topic/Comment structure. In IT, differently, the dataset also includes two utterances presenting the dislocated constituent in the Appendix of Topic (APT). The Appendix of Topic integrates the text from the previous TOP; in (15) it is realized as a *suffix* falling prosodic unit and does not carry functional prosodic prominence. It serves a different information function than the Topic in the utterance and is functionally hierarchically subordinated to the previous unit.

(15) *ALD: *è* *per questo* *che* *io*/TOP ***il rappresentante***/APT ***lo***/
be.PRS.3SG for this that I the representative 3SG.OBJ.M
rifarei *cento volte*//COM (ifammn14_117)
do_again:COND.1SG a hundred times
'*ALD: that's why I/TOP the salesman/APT would do it again/ a hundred times//COM'

The AE mini-corpus presents 6 Topic/Comment structures, one utterance in which the dislocated constituent is in an Appendix of Topic (16) and one in a Locutive Introducer unit (INT) (17). Both APT and INT are textual units, according to L-AcT, and always located before the Comment. The Locutive Introducer serves, specifically, to introduce reported speech or spoken thoughts, indicating that what follows refers to a different hic et nunc than the context (Moneglia & Raso 2014) and does not produce a functional prosodic prominence. In other words, it lacks both acoustic and perceptual salience. The overall prosodic form of the movement is described as brief and falling, and is characterized by a marked increase in speech rate.[15]

[15] In L-AcT studies, speech rate is measured using language-specific automatic normalization systems based on z-scores (Barbosa 2019).

(16) *ALC: well/ when you said that though/ my [/] **my new boss**/[APT] **she** came [/] she told M ike yesterday/ she's/ I wanna be there at seven o'clock to go/ to community meeting//[COM] (afamdl03_1)[16]

(17) *ALA: so **mom**/[INT] **she** just goes/[INT] I feel like you've got a whole other world outside of us//[COM] (afammn03_92)

In the non-partitioned cases, the entire LD structure usually occupies a Comment unit as in (18) and (19):

(18) *NIC: *allora* ***il rosso*** ***lo*** *scarto*//[COM] (ifamcv09_95)
then the red 3SG.OBJ.M discard:1SG
'*NIC: so **the red** I discard **it**//[COM]'

(19) *VIG: ***eso*** *sí* ***lo*** *puede* *hacer*//[COM] (epubdl07b_87)
that yes 3SG.OBJ.M can:3SG do:INF
'*VIG: **this** you can do **it**//[COM]'

Non partitioned LDs can also occur in TSs with more than one illocutionary unit, as in the structures defined by L-AcT as Multiple Comments (20) and Bound Comments (21). The former (tag: CMM) are typically dyadic structures in which multiple illocutions are held together by a single terminated intonation pattern. They reflect codified discourse models, such as list, comparison, alternative, and functional recall (Saccone & Panunzi 2020). The latter (tag: COB) are Comment units with homogeneous illocution, linked in prosodically continuous and functionally cohesive chains, typical of monologic discourse (Panunzi & Scarano 2009; Saccone 2022).

(20) *OLG: *la* *&direc [/] pero* ***la*** ***dirección*** *no* ***la*** *tenemos*/[CMM]
the dir [/] but the address NEG 3SG.OBJ.F have:1PL
solo tenemos el nombre y el teléfono//[CMM] (epubdl07a_139)[17]
'*OLG: the addre [/] but **the address** we don't have **it**/[CMM] we only have the name and the phone number//[CMM]'

[16] "[/]"marks a tone unit boundary following a disfluency phenomenon, with or without words repetitions.

[17] "&" marks the beginning of interrupted words or vocalization/hesitation phenomena.

(21) *CLA: *loro* ***il*** ***confine*** ***sudanese*** ***lo*** *vogliono/* *tranquillo/*[COB]
they the border Sudanese 3SG.OBJ.M want:3PL / calm
e questi gli creano una turbativa//[COM] (ifammn03_13)
'*CLA: **the Sudanese border** they want **it**/ quiet/[COB] and these create a disturbance//[COM]'

In (20), the TS is structured in two CMM-illocutionary units "pero la dirección no la tenenos" and "solo tenemos el nombre y el teléfono", marked by prosody to form a comparison intonation pattern. In (21), the chain of two COBs ("loro il confine sudanese lo vogliono tranquillo", and "e questi gli creano una turbativa")[18] builds a segment of a travel story, incrementally adding new information. The COBs are thus constructed as a work in progress shaped by the speaker in the unfolding of discourse.

The lack of a boundary isolating the dislocated constituent to the left of more than one Comment unit signals the scope of the constituent "la dirección", limiting its range to a single illocutionary unit "pero la dirección no la tenenos".

The opposite occurs when the boundary is present, as in (22), where the scope of the constituent "i fiori" extends over both CMM illocutionary units ("li hai aggiunti ora" and o "c'erano anche l'altra volta":

(22) *CLA: ***i*** ***fiori/***[TOP] ***li*** *hai* *aggiunti* *ora/*[CMM] *o*
ART.M.PL flower:PL 3PL.OBJ.M.PL have:2SG add:PTCP.M.PL now or
c' *erano* *anche* *l'* *altra* *volta*?[CMM] (ifamdl15_202)
there be:IPFV:3PL also the other time
'*CLA: **flowers**/[TOP] did you add **them** now/[CMM] or were they there last time too?[CMM]'

Therefore, the examples suggest that the partitioning of LD structures in two distinct units is not explained by the cognitive load of the information structure, since both articulations (partitioned and non-partitioned) can also be found in cases where the illocutionary structure is complex and composed by more than one root unit, as in Multiple Comments and Bound Comments. Indeed, where more than one illocution is expressed within the same TS, a prosodic prominence of the dislocated constituent is possible if its scope extends beyond the first illocution of the TS.

[18] The final Bound Comment is consistently tagged as COM instead of COB.

To summarize, partitioned LDs exhibit varied information structures, indicating that the information features of the dislocated constituent may – but do not necessarily – align with those of Topics. However, it is important to note that when the LD is partitioned, the dislocated constituent is placed in a dedicated tone unit with its own information function. That is to say that, when present, the prosodic boundary separating the dislocated constituent from the co-referential pronoun does not merely scan the locutive content of the unit for physiological reasons – such as breathing – or disfluency phenomena – such as syllables or words retracing.[19]

3.1.2 Syntactic Weight of LDs

Notably, the partition of the analyzed items – the dislocated constituent and the pronoun – into distinct information units does not correlate with the syntactic weight of the locutive content,[20] cross-linguistically.

In Italian, the majority of dislocated constituents in our dataset consist of simple nominal syntagms (NSs), with 44 out of 56 occurrences falling into this category. These NSs are found in both partitioned and non-partitioned information structures, with 13 occurrences of dislocated NSs without partition and 31 occurrences with partition. Additionally, a similar trend is observed for nominal syntagms that include an adjective, where 3 instances occur without partition and 4 occur with partition.

In Spanish, a similar pattern emerges, with simple NSs comprising the majority of dislocated constituents (61 out of 73 occurrences). These NSs also appear in both partitioned and non-partitioned information structures, with 42 occurrences without partition and 19 with partition. As in Italian, nominal syntagms containing an adjective show a similar distribution, with 4 occurrences without partition and 2 with partition.

In American English, as previously mentioned, all LDs occur in dedicated tone units. Among these, the majority involve NSs with an adjective (5 out of 8 occurrences), followed by simple NSs (2 out of 8).

[19] In cases where tone units lack an independent information function and prosodic prominence, L-AcT identifies a Scanning unit (SCA). When the locutive content of a unit is abandoned – interrupted, substituted, retraced, self-corrected, or otherwise – the corresponding L-AcT tag is Empty (EMP).

[20] Cf. Crocco & Savy (2007); Alfano et al. (2021).

The evidence presented above indicates that, in the cases involving simple NSs or NSs with adjectival modification there is no correlation between syntactic weight and the presence or absence of prosodic partition.

The only context in which an association emerges is with structurally more complex – thus syntactically heavier – constituents, which, however, represent a minimal portion of the dataset. In fact, across all the three languages, only a small subset of LDs exhibits structures such as a nominal syntagm combined with a prepositional phrase or a relative clause. These more complex structures are always realized in separate information units, and this pattern is consistent across all three languages. For example, in Italian, there are 5 instances of such structures out of 56 total LDs; in Spanish, 6 out of 73; and in AE, only 1 out of 8.

3.1.3 Prosodic features of LDs in Topic/Comment pattern

We then conducted a prosodic analysis of LDs with partition, where a boundary tone separates the dislocated constituent from the pronoun, with the dislocated constituent functioning as the Topic. The aim of this analysis is to determine whether any acoustic-prosodic feature systematically marks left dislocation (LD).

For these occurrences, we analyzed the acoustic correlates of both the unit and boundary, along with the nuclear prosodic configuration of the LD constituent.
Regarding prosodic features at the boundary, we observed vowel lengthening before the boundary and either a pause or an f0 shift (upward or downward) following it. Lengthening occurred in 4 cases (3 in SP and 1 in IT) across different prosodic configurations. 6 Topics were followed by a pause $>$150 ms,[21] with pauses varying in duration (min: 165 ms; max: 1778 ms). An f0 shift after the boundary was observed in 11 Topics (4 in AE; 2 in SP; 5 in IT), with 6 shifts upward and 5 downward.[22] The three observed features did not co-occur. Such prosodic variability suggests a high degree of heterogeneity in the sample, and points to the lack of a systematic correlation between these features and the phenomenon under investigation.

[21] We observed pauses, consistent with their documented role in LD structures across languages (e.g., Delais-Roussarie et al. 2004 for French; Feldhausen 2016 for Spanish). A 150 ms threshold was adopted since it aligns with the average duration of Italian stop consonants (Giannini 2008; Dovetto & Gemelli 2013), captures short planning-related pauses reported from 100 ms (Matzinger et al. 2020), and remains below the 300 ms perceptual threshold for reliable detection (Duez 1985; Wang et al. 2012).

[22] The f0 shift indicates differences in pitch range between two adjacent intonation units. It was calculated as the difference between the means of the five f0 points before and after the boundary.

As for the prosodic type of Topics, they were classified based on their contour, following the features discussed in §2.2.1.

It is first necessary to consider that the distribution of Topic prosodic types varies cross-linguistically. Following corpus-based analysis in Firenzuoli & Signorini (2003), Type 1 results the most occurring in Italian spontaneous speech[23] (55.2%), while Type 3 is the most common in American English[24] (72%) (Cavalcante 2020). To date, no such analysis has been conducted on Spanish, which prevents a direct comparison with the other languages considered.

With regard to our dataset, we identified three types of Topic contours, as shown in Table 4: Type 1, Type 2, and Flat Type.

	Type 1	**Type 2**	**Flat**	**Total**
IT	19	11	3	33
SP	2	19	3	24
AE	3	0	3	6

Table 4: Topic prosodic types of LDs.

The distribution of prosodic types across the three languages is clearly uneven. Italian shows a predominance of Type 1 (19/33), which aligns with patterns already observed in the language's general prosodic tendencies (23) (see Figure 4).

(23) *CLA: ***i*** ***fiori***/TOP ***li*** *hai* *aggiunti* *ora*/CMM *o*
ART.PL flower:PL 3PL.OBJ.M.PL have:2SG add:PTCP.M.PL now or
c' *erano* *anche* *l'* *altra* *volta?*CMM (ifamdl15_202)
there be:IPFV:3PL also the other time
'*CLA: **flowers**/TOP did you add **them** now/CMM or were they there last time too?CMM'

[23] The Italian reference corpus corresponds to the complete version of the mini-corpus employed here.

[24] The reference corpus for American English is the corpus currently being analyzed in this study.

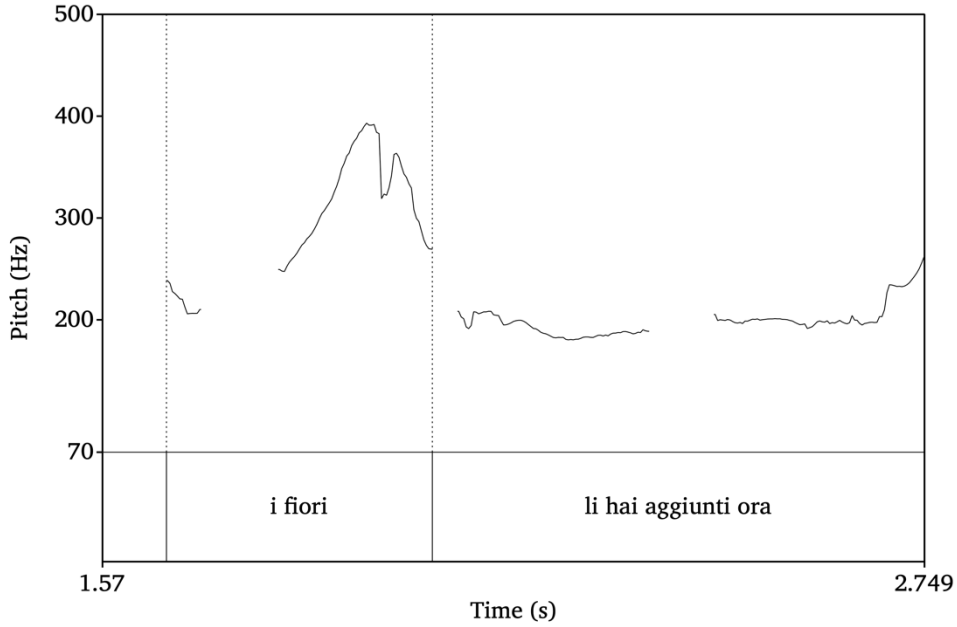

Figure 4: Prosodic contour of (23). Type 1 Topic: *i fiori* (female speaker).

In contrast, Spanish shows a clear preference for Type 2 (19/24), suggesting a different intonation strategy in marking Topics (24) (see Figure 5).

(24) *LOL: [145] ***a Bretón***/^{TOP} ***le*** *han &can [/] condenado* *a cuarenta*//^{COM}
to Bretón to.him have.3PL condemn:PTCP to forty
(efamcv06_145)
'*LOL: **Bretón**/^{TOP} they've &sen [/] sentenced **him** to forty//^{COM}'

American English displays no clear preference, with an equal distribution between Type 1 and Flat realizations (3 each), and no occurrences of Type 2 (see (25) for an example of Flat Topic; Figure 6).

(25) *RAN: **Horizon**/^{TOP} you can't do nothing with **him**/^{CMM} so just have him hold there//^{CMM} (apubdl01_61)

These patterns further support the idea that prosodic realization of LD Topics is language-dependent and shaped by language-specific prosodic norms.

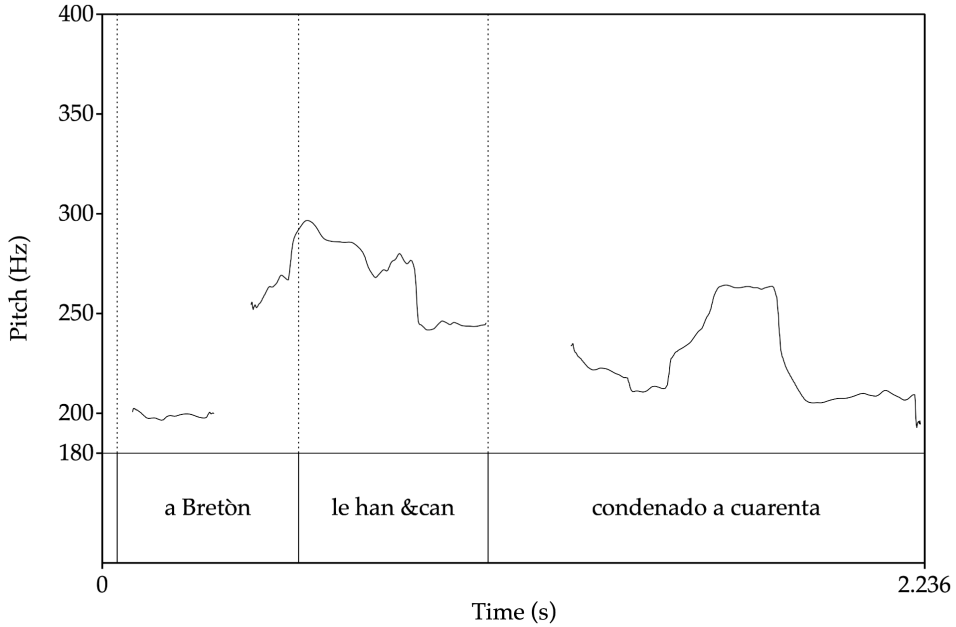

Figure 5: Prosodic contour of (24). Type 2 Topic: *a Bretón* (female speaker).

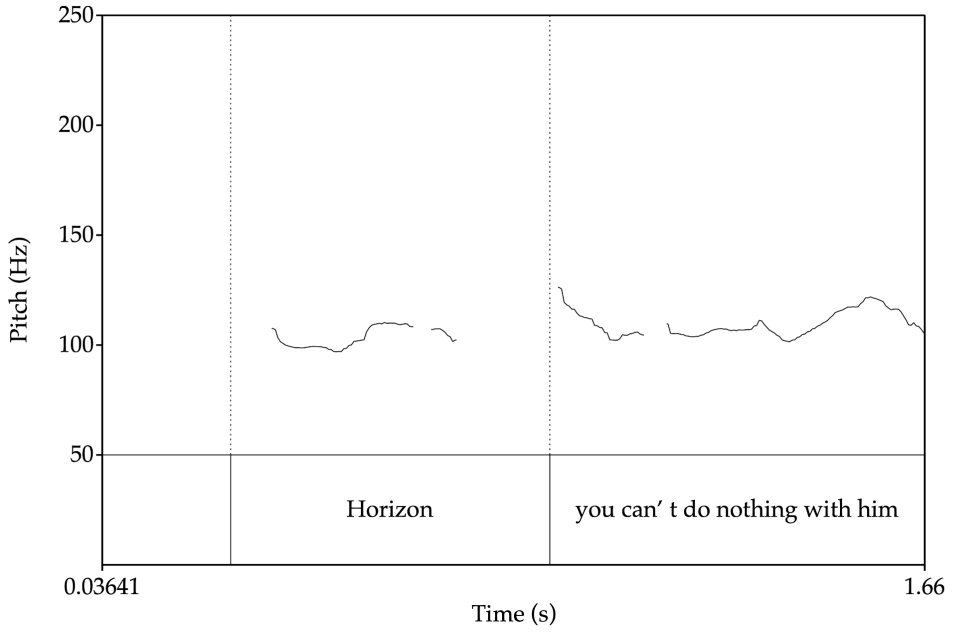

Figure 6: Prosodic contour of (25). Flat type Topic: *Horizon* (male speaker).

As for the acoustic features of prominence movements, both Type 1 (rising + falling) and Type 2 (rising) in our dataset are characterized by significant intonation excursions. These movements are perceptible in the speech flow, with a pitch range >3 semitones (st) – the threshold corresponding to the Just Noticeable Difference (JND) for f0 variation detectable by listeners and relevant for communicative events,

as identified by 't Hart (1981) in his analysis of Dutch spontaneous speech.[25] In the case of Type 1, both the rising and falling movements exceed this perceptual threshold. The third detected pattern, the Flat type, lacks perceptible f0 movements, remaining below the 3 st JND.

Analyzing the f0 movements in correlation with stress position, we observed that Type 1 consistently places stress at the highest point of the rising movement, regardless of the language. In contrast, two different configurations were found for Type 2 and Flat.

Type 2 appears in the SP and IT datasets, typically with the stressed syllable positioned at the starting (lower) point of the rising configuration. In instances where this is not the case (5 occurrences in SP and 1 in IT), the stress falls on the end of the tone unit, the highest point of the rising, and always corresponds to oxytone words.

The Flat type is present in all three datasets, exhibiting a high f0 level in SP and IT, while in AE, it occurs with a lower f0 level, sometimes close to the baseline of the utterance. As a result, the IT and SP Flat Topics are more prosodically prominent in the speech flow than those in AE.

3.2. (Mis)matches between prosodic-information traits and discourse functions

The occurrences of LD retrieved in our mini-corpora were classified according to the four discourse functions outlined in our approach:

(I) Topicalization of new or non-active referents,
(II) Topicalization of old referents,
(III) Highlighting of a contrastive relation within the discourse, and
(IV) Highlighting of the speaker's personal stance.

We examined whether all instances of LDs matched at least one of the defined functions and assessed the extent of intra- and cross-linguistic variation. Furthermore,

[25] Specific JND values for Italian are not well-established, and the literature shows no consensus, with studies referencing JND both in Hz and in st. In addition to the reference chosen here, other studies report JNDs of 1 st (Lehiste 1970; Hermes & van Gestel 1991; Barbosa 2019). We use st to normalize variations due to gender or contextual factors (Sorianello 2006; Simpson 2009; Stanford 2016). More specifically, regarding AE speech, studies show that when the data are converted into semitones, the cross-gender differences disappear esntirely (Pépiot 2014).

we investigated potential interactions between information traits and the prosodic contour of the left-dislocated constituents, on the one hand, and the discourse functions of LDs, on the other.

In the three mini-corpora, we observed considerable cross-linguistic variation in both the frequency and nature of discourse functions performed by LDs. While American English LDs fulfill only functions (I) and (II), Italian and Spanish LDs – though asymmetrically – also perform additional functions. The data in §3.2.1 show the intra- and cross-linguistic variation in LDs discourse functions, while §3.2.2 shows that the discourse functions performed do not match LDs prosodic-information patterns. Each function described can be fulfilled by LDs independently of the extra-clausality of the left-dislocated constituent: all three languages display similar functions with partitioned and non-partitioned LDs.

3.2.1. Intra- and cross-linguistic variation in LDs discourse functions

Table 5 presents the classification of discourse functions by language (rows) and function type (columns), including the absolute number of occurrences for each function and the percentage relative to the total LDs in each language.

Language	**Intro referents**	**Resume referents**	**Highlight contrast**	**Highlight stance**	**Other**	**Total**
IT	12 (21.5%)	7 (12.5%)	14 (25%)	5 (8.9%)	18 (32.1%)	56 (100%)
SP	15 (20.5%)	0	17 (23.3%)	22 (30.1%)	19 (26.1%)	73 (100%)
AE	6 (75%)	2 (25%)	0	0	0	8 (100%)

Table 5: Discourse Functions of LD Occurrences.

American English LDs display less functional variation compared to Italian and Spanish. This cross-linguistic variation is both quantitative and qualitative. In our dataset, American English LDs are almost exclusively used for function (I), introducing a new referent in the text as a Topic. In the following AE monologue, the speaker

comments on the impossibility of spending thousands of dollars on art, while he recalls the first painting he bought in Mexico ("the first painting I ever bought/ my father-in-law bought it from Geri [/] Geri Rae//", TS [87]). The introduction of this new referent ("the first painting") disrupts the topic chain previously established in the text, which had been centered on a different place ("while I was down there", i.e., in Mexico City, TS [84]) and another concept ("to spend a thousand dollars [...] for a painting", TS [86]). After introducing this new referent, the discourse shifts to focus on it for two subsequent utterances.

(26) *ALN:[83] well/ we bought a painting/ &he/ of a little Mexican woman/ still got it/ with a watermelon on her head/ like it a lot/ think I paid/ thirty dollars for it/ well that was about my price range// [84] &he/ while I was down there/^{TOP} we/ may have bought a couple of other &i [/] inexpensive paintings/^{COB} maybe collectively we bought [/] maybe spent a hundred and fifty dollars// [85] &he/ hell I was not an art [/] we weren't art collectors// [86] and/ &he/ to spend a thousand dollars/ to me/ for a painting/ was unthinkable// [87] &he/ **the first painting I ever bought**/^{TOP} my father in law bought **it** from Geri [/] Geri Rae//^{COM} [88] it was one of Geri's original early early paintings/ which I thoroughly enjoy// (afammn02_ 83-90)

In our American English data, LDs performing function (I) consistently signal that the left-dislocated constituents are new, relevant, and persistent in the subsequent topic chain. In Italian and Spanish, LDs performing function (I) also involve already given, but non-active referents. The following excerpt from a Spanish dialogue illustrates an LD "la grasa/ la vamos echando según va saliendo/ así?" (TS [106]) that introduces a Topic represented by a referent previously mentioned "grasa" (TS [100]), but which had been competing with other referents (see Givón 1983). "La grasa" has now become the most relevant referent and will remain active in the following utterances.

(27) *DOR: [100] *sí/ sí sí/ tenemos grasa//* [101] *he echado eso/ pero también viene bien/ porque si no se [/] se agarra a la sartén//* [102] *te digo* + [103] *vamos a ver//* [104] *una plato/ para irlo echando//*

*CAR: [105] *ya//*

[106] *y ahora/* ***la*** ***grasa***/^{TOP} ***la*** *vamos* *echando*

ART.SG.F fat 3SG.OBJ.F go:PRS.1PL add:GER

según *va* *saliendo/*[COM] *así* ?
as go.3SG come.out:GER thus
*DOR: [107] *no//*
*CAR: [108] *no//*
*DOR: [109] ***la*** ***grasa/***[TOP] ***la*** *dejamos* *aquí/*[COB]
ART.SG.F fat 3SG.OBJ.F leave:PRS.1PL here
que vaya saliendo una/[COB] *y otra/*[COB] *y otra//*[COM]
that go:SBJV.PRS.3SG come:GER one:F and another:F and another.F
(efamdl02_100-109)

'*DOR: [100] yes/ yes yes/ we have fat// [101] I added that/ but it's also good/ because if not[/] it sticks to the pan// [102] I'll tell you + [103] let's see// [104] a plate/ to add it//
*CAR: [105] already// [106] and now/ **the fat**/[TOP] are we adding **it** as it comes out/[COM] like this?
*DOR:[107] no//
*CAR: [108] no//
*DOR: [109] **the fat**/[TOP] we leave **it** here/[COB] so that it comes out again/[COB] and again/[COB] and again//[COM]'

The lack of functional variation in American English LDs suggests their high syntactic and pragmatic markedness.[26] In contrast, the variability in the activation status of the left-dislocated constituents in Italian and Spanish LDs performing function (I) suggests that LDs are flexible cohesive devices in discourse. In summary, the differences among the three languages are striking both cross-linguistically and intra-linguistically.

The classification shown in Table 5 reveals that asymmetries can also be observed between Italian and Spanish. Spanish LDs may be exploited to perform function (III) – that is, to highlight a contrastive relation. In example (27) above, the second LD occurring in the excerpt ("la grasa / la dejamos aquí/ que vaya saliendo una/ y otra/ y otra") resumes the referent "grasa" to provide a negative response to the preceding question, thus denying speaker CAR's expectations. In this case, therefore, the resumption of the referent is exploited at a logic level, performing

[26] Here we use the theoretically obsolete (Haspelmath 2006), but operationally clear notion of markedness. We refer to syntactic markedness with respect to the word order of the structures, diverging from the most frequent SVO order in Italian and English. Pragmatic markedness, instead, refers to the fact that LDs perform their discourse functions only in specific given contexts.

function (III). Indeed, LDs performing function (III) signal the textual relevance of the logic relation encoded in the Comment of the LD. This aligns perfectly with observations of Italian LDs in our mini-corpus, which can perform the same function in similar contexts.[27]

However, in our dataset, Italian LDs also perform function (II), which does not involve the exploitation of the LD on a logic level and is absent in our Spanish data. In the following Italian example, speaker ZIA resumes the referent "Cinturini", with the LD in TS [4], after a digression (TS [2]) in which she introduced the competing referent "il libretto del Lavoro".

(28) *ZIA: [1] *poi so' entrata da Cinturini appena quattordic'anni//* [2] *perché a quei tempi/ il libretto del lavoro/ te lo facevano a quattordic'anni//*
*ALE: [3] *mh mh//*
*ZIA: [4] ***da Cinturini**/TOP **c'** *ho lavorato du' anni//*COM
at Cinturini there have:1SG work:PTCP two year:PL
[5] *poi a me da Cinturini non me piaceva/ facevo sempre li dispetti/ allora m'hanno licenziato//* (ifammn05_1-4)

'*ZIA: [1] then I joined Cinturini when I was only fourteen years old// [2] because in those days/ the employment card/ they made it for you at fourteen years old//
*ALE: [3] mh mh//
*ZIA: [4] **at Cinturini**/TOP I worked **there** for two years//COM [5] then as for me I didn't like Cinturini/ I was always teasing/ then they fired me//'

In turn, Spanish LDs are used more frequently than Italian LDs for the dialogic management of discourse – specifically, to highlight a speaker's personal stance (function IV). In the following excerpt, CAR and DOR are discussing the quality of garlic soup. CAR asks if DOR has ever made it at home. DOR replies she has not, but recalls preparing and eating it on a farm in Seville; she concludes by expressing her personal stance on the experience. In this case, the LD serves to emphasize the speaker's personal stance within the context of a complex utterance (with Multiple Comments), which reports both actions and a metatextual reflection.

[27] For a detailed discussion of LDs contrastive function in Italian both in spoken and written texts, the reader is referred to Cimmino (2024).

(29) *DOR: *y allí las hice yo/CMM y las comí*
3PL.OBJ.F make:PST.1SG I and 3PL.OBJ.F eat:PST.1SG
yo/CMM **a mí**/TOP_p **me** gustaron//PAR28
I to me to.me like:PST.3PL (efamdl02_54)
'*DOR: and I made them there/ CMM I ate them/ CMM **me**/ TOP_p **I** liked them// PAR'

In Italian LD occurrences, it is possible to find similar examples – namely, LDs performing function (IV) introduced by the fixed formula "a me mi", functionally equivalent to "a mí me" as seen in the example above. However, they occur less frequently and differ qualitatively, as in Italian they highlight a personal stance only in dialogic contexts, and typically in (strong or mild) opposition to other speakers.

Once again, our closer look at the interaction between LDs and their context of occurrence revealed both quantitative and qualitative differences among the languages under scrutiny. We observed that both Italian and Spanish LDs go far beyond the topicalizing function described in the literature (e.g., Lambrecht 1994), since they do not only fulfill referent management functions (I and II, in our taxonomy), but also functions linked to the logic or interpersonal dimension of the text (III and IV). This demonstrates that both languages are characterized by a high pragmatic and syntactic flexibility in the use of LDs (unlike English). Notwithstanding this similarity, the exploitation of LDs still remains context-based (and not language-based), since LDs potential is exploited following the speaker's expressive needs, which change based on the text and the interaction with the interlocutors. Another piece of evidence retrievable from Table 5 supports the unpredictability of LD usage in context. In our dataset, approximately 30% of both Italian and Spanish LD functions cannot be accounted for by our functional classification, despite the variety of textual dimensions considered (topic, logic, or dialogic – see §2.2.2 for details on this aspect).

Indeed, some occurrences represent a blend of more than one function, as in the following Italian LD, which operates on both a logical and dialogic level. Speaker ART is describing his work as a leatherworker when DAN interrupts to ask for confirmation. She hypothesizes that the leatherworker does not produce the lining for the bags himself and must purchase it elsewhere ("gliela forniscono"). The syntactic and informational prominence created by the LD serves both to introduce a personal

[28] PAR marks parenthetical units, which insert supplementary information into the utterance. The subscript "_p" in TOP_p signals that the tagged unit (TOP) is embedded within a prosodic sequence of unit functioning as a parenthetical.

stance – disruptive of the ongoing discourse – and to highlight a shift in the coherence relations of the text (which we call here logic, cfr. §2.2.2): DAN interrupts the descriptive sequence of ART with a request for confirmation. The cohesive anchor for this complex operation is the repetition of the active referent "la fodera", while the COM "gliela forniscono" inquires about the polarity of the verb (it could be paraphrased by: *do you purchase the lining or not?*).

(30) *ART: [136] *naturalmente c'è una fodera//* [137] *questo è chiaro//* [138] *una fodera//*
*DAN: [139] ***la*** ***fodera***/TOP *glie=**la*** *forniscono//*COM
ART.SG.F lining to.him/her=3SG.OBJ.F provide:PRS.3PL
*ART: [140] *sì sì/ le fodere/ sono cose//* [141] *una fodera//* [142] *si foderano//* (ifamdl04_136-142)
'*ART: [136] of course there's a lining// [137] that's obvious// [138] a lining.
*DAN: [139] **the lining**/TOP they provide **it**//COM
*ART: [140] yeah yeah/ linings/ are stuff [141] a lining/ [142] they get lined//

Other occurrences seem to have no impact on the ongoing discourse, simply highlighting the speaker's expressive needs. For example, in the following Spanish LD (TS [161]), the textual prominence acquired through the topicality of the left-dislocated constituent is not exploited in any of the topic, logic, or dialogic dimensions, but rather serves to expressively emphasize the speaker's deep knowledge of the norms. In fact, the LD comments on the previous utterance constituting a digression from the topic progression and logic architecture of the text.

(31) *SIN: [158] *yo entonces era opositor//*
*CAR: [159] *mh//*
*SIN: [160] *y por lo tanto/ me sabía muy bien esos decretos//* [161] *porque*
las ***normas***/TOP *me* ***las*** *tenía* *que* *saber*
ART.PL.F rule:PL to.me 3PL.OBJ.F have.to:IPFV.3SG COMP know:INF
*muy bien//*COM (epubmn03_158-161)
very well
'*SIN: [158] back then I did the civil service//
*CAR: [159] mh//

*SIN: [160] so/ I knew those decrees really well// [161] because **the rules**/ TOP I really had to know **them**// COM'

The intra- and cross-linguistic variability observed in our data, along with the gaps in our description, supports the idea of the contextual variability of LDs, disentangling them from the traditionally assumed topicalizing function. In what follows, we demonstrate that LD occurrences are not always linked to the concept of Topic either.

3.2.2 Information partition, Topic prosodic types and discourse functions of LDs

As shown in Table 3 (§3.1.1), American English LDs are always characterized by a partitioned information profile; nonetheless, in our dataset, they can perform two different functions. Italian and Spanish LDs can exhibit either a partitioned or non-partitioned information structure and perform four different functions, irrespective of their form. For example, in both Italian and Spanish, function (IV) is introduced by the fixed expressions "a me mi"/ "a mí me". Despite having the same syntactic structure and fulfilling the same discourse function, these fixed expressions can be delivered in either a partitioned or non-partitioned manner, as shown in excerpts (32) to (35).

(32) *VAL: ***a me***/TOP ***mi*** *faceva* *schifo* *camminare* *&ne* [/]
to me to.me make:IPFV.3SG disgust walk:INF in
nelle *camere* *degli* *alberghi* *senza* [/] *cioè/*
in:ART.F.PL room:PL of:ART.M.PL hotel:PL without that.is
scalza//COM (ifammn08_118)
barefoot:F.SG
'*VAL: **as for me**/TOP **I** hated walking &aro [/] around hotel rooms without [/]I mean/ barefoot//COM'

(33) *LAU: *cioè/* ***a me*** ***mi*** *interessa* *molto/* *cioè/* *costrui'* *la*
that.is to me to.me interest:PRS.3SG much that.is build.INF ART
tre *F/*CMM *non* *mi* *interessa* *affatto/* *essere* *marxista*
three F not to.me interest:PRS.3SG at.all be:INF marxist
rivoluzionaria//CMM
revolutionary.F.SG (ipubcv01_145)

‘*LAU: I mean/ **as for me I** am really interested in/ I mean/ building the three F/CMM I am not Interested at all in/ being a Marxist revolutionary//CMM’

(34) *PIU: ***a*** ***mí*** *el* *vino* *sí* ***me*** *gusta//*COM (efamcv04_6)
to me ART.M.SG wine yes to.me like:PRS.3SG
‘*PIU: **me I** do like wine//COM’

(35) *DOR: *y* *allí* *las* *hice* *yo/*CMM *y* *las*
and there 3PL.OBJ.F make:PST.1SG I and 3PL.OBJ.F
comí
eat:PST.1SG
*yo/*CMM ***a*** ***mí/***TOP_P ***me*** *gustaron//*PAR (efamdl02_54)
I to me to.me like:PST.3PL
‘*DOR: and there I made them/CMM I hate them/CMM **me**/TOP_P **I** liked them//PAR’

Irrespective of their information structure, LD occurrences performing function (IV) signal that the speaker is about to express a personal stance, which may or may not align with the preceding interlocutors’ points of view.

In conclusion, in our data, information partition does not systematically correlate with a specific discourse function; furthermore, in what follows, we show that we found no direct correlation between the prosodic type of Topics and their discourse function. Each Topic prosodic type supports different discourse functions and varies in frequency at both intra- and cross-linguistic levels.

Building on our prosodic analysis, we can compare our results with the findings of Frascarelli (2007), who argues for a direct relationship between prosodic form and function in Italian LD. Despite the theoretical differences in the description of prosodic forms, we observe a correspondence between some of the Topic types identified in our study and the contours defined by Frascarelli through the ToBI transcription system of prosodic labelling (Silverman et al. 1992). Specifically, our Type 2 corresponds to the (L* + H) Topic, while the Flat Topic with high f0 level aligns with the (H*) Topic[29]. According to Frascarelli, there is a biunivocal correspondence between (L* + H) contour and the function of Aboutness-shift Topic, and between (H*)

[29] This correlation is based on the descriptive criteria provided in Frascarelli (2007) and the alignment between tonic vowel and the f0 contour in our sample. A more detailed prosodic analysis is not included here due to space constraints.

contour and the Contrastive Topic. Accordingly, in IT, we investigated whether Type 2 correlates with the Aboutness-Shift function, i.e., our function (I), introduction of a new referent, and whether the Flat type correlates with the Contrastive function, i.e., our function (III), which highlights a contrastive relation.

As for the Italian sample,[30] we found occurrences of Type 2 with both the function of introduction of a new or non-active referent as a Topic, that is, function (I), as exemplified in example (36) (Figure 7), and the highlighting of a contrastive relation, that is function (III), as exemplified in example (37) (Figure 8). In example (36), the new referent "altre borse", 'other bags' is introduced, which is part of the set of referents established in TS [144-145] ("ci sono varie forme di borse essenzialmente sono due", 'there are different kinds of bags basically they are two'). In example (37), a contrastive relation is established between "prendiamo un appuntamento", 'let's set up a meeting' and "almeno il giornale glielo offri" 'at least the newspaper you offer it'; in the context of political dissemination those are in fact two opposing stances.

(36) *ART: [144] *ci sono/ varie &f* + [145] *forme di borse/ essenzialmente/ sono due//* [146] *&he diciamo/ come/ tipo di lavorazione//* [147] *questa si chiama* +[148] *una struttura rigida/ in pratica//* [149] *si chiama a sascé//* [150] *noi si chiama sascé/ o [/] o a scatolina//* [151] *praticamente si forma una parte centrale/ e ci si applica/ le parti laterali//*

*DAN: [152] *mh mh//*

*ART: [153] *mentre/* ***altre*** ***borse/***[TOP] ***le*** *chiamano*
while other:F.PL bag:PL 3PL.OBJ.F call:PRS.3PL
a filetto//[COM] (ifamdl04_144-153)
a filetto

'*ART: [144] there are different k + [145] the kinds of bags/ basically/ they are two// [146] let's say/ like/ the type of craftsmanship// [147] this one's called + [148] a rigid structure/ basically// [149] it's called sascé// [150] we call it sascé/ or/ or little box// [151] basically you form a central part/ and then you attach/ the sides//

*DAN: [152] mh mh//

*ART: [153] while/ **other bags/**[TOP] they call **them** filetto//[COM]'

[30] To better illustrate the functions of LDs, the examples include preceding and, when necessary, following co-text. The terminated sequences are consecutively numbered, and the LD constituent is highlighted in bold. Each example is accompanied by its prosodic and functional annotation.

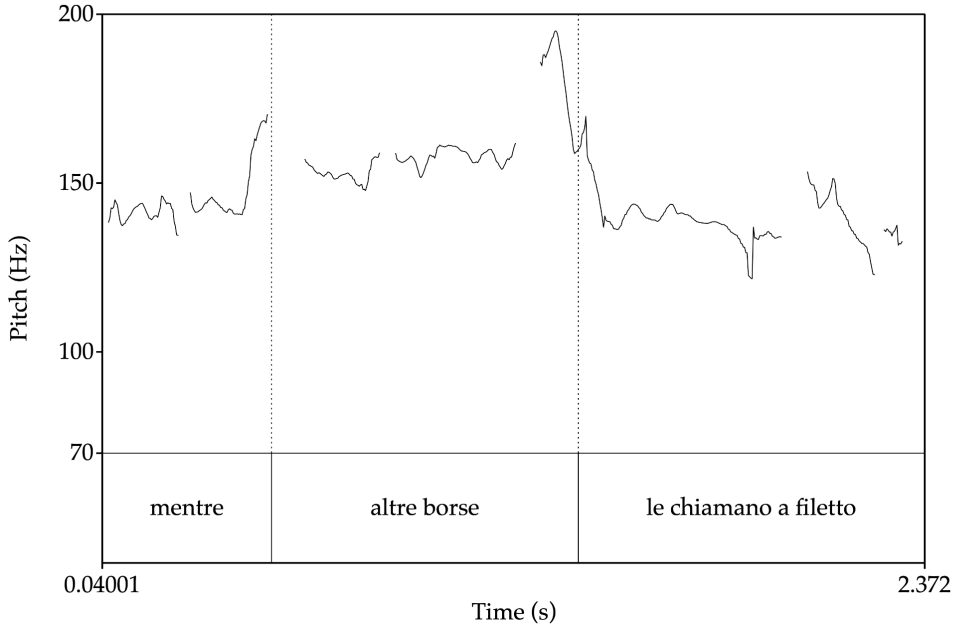

Figure 7: Prosodic contour of TS [153] in (36). Type 2 Topic: *altre borse* (‘*other bags*’); function (I) (male speaker).

(37) *PAO: [492] *questo io/ non lo capisco//* [...] [497] *nello stesso momento/ dica/ &he/ piglia la tessera &de [/] del [/] della tre F/ e costruisca la tre F/ e poi dica/ io sono di Socialismo rivoluzionario/ e non dica per esempio/ ascolta/ ci prendiamo un appuntamento/ ci si trova un’altra volta/ se ne parla/ perché secondo me/ è molto importante//*

*OTT: [498] *sì ma appena hai l’opportunità/*

almeno	***il***	***giornale/***$^{\text{TOP}}$	*glie=**lo***	*offri/*$^{\text{COM}}$
at.least	ART.M.SG	newspaper	to.him=3SG.OBJ.M	offer:PRS.2SG

a una persona/ scusami// (ipubcv01_492-498)
to a person excuse.me

‘*PAO: [492] this I/ don’t get it// [...] [497] at the same time/ say/ &he/ takes the card &of [/] the [/] the three F/ and builds the three F/ and then says/ I am from Revolutionary Socialism/ and doesn’t say, for example/ listen/ let’s set up a meeting/ we’ll meet another time/ we’ll talk about it/ because in my opinion/ it’s very important//

*OTT: [498] yeah but as soon as you get the chance/ at least **the newspaper/**$^{\text{TOP}}$ you offer **it/**$^{\text{COM}}$ to someone/ sorry//’

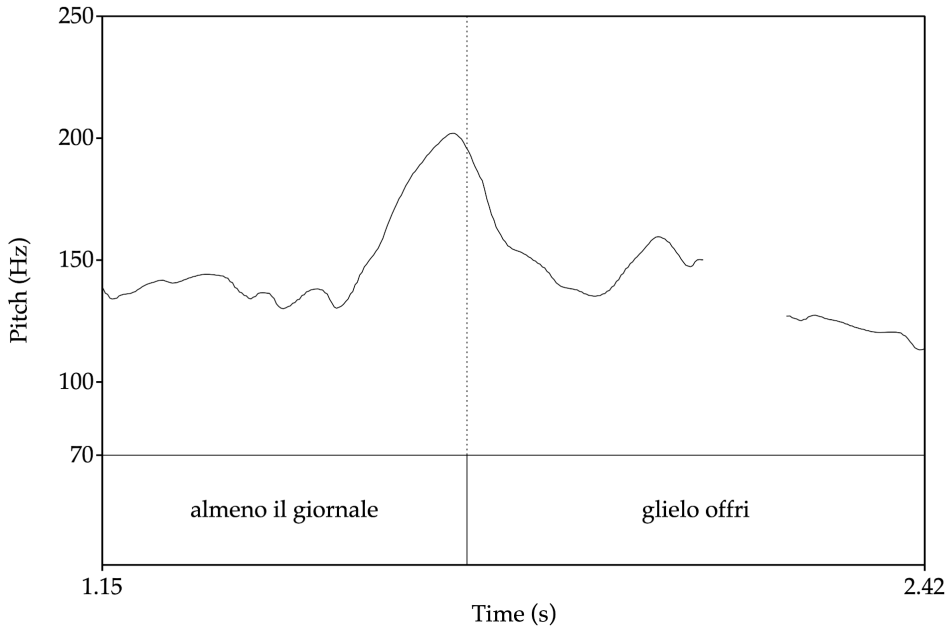

Figure 8: Prosodic contour of TS [498] in (37). Type 2 Topic: *almeno il giornale* (‘*at least the newspaper*’); function (III) (male speaker).

Therefore, no direct correlation was found, in contrast to the findings of Frascarelli (2007).

Similarly, the same check for a bidirectional relationship between Topic prosodic type and discourse function was conducted for both Spanish and American English. However, the findings were consistent, revealing no correlation in either language.

Additionally, we explored potential cross-linguistic links between discourse functions and prosodic types. The result, once again, is negative, as our dataset reveals that the same discourse function is associated with multiple Topic prosodic forms across the three languages. The following examples illustrate occurrences of LD with function (I) introduction of a new or non-active referent as Topic, such as (37), in type 2, presented above for Italian (see Figure 8). Example (38) (see Figure 9), from Spanish, is a type 1 Topic, while (39) (see Figure 10), from American English, is a Flat Topic.

(38) *MOJ: [105] *pues eso/ pues [/] para empezar/ ahí donde hay ese bloque grandísimo/ de casas y el jardín/ eso era todo los escolapios//*
*POH: [106] *&a [/] &ais [/] allí estudié yo/ y nosotros/ ahí estuvimos todos//*
*MOJ: [107] *era todo/ los escolapios//* [108] *o sea que ahí era todo escolapios/ sin [/] el jardín incluido//*

[109] ***el*** ***jardín/***[TOP] ***lo*** *regalaron* *al*
ART.M.SG garden 3SG.OBJ.M give:PST.3PL to.the
ayuntamiento/[COM] *los* *escolapios//*
town.hall ART.PL:M escolapios
[110] *que eso no [/] tampoco lo dice nadie/ pero bueno//* (efammn04_105-110)

'*MOJ: [105] so that/ well [/] to begin/ where there's that huge block/ of houses and the garden/ that was all the Escolapios Piarists//
*POH: [106] &a [/] &ai [/] I studied there/ and we/ we all studied there//
*MOJ: [107] it was all/ the Escolapios// [108] I mean that there was the Escolapios/ without [/] the garden included// [109] **the garden**/[TOP] **it** was donated to the town hall/[COM] by the Escolapios// [110] this is something no one [/] ever says either/ but anyway//'

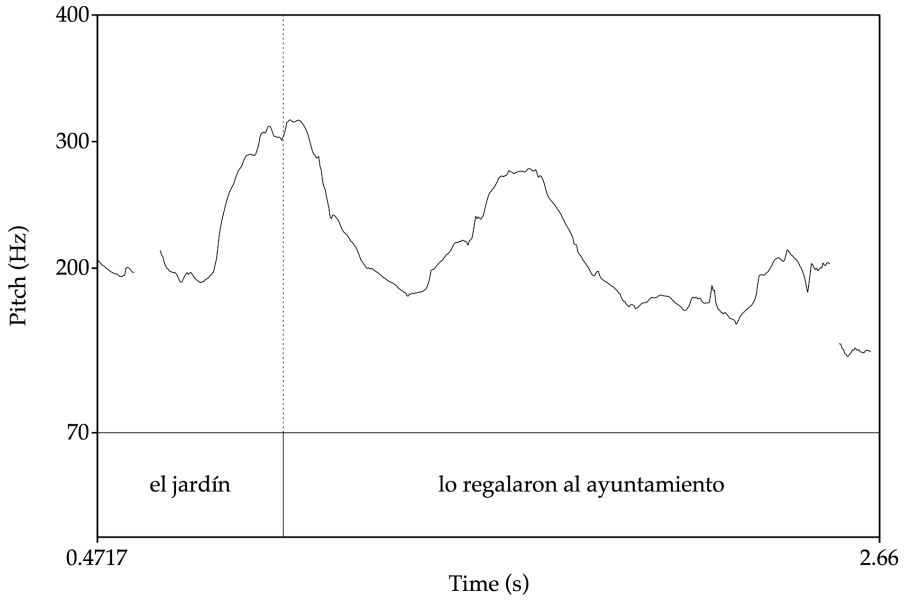

Figure 9: Prosodic contour of TS [109] in (38). Type 2 Topic: *el jardin* ('*the garden*'); function (I) (female speaker).

(39) *COR: [16] and it's like/ this one/ guy was &t [/] say/ it was like/ they got sent to the village//[17] which you don't wanna get sent to the village// [18] but they do// [19] and/ they go +[20] it was like/ **this one guy**/[TOP] **he** gets this master//[COM] [21] and/ he winds up +[22] the master/ &fi [/] falls in love with him/ but he was like +[23] like// the day before// he makes him walk around// he's like a &m [/] human// pony// (afammn05_16-23)

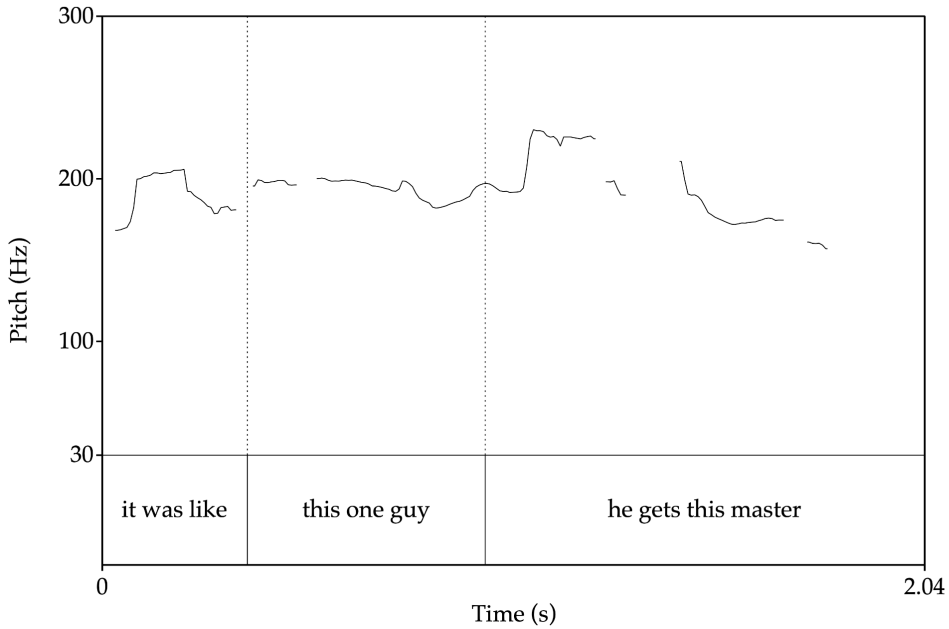

Figure 10: Prosodic contour of TS [20] in (39). Flat type Topic: this one guy; function (I) (female speaker).

In conclusion, the examples above have illustrated a lack of correlation between information structure, Topic prosodic types and discourse functions, which we observed both intra-linguistically and cross-linguistically across the three languages considered.[31]

4. Conclusions: LDs as prominence cues

Based on a functional and textual theoretical approach, and actual data from spoken Italian, Spanish, and American English LDs, we have disentangled the functional account of LDs from the notion of topicality. In fact, we have shown that the concept of Topic is not always relevant to the description of the discourse functions of LDs in the languages under scrutiny both at an information structure level or at a discourse level and that other information structure configurations are possible, as well as other discourse functions in contexts.

[31] A comparison with Brunetti et al. (2010) may offer useful insights, as their study suggests that in Neapolitan Italian, syntax and prosody contribute differently to the marking of informational categories.

We defined LDs as syntactic structures in which the dislocated constituent is extra-clausal and is reduplicated by a co-referential pronoun. This (linear) syntactic definition allowed us to observe the functional behavior of LDs without bias from categorical concepts. To describe the discourse functions of Italian, Spanish, and American English LDs in spontaneous speech, we separately analyzed their syntactic, prosodic, and information features, and then examined how these structures interact with their contexts of occurrence.

In the first step, we demonstrated that there is no systematic correspondence between LDs' syntax and prosodic-information structure. More precisely, although LDs in our data are all syntactically partitioned by definition, not all of them exhibit an information partition, nor are they necessarily characterized by a prosodic boundary between the dislocated element and the co-referential pronoun. In particular, American English LDs are all partitioned; Italian LDs show a prevalence of partitioned LDs; and Spanish LDs show a prevalence of non-partitioned LDs. When LDs are partitioned, the dislocated constituent is placed in a dedicated tone unit with its own information function, which is not necessarily a Topic one. Moreover, the information partition does not depend on the syntactic weight of the left-dislocated constituent or on cognitive load. Last but not least, no acoustic-prosodic feature systematically marks left dislocation; that is, the prosodic contour of a left-dislocated topic constituent varies both within and across languages.

In the second step, we demonstrated that there is no systematic correspondence between prosodic-information traits and discourse functions. More precisely, discourse functions across Italian, Spanish, and American English show both high intra-linguistic and cross-linguistic variation. In particular, American English LDs are syntactically and pragmatically marked, and therefore display low frequency and low flexibility in their textual exploitation. In contrast, the discourse functions of Italian and Spanish LDs are varied and unpredictable, as they are context-dependent. In fact, the discourse functions of Italian and Spanish LDs go far beyond the assumed topicalizing function, and their use is asymmetrical within the dataset analyzed. Moreover, we demonstrated that the presence of an information partition does not systematically correlate with a specific discourse function, and that there is no direct correlation between the prosodic type of Topics and their discourse function, as each Topic prosodic type supports different discourse functions and varies in frequency both within and across languages.

In our view, the data presented argue against a fixed correlation between form and function in LDs. We have shown that, notwithstanding the presence of the pronoun resumption in all the occurrences considered, the discourse functions largely varied intra- and cross-linguistically. In other words, the discourse functions of LDs cannot be predicted based on their syntactic or prosodic-information traits. Instead, describing LDs' functions requires moving beyond an utterance-centered approach in favor of a discourse-centered one – an approach that considers, as essential factors, their interaction with the topic, logic, and dialogic dimensions of the text. Overall, regardless of the specific textual dimension involved, the discourse functions of LDs always involve the discourse prominence of the left-dislocated constituent or its related elements. Indeed, LDs signal to the interlocutor a disruption in the ongoing discourse. The discourse prominence acquired syntactically, informationally, and prosodically by the structure is thus exploited at the discourse level in a variety of contextually relevant ways.

From our perspective, LDs' discourse functions are better understood in light of the notion of discourse prominence (Himmelmann & Primus 2015), defined as a relational and dynamic structure-building principle (von Heusinger & Schumacher 2019: 117). More precisely, LDs can be described as prominence cues used by speakers to signal a disruption in the ongoing discourse, not only at a topic level, but at all levels of the architecture of texts. As we have shown, their discourse functions must be accounted for in relation to at least the topic, logic, and dialogic dimensions of the text. This means that LDs can be considered prominent not only in relation to other topic, but also to other logic relation, or dialogic textual entities. Last but not least, the nature of the discourse prominence acquired by LDs must be considered context-dependent – as it is dynamic and evolves as the text unfolds.

Our approach could be applied to the analysis of LD functions in potentially all languages of the world. Typological studies have shown that syntactic structures corresponding to those defined in this study are universally attested (Maslova & Bernini 2006). Further research is needed to establish intra-linguistic features and discourse functions, which could, in turn, enhance cross-linguistic descriptions of the discourse functions of LDs.

Acknowledgements

We are grateful to the participants of SLE 2022 for the stimulating discussion and valuable insights that contributed to the development of this work, and to the two anonymous reviewers for their comments and suggestions. We also wish to thank the editors, in particular Andrea Sansò and Eleonora Zucchini, for their careful attention and helpful feedback.

Abbreviations

1 = 1st person	DEF = definite	PL = plural
2 = 2nd person	F = feminine	PRS = present
3 = 3rd person	GER = gerund	PST = past
ACC = accusative	INF = infinitive	PTCP = participle
ART = article	IPFV = imperfective	REFL = reflexive
COMP = complementizer	M = masculine	REL = relative
COND = conditional	NEG = negation	SBJV = subjunctive
DAT = dative	OBJ = object	SG = singular

References

Alfano, Iolanda & Violetta Cataldo & Riccardo Orrico & Loredana Schettino. 2021. Sentence topics in Italian: An analysis on the CHROME Corpus. *Loquens* 8(1–2).

Amir, Noam & Vered Silber-Varod & Shlomo Izre'el. 2004. Characteristics of intonation unit boundaries in spontaneous spoken Hebrew: Perception and acoustic correlates. *Proceeding of the 2nd International Conference on Speech Prosody – SP2004*. 677–680. Nara

Asher, Nicolas & Alex Lascarides. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.

Barbosa, Plinio A. 2008. Prominence-and boundary-related acoustic correlations in Brazilian Portuguese read and spontaneous speech. *Proc. speech prosody*. 257–260. Campinas.

Barbosa, Plinio A. 2019. *Prosódia*. São Paulo: Parábola Editorial.

Barth-Weingarten, Dagmar. 2016. *Intonation units revisited. Cesura in talk-in-interaction*. Amsterdam/Philadelphia: John Benjamins.

Baumann, Stefan & Ardnt Riester. 2012. Referential and Lexical Givenness: Semantic, Prosodic and Cognitive Aspects. In Elordieta Gorka & Prieto Pilar (eds.), *Prosody and Meaning*, 119–162. Berlin: Mouton de Gruyter.

Benincà, Paola & Giampaolo Salvi & Lorenza Frison. 1988. L'ordine degli elementi della frase e le costruzioni marcate. In Lorenzo Renzi & Giampaolo Salvi & Anna Cardinaletti (eds.), *Grande Grammatica Italiana di Consultazione*, vol. I, 129–239. Bologna: Il Mulino.

Berretta, Monica. 2002. Ordini marcati dei costituenti maggiori di frase: una rassegna. In Silvia Dal Negro & Bice Mortara Garavelli (eds.), *Temi e percorsi della linguistica. Scritti scelti*, 149–199. Vercelli: Mercurio.

Berruto, Gaetano. 1985. 'Dislocazioni a sinistra' e 'grammatica' dell'italiano parlato. In Annalisa Franchi De Bellis & Leonardo M. Savoia (eds.), *Sintassi e morfologia della lingua italiana d'uso. Teorie e applicazioni descrittive*, 59–82. Roma: Bulzoni.

Biber, Douglas & Stig Johansson & Geoffrey Leech & Susan Conrad & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. New York: Longman.

Birner, Betty J. & Gregory Ward. 1998. *Information Status and Noncanonical Word Order in English*. Amsterdam/Philadelphia: John Benjamins.

Boersma, Paul & David Weenink. 2024. *Praat: doing phonetics by computer* [Computer program]. Version 6.4.13. http://www.praat.org/

Brunetti, Lisa. 2024. Contrast in a QUD-based information-structure model. In Jorna Brysbaert & Jaren Lahousse (eds.), *On the role of contrast in information structure*, 191–224. Berlin: Mouton de Gruyter.

Büring, Daniel. 2016. (Contrastive) Topic. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford Handbook of Information Structure*, 64–85. Oxford: Oxford University Press.

Casielles Suárez, Eugenia. 2003. Left-dislocated Structures in Spanish. *Hispania* 86(2). 326–338.

Cavalcante, Frederico A. 2015. *The topic unit in spontaneous American English: A corpus-based study*. Master's Dissertation, Federal University of Minas Gerais, Belo Horizonte.

Cavalcante, Frederico A. 2020. *The information unit of topic: A crosslinguistic, statistical study based on spontaneous speech corpora.* Doctoral Dissertation, Universidade Federal de Minas Gerais, Belo Horizonte.

Cavalcante, Frederico A. & Adriana C. Ramos 2016. The American English spontaneous speech minicorpus: Architecture and comparability. *CHIMERA: Romance Corpora and Linguistic Studies* 3(2). 99–124.

Cavalcante, Frederico A. & Tommaso Raso & Plinio A. Barbosa 2023. The topic information unit: Modeling prosodic forms in a crosslinguistic perspective. In Marina Castagneto & Miriam Ravetto. (eds.), *La comunicazione parlata. Atti del Congresso SLI–GSCP – Università del Piemonte Orientale (Vercelli, 5–7 maggio 2021)*, 427–460. Roma: Aracne.

Cavalcante, Frederico A. & Tommaso Raso & Giulia Bossaglia & Maryualê M. Mittmann & Bruno Rocha. 2020. Topic unit detection in spontaneous speech: Measuring reliability using the Kappa statistic. *Chimera: Romance Corpora and Linguistic Studies* 7. 69–106.

Chafe, Wallace. 1987. Cognitive constraints on information flow. In Russell S. Tomlin (ed.), *Coherence and grounding in discourse*, 21–51. Amsterdam/Philadelphia: John Benjamins.

Chafe, Wallace. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.

Cimmino, Doriana. 2016. Dislocazioni a sinistra. In Anna-Maria De Cesare & Davide Garassino, Rocío Agar Marco, Ana Albom, Doriana Cimmino, *Sintassi marcata dell'italiano dell'uso medio in prospettiva contrastiva con il francese, lo spagnolo, il tedesco e l'inglese. Uno studio basato sulla scrittura dei quotidiani online*, 379–416. Frankfurt am Main: Peter Lang.

Cimmino, Doriana. 2017. La topicalizzazione in italiano in prospettiva contrastiva con l'inglese. Il caso della scrittura giornalistica online. Doctoral Dissertation, University of Basel.

Cimmino, Doriana. 2020. Proprietà pragmatiche di dislocazioni a sinistra e anteposizioni nel giornalismo online. In Jacqueline Visconti, Manuela Manfredini & Lorenzo Coveri (eds.), *Linguaggi settoriali e specialistici. Sincronia, diacronia, traduzione, variazione. Atti del XV Congresso SILFI* (Genova, 28-30 maggio 2018), 485–492. Firenze: Franco Cesati.

Cimmino, Doriana. 2023. On the topic-marking function of Left Dislocations and Preposings: Variation across spoken and written Italian and English. In Simone Mattiola & Alessandra Barotto (eds.), *Discourse phenomena in typological perspective*, 337–368. Amsterdam/Philadelphia: John Benjamins.

Cimmino, Doriana. 2024. Contrast and left dislocations: Beyond contrastive topics. In Jorina Brysbaert & Karen Lahousse (eds.), *On the place of contrast in information structure: definition, types, encoding and annotation*, 41–68. Berlin: Mouton De Gruyter.

Cimmino, Doriana & Alessandro Panunzi. 2017. La variazione funzionale delle strutture marcate a sinistra in italiano: Uno studio su corpora tra parlato e scritto. *Studi di Grammatica Italiana* 36. 117–179.

Contreras, Heles. 1978. *El orden de palabras en español*. Madrid: Cátedra.

Cresti, Emanuela. 2000. *Corpus di Italiano parlato*. Firenze: Accademia della Crusca.

Cresti, Emanuela & Alessandro Panunzi & Antonietta Scarano. 2005. The Italian corpus. In Emanuela Cresti & Massimo Moneglia (eds.), *C-ORAL-ROM: Integrated reference corpora for Spoken Romance Languages*, 71–110. Amsterdam/Philadelphia: John Benjamins.

Cresti, Emanuela & Massimo Moneglia. (eds.). 2005. *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam/Philadelphia: John Benjamins.

Cresti, Emanuela & Massimo Moneglia. 2010. *Informational patterning theory and the corpus based description of spoken language*. Firenze: Firenze University Press.

Cresti, Emanuela & Massimo Moneglia. 2018a. The illocutionary basis of Information Structure: Language into Act Theory. In Evangelia Adamou & Katharina Haude & Martine Vanhove (eds.), *Information structure in lesser described languages: Studies in prosody and syntax*, 359–401. Amsterdam/Philadelphia: John Benjamins.

Cresti, Emanuela & Massimo Moneglia. 2018b. The definition of the TOPIC within Language into Act Theory and its identification in spontaneous speech corpora. *Revue Romane* 53(1). 30–62.

Cresti, Emanuela & Massimo Moneglia & Alessandro Panunzi. 2018. The LABLITA Corpus & the Language into Act Theory: analysis of Viterbo excerpts. In Amedeo De Dominicis (ed.), *Speech audio archives: preservation, restoration, annotation, aimed at supporting the linguistic analysis. Contributi del Centro Linceo "Beniamino Segre"*, 137, 47–63. Roma: Bardi Edizioni.

Cresti, Emanuela & Massimo Moneglia. 2023. The role of prosody for the expression of illocutionary types: The prosodic system of questions in spoken Italian and French according to Language into Act Theory. *Frontiers in Psychology of Language* 8. 1–28.

Crocco, Claudia & Renata Savy. 2007. Requesting in Italian as a foreign language: Pragmatic and prosodic features. In Anna De Meo & Massimo Pettorino (eds.), *Prosodic and rhythmic aspects of L2 acquisition: the case of Italian*, 153–175. Newcastle upon Tyne: Cambridge Scholars Publishing.

Daneš, František. 1974. Functional sentence perspective and the organisation of the text. In František Daneš (ed.), *Papers on functional sentence perspective*, 106–128. The Hague: Mouton.

Delais-Roussarie, Elisabeth & Jenny Doetjes & Petra Sleeman. 2004. Dislocations in French. In Francis Corblin & Henriëtte de Swart (eds.), *Handbook of French semantics*, 501–528. Stanford: CSLI.

Dilley, Laura C. & Stefanie R. Shattuck-Hufnagel & Mari Ostendorf. 1996. Glottalization of word initial vowels as a function of prosodic structure. *Journal of Phonetcs*, 24(4). 423–444.

Dovetto, Francesca M. & Monica Gemelli. 2013. *Il parlar matto. Schizofrenia tra fenomenologia e linguistica. Il corpus CIPPS*. Roma: Aracne.

Downing, Angela. 1997. Discourse pragmatic functions of the Theme constituent in spoken European Spanish. In John H. Connolly & Roel M. Vismans & Christopher S. Butler & Richard A. Gatward (eds.), *Discourse and pragmatics in functional grammar*, 137–161. Berlin: Mouton de Gruyter.

Du Bois, John W. & Wallace Chafe & Charles Meyer & Sandra Thompson. 2000–2005. *Santa Barbara corpus of spoken American English*. Washington, DC: Linguistic Data Consortium.

Duez, Danielle. 1985. Perception of silent pauses in continuous speech. *Language and Speech* 28(4). 377–389.

Duranti, Alessandro & Elinor Ochs. 1979. Left dislocation in Italian conversation. In Talmy Givón (ed.), *Discourse and syntax*, 377–418. New York: Academic Press.

Feldhausen, Ingo. 2016. The Relation between Prosody and Syntax: The case of different types of Left-Dislocations in Spanish. In Meghan Armstrong & Nicholas Henriksen & Maria del Mar Vanrell (eds.), *Intonational Grammar in Ibero-Romance: Approaches across linguistic subfields*, 153–180. Amsterdam/Philadelphia: John Benjamins.

Ferrari, Angela. 2003. *Le ragioni del testo: Aspetti sintattici e interpuntivi dell'italiano contemporaneo*. Firenze: Accademia della Crusca.

Ferrari, Angela. 2014. The Basel model for paragraph segmentation: The construction units, their relationships and linguistic indication. In Salvador Pons (ed.), *Discourse*

segmentation in Romance languages, 23–54. Amsterdam/Philadelphia: John Benjamins.

Ferrari, Angela & Luca Cignetti & Anna-Maria De Cesare & Letizia Lala & Magda Mandelli & Claudia Ricci & Enrico Roggia. 2008. *L'interfaccia lingua-testo: Natura e funzioni dell'articolazione informativa dell'enunciato.* Alessandria: Edizioni dell'Orso.

Ferrari, Angela & Anna-Maria De Cesare. 2009. La progressione tematica rivisitata. *Vox Romanica* 68. 98–128.

Fernández Ramírez, Salvador. 1986. *Gramática española*. Madrid: Arco/Libros.

Firenzuoli, Valentina. 2003. *Le forme intonative di valore illocutivo dell'italiano parlato: Analisi sperimentale di un corpus di parlato spontaneo (LABLITA)*. Doctoral Dissertation, University of Firenze.

Firenzuoli, Valentina & Sabrina Signorini. 2003. L'unità informativa di topic: Correlati intonativi. In Giovanna Marotta (ed.), *La coarticolazione. Atti delle XIII giornate di studio del Gruppo di Fonetica Sperimentale*, 177–184. Pisa: ETS.

Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential pro: An interface approach to the linking of (null) pronouns. *Natural Language and Linguistic Theory* 25. 691–734.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form*, 87–116. Amsterdam/Philadelphia: John Benjamins.

Garassino, Davide & Jacob Daniel. 2018. Polarity focus and non-canonical syntax in Italian, French and Spanish: Clitic left dislocation and sì che / sí que-constructions. In Christine Dimroth & Stefan Sudhoff (eds.), *The Grammatical Realization of Polarity Contrasts: Theoretical, Empirical, and Typological Approaches*, 227–254. Amsterdam: John Benjamins Publishing.

Giannini, Antonella. 2008. I silenzi del telegiornale. In Massimo Pettorino & Antonella Giannini & Antonella & Marianna Vallone & Renata Savy (eds.), *La comunicazione parlata – Atti del congresso internazionale (Napoli 23-25 febbraio 2006)*, 97–108. Napoli: Liguori editore.

Gili Gaya, Samuel. 1961. *Curso superior de sintaxis española*. Barcelona: Vox.

Givón, Talmy (ed.). 1983. *Topic continuity in discourse: A quantitative crosslanguage study*. Amsterdam/Philadelphia: John Benjamins.

't Hart, Johan. 1981. Differential sensitivity to pitch distance, particularly in speech. *The Journal of the Acoustical Society of America* 69(3). 811–821.

't Hart, Johan & René Collier & Antonie Cohen. 1990. *A perceptual study on intonation: An experimental approach to speech melody*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1). 25–70.

Hermes, Dik J. & Joost C. van Gestel 1991. The frequency scale of speech intonation. *Journal of the Acoustical Society of America* 90. 97–102.

von Heusinger, Klaus & Petra B. Schumacher. 2019. Discourse prominence: definition and application. *Journal of Pragmatics* 154. 117–127.

Hidalgo Downing, Raquel. 2006. *La "dislocación a la izquierda" en el español hablado: Funciones discursivas.* Doctoral Dissertation, Universidad Complutense de Madrid.

Himmelmann, Nikolaus & Beatrice Primus. 2015. Prominence beyond prosody: A first approximation. In Amedeo De Dominicis (ed.), *Prominence in linguistics. Proceedings of the pS-prominenceS international conference*, 38–58. Viterbo: DISUCOM Press, University of Tuscia.

Hirst, Daniel & Albert Di Cristo. (eds.) 1998. *Intonation systems. A survey of twenty languages*, Cambridge: Cambridge University Press.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: The Macmillan Company.

Huddleston, Rodney & Geoffrey Pullum 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Izre'el, Shlomo & Heliana Mello & Alessandro Panunzi & Tommaso Raso. (eds.). 2020. *In search of basic units of spoken language: A corpus-driven approach*. Amsterdam/Philadelphia: John Benjamins.

Kecskes, Istvan. 2012. Is there anyone out there who is interested in the speaker? *Language and Dialogue* 2(2). 285–299.

Lambrecht, Knud. 1994. *Information structure and sentence form: topic, focus and the mental representation of discourse referents*. Cambridge: Cambridge University Press.

Lambrecht, Knud. 2001. Dislocation. In Martin Haspelmath & Ekkeard König & Wulf Oesterreicher & Raible, Wolfgang (eds.), *Language typology and language universals: An international handbook 2*, 1050–1078. Berlin: Mouton de Gruyter.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge: MIT Press.

MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*. 3rd edn. Mahwah: Lawrence Erlbaum Associates.

Mann, William C. & Sandra A Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8(2). 243–281.

Maslova, Elena & Giuliano Bernini. 2006. Sentence topic in the languages of Europe and beyond. In Giuliano Bernini & Marcia L. Schwartz (eds.), *Pragmatic organization of discourse in the languages of Europe*, 67–120. Berlin: Mouton de Gruyter.

Matzinger, Theresa & Nikolaus Ritt & W. Tecumseh Fitch. 2020. Non-native speaker pause patterns closely correspond to those of native speakers at different speech rates. *PLoS ONE 15*: e0230710.

Mello, Heliana & Tommaso Raso. 2011. Illocution, modality, attitude: different names for different categories. In Heliana Mello & Alessandro Panunzi & Tommaso Raso (eds.), *Pragmatics and prosody: Illocution, modality, attitude, information patterning and speech annotation*, 1–18. Firenze: Firenze University Press.

Mello, Heliana & Tommaso Raso & Maryualê M. Mittmann & Heloísa P. Vale & Priscila O. Côrtes. 2012. Transcrição e segmentação prosódica do corpus C-ORAL-BRASIL: Critérios de implementação e validação. In Tommaso Raso & Heliana Mello (eds.). *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal*, 125–174. Belo Horizonte: UFMG.

Mittmann, Maryualê M. 2012. *C-ORAL-BRASIL e o estudo da fala informal: Um novo olhar sobre o tópico no português brasileiro*. Doctoral Dissertation, Universidade Federal de Minas Gerais.

Mittmann, Maryualê M. & Plinio A. Barbosa. 2016. An automatic speech segmentation tool based on multiple acoustic parameters. *CHIMERA: Romance Corpora and Linguistic Studies*, 3(2). 133–147.

Mo, Yoonsook & Jennifer Cole. 2010. Perception of prosodic boundaries in spontaneous speech with and without silent pauses. *The Journal of the Acoustical Society of America* 127(3).

Moneglia, Massimo & Emanuela Cresti. 1997. L'intonazione e i criteri di trascrizione del parlato adulto e infantile. In Umberta Bortolini & Elena Pizzuto (eds.), *Il progetto CHILDES: Strumenti per l'analisi del linguaggio parlato*, vol. II, 57–90. Pisa: Pacini Editore.

Moneglia, Massimo & Tommaso Raso & Maryualê M. Mittmann & Heliana Mello. 2010. Challenging the perceptual relevance of prosodic breaks in multilingual spontaneous speech corpora: C-ORAL-BRASIL/ C-ORAL-ROM. *Speech prosody 2010 conference proceedings*, 1–4. Chicago.

Moneglia, Massimo & Tommaso Raso. 2014. Notes on Language into Act Theory (L-AcT). In Raso, Tommaso & Heliana Mello. (eds.), *Spoken corpora and linguistic studies*, 468–495. Amsterdam/Philadelphia: John Benjamins.

Nicolas Martinez, Carlota. 2012. *Cor-DiAL (Corpus oral didáctico anotado lingüísticamente)*. Madrid: Liceus.

Nicolás Martínez, Carlota & Marina Lombán Somacarrera. 2018. Mini-corpus del español para DB-IPIC. *CHIMERA: Romance Corpora and Linguistic Studies* 5(2). 197–215.

Panunzi, Alessandro & Lorenzo Gregori 2012. DB-IPIC: An XML database for the representation of information structure in spoken language. In Heliana Mello & Alessandro Panunzi & Tommaso Raso (eds.), *Pragmatics and prosody: illocution, modality, attitude, information patterning and speech annotation*, 133–150. Firenze: Firenze University Press.

Panunzi, Alessandro & Maryualê M. Mittmann 2014. The IPIC resource and a cross-linguistic analysis of information structure in Italian and Brazilian Portuguese. In Tommaso Raso & Heliana Mello. (eds.), *Spoken corpora and linguistic studies*, 129–154. Amsterdam/Philadelphia: John Benjamins.

Panunzi, Alessandro & Antonietta Scarano. 2009. Parlato spontaneo e testo: analisi del racconto di vita. In Luisa Amenta & Giuseppe Paternostro (eds.), *I parlanti e le loro storie. Competenze linguistiche, strategie comunicative, livelli di analisi. Atti del convegno (Carini-Valderice, 23-25 ottobre 2008)*, 121–132. Palermo: Centro di studi filologici e linguistici siciliani.

Pépiot, Erwan. 2014. Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. *Speech Prosody* 7. 305–309.

Prandi, Michele. 2006. *Le regole e le scelte. Introduzione alla grammatica italiana*. Torino: UTET.

Prince, Ellen F. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. In Peter Culicover & Louise McNally (eds.), *The limits of syntax*, 281–302. San Diego: Academic Press.

Ramos, Adriana C. 2015. *The use of certainty adverbs in Brazilian Portuguese and American English: A semantic/pragmatic approach*. Master's Dissertation, Universidade Federal de Minas Gerais.

Raso, Tommaso & Maryualê M. Mittmann 2009. Validação estatística dos critérios de segmentação da fala espontânea no corpus C-ORAL-BRASIL. *Revista de Estudos da Linguagem*, 17(2). 73–91.

Raso, Tommaso & Heliana Mello (eds). 2012. *C-ORAL-BRASIL I: Corpus de referência de português brasileiro falado informal.* Belo Horizonte: UFMG Press.

Raso, Tommaso & Bruno Rocha. 2015. Como a categoria de atitude condiciona a metodologia para o estudo das ilocuções. *Revista Diadorim*. 173–197.

Raso, Tommaso & Frederico A. Cavalcante & Maryualê M. Mittmann. 2017. Prosodic forms of the Topic information unit in a cross-linguistic perspective: A first survey. In Anna De Meo & Francesca M. Dovetto (eds.), *Proceedings of the SLI-GSCP International Conference (13–15 June 2016)*. Roma: Aracne Editrice.

Reinhart, Tanya 1981. Pragmatics and linguistics: an analysis of sentence topic. *Philosophica* 27. 53–94.

Rocha, Bruno. 2012. *Características prosódicas do tópico em PE e o uso do pronome lembrete*. Master's Dissertation, Universidade Federal de Minas Gerais.

Saccone, Valentina. 2022. *Le unità del parlato e dello scritto mediato dal computer a confronto: la dimensione testuale della comunicazione spontanea*. Alessandria: Edizioni dell'Orso.

Saccone, Valentina & Panunzi, Alessandro. 2020. Le unità di comment multiplo. Analisi secondo la Teoria della Lingua in Atto. In Anna De Meo & Francesca M. Dovetto (eds.), *La comunicazione parlata. Atti del Congresso SLI–GSCP – Università degli Studi di Napoli "L'Orientale" (Napoli, 12-14 dicembre 2018)*, 263–286. Roma: Aracne.

Saccone, Valentina & Simona Trillocco & Massimo Moneglia. 2023. Markers of schizophrenia at the prosody/pragmatics interface: Evidence from corpora of spontaneous speech interactions. *Frontiers in Psychology*, *Psychology of Language* 14.

Sedano, Mercedes. 2012. Dislocación a la izquierda y a la derecha: semejanzas y diferencias. In Tomás E. Jiménez Juliá & Belén López Meirama & Victoria Vázquez Rozas & Alexandre Veiga (eds.), *Cum corde et in nova grammatica: estudios ofrecidos a Guillermo Rojo*, 791–805. Santiago de Compostela: Universidade de Santiago de Compostela.

Silva Corvalán, Carmen. 1984. Topicalización y pragmática en español. *Revista Española de Lingüística* 14(1). 1–20.

Silverman, Kim & Mari Beckman & John Pitrelli & Mari Ostendorf & Colin Wightman & Patti Price & Janet Pierrehumbert & Julia Hirschberg. 1992. ToBI: a standard for labelling English prosody. *Proceedings of ICSLP92* 2, 867–870.

Simpson, Adrian P. 2009. Phonetic differences between male and female speech. *Language and Linguistics Compass* 3. 621–640.

Sorianello, Patrizia. 2006. *Prosodia. Modelli e ricerca empirica*. Roma: Carocci.

Sornicola, Rosanna. 1981. *Sul parlato*. Bologna: Il Mulino.

Stanford, James. 2016. Sociotonetics using connected speech. *Asia-Pacific Language Variation* 2(1). 409–450.

Tizón-Couto, David. 2012. *Left dislocation in English: a functional-discoursal approach.* Bern: Peter Lang.

Wang, Yu-Tsai & Ignatius Nip & Jordan Green & Ray Kent & Jane Kent & Cara Ullman, Cara. 2012. Accuracy of perceptual and acoustic methods for the detection of inspiratory loci in spontaneous speech. *Behavior Research Methods* 44. 1121–1128.

CONTACT

dorianacimmino7@gmail.com
valentina.saccone@unifi.it