

## Sentence topics in Italian: An analysis on the CHROME Corpus<sup>1</sup>

Iolanda Alfano<sup>1</sup>, Violetta Cataldo<sup>1,2</sup>, Riccardo Orrico<sup>1,3</sup>, Loredana Schettino<sup>1</sup>

<sup>1</sup> University of Salerno, Italy

<sup>2</sup> Ghent University, Belgium

<sup>3</sup> University of Naples "Federico II", Italy

ialfano@unisa.it ORCID: <https://orcid.org/0000-0001-6141-327X>

vcataldo@unisa.it ORCID: <https://orcid.org/0000-0002-3915-1143>

riccardo.orrigo@unina.it ORCID: <https://orcid.org/0000-0001-9260-7210>

lschettino@unisa.it ORCID: <https://orcid.org/0000-0002-3788-3754>

Submitted: 30/09/2021; Accepted: 10/12/2021; Published on line: 21/11/2022

**Citation / Cómo citar este artículo:** Iolanda Alfano, Violetta Cataldo, Riccardo Orrico, Loredana Schettino (2021). Sentence topics in Italian: An analysis on the CHROME Corpus. *Loquens*, 8(1-2), e083. <https://doi.org/10.3989/loquens.2021.083>.

**ABSTRACT:** The present study deals with the phonetic description of sentence topics in Italian tourist guides' speech. Topical coherence characterizes the communicative strategies that human experts adopt when delivering contents to the visitors of cultural sites. Topical progression, which ensures temporal, spatial, and referential continuity, is frequently expressed by sentence topics as well.

The relevant literature generally supports the idea of a topic accent and a rising-falling (or "hat") contour is described as the most frequent for the unmarked topic in Italian utterance structures, but other realizations are also possible. The hypothesis that we want to test in this work is whether this variability is due to specific factors. Hence, we investigate phonetic realization of sentence topics as a function of syntactic features -structure, function and weight- and textual-pragmatic features -discourse role considering ±aboutness, ±contrastiveness, ±givenness-. Specifically, tonal events, i.e., accents and boundaries, phonetic phrasing, and disfluency phenomena were investigated. Results show that both syntactic and pragmatic factors play a role in the phonetic realization of topics, though they act at different levels. In particular, disfluencies are found to be affected by syntactic weight and givenness, while tonal events seem to depend mainly on the discourse role.

**Keywords:** sentence topic, Italian, prosody, discourse role, syntactic structure.

**RESUMEN:** *Los tópicos oracionales en italiano: análisis del Corpus CHROME.*- El presente estudio se plantea un análisis fonético de los tópicos oracionales en italiano, examinando habla de guías turísticos. La coherencia temática caracteriza las estrategias comunicativas que adoptan los expertos humanos al transmitir contenidos a los visitantes de los sitios culturales. La progresión temática, que garantiza la continuidad temporal, espacial y referencial, se expresa también con frecuencia mediante tópicos oracionales. Por lo tanto, el corpus examinado ofrece la posibilidad de analizar la realización de estas entidades.

La bibliografía pertinente apoya, en general, la idea de un acento de tópico e indica un contorno ascendente-descendente como el más frecuente para el tópico no marcado, aunque resultan posibles otras realizaciones.

La hipótesis que queremos comprobar en este trabajo es si la variabilidad encontrada en la bibliografía se debe a factores sintácticos y pragmáticos específicos. Por lo tanto, investigamos la realización fonética de los tópicos oracionales en función de características sintácticas -estructura, función y "peso"- y de factores textuales y pragmáticos

<sup>1</sup> This article is the result of a continuous collaboration between the authors. However, for academic purposes only, Iolanda Alfano is responsible for sections 1, 3.2 and 4.1, Violetta Cataldo for sections 3.1, 3.3 and 4.2, Riccardo Orrico for sections 3.4 and 4.3, Loredana Schettino for sections 2, 3.5 and 4.4. All the authors are responsible for sections 5 and 6.

-rol discursivo considerando los siguientes rasgos:  $\pm$ aboutness,  $\pm$ contrastiveness,  $\pm$ givenness-. En concreto, se investigaron los eventos tonales, es decir, los acentos y las fronteras, el fraseo prosódico y los fenómenos de disfluencia. Los resultados muestran que tanto los factores sintácticos como los pragmáticos desempeñan un papel en la realización fonética de los tópicos oracionales, aunque actúan a diferentes niveles. En particular, las disfluencias se ven afectadas por el peso sintáctico y el estatuto informativo, mientras que los eventos tonales parecen depender principalmente del rol discursivo.

**Palabras clave:** tópico oracional, italiano, prosodia, rol discursivo, estructura sintáctica.

## 1. INTRODUCTION

The notion of sentence topic is one of the most controversial linguistic ideas. From a semantic perspective, Maslova & Bernini (2000) identify two central pieces of evidence against a unified approach to sentence topic: the vagueness of the notion of “aboutness” and the existence of multiple “topic constructions” with different functions, both within and across languages. Yet, they argue for a universal phenomenon of sentence topic, which would allow accounting for language-internal and cross-linguistic variation in topic encoding and for universal constraints on this variation.

By default, a topic is usually defined as what the sentence is about (Reinhart, 1981; Gundel, 1988; Lambrecht, 1994; Krifka, 2008b) and is frequently identified as the single most salient given referent in an utterance. The introduction of the term is due to Hockett (1958, p. 201), who defined it as “What the speaker announces, in a sentence, before proceeding to say something about it, in the Comment”. However, one of the most widely accepted definition of sentence topic and its complement comment is Gundel’s topic definition:

An entity, E, is the topic of a sentence, S, iff in using S the speaker intends to increase the addressee’s knowledge about, request information about, or otherwise get the addressee to act with respect to E. A predication, P, is the comment of a sentence, S, iff in using S the speaker intends P to be assessed relative to the topic of S. (Gundel, 1988, p. 210).

Other proposals consider topicality as a general organizing principle in discourse, where the topic associated with a discourse unit is provided by the (explicit or implicit) question it answers and the relation between discourse units is determined by the relation between these topic-providing questions (van Kuppevelt, 1995).

Within the “Language into Act Theory” (Cresti & Moneglia, 2018), topics are seen as units that develop the information function of field of application for the illocutionary force (expressed by the comment unit). Therefore, they do not convey the illocution of the utterance; they always precede the comment and can be identified in speech only considering their prosodic performance.

Despite the differences between the various theoretical proposals, we can state that they all consider the topic as the basis for what is said or the frame for the most relevant part of the message.

Additionally, the various notions of topic that exist in the literature refer to domains of different extent, e.g., sentence, speech act, dialogue, sub-dialogical parts. This study deals with sentence topics in Neapolitan Italian occurring in the left periphery of declarative sentences.

For the identification of topics in this work, we built on several debated issues in the relevant literature, assuming that sentence topics:

- do not have to be referential, since they can also express situations or states of affairs,
- are optional,
- do not necessarily occur in a fixed position in the utterance,
- are not necessarily given,
- there can be several topics in one utterance.

As for Italian sentence topic intonation, the relevant literature generally supports the idea of a topic accent and a rising-falling (or “hat”) contour is described as the most frequent for the unmarked topic, but other realizations are also possible.

On Florentine Italian, Cresti & Firenzuoli (2002) describe topic contours as composed of two tonal events, namely a rise-fall, situated on the stressed syllable of the topic, and a fall-rise placed on the last syllable of the topic unit. Firenzuoli & Signorini (2003) identify three contours with different frequency of occurrence, which have in common a rise movement on the last tonic syllable of the nucleus.

According to Mereu & Trecci (2004) and Mereu & Frascarelli (2006), topics in Rome Italian in utterance initial position are prosodically marked by a rise of  $f_0$  on the last tonic of the entire constituent, regardless of the more or less complex structure of the topic or its syntactic function.

In another study on the same variety, Giordano & Crocco (2005) examine topics realized as a tonal unit. The authors find that these units are mostly characterized by a “high accent” (44%) but may also show a “low accent” (34%) and other minor realizations. Interestingly, topics with specific syntactic structures, i.e., left dislocations or thematizations, are always realized by the “high accent” configuration.

Crocco & Savy (2007) investigate phonetic phrasing, tonal pattern, and phrase structure in left peripheral sentence topic in dialogues. They indicate that high/rising tones are frequently associated to topics, but their results also show a widespread presence of falling tones (42%), which occurs when topic and tone units are coextensive.

Crucially, different intonational properties as a function of the discourse role played by the topic have been pointed out by Frascarelli & Hinterhölzl (2007) in German and Rome Italian. In both languages, different pitch accents mark different types of topic. In particular, they identify three tonal events associated with topic expressions, i.e., L\*+H for (shifting) aboutness topics, H\* for contrastive topics and L\* for familiar topics. Other prag-

matic contrasts have been investigated in Neapolitan Italian, namely the contrast between regular topics, in exhaustive answers, and contrastive topics, in non-exhaustive answers, i.e., when the topic is a subset of a relevant topical entity (see §2; Büring, 2016). A number of prosodic differences related to the information structure have been noticed. Specifically, the presence of a phrase break and a downstep in the register level after the subject topic distinguish contrastive from non-contrastive topics in SVO constructions (D’Imperio & Cangemi, 2011); in the same way, such prosodic features appear to differentiate also regular *vs.* partial object topics in clitic left dislocation constructions (Brunetti et al., 2010).

Recently, the investigation of these information categories has been extended to other varieties of Italian spoken in Campania (Salerno and Cilento Italian; Cataldo et al., 2021). These results highlight the presence of specific prosodic properties in the realization of contrastive *vs.* non-contrastive topic expressions. Indeed, in both varieties, topics in sentence-initial position are realized as a rise, but partial topics show a wider span and steeper slope of the rise than non-contrastive topics.

Furthermore, the phonetic realization of topic units in speech should take into account the presence of disfluency phenomena. Indeed, disfluency rates and types were found to vary according to different sociolinguistic factors, such as (discourse) topic under discussion (Bortfeld et al., 2001).

As far as we know, analyses involving all these factors have not been carried out systematically. Moreover, although most of the studies refer to declarative utterances, in some cases the modality is not considered, since topics are examined in both declarative and interrogative utterances (Giordano & Crocco, 2005; Crocco & Savy, 2007). Whereas, it has been widely observed that a sharp contrast exists between hanging topic left-dislocations in interrogatives and declaratives: only the former are obligatorily realized with a pause and may have a low edge tone (for example, for Spanish, Feldhausen, 2016). Further evidence in this sense has been collected for Neapolitan Italian (Petroni & D’Imperio, 2011) and German (Petroni & Niebuhr, 2014). In these studies, the cues that allowed the sentence modality discrimination were placed as early as in the prenuclear region, e.g., shape of the  $f_0$  curve, peak alignment of the prenuclear accent as well as the boundary type at the end of the word bearing the accent (incidentally coinciding with sentence topics in those studies).

Hence, our goal is to investigate phonetic realization of sentence topics in declaratives as a function of syntactic features (structure, function, and “weight”) and textual-pragmatic features (discourse role considering  $\pm$ aboutness,  $\pm$ contrastiveness,  $\pm$ givenness).

The hypothesis we want to test is whether the variability found in the literature is due to specific syntactic or pragmatic factors.

The paper is composed of three main parts. The second section (§2) provides the theoretical background, necessary to explain how we considered discourse role and givenness, and the third (§3) is devoted to the descrip-

tion of the method. Finally, in the last sections, the results are summarized (§4) and discussed (§5) and the general conclusions are drawn (§6).

## 2. BACKGROUND

Investigations on sentence topic have concerned different levels of analysis in order to describe its structural and functional characteristics.

As for the syntactic features of topical elements, Reinhart (1981, p. 56) points out that no specific syntactic structure exclusively defines sentence topics because “different expressions of the same sentence can serve as topics in different contexts of utterance”; there is however a tendency in discourse to interpret the grammatical subject of a sentence as its topic. Among others, Brunetti (2009) argues that this tendency may be due to the “agent-like” properties both entities may have. In their analysis of the topic-comment structure of sentences, Gundel and colleagues (1997, p. 2) assume, among other things, that topics need not be sentence initial, nor be represented by noun phrases. Furthermore, in a cross-linguistic study (Gundel et al., 1993), the same authors explore a correlation between the form of referring expressions in discourse and the assumed cognitive status of the referent, i.e., whether it has been already introduced or is somehow accessible. More specifically, they start from the assumption that different determiners and pronominal forms imply different cognitive statuses and identify a hierarchy of givenness statuses that are relevant to the expression of referents, e.g., a definite article conventionally signals the uniquely identifiability of the referent, whereas a demonstrative determiner signals familiarity and identifiability.

As for left peripheral topic structures, Crocco & Savy (2007) in their syntax-prosody interface analysis, find that the great majority of topics are realized as Noun Phrases, followed by a small but substantial number of Prepositional Phrases, and rare cases of Adverbial Phrases. Among the structural features, they also consider the syntactic “weight” of topical constituents (Voghera & Turco, 2008) which they find to correlate with the prosodic pattern of topics, i.e., the syntactic and the prosodic head of the topical constituents coincide in light structures, but they do not in heavy ones (see §3.2).

Especially important in the characterization of sentence topics are the specific functional roles they may play in discourse. Krifka (2008a) identifies two information-structure functions: “addressation” and “delimitation”.

Addressation is the function related to the default definition of topic or “aboutness topic” as “what the sentence is about” (Reinhart, 1981; Gundel, 1988; Lambrecht, 1994; Krifka, 2008b). Specifically, “the topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the Common Ground content.” (Krifka, 2008b, p. 41). So just like information stored in a file card system, new information is not added to the common ground in an unstructured way, but rather associated with entities.

For example: “*Peter* fell asleep.” (Krifka, 2008a, p. 1). In this sentence, ‘Peter’ is the entity pointed out, while that he ‘fell asleep’ is the information added about the entity.

Götze and his colleagues (2007) point out a number of features that an aboutness topic needs to possess: a topic of a sentence is an aboutness topic (AT) if the sentence would be a natural continuation of “Let me tell you something about AT?”, a good answer to the question “What about AT?”, and could be transformed into the sentence “Concerning AT, ...”, (Götze et al., 2007, p. 165).

Delimitation is the function of “frame-setting”, which means defining the domain under which the predication should be interpreted. For example: [How is Bill doing?] “*Financially*, he’s doing fine.” (Krifka, 2008a, p. 1). In this example, ‘Financially’ has the function of delimiting the condition under which the predication (i.e., that Bill is doing fine) holds. It evokes alternatives (e.g., his health, his love life, and so on) relevant to the big issue (i.e., how Bill is doing) and for which other predications might hold, as in “*Financially*, he’s doing fine, but he had a heart operation last month.”

Topical entities may also perform both an addressation and delimiting function, meaning that an aboutness topic may evoke contrastive alternatives. In literature, such cases are referred to as “contrastive topics” (CT; Büring, 2016). For example: [How are your parents doing?] “*My father* is doing fine.” [alternative evoked: my mother] (Krifka, 2008a, p. 2).

Crucially, the CT-alternatives are relevant to the sentence containing the CT, though they are not answered. More specifically, for a sentence containing a contrastive topic to be felicitous, there must be at least one question meaning which is (i) currently pertinent, (ii) logically independent, and (iii) identifiable (Büring, 2016). According to Büring’s formalization, the marking of a CT triggers these requirements to be understood as conventional implicatures. If we apply the rules to the previous example ‘*My father* is doing fine.’, we could say that it is possible to identify at least one question that instantiates ‘who is doing fine?’ that is pertinent and independent of the sentence itself (i.e., my mother). Büring further distinguishes between two uses of contrastive topics: “partial topic” and “purely implicational topic”. When the topic in an answer is a subset of the topical entity in the question, it is acknowledged as partial topic, which is typically compatible with multiple wh-questions, or single wh-question containing plurals as in the following examples (Büring, 2016, p. 68): “Which guest brought what?” “*Fred* brought the beans”. “Where do your siblings live?” “*My sister* lives in Stockholm”. Instead, when the topic answers to the question but evokes potentially relevant alternatives to the question, it is referred to as purely implicational topic. For example, in “Where was the gardener at the time of the murder? The *gardener* was in the house” the speaker wants to highlight that there are other people potentially relevant to the question (who is the murderer), e.g., “where was the chauffer?”, “where was the cook?”, and so on.

From this picture emerges that “aboutness” and “contrastiveness” are the two relevant dimensions to distin-

guish three types of topics: “Aboutness Topic”, aboutness, non-contrastive topics; “Contrastive Topics”, aboutness, contrastive topics; “Frame-setting Topics”, non-aboutness, contrastive topics.

Another textual-pragmatic aspect which may be relevant to the expression of topical entities is the information status, i.e., or the degree of givenness in context.

The notion of givenness has been approached from different perspectives (cf. among others, Mathesius, 1929; Halliday, 1967; Sgall, 1972; Chafe, 1976; Firbas, 1987; Gundel et al., 1993; Gundel, 2003; for an overview on information structure units, see von Heusinger, 2002). A relevant account for the purpose of the present study is proposed by Baumann & Riester (2012). The authors elaborate on the notion of information status to provide a possibly complete definition that is able to take into account the important aspects related to an item’s givenness or novelty and is useful for investigations on the interface between information structure and prosody. In their proposal, they consider the level of cognitive activation of the discourse referents a central aspect for the analysis of an item’s givenness. In this view, consciousness is determined “first and foremost” by the dynamic discourse context that is regarded as “a cognitive dimension shared by the interlocutors at the time and place of utterance” (Baumann & Riester, 2012, p. 123). Furthermore, the authors argue that two levels of givenness should be considered to account for an item’s information status: referential and lexical givenness. On the referential level, an item is given (to a certain degree of activation) if there is a coreferential antecedent, meaning that a reference to it can be found in the previous context; on the lexical level, an item is given if there is an identical expression, a synonym, or a hyponym in the previous context.

### 3. METHOD

#### 3.1. Corpus

The corpus we analysed was collected within the Italian national project CHROME (Cultural Heritage Resources Orienting Multimodal Experiences), which aims at developing a data collection and annotation procedure to support the development of new interactive technologies for cultural heritage. The audiovisual recordings involve three art historians. Each recruited expert guide accompanies four groups of four people in an hour-long guided tour at the San Martino Charterhouse in Naples (for more details, see Origlia et al., 2018).

The dataset under investigation consists of 80’ of speech (about 27’ per guide). A total amount of 228 topic items was found and annotated according to syntactic and pragmatic features.

#### 3.2. Syntactic features

Both structural and functional aspects were considered, according to the type of phrases and their syntactic Weight.

As for syntactic Weight, Voghera & Turco (2008) provide a scale for verbal phrases and for nominal phrases (Noun, Predicative Noun and Prepositional Phrases). We refer to the nominal scale since Crocco & Savy (2007) found that topics are mostly realized as Noun, Prepositional and Adverbial Phrases. The scale takes into account both phrases' structure and expansion and considers five levels of weight according to the presence/absence of determiners (det) and modifiers (mod) and whether the head is a noun or a pronoun (pro):

1. [+ det] [+ mod]
2. [+ det] [- mod]
3. [- det] [+ mod]
4. [- det] [- mod]
5. [+ pro] [- det] [- mod].

We adapted the scale observing not only the presence/absence of modifiers, but also the type of modifier (adjective, Prepositional Phrase, relative clause). On the other hand, as proposed by Crocco & Savy (2007), we did not consider determiners. Accordingly, we established four levels of weight, measured as a function of the presence and type of modifiers, classifying Light (L), Medium (M), Heavy (H) and very Heavy (H+) phrases (Table 1).

**Table 1:** Weight scale and examples from the CHROME corpus.

Weight	Modifier	Example
L	-	<i>la certosa</i> ['the charterhouse']
M	+ adj.	<i>la nostra certosa</i> ['our charterhouse']
H	+ PP	<i>la Certosa di Napoli</i> ['the Charterhouse in Naples']
H+	+ relative clauses	<i>la Certosa che vediamo oggi</i> ['the Charterhouse we see today']

### 3.3. Pragmatic features

Topics were classified according to the features of aboutness, which defines the topic as the entity that the sentence is about (Reinhart, 1981; Gundel, 1988; Lambrecht, 1994; Krifka, 2008b) and contrastiveness, which evokes alternative topics for which other predications might hold (Büring, 2016).

The dimension of aboutness was evaluated on the basis of the test proposed by Götze and his colleagues:

- X is the Aboutness Topic of a sentence S containing X if
- a. S would be a natural continuation to the announcement Let me tell you something about X
  - b. S would be a good answer to the question What about X?
  - c. S could be naturally transformed into the sentence Concerning X, S', where S' differs from S only in-

sofar as X has been replaced by a suitable pronoun. (Gotze et al., 2007, p. 165).

As for the dimension of contrastiveness, the following test was elaborated based on Büring's (2016) formalization: X is the Contrastive Topic in a sentence S containing X if

- a. There is an easily identifiable alternative X' evoked by X
- b. X' is independent from S (i.e., the information about X' is not resolved in S)
- c. X' is pertinent with reference to S (i.e., it would contribute to address a bigger issue that is relevant for the current discourse).

Three discourse roles were identified on the textual transcription of the recordings only (without hearing them, in order to avoid circularity): *Neutral* (N-Topic), *Frame Settings* (FS-Topic) and *Contrastive* (C-Topic), see Table 2.

**Table 2:** Discourse roles identified according to  $\pm$ aboutness and  $\pm$ contrastiveness features.

Aboutness	Contrastiveness	Discourse function
+	-	N-Topic
+	+	C-Topic
-	+	FS-Topic

N-Topic functions to address, i.e., it points out an entity or a reference point, adding an information about it. FS-Topic functions to delimitate the domain under which the predication should be interpreted (Krifka, 2008a). C-Topic functions to address a bigger issue that is relevant for the discourse, when it evokes an easily identifiable alternative (see § 2). The example (1) illustrates the three types:

(1) *La Certosa di San Martino (N-Topic) ha almeno due anime. Una (C-Topic) racconta la storia della città. All'epoca (FS-Topic) i certosini (N-Topic) vivevano in un luogo isolato*

[*San Martino Charterhouse (N-Topic) has two souls at least. One (C-Topic) tells the story of the city. At the time (FS-Topic) Carthusian monks (N-Topic) lived in an isolated place.*]

Furthermore, topic Givenness was considered. In order to reduce the degree of arbitrariness in the annotation of the informative status, we considered givenness basing on a textual analysis, being impossible to make assumptions about the state of activation of the information in the mind of the addressee (see §2). Referring to denotations and not to linguistic expressions, -therefore, considering referential and not lexical givenness- we distinguished among *New* (not mentioned before and not being recov-

erable from the preceding discourse), *Given* (mentioned in the immediate common ground content) and *Resumed*, already mentioned, but not in the immediate common ground content (Baumann & Riester, 2012). The example (2) shows the three types:

(2) [...] *cercherò di farvi immaginare insomma di com'era com'era la vita in questo quartiere, quando non era ancora di fatto un quartiere, settecento anni fa, nel 1300. Non so se ti è capitato, ti do del tu, di passeggiare per le strade del Vomero oggi, ma adesso è un quartiere appunto molto animato, ci sono molti negozi, molti locali, ma in realtà questa è una trasformazione anche questa relativamente recente. Nel '300 la collina del Vomero (Given) era pressoché disabitata, quindi si trattava di campagna, un luogo perfetto per i certosini che naturalmente cercavano una vita di solitudine, una vita appartata. E il duca di Calabria Carlo, figlio del re Roberto il Saggio (New), incoraggia appunto i certosini a stabilirsi in questo luogo.*

[...]  
*Carlo, Duca di Calabria (Resumed), in realtà non vede la fine dei lavori perché ufficialmente si concludono nel 1368 quando lui era già morto.*

[I will try to make you imagine what life was like in this neighbourhood, when it was not yet a neighbourhood, seven hundred years ago, in 1300. I don't know if you've ever walked through the streets of Vomero today, but now it's a very lively neighbourhood, there are lots of shops, lots of bars and restaurants, but it's actually a relatively recent transformation. In the 14th century, the Vomero hill (Given) was almost uninhabited, so it was the countryside, a perfect place for the Carthusians who naturally sought a life of solitude, a secluded life. And the Duke of Calabria Charles, son of King Robert the Wise (New), encouraged the Carthusians to settle there.

[...]  
*Charles, Duke of Calabria (Resumed), actually did not see the end of the works because they were officially completed in 1368 when he was already dead]*

The first occurrence of the topics “In the fourteenth century” and “the Vomero hill” were considered *Given*, since both time and place had already been introduced in the discourse. On the contrary, the topic “the Duke of Calabria, Charles” had not been previously mentioned, nor was recoverable; therefore, was considered *New*. Finally, the same discourse topic was tagged as a *Resumed* sentence topic during the same visit with the same group of people.

### 3.4. Intonation analysis

Firstly, a phrasing level was labelled, isolating tone units (TUs), considering a number of phonetic boundary markers, not necessarily co-occurring, i.e., presence of a final pause;  $f_0$  declination of both  $f_0$  and energy; parametrical reset at the beginning of a new TU; prepausal lengthening.

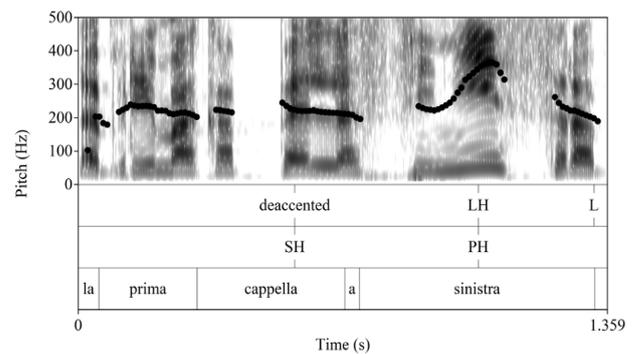
Then, we analysed:

- Pitch movements on the syntactic head (SH) and on the prosodic head (PH), i.e., the stress that may be considered hierarchically higher than any other prominence in the topic. It corresponds to SH in light phrases;
- Boundary (B) of topical entity;
- General pitch span information, measured as global maxima minus minima in semitones (ST).

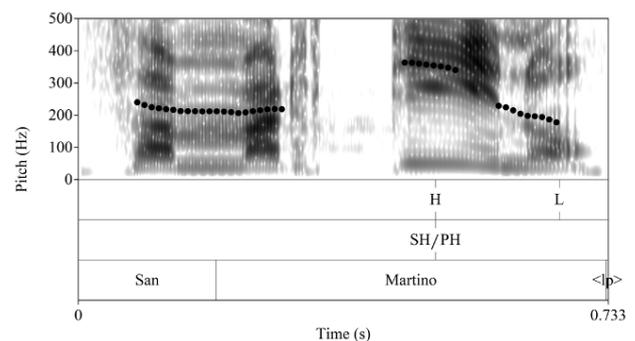
Only in topics occurring as separate TUs, PHs and Bs were considered.

These parameters were phonetically described as follows. Pitch movements on SHs and PHs were grouped into rising, falling, high, low tones and deaccented (non-prominent realizations). Figure 1 shows an example of deaccented SH and rising PH, whereas Figure 2 shows an example of coincident SH and PH.

**Figure 1:** Deaccented SH and rising PH in the topic “La prima cappella a sinistra” [The first chapel on the left].



**Figure 2:** Coincident SH and PH in the topic “San Martino”.



Bs, on the other hand, were classified as high (above the baseline) or low (baseline level).

### 3.5. Disfluencies

As for disfluencies, we considered cases of repair (deletions, substitutions, insertions) and hesitation (silent, filled, lexicalized filled pauses, and prolongations; Shriberg, 1994; Eklund, 2004). For example:

- (3) <ehm>quindi <eeh> il<ll> / Carlo diciamo  
[‘<ehm> so <eeh> the<ee> / Carlo like’]

Each disfluent item was labelled according to its position, namely before, within and after each topic item and was further classified according to its disfluency complexity, distinguishing among simple instances, when just one phenomenon occurred, see example (4), or complex ones, for two or more phenomena, see example (5).

- (4) la <aa> Certosa di San Martino qui a Napoli ha almeno due anime  
[the<ee> San Martino Charterhouse here in Naples has at least two souls]

- (5) questi <eeh> lavori <sp> <eeh> di ammodernamento quindi incominciano alla fine del millecinquecento  
[these <eeh> works <sp> <eeh> of modernization therefore begin at the end of the 16<sup>th</sup> century]

## 4. RESULTS

In this section we report the results from the analysis of sentence topics within our corpus. The results are presented according to the different levels of analysis: topics’ syntactic features (§4.1) and the pragmatic role they play within the discourse (§4.2), their prosodic realization (§4.3) and whether or not they include disfluent phenomena (§4.4).

### 4.1. Syntactic features of sentence topics

A total of 228 items of sentence topic were found in the corpus. While in most cases the items identified were the only topic within the sentence, in 29 occurrences they were found to be combined with each other within the same sentence; of these topic combinations, 26 were made of two distinct topical items, and the remaining 3 were made of 3 items.

The analysis performed at the syntactic level shows that, in our corpus, topics can be realized as Noun Phrases (NPs), Prepositional Phrases (PPs), Adverbial Phrases (AdPs), as shown in Table 3.

**Table 3:** Count and percentage of topic occurrences according to type of phrases.

Type	Count	Perc.
NP	150	66%
PP	56	25%
AdP	22	9%

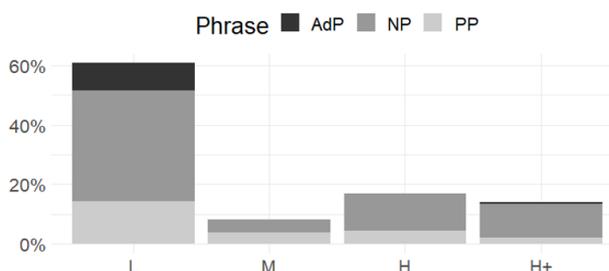
NPs were attested to be the most frequent, amounting to 150 items (around 70% of total occurrences), followed by PPs and, finally, by AdPs. As far as their syntactic function is concerned, in 99% of cases, NPs

serve the function of subject of the sentence, while PPs and AdPs are mainly employed as adjuncts of time or place.

As for the syntactic Weight, we found that the most frequent occurrences of topic items are made of one phrase in which the head is not accompanied by modifiers, i.e., light topics (L; over 60% of the total occurrences). More complex topics were also attested, though with a much smaller frequency: specifically, M topics (those with an adjectival modifier) amount to 8% of the total occurrences, H topics (with a Prepositional Phrase modifier) were found in the 17% of the cases, and H+ topics (in which the modifier is realized as a relative clause) account for the 14% of the cases. No specific relationship appears to be there between syntactic Weight and type of phrase: the distributions of weight described above are maintained across the phrase types, with the exception of AdPs, for which only one non-L occurrence was found.

Figure 3 reports the percentages of occurrence of the Weights and the type of phrase.

**Figure 3:** Percentage of the topic occurrences according to the syntactic Weight and phrase type.



### 4.2. Pragmatic features of sentence topics

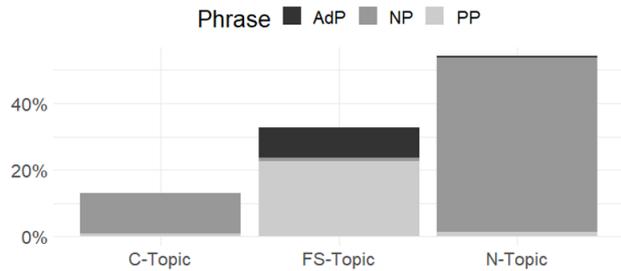
As for pragmatic features, the analyses that we performed were concerned with identifying topics’ discourse role and their degree of givenness. Results show that the topic items were not evenly distributed according to the discourse role they play. Most of the topics analysed (54%) were found to be neutral topics (N-Topic); frame-setting topics (FS-Topic) were found to account for the 32% of the topic occurrences and, finally, contrastive topics (C-Topic) account only for the remaining 14%. Additionally, it might be noteworthy to mention that in case of topic combinations, at least one of the isolable topic items is always a FS-Topic, as shown in example (6).

- (6) *Alla fine del Millecinquecento*<sub>FS-Topic</sub> *con la Controriforma*<sub>FS-Topic</sub> *i certosini*<sub>N-Topic</sub>

[‘at the end of 1500 / with the Counter-Reformation / the Carthusian monks’].

The example reported above also shows another tendency that we attested. In particular, the discourse role classification reveals a correlation with the type of phrase topics were realized by, as shown in the bar plot in Figure 4.

Figure 4: Percentage of phrase types in the three discourse roles.



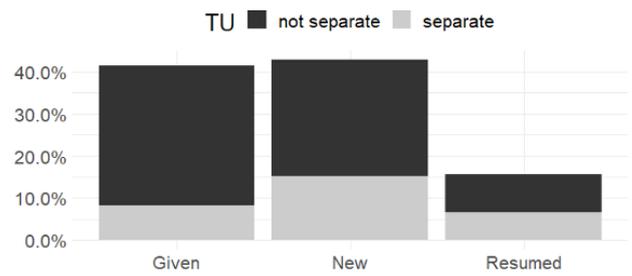
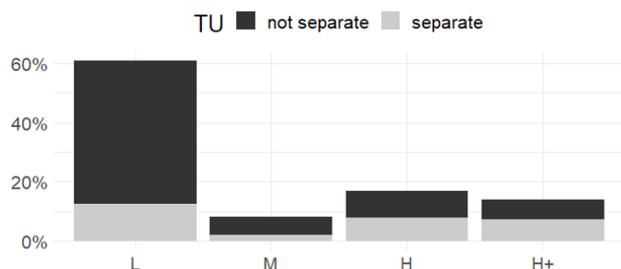
NPs are primarily used to express an aboutness topic, i.e., N-Topic and C-Topic, and these categories are almost never realized by other types of phrases. FS-Topics, on the other hand, are mainly realized by PPs and, in a smaller percentage of cases, by AdPs.

Finally, as far as Givenness is concerned, we found that nearly half of topic items are *New* (43%), while the remaining 57% shows different degrees of givenness: *Given* (41.2%) and *Resumed* (15.8%).

### 4.3. Intonation

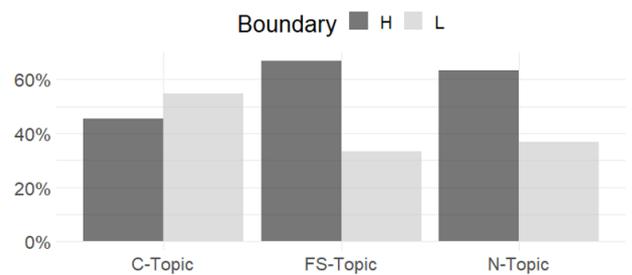
The first result concerns the separation of the items analysed in TUs. We found that 34.8% of topics are tonally separated from the rhematic part of the utterance, therefore most of the items are not contained within an isolated unit. A general inspection of these cases shows that the realization of topic items as belonging or not to the same tonal unit might be linked to factors that are of both syntactic and pragmatic nature. Specifically, distributional data show that L topics are far less frequently realized as a separate TU; similarly, tonal separation of topics appears to be also linked to Givenness, with a slightly higher frequency of separation in case of *New* topics. These distributions are showed in Figure 5.

Figure 5: Frequency of separate vs. non-separate tone units as a function of Weight (top) and Givenness (bottom).



Among the realizations of topics as a separate TU, we found that the 62.3% of the occurrences displays a high boundary tone. The distribution of boundary type, however, was found to be linked to the discourse role (see Figure 6); statistically, the significance of this relationship was measured using the Chi-Square, which shows that the distribution of boundary type is indeed dependent on the discourse role ( $\chi^2 = 11.56, p = .003$ ).

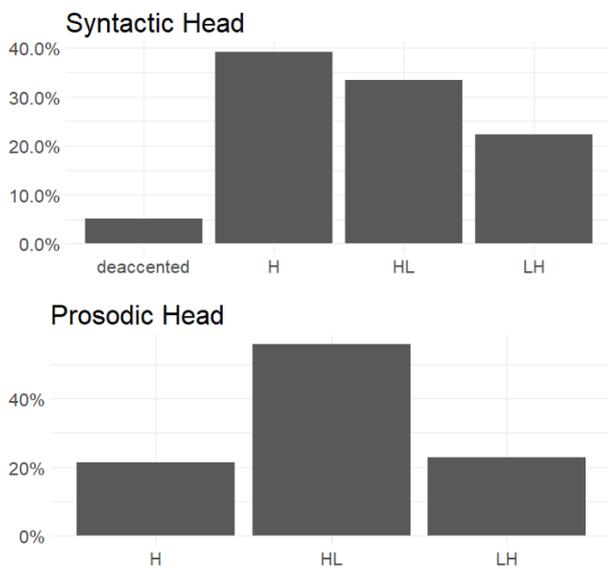
Figure 6: Boundary realization as a function of discourse role.



More specifically, and as shown in Figure 6, N- and FS-Topic items were found to show a higher frequency of high boundaries, while C-Topics show a slightly higher percentage of low boundaries, both as opposed to the occurrence of H boundaries for the same category and as opposed to the frequency of occurrence of Ls in the other topic types.

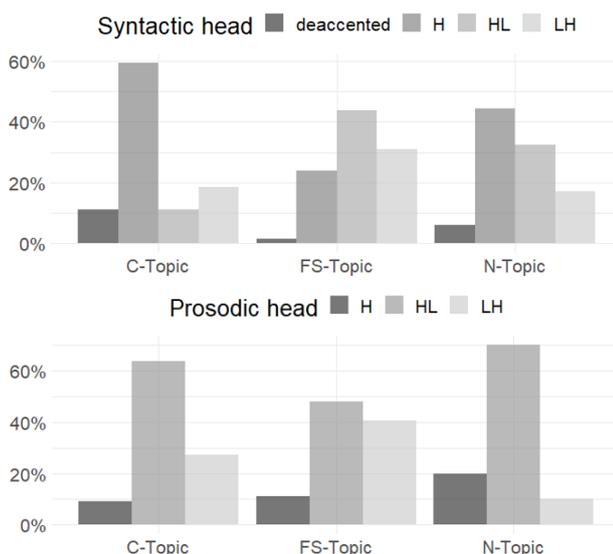
As for the tonal movements associated with strong syllables, we analysed accents occurring on SHs and those occurring on the TU's PH (in case PH was within the topic). As shown in section 3.4, three different accent types were attested, i.e., high, fall, and rise accents (plus, we included cases in which SH was de-accented). The general distributions of the different tonal conditions for PH and SH are shown in Figure 7. In general, we found that, in both positions, falling accents (HL) are very frequent, while rising accents (LH) were attested to be the least frequent within the corpus. A difference, however, was also found for the position (SH or PH), namely the frequency of occurrence of high accents (H), which appears to be the most frequent category used for SH, though the least used for prosodic heads.

**Figure 7:** Frequency of accent type in topics' syntactic head (top) and prosodic head (bottom).



As for accents' distribution according to pragmatic features of the topic, we found a correlation between the discourse role played by the topic and tonal movements associated with both SH ( $x^2 = 56.33$ ,  $p < .001$ ) and PH ( $x^2 = 34.16$ ,  $p < .001$ ), reported in Figure 8.

**Figure 8:** Frequency of accent type in topics' syntactic head (top) and prosodic head (bottom) as a function of discourse role.



More specifically, as for SH, C-Topics and N-Topics (those characterized by the feature +aboutness) show a similar behaviour: they are more likely realized with high accents

on their SH, which occur, respectively, in 56.3% and 44.4% of the cases. On the contrary, FS-Topics show a much more variable picture, though there appears to be a tendency to be realized mainly with a fall on the syntactic head. Additionally, FS-Topics also present a higher percentage of LH accents as opposed to the other two categories. These results, however, might be due to the fact that, in some cases, SH coincide with PH. For this reason, the same computations were made in heavy constituents only (H and H+), in which SH is clearly separated from PH, as in example 7.

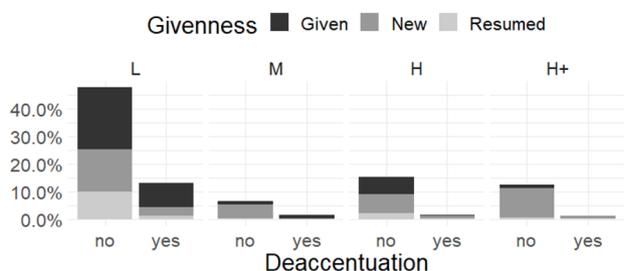
(7) *Il Duca di Calabria, Carlo, figlio di Re Roberto il Saggio*

[The Duke of Calabria, Charles, son of King Robert the Wise]

In these cases, high accents appear to be the most frequent type in all discourse roles. As for the tonal movement associated with PH, it was found that falling accents are the most frequent in all the conditions, though an interesting tendency appears to be there linking rising accents with C- and FS-Topics (hence, +contrastive topics): the frequency of LH accents is pretty high in these two categories as opposed to what is observed for N-Topics.

As for Givenness, we did not find any specific correlation with the type of accent or boundary used by the speakers, though this feature appears to be linked on the one hand to deaccentuation of SH and, on the other, to general pitch span information (see below). It is shown in Figure 9 that some of the topics that are not realized as a separate TU were found to be deaccented i.e., topic constituents that do not bear any prosodic prominence, which occurs in 19.5% of the cases. The vast majority of deaccented topics (93.5%) is either a pronoun or a deictic element, which are often anaphoric and hence *Given*. In general, however, it appears to be the case that +Givenness is a necessary condition, interacting with syntactic Weight, for deaccenting a topic.

**Figure 9:** Frequency of deaccented items as a function of Weight and Givenness.



Deaccentuation therefore is much more frequent when a topic constituent is light and, at the same time, *Given*. Deaccented cases, indeed, occur in the following situations: a) when SH is a pronoun or a deictic element characterized by givenness and “phonetic lightness” (*lui* [‘he’], *qui* [‘here’]),

and b) in only two occurrences of *Given* light NPs, such as *I certosini* [‘Carthusian monks’]. Moreover, we found deaccented C-Topics, when SH is modified by a PP, which bears the main prominence (*la prima cappella a SINISTRA* [‘the first chapel on the LEFT’]).

Finally, we found that Givenness is a strong predictor of phonetic information linked to pitch span. The correlation between span and pragmatic features of utterances was modelled in R using linear regression. The model was built with Span (in ST) as dependent variable, while Givenness, Discourse role, and their interaction were employed as independent variables. Table 4 shows a detail of the results of the model, indicating that Givenness is the only predictor for pitch span variability, whose effect appears to be independent of discourse role, since the interaction between these two variables did not yield significance. Additionally, and contrarily to previous investigations, the main effect of Discourse role also yielded non-significant results.

**Table 4:** Details of the outcome of the linear regression model.

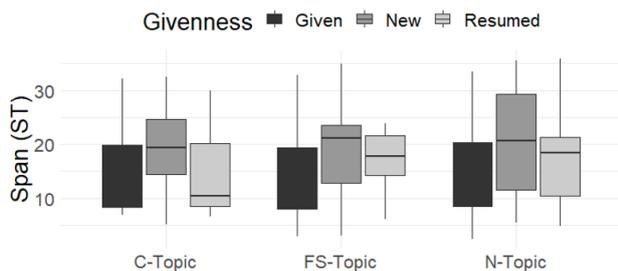
Predictor	Df	F	p-value
Givenness	2	5.855	0.003*
Discourse Role	2	0.395	0.674
Givenness: Discourse Role	4	0.069	0.991

To further investigate the effect of Givenness, we also performed a pairwise comparison across the three levels (*Given, New, Resumed*) within the variable; this will also provide information about both direction and magnitude of the effect of Givenness. The data are reported in Table 5 and Figure 10.

**Table 5:** Pairwise comparisons of levels within the variable Givenness.

Pair	Estim.	Std. Err	t-value	p-value
Given–New	-4.71	1.638	-2.876	0.012*
Given–Resumed	-1.61	2.406	-0.668	0.779
New–Resumed	3.10	2.240	1.386	0.345

**Figure 10:** Pitch span (ST) as a function of discourse role and Givenness.



#### 4.4. Disfluencies

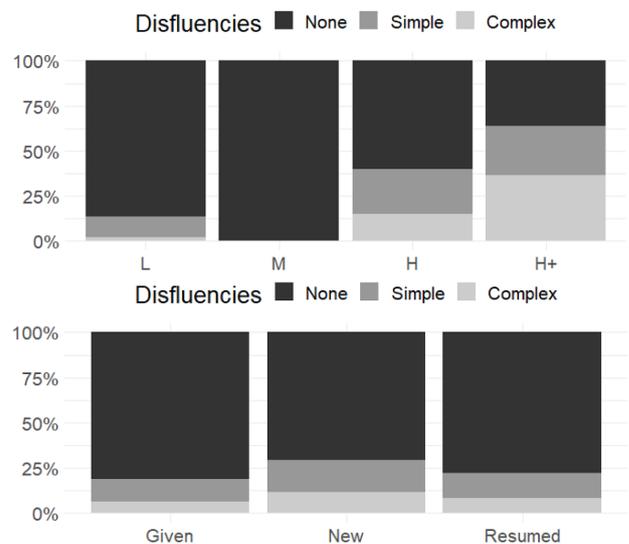
The last result we present concerns the role of disfluencies in the production of topic items, by analysing the occurrence of disfluencies, their complexity, and their position with respect to the topic.

The analysis of disfluencies showed that around half of the topic items under investigation (54%) was found to be realized with disfluent phenomena, mostly occurring within the topic unit (in 46% of the cases) and/or before it (33%). As for disfluency complexity, we found that more than half of disfluent topics show simple disfluencies, i.e., with only one disfluent phenomenon (more specifically, in 56% of the cases) while the remaining 44% presents complex disfluent sequences.

Moreover, the presence of disfluencies was found to be significantly related to topics’ syntactic Weight ( $x^2 = 17.09$ ;  $p < .001$ ). Indeed, unlike L and M phrases, H and H+ ones are most likely realized as disfluent sequences. Such significance regards the disfluencies occurring within topics, indicating that more complex topic items are more frequently realized with a disfluent phenomenon located within the item itself. Furthermore, syntactic Weight was also found to have a significant correlation with disfluency complexity ( $x^2 = 27.57$ ;  $p < .001$ ). As a matter of fact, in L, M and H phrases only simple disfluent sequences occur, whereas around half of H+ phrases are realized with a higher number of disfluent phenomena.

The occurrence of disfluent phenomena was also found to vary as a function of the topic’s pragmatic features. Specifically, the presence of disfluencies increases within new topics. However, similarly to the correlation with Weight, significance was only found for disfluencies occurring within topics ( $x^2 = 6.68$ ;  $p < .01$ ). Figure 11 shows the distributions of disfluencies according to both Weight and Givenness.

**Figure 11:** Presence and complexity of disfluencies as a function of Weight (top) and Givenness (bottom).



## 5. DISCUSSION

Topical coherence characterizes the communicative strategies that experts adopt when delivering contents to the visitors of cultural sites. Topical progression, which ensures temporal, spatial, and referential continuity, in the tourist guides' type of speech is frequently expressed by sentence topics as well.

Starting from the syntactic features, in most cases the topic is constituted by Noun Phrases performing the function of subject and with light weight, in line with previous studies (Crocco & Savy, 2007; Voghera & Turco, 2008), even if both phrase structure and weight are quite variable (Table 3). Among the syntactic features considered, Weight appears to have the most effect on the phonetic realization of topical expressions. Specifically, heavy constituents are more likely realized as separate tonal units, as compared to lighter phrases which are more often incorporated in the following tonal sequence. Furthermore, we found a correlation between Weight and the presence of disfluencies, which mainly occur in complex combinations, within syntactically heavy and very heavy phrases. This result confirms the studies reported in Lickely (2015) which highlight that the production of long and/or complex constituents triggers the employment of hesitations as a means of online planning devices, hence assuming the link between the occurrence of disfluent phenomena and cognitively demanding utterances. Accordingly, both results unveil a greater effort affecting the phonetic realization of heavy topical constituents, which, on the one hand, need to be tonally stand-alone for rhythmical reasons, and on the other hand, require more hesitations for planning reasons.

As for pragmatic features, the prevalent information-structure function in the communicative situation under analysis is addressation (N- and C-Topics), even if delimitation is clearly expressed as well (FS- and C-Topics). Turning to the information status of topics, different degrees of givenness were considered. Indeed, sentence topics can denote an expression that is not present in the immediate common ground in nearly half of the cases. This implies that, in line with expectations, *New* (discourse) topics are frequently introduced by means of sentence topics in this kind of speech.

Looking at the effect of pragmatic features on the phonetic realization, we found a general correlation between discourse role and tonal events, as already stated by Frascarelli and Hinterhölzl (2007) in a different theoretical framework. In topics uttered as a separate TU, boundaries are realized as low in C-Topics and as high in N- and FS-Topics. Accent choice is also modulated by discourse role and, in particular, nuclear accents seem to convey this information. Indeed, the feature +aboutness correlates with high or falling accents, while in +contrastive topics a higher occurrence of rising accents was registered, especially when they carry the feature -aboutness (i.e., FS-Topics). Such a distinction based on the accentual realization is in line with previous investigations on the same variety examined in the present work, namely Neapolitan Italian, and other geographically close Italian

varieties. In particular, contrastive topics were found to be prosodically different from non-contrastive ones (Brunetti et al., 2010; D'Imperio & Cangemi, 2011; Cataldo et al., 2021). As a matter of fact, unlike these studies, in our data pitch range does not play a role in discriminating different discourse roles. However, the feature of  $\pm$ contrastiveness – the category of “Contrast”, as suggested by Brunetti and colleagues – appears to be clearly represented by prosodic means, although via different prosodic cues and at least as regards Campania Italian varieties.

Correlations between Givenness and phonetic realization were also detected. Firstly, the presence of disfluencies is much more likely in *New* entities, which imply a greater cognitive effort. This result is indeed in line with both Barr's (2001) experimental results showing that speakers tend to be more disfluent when introducing new information, and Arnold et al. (2003), who found that hesitation phenomena occur in production when referring to discourse-new items and help speech comprehension of discourse (given) status. Secondly, along the lines of Avesani & Vayra (2005) who found deaccentuation of given constituents in task-oriented dialogues, we found that topics can be deaccented only if they are *Given*. However, in both studies, speakers are more likely to accent *Given* information (93% of instances in Avesani & Vayra, about 80% in our dataset). Moreover, our findings are also in line with Sbranna and colleagues (2021), who did not detect strategies of deaccentuation in sentence-final given items in Neapolitan Italian.

Lastly, global span information correlates with Givenness: *New* topics show a wider span. Similarly, Féry & Ishihara (2010) showed that givenness has an effect on different prosodic domains, including pitch range; indeed, items with different degrees of givenness are ordered along a hierarchy of pitch height. In particular, first occurrences of a word are realized higher than its second occurrences which in turn are higher than all the other following occurrences.

Finally, our results suggest that, despite the variability detected, general prosodic information is an essential part of the definition of the discourse role of sentence topics which, in some cases, might represent the only way for the speaker to encode that meaning and, for the listener, to take up the specific topic as either exhaustive or a smaller part of a bigger issue. This brings out a critical issue from the methodological point of view. The identification of the pragmatic features of our topic items was carried out on a textual basis only, i.e., only relying on the textual transcription of the speech data; such a criterion had the aim of avoiding circularity between the identification of the parameters and the phonetic realization of the related topic expressions. Crucially, in light of our results, the phonetic features we found to correlate with pragmatics are the cues that should be considered for the topic characterization in that they might represent the only specification of the role of topics in discourse. In particular, prosodic features, i.e., accentual and boundary realization, can act as discriminating factors with regard to discourse role, as to say to identify textual features of  $\pm$ aboutness and  $\pm$ contrastiveness.

## 6. CONCLUSIONS

The aim of this study was to investigate the realization of sentence topics in Italian in order to explore whether the variability detected in previous descriptions of topic realization could be ascribed to specific factors. Specifically, we deepened the role of syntactic and pragmatic factors. To this end, phonetic realization, considering accents and boundaries, prosodic phrasing, pitch span and disfluency phenomena, was investigated as a function of topics' syntactic (Phrase structure, Function, and Weight) and textual-pragmatic (Discourse role and Givenness) features. For this purpose, a dataset of semi-spontaneous and semi-monologic speech was selected; a total of 228 topic items uttered by female speakers was found and analysed.

Our findings suggest that textual-pragmatic features,  $\pm$ aboutness,  $\pm$ contrastiveness,  $\pm$ givenness, and syntactic weight covary with phonetic properties. In particular, the intonational features, namely accent and boundary type, correlate with the topics' discourse role, the global span with the information status, namely *New* topics, and the presence of disfluent phenomena with syntactic heavy constituents and *New* instances.

## 7. REFERENCES

- Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal three, um, new information. *Journal of Psycholinguistic Research*, 32(1), 25-36.
- Avesani, C., & Vayra, M. (2005). Accenting, deaccenting and information structure in Italian dialogue. In L. Dybkjaer & W. Minker (Eds.), *6th SIGdial Workshop on Discourse and Dialogue* (pp. 19-24). Lisbon, Portugal, September 2-3. ISCA Archive, <http://www.isca-speech.org/archive/sigdial6>
- Barr, D. J. (2001). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In C. Cavé, I. Guaitella, & S. Santi (Eds.), *Oralité et gestualité: Interactions et comportements multimodaux dans la communication* (pp. 597-600). Paris: L'Harmattan.
- Baumann, S. & Riestter, A. (2012). Referential and lexical givenness: Semantic, prosodic and cognitive aspects. *Prosody and Meaning*, 25, 119-162. <http://dx.doi.org/10.1515/9783110261790.119>
- Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, 44(2), 123-147. <http://dx.doi.org/10.1177/00238309010440020101>
- Brunetti, L. (2009). On the semantic and contextual factors that determine topic selection in Italian and Spanish. *The Linguistic Review*, 26(2-3), 261-289. <http://dx.doi.org/10.1515/tlir.2009.010>
- Brunetti, L., D'imperio, M. & Cangemi, F. (2010). On the prosodic marking of contrast in Romance sentence topic: evidence from Neapolitan Italian. In M. Hasegawa-Johnson (Ed.), *Proceedings of the 5th International Conference on Speech Prosody 2010* (pp. 1-4). Chicago, IL, USA, May 10-14. ISCA Archive, <http://www.isca-speech.org/sp2010/ITRW>
- Büring, D. (2016). (Contrastive) topic. In C. Féry & S. Ishihara (Eds.), *The Oxford Handbook of Information Structure* (pp. 64-85). Oxford: Oxford University Press.
- Cataldo, V.; Orrico, R. & Crocco, C. (2021, June 21-23). Topic and focus accents in closely related varieties of Campania Italian [Poster presentation]. *4th Phonetics and Phonology in Europe (PaPE 2021)*, Barcelona.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (Ed.), *Subject and Topic* (pp. 25-55). Cambridge, MA: Academic Press.
- Cresti, E. & Firenzuoli, V. (2002). L'articolazione informativa topic-comment e comment-appendice: correlati intonativi. In A. Regnicoli (Ed.), *La fonetica acustica come strumento di analisi della variazione linguistica in Italia. Atti delle XII Giornate di studio del Gruppo di Fonetica Sperimentale* (pp. 153-160). Roma: Il Calamo.
- Cresti, E. & Moneglia, M. (2018). The illocutionary basis of Information Structure. Language into Act Theory (L-Act). In E. Adamou, K. Haude, & M. Vanhove (Eds.), *Information Structure in Lesser-described Languages: Studies in Prosody and Syntax* (pp. 359-401). Amsterdam: John Benjamins.
- Crocco, C. & Savy, R. (2007). Topic in dialogue: prosodic and syntactic features. In G. Murray & S. Renals (Eds.), *Proceedings of the 8th Annual Conference of the International Speech Communication Association* (pp. 114-117). Antwerp, Belgium, August, 27-31. International Speech Communication Association (ISCA)
- D'Imperio, M. & Cangemi, F. (2011). Phrasing, register level down-step and partial topic constructions in Neapolitan Italian. In C. Gabriel & C. Lléo. (Eds.), *Hamburg Studies on Multilingualism*, 10 (pp. 75-94). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/hsm.10.05dim>
- Eklund, R. (2004). *Disfluency in Swedish Human-human and Human-machine Travel Booking Dialogues*. [PhD thesis, Linköping University]. Electronic Press.
- Feldhausen, I. (2016). The relation between prosody and syntax: The case of different types of left-dislocations in Spanish. In M. Armstrong, N. Henriksen, & M. M. Vanrell (Eds.), *Interdisciplinary Approaches to Intonational Grammar in Ibero-Romance* (pp. 153-180). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/ihll.6.08fel>
- Féry, C. & Ishihara, S. (2010). How focus and givenness shape prosody. In M. Zimmerman & C. Féry (Eds.), *Information Structure* (pp. 36-63). Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199570959.003.0003>
- Firbas, J. (1987). Il funzionamento del dinamismo comunicativo nella prospettiva funzionale della frase. In R. Sornicola & A. Svoboda (Eds.), *Il campo di tensione. La sintassi della Scuola di Praga* (pp. 195-209). Naples: Liguori.
- Firenzuoli, V. & Signorini, S. (2003). L'unità informativa di topic: correlati intonativi. In G. Marotta & N. Nocchi (Eds.), *La coarticolazione. Atti delle XIII Giornate di studio del Gruppo di Fonetica Sperimentale* (pp. 177-184). Pisa: Edizioni ETS.
- Frascarelli, M. & Hinterhölzl, R. (2007). Types of topics in German and Italian. In S. Winkler & K. Schwabe (Eds.), *On Information Structure, Meaning and Form* (pp. 87-116). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/la.100.07fra>
- Giordano, R. & Crocco, C. (2005). Sul rapporto tra intonazione e articolazione informative. In F. Albano Leoni & R. Giordano (Eds.), *Italiano parlato. Analisi di un dialogo* (pp. 159-188). Naples: Liguori.
- Götze, M., Weskott, T., Endriss, C., Fiedler, I., Hinterwimmer, S., Petrova, S., Schwarz, A., Skopeteas, S., & Stoel, R. (2007). Information structure. In S. Dipper, M. Götze, & S. Skopeteas (Eds.), *Interdisciplinary Studies on Information Structure, Working Papers of the SFB 63*, 7 (pp. 147-187). Postdam: Universitätsverlag.
- Gundel, J. (1988). Universals of topic-comment structure. In M. Hammond, E. Moravcsik, & J. Wirth (Eds.), *Studies in Syntactic Typology* (pp. 209-239). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/tsl.17.16gun>
- Gundel, J., Hedberg, N. & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69(2), 274-307. <http://dx.doi.org/10.2307/416535>
- Gundel, J., Hedberg, N., & Zacharski, R. (1997). Topic-comment structure, syntactic structure and prosodic tune. *Workshop on Prosody and Grammar in Interaction* (pp. 13-15). Helsinki, Finland.
- Gundel, J. (2003). Information structure and referential givenness/newness. How much belongs in the grammar? *Journal of Cognitive Science*, 4, 177-199. <http://dx.doi.org/10.21248/hpsg.2003.8>
- Halliday, M. A. (1967). Notes on transitivity and theme in English. *Journal of Linguistics*, 3, 199-244.
- Hockett, C.F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.
- Krifka, M. (2008a). What do contrastive topics and frame setters have in common? The role of addressing and delimitation in informa-

- tion structure. *Conference on Contrastive Information Structure Analysis (CISA 2008)*. Mars, 18-19. University of Wuppertal.
- Krifka, M. (2008b). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3-4), 243-276. <http://dx.doi.org/10.1556/ALing.55.2008.3-4.2>
- Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic Focus and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.
- Maslova, E. & Bernini, G. (2000). Sentence topics in the languages of Europe and beyond. In G. Bernini & M. Schwarz (Eds.), *Pragmatic Organization of Discourse in the Languages of Europe* (pp. 67-120). The Hague: Mouton de Gruyter. <http://dx.doi.org/10.1515/9783110892222.67>
- Mathesius, W. (1929). *Sulla cosiddetta articolazione attuale della frase*. In R. Sornicola & A. Svoboda (Eds.), *Il campo di tensione. La sintassi della Scuola di Praga* (pp. 181-194). Liguori.
- Mereu, L. & Frascarelli, M., (2006). L'interfaccia Sintassi-Fonologia. Interpretazione e implicazioni teoriche. In R. Savy & C. Crocco (Eds.), *Analisi prosodica. Teorie, modelli e sistemi di annotazione. Atti del II Convegno Nazionale dell'Associazione Italiana di Scienze della Voce* (pp. 256-285). Mantova: EDK Editore.
- Mereu, L. & Trecci, A. (2004). Focus sul topic. In F. Albano Leoni, F. Cutugno, M. Pettorino, & R. Savy (Eds.), *Atti del Convegno Il Parlato italiano* [CD-ROM]. Naples: M. D'Auria Editore.
- Origlia, A., Savy, R., Poggi, I., Cutugno, F., Alfano, I., D'Errico, F., Vincze, L., & Cataldo, V. (2018). An audiovisual corpus of guided tours in cultural sites: Data collection protocols in the CHROME Project. In B. N. De Carolis, C. Gena, T. Kufflik, A. Origlia, & G. E. Raptis (Eds.), *Proceedings of the 2018 AVI-CH Workshop on Advanced Visual Interfaces for Cultural Heritage* (pp. 1-4). <https://dl.acm.org/doi/10.1145/3206505.3206597>
- Petrone, C., & D'Imperio, M. (2011). From tones to tunes: Effects of the f0 prenuclear region in the perception of Neapolitan statements and questions. *Prosodic Categories: Production, Perception and Comprehension* (pp. 207-230). Berlin: Springer. [http://dx.doi.org/10.1007/978-94-007-0137-3\\_9](http://dx.doi.org/10.1007/978-94-007-0137-3_9)
- Petrone, C., & Niebuhr, O. (2014). On the intonation of German intonation questions: The role of the prenuclear region. *Language and Speech*, 57(1), 108-146. <http://dx.doi.org/10.1177/0023830913495651>
- Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics in pragmatics and philosophy. *Philosophica*, 27(1), 53-94.
- Sbranna, S., Ventura, C., Albert, A., & Grice, M. (2021). Prosodic expression of information status in Italian. [Poster presentation]. *4<sup>th</sup> Phonetics and Phonology in Europe (PaPE 2021)*, June, 21-23, Barcelona.
- Sgall, R. (1972). Topic, fuoco e ordine degli elementi nelle rappresentazioni semantiche. In R. Sornicola & A. Svoboda (Eds.), *Il campo di tensione. La sintassi della Scuola di Praga* (pp. 195-209). Liguori.
- Shriberg, E.E. (1994). *Preliminaries to a Theory of Speech Disfluencies* [Unpublished PhD thesis]. University of California.
- van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics*, 31, 109-149.
- Voghera, M. & Turco, G. (2008). Il peso del parlare e dello scrivere. In M. Pettorino, A. Giannini, M. Vallone & R. Savy (Eds.), *La comunicazione parlata* (pp. 727-760). Naples: Liguori.
- von Heusinger, K. (2002). Information structure and the partition of sentence meaning. In E. Hajičová, P. Sgall, J. Hana & T. Hoskovec (Eds.), *Prague Linguistic Circle Papers: Travaux du Cercle Linguistique de Prague nouvelle série*, 4 (pp. 275-305). Amsterdam : John Benjamins. <http://dx.doi.org/10.1075/plcp.4.14heu>