Discourse markers and turn-planning at the pragmatics-prosody interface: the case of *allora* in spoken Italian

1. Introduction

Discourse Markers (DMs) are “those elements that –partly drained from their original meaning– serve to connect phrasal, inter- and extra-phrasal elements, to underline the discourse structure, to convey the utterance’s placement in the interpersonal dimension, and to signal ongoing cognitive processes”[1] (Bazzanella 1995: 225). A defining property of discourse markers (DMs) is that they are interpreted at the interface between syntax, pragmatics, and prosody (Bazzanella 1995; Schiffrin 1987 a.o.). Despite a general agreement on this in the literature, an integrated approach to the study of DMs putting together the different angles is still to be developed. In particular, while it is generally agreed that one should take prosody into account when dealing with DMs (Bazzanella 1995; Schiffrin 1987), much work is still needed to understand how the intonation of the markers interacts with syntactic position and discursive interpretations. We aim at giving a contribution in that direction, by analyzing the Italian DM *allora* in its prosodic and discursive relationship with the turn where it occurs in spontaneous speech. *Allora* is originally a temporal adverb (meaning ‘then’, ‘at that time’), which has developed also a discursive use (Molinelli 2018; Miecznikowski, Gili Fivela & Bazzanella 2008; Ghezzi & Molinelli 2020). In spoken Italian, *allora* is indeed more often used as a DM than as an adverb, and it displays several “speaker-related, textual and interaction-managing functions” (Bazzanella et al. 2007: 13). Its main function is to take the floor to open

[1] Our translation. Among the many definitions given for the term “discourse markers” (cf. Fraser 1999; Schiffrin 1987), we follow the definition from Bazzanella (1995) because, in our view, it captures at best the multifunctional nature of DMs.

the turn in a dialogical exchange, but also to introduce a new move or action in the interactional sequence (Bazzanella 1995; Bazzanella et al. 2007).

In this article, we analyze a number of phonetic features of DM *allora* in turn-initial position. The paper is organized as follows: in the next section, we briefly review the main findings on the syntax and semantics-pragmatics of DMs, which will serve as the baseline for our investigation. In Section 3 we present our experiment: Section 3.1 illustrates the methodology, and Section 3.2 presents the results of the prosodic analysis. Section 3.3 offers an overview of our findings, which are discussed in Section 4. Finally, Section 5 draws the conclusions.

2. Discourse markers at pragmatics - prosody interface

The classic studies on DMs (Bazzanella 1995; Schiffrin 1987 a.o.) show that these elements do not contribute to the truth-compositional semantics of the sentence. The meaning of DMs is bleached as a result of a grammaticalization (Roberts & Roussou 1999) or pragmaticalization process (see Degand & Evers-Vermeul 2015 for an overview). Therefore, DMs are "external to the propositional content" (Bazzanella 2006: 456): they rather convey the pragmatic flavors connected to it, and they are the direct expression of the speaker's attitudes and his/her epistemic and emotional world (Zimmermann 2011: 2027; Corver 2015: 1). As for syntax, most formal approaches ascribe DMs to the left periphery of the sentence, above Rizzi's (1997) ForceP (Bayer, Hinterhölzl & Trotzke 2015; Cardinaletti 2015; Zimmermann 2008) In such works, DMs are taken to realize syntactic projections within Speas and Tenny's (2003) Speech Act Layer (see also Hill 2007 and much subsequent work: Coniglio & Zegrean 2012; Haegeman & Hill 2014 a.o.), in which the points of view and the attitudes of the participants to the conversation are encoded.

The idea that the prosodic contour with which the marker is realized has a role in its discursive-pragmatic interpretation is a red line crossing over the theoretical background sketched so far (see also Aijmer 2002; Dehé & Wichmann 2010; Maschler 2009). Nevertheless, despite a rich body of work demonstrating that prosodic features can convey different semantic-pragmatic interpretations (Author & Author 2019; Bocci 2013; Frascarelli 2008; Frascarelli 1999; Watson & Gibson 2004; Orrico & D'Imperio 2020; Büring 2016; Elordieta & Prieto 2013), more work is needed to understand how and which prosodic features are relevant for the interpretation of DMs. In fact, the particles' intonation not only can be "used" by the speakers to convey distinct pragmatic or interactional functions (Savino 2011; Cerrato & D'Imperio 2003; Didirková, Christodoulides & Simon 2018; Raso & Vieira 2016), but it also disambiguates between the "sentential" and the "discursive" use of the same lexical item (Hirschberg & Litman 1993; Wennerstrom 2001)[2]. Furthermore, by "rethinking prosody dialogically" (Couper-Kuhlen 2014: 222), much work in the framework of Conversation Analysis (CA) has shown that both prosody and DMs have a central role also beyond the sentence[3], at the level of turn-taking and topic-managing (Ogden 2006; Couper-Kuhlen 2004, 2014; Local 2003; Schegloff & Lerner 2009; Walker 2012). In this view, DMs "collaborate" – on a par with

[2] That is, the use through which DMs convey a semantic value (according to the grammatical category they belong to, e.g. adverbs, verbs etc.) *vs* their use as grammaticalized particles.

[3]Different research approaches take chunks of different length and nature as units of analysis (Zellers & Post 2012:120–121), and the very same notion of sentence is still under debate (Cresti 2001, 2018; Moneglia 2006; Martin 2015). We thank an anonymous reviewer to have brought this point to our attention. Throughout the paper, we refer to the notion of sentence as defined in the formal syntactic literature, i.e. as the unit composed by the clause and the left periphery on top of it (Rizzi 2004 and much related work). The relationship between the syntactic, pragmatic and prosodic units of analysis would deserve a dedicated work. Nevertheless, we believe that our findings on the prosodic properties of turn-initial *allora* could help to characterize the left periphery of the sentence, which has been broadly discussed in the syntactic literature (Rizzi 1997 and much subsequent work).

other linguistic (interjections, repairs…) and extra-linguistic (breathing, gestures, and gaze) cues – with phonetic features (pitch, rhythm, timing, loudness…) to express different attitudinal, discursive and sociolinguistic meanings (Prieto & Roseano 2021; Wennerstrom 2001).

Our work is placed against this background, in which prosody and DMs are conceived as being the two sides of the same coin. They collaborate to give pivotal information about the organization of the discourse, both at the macro-level of the turn organization (Bolden 2009; Couper-Kuhlen 2004; Local 2003; Wiklund 2014; Wichmann 2007) and the micro-level of a single turn (Walker 2012; Local & Walker 2004). In this work, we demonstrate experimentally how the two levels are interconnected. For this study we draw from both the Phonetics of Talk-in-Interaction (PTI) and the Empirical Phonology (EP) frameworks (Zellers & Post 2012:120–121)[4]: we analyze dialogical turns from spontaneous task-oriented interactions, taking the Intonational Phrase (IP) as the unit of analysis.[5] The IP has been theorized by the most prominent representative of the EP approach, namely the Autosegmental–Metrical phonology (AM). In this framework, the IP is the domain of the intonational contour made of (at least one nuclear) pitch accent and the edge tones (Pierrehumbert 1980; Ladd 1996; Beckman & Pierrehumbert 1986), and it is related to the morpho-syntactic structure (i.e. to the sentence) via mapping rules (Selkirk 1984; Truckenbrodt 1999; Bennett & Elfner 2019; Elordieta 2008; Nespor & Vogel 1986). We investigate the Italian DM *allora* occurring at turn openings and

[4] The convenience of combining different approaches and methodologies is discussed in Zellers & Post (2012).

[5] The dataset of this paper is constituted by turns, i.e. the dialogical chunks initiated by one speaker in the conversation, relying on the transcription of the CLIPS corpus (www.clips.unina.it. The methodology for the dialogue transcription is described in Savy 2007). However, since the focus of the present work is the prosodic and discursive relationship between *allora* and the rest of the turn, the unit of our analysis is a prosodic unit, i.e. the intonational phrase (IP). For each turn, we consider for the analysis the following units: the discourse maker *allora*, produced as an independent IP + the following IP (see main text for more details).

being uttered as an independent IP, to inquire its relationship with the IP that follows within the same turn. We show how the DM's prosodic features can tell us in what discursive relationship the marker is with the rest of the turn and, consequently, with the whole dialogical set-up.

It has been shown that the first position of the turn has an essential role in the organization of the discourse, as it hosts elements that at the same time connect the sentence to the preceding context, and function as the harbinger of what follows (Schegloff 1996: 77; Heritage 2013: 333; García García 2021: 40). This seems to hold in various languages, and some studies have demonstrated (more or less explicitly) that turn-initial elements are often "outside" the core illocution, from both a discursive and a prosodic point of view (Raso & Vieira 2016; Leemann & Siebenhaar 2006). Furthermore, there is evidence of the fact that the core proposition generally presents a higher intonational contour, and/or it displays a pitch reset compared to the elements that introduce it. Couper-Kuhlen (2001, 2004) shows that a high onset or a sudden upward shift of the pitch can signal the beginning of a new action in the conversation. What also emerges from these works, however, is that the pitch peaks can be preceded by elements – DMs or "lexical prefaces", in the author's terminology – which "serve to alert the interlocutor that something is about to happen" (Couper-Kuhlen 2001: 338). Crucially, these particles are found outside the high onset, and their f0 is generally low. The same emerges in Wennerstrom (2001: 100), who considers prosody as a (type of) DM: "Because many discourse markers perform organizational and interactional functions in text, they are extraneous to the propositional content and likely to have L* pitch accents. This does not exclude the association of lexical discourse markers with H* or L+H* pitch accents in cases where the marker does play a role in the information structure of the discourse."

As it appears, DMs function as "satellites" of the host sentence, giving relevant information for both the speaker and the hearer on the dialogical organization of the turn, but at the same time being outside the principal grammatical and discursive-prosodic planning implanted by the speaker. Our work aims to give a prosodic characterization of this insight. To do so, we analyze the Italian DM *allora* occurring in turn-initial position in Map Task dialogues. Based on the works discussed so far, we assume that the “satellite” nature of DMs should be reflected in their pitch contour. In particular, we expect that the turn-initial *allora*, when realized as an independent IP:

(i) is often reduced and uttered at a faster speech rate compared to the rest of the sentence

(ii) is uttered at a globally lower pitch level compared to the following IP

(iii) the IP that follows the *allora* presents a pitch reset

The first prediction is grounded on some evidence of the fact that DMs are normally phonetically reduced (Frank-Job 2005), as they undergo phonetic erosion as a consequence of grammaticalization (Lehmann 1985; Roberts & Roussou 1999). The second and third predictions are strictly related to one another, and they are based on the already mentioned literature showing that the beginning of a new topic and/or a new dialogical move is usually marked with an upward shift of the pitch at the beginning of the sentence (see also Brazil, Coulthard & Johns 1980; Sluijter & Terken 1993; Wichmann 2007).

3. The experiment

3.1 Method

3.1.1 The data

Our database consists of task-oriented dialogical exchanges between native speakers of Italian, elicited with the Map Task technique (MT) (Anderson et al. 1991) for the online corpus CLIPS.[6] The MT consists of a game in which the two speakers have two different versions of the same map. The task of each participant is to duplicate the other speaker's route on the map. To do so, the speakers –who cannot see each other– have to ask and give directions to the interlocutor. They alternate in the roles of "instruction giver" –the speaker who guides the interlocutor to draw the portion of the route– and "instruction follower" –the interlocutor who follows the indications of the giver, to draw the missing part of the route. The resulting interaction consists of structured dialogues in which the implemented pragmatic moves and the alternation of the turns are easily individuated and controlled for the analysis. While MTs are pragmatically constrained by the type of task performed by the speakers, they are spontaneous as far as syntax and prosody are concerned, since the speakers can speak freely to execute the task. The DM *allora* occurs in the corpus 919 times in total. From these occurrences, we extracted 95 turns prefaced by *allora* followed by a silent pause (>50ms). The reason why we chose to focus on the cases in which *allora* is followed by a pause is twofold: first, the silent break helps to individuate the boundaries between the DM and the rest and to isolate the two for the prosodic analysis. Second, the separated marker represents in our opinion the most typical case of the "introducing" DM (as discussed in Section 2). The choice of *allora* followed by a pause seemed then the best option to investigate the prosodic relationship between the DM and the turn in which it occurs. The reader should anyways bear in mind that this is only one of the configurations in which the DM can be found[7] (for a more

---

[6] See Sobrero & Tempesta (2006) for a detailed description of both the informants and the places of recordings of the CLIPS corpus.

[7] Notice that we also excluded the cases where the turn-initial *allora* was clustered with other DMs, like in the example below, in which *allora* occurs in a sequence with *vabbè* 'whatever' and *comunque* 'anyways':

i. *p1g#97: no vabbe' allora <ehm> comunque <sp> a sinistra della<aa> torta*
   p1g#97: no whatever so <ehm> anyways <sp>to the left of the cake

detailed discussion on this point, see Section 4). For each of the *allora*-prefaced turns, we have first individuated the pragmatic function of the marker by looking at the dialogical context. We have then compared the prosodic characteristics of the DM itself (henceforth *allora*-IP) and the subsequent IP (henceforth sub-IP).

3.1.2 Conversational move coding

For the classification of the pragmatic function of the *allora*-prefaced turns, we referred to the HCRC coding system (Carletta et al. 1996). For this study, we applied the HCRC's lower level of coding, the Move Coding Scheme, in which the dialogues are analyzed as being made up of conversational moves of various kinds, as summarized in Table 1 below. The conversational move of each turn prefaced by *allora* was tagged on the transcribed dialogues using the Nvivo12 software (QSR International 2018).

Initiating Moves	Move
	align
	check
	explain
	instruct
	query-w
	query-yn
Response Moves	acknowledge
	clarify
	reply-y
	reply-n
	reply-w

Table 1: the HCRC Move Coding Scheme

As Table 1 shows, the moves can pertain to two categories[8]: the Initiating Moves are the moves used by the speakers to begin a new action or to introduce a new discourse. They are often found at the beginning of the game, and they are mostly used by the speaker in the role of instruction-giver. The Response Moves are moves through which the speaker replies to a previous dialogical move, and they are mostly used by the speakers in the instruction-follower role. For a detailed description of all the moves, the reader is referred to the above-cited work.

3.1.3 Prosodic analysis

The prosodic analysis was conducted using Praat (Boersma & Weenink 2021). Different measures were taken comparing *allora*-IP and sub-IP of each of the *allora*-prefaced turns. First, the absolute duration (in seconds) of each IP was extracted and the syllables of each IP were counted based on orthography. Then, the speech rate (SR) of the two IPs was calculated in syllable/seconds. We also measured the duration of the syllables of the DM, to investigate potential reduction. Pitch Span (PS) was calculated for the two IPs separately as the difference between the maximum and the minimum f0 levels (DeLooze et al. 2014). Then, the *allora*-IP and the sub-IP of each turn were compared through the measure of their maximum f0 peak, i.e. Pitch Height (PH). Finally, we checked for the presence of a reset of the pitch (Ladd 1988) between the two IPs in each turn. To do so, we took the measure labeled Pitch First Syllable (PFS), as follows (adapted from Swerts 1997; see also Chow 2005): the highest f0 value on a stressed syllable at the vowel's maximum intensity was taken in two points, that is the stressed syllable of *allora* and the first stressed syllable following the silent

[8] In the original coding scheme, a third category is present: the Ready Move. It encompasses moves that signal the end of a game and/or prepare for a new one. This category was not relevant for our analysis, since none of the turns in our data fell under it. We then decided to exclude it whatsoever from the classification.

pause. The pitch reset was then estimated as the difference between the pitch values after and before the boundary. All the measures are summarized in Table 2 below:

Name	Measure	Abbreviation
Speech Rate	Syllable/sec	SR
Pitch Span	f0 maximum – f0 minimum	PS
Pitch Height	f0 maximum	PH
Pitch First Syllable (pitch reset)	f0 first stressed syllable of the IP	PFS

Table 2: prosodic measures taken for each of the *allora*-prefaced turns

All the measures were taken in Hz and then converted to the ERB scale (Nolan 2003) with Matlab (The MathWorks Inc 2021).[9] The data on the prosodic properties of the compared IPs were then fit in linear mixed models using the *lme4* package (Bates et al. 2015) in RStudio (R Core Team 2021). The *afex* package (Singman et al. 2016) was then used to obtain p-values (the level of significance was set to $p < 0.05$). The models were run as follows: each of the measures described in Table 2 formed a dependent variable. The variable "turn length" (computed as the number of syllables of each *allora*-prefaced turn) was included as an independent variable (De Looze et al. 2014, 2015) to control for the difference in length of the compared speech chunks, and its possible effect on the pitch natural declination (Fuchs et al. 2016; Ladd 1988; Pierrehumbert 1979). The variable "conversational move" (Section 3.1.2) was also an independent variable. Because some moves (e.g. "acknowledge", "clarify" and "reply-y") displayed a very low frequency in our corpus (see section 3.2.1 below), we only looked at the (binary) difference between the Initiating and the Response moves in our analysis. Being more frequent, the Initiate group is taken as the reference level. The variable

[9]The Matlab frequency conversion from Hz to the ERB scale uses the following formula: $erb = A\log_{10}(1+hz(0.00437))$ where $A = 1000\log_e(10)/(24.7)(4.37)$

intonational phrase (i.e. the binary distinction between *allora* and the rest, coded as "IP") was also included as an independent variable. "Speaker" was treated as a random factor, to control for inter-speaker variability (De Looze et al. 2014). Finally, also "turn" was included as a random factor, to pair the observation on *allora* and the following IP belonging to the same dialogical turn. The resulting mixed models can be expressed with the following R formula (in which *DV* stands for one of the four dependent variables):

fm*x* <- lmer(DV ~ turn_length + conversational_move +IP + (1|turn) + (1|speaker)

However, to account for pitch reset (see above), we had to look at how the effects were different for the *allora*-IP and the sub-IP. Hence, we fitted mixed models with interaction terms (of "IP" on the one hand and "turn length" or "conversational move" on the other) for that analysis.

3.2 Results

3.2.1 Descriptive Statistics

In this section, we offer a descriptive analysis of the conversational moves realized by the *allora*-prefaced turns, and of the prosodic measures taken on the DM and the following IP. As mentioned in Section 3.1.2, the selected turns were analyzed by looking at the context (i.e. the preceding and following turns) to individuate what conversational action, or move, the speaker was realizing in the turn introduced by the DM *allora*. The results of the coding are summarized in Table 3:

	Move	Occurrences	Frequency
Initiating Moves	align	12	13%
	check	11	12%
	explain	13	14%
	instruct	49	52%
	query-w	2	2%

	query-yn	2	2%
Response Moves	acknowledge	4	4%
	clarify	1	1%
	reply-y	1	1%

Table 3: absolute occurrences and relative frequencies of the conversational moves coded in the corpus

As shown in the table, the “instruct” move is the most represented (52%). This is the move by which the speaker commands the interlocutor to take an action. In a game like the MT, it is really common for participants to give explicit instructions to the other one, as shown in the examples below[10]:

(1)	*p1G#1:allora <sp> cerca di fare una<aa> un cerchio intorno ai limoni*	so<sp>try to make a<aa>turn around the lemons

(2)	*p1G#67:<eh> io ho due macchine*	<eh> I have two cars
	p2F#68:<eh>	<eh>
	p1G#69:va bene?	is it all right?
	p2F#70:sì	yes
	p1G#71:allora <sp> passi sotto la macchina rossa	so<sp>you pass under the red car

In example (1) the game has just started, and the “instruction giver” takes the floor with *allora*, to then provide the first direction to the follower. In example (2), instead, the speaker opens the turn with *allora* to instruct the interlocutor after answering a question of clarification. Examples (3) and (4) below show further cases in which *allora* initiates turns realizing the “align” and “check” moves respectively:

(3)	*p1G#105:allora<sp>sei arrivata all'altezza del parafango anteriore?*	so<sp>you are arrived up to the anterior fender?

[10] Following the transcription rules of the CLIPS (Savy 2005), pauses (short: “sp”, or long: “lp”), interjections and vowel lengthening are reported in squared quotation marks (< >).

(4) *p2G#286:sì sì<lp>okay<lp>allora <sp> c'è il camion vero?*	yes yes<lp>okay<lp>so<sp>there is the truck right?

With the "align" move in (3), the speaker makes sure that the partner is on track and he/she agrees on the progression of the conversation. In (4), the speaker "checks" with the interlocutor about some uncertain beliefs. As the examples show, speakers preface their turns with *allora*, across the different types of conversational actions. The figures below provide some examples of the f0 contour typically produced by the speakers on the *allora*-prefaced turns:

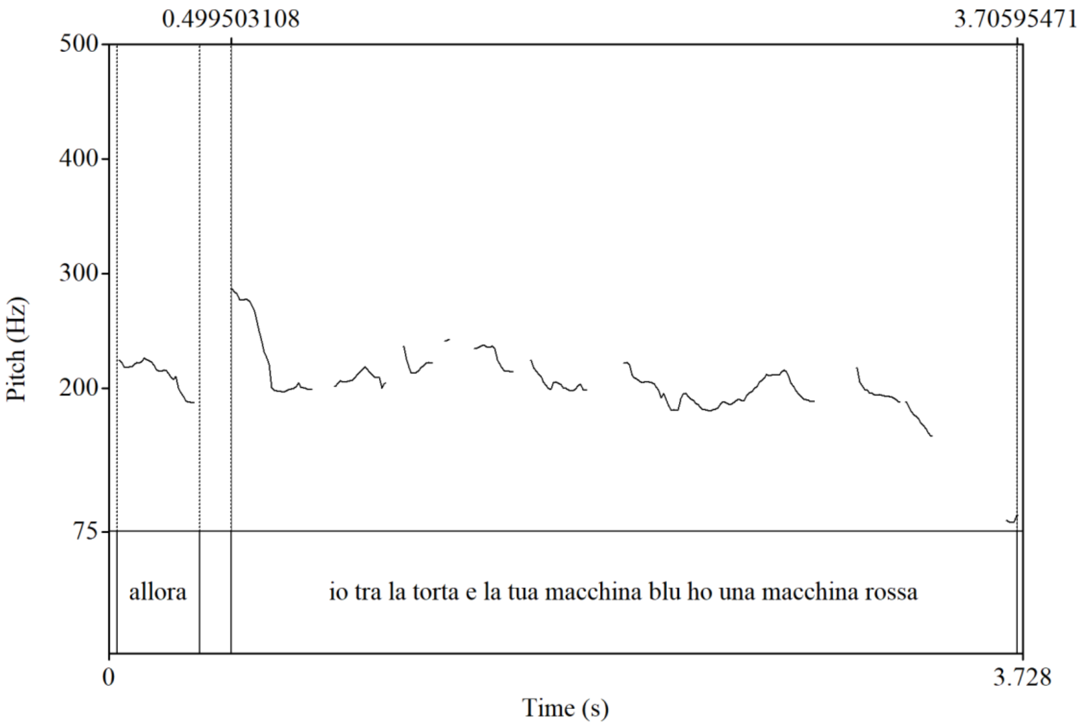

Figure 1: f0 contour of the *allora*-prefaced turn: *allora io tra la torta e la tua macchina blu ho una macchina rossa*, 'then between your cake and your blue car I have a red car'

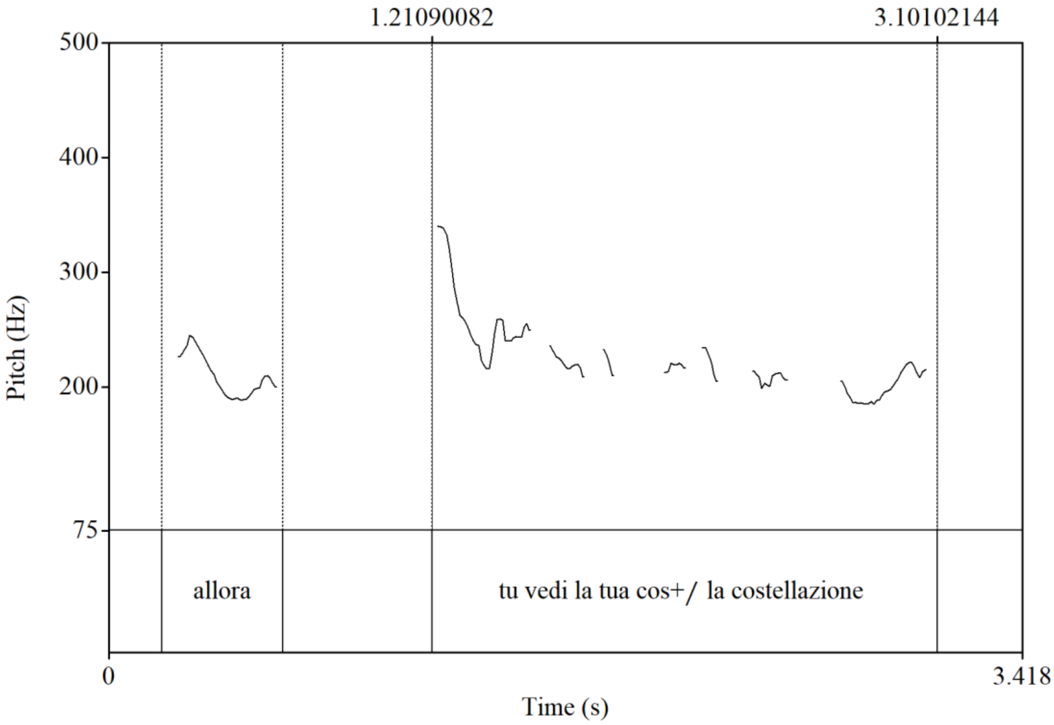

Figure 2: f0 contour of the *allora*-prefaced turn: *allora tu vedi la tua/ la costellazione*, ‘then you see your/the constellation’

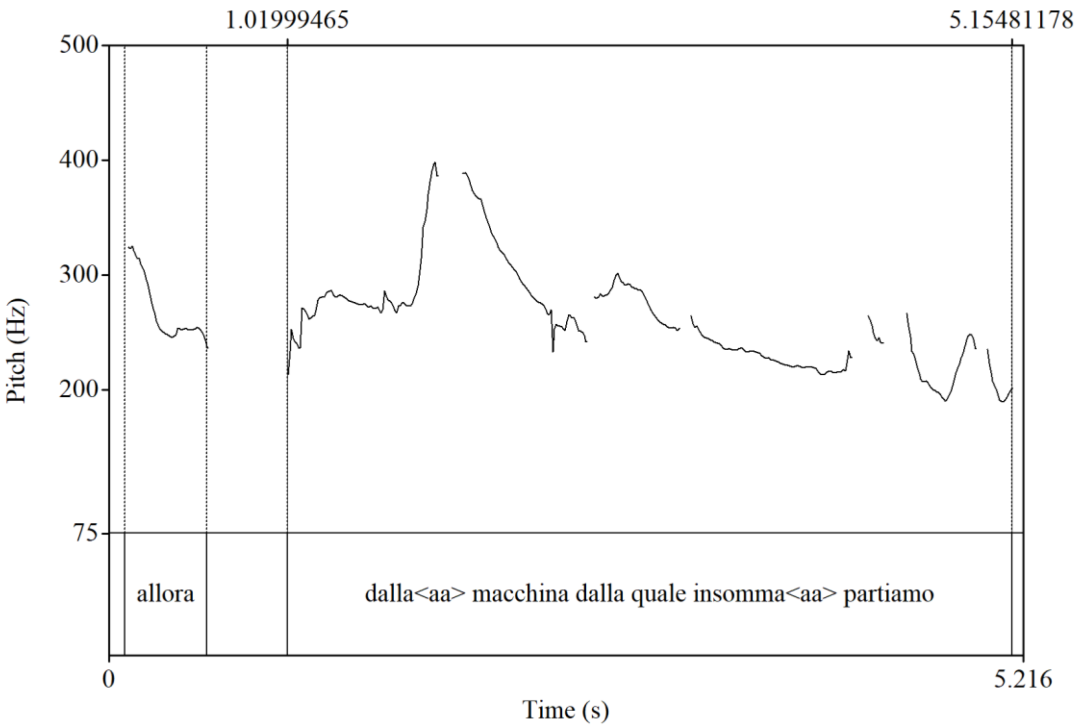

Figure 3: f0 contour of the *allora*-prefaced turn: *allora dalla macchina dalla quale insomma partiamo*, 'then from the car from which we hence start'

As shown in Figures 1-3, the DM presents a falling f0 profile. Its f0 contour appears at a lower global pitch level compared to the following sentence, and a reset of the pitch is visible after the silent pause.

The descriptive statistics of the prosodic analysis is given in Table 4. For each of the four variables, the mean is close to the median so the variables do not appear to be very skewed.

Name	Mean	SD	Median	Min	Max
Speech Rate	6,813	3,077	6,466	0	30,221
Pitch Span	2,4136	1,729	1,745	0,269	8,634
Pitch Height	6,520	1.970	6,411	3,260	11,682
Pitch First Syllable (pitch reset)	5,297	1.40	5,387	2,841	9,051

Table 4: descriptive statistics of the prosodic measures taken, for each of the *allora*-prefaced turns, on the DM and the following IP

In the next section, we offer the results of the linear mixed models fitted with each of the prosodic measures.

3.2.2 Prosodic properties: statistical analysis

3.2.2.1 Speech Rate

The results of the LMM with SR as the dependent variable are summarized in Table 5 below (the variance of the "turn" random effect is 1.332, that of the "speaker" random effect is 0.689 and the residual variance is 5.770):

Fixed Effect	Estimate	Standard Error	Degree of Freedom	t-value	p-value
Intercept(*allora*-IP)	7.973	0.467	108.228	17.085	<2e-16***
turn length	0.025	0.038	61.739	0.646	0.521
conv.move:reponse	-0.326	0.906	41.636	-0.360	0.721
Sub-IP	-2.691	0.355	90.421	-7.572	3.02e-11***

Table 5: statistics of the fixed factors fit in the LMM with SR as the dependent variable

As the table shows, there is no significant effect of "turn length" on SR, nor is there a difference between the Initiate and Response conversational moves. However, the "IP" variable has a significant effect on SP: that is, the SR is significantly different between the two IPs. The effect graph in Figure 1 shows the difference: the *allora*-IP is produced at a faster speech rate than the sub-IP (the non-overlapping confidence intervals of the two estimated speech rates are another indication of the significant difference).

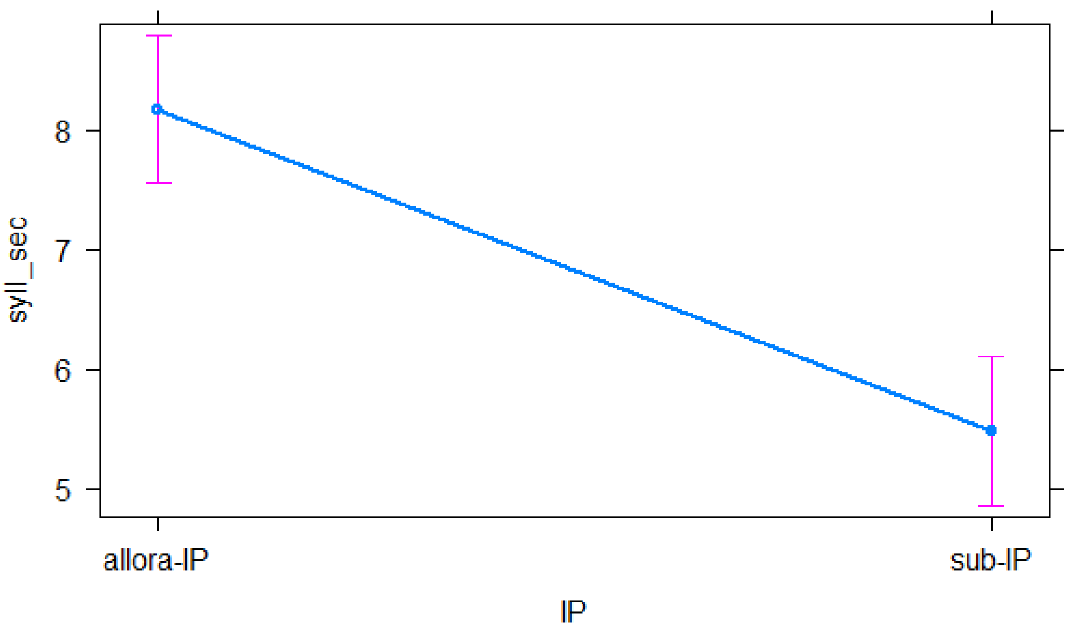

Figure 4: effect of the “IP” variable on SR

In order to have a more precise characterization of the articulation of the DM in our data, we compared the duration (in seconds) of the three syllables of *allora* (/al.'lo:.ra/). In Italian, the vowel bearing lexical stress is longer than unstressed vowels (Bertinetto & Loporcaro 2005; Krämer 2009; Bertinetto 1981; D’Imperio & Rosenthall 1999). It has also been shown that vowels preceding a prosodic boundary are normally lengthened in various languages including Italian (Petrone et al. 2014). Therefore, we expect the middle, stressed syllable of *allora* to be longer than the others and, since the DM in our corpus is produced as an independent IP (and it precedes a silent break) we could expect lengthening on the last syllable as well. All instances of *allora* were annotated by hand in Praat at the level of segment and syllables. A linear mixed model was fit with “duration” (in seconds) as the dependent variable, “syllable” as the predictor variable, “turn” and “speaker” as random factors. Table 6 below reports the results (the variances of the “turn” and the “speaker” random effects are round off 0, and the residual variance is 0.001)[11]:

Fixed Effect	Estimate	Standard Error	Degree of Freedom	t-value	p-value
Intercept (final syllable)	0.119	0.005	63.134	24.284	<2e-16***
first syllable	-0.004	0.005	188	-0.717	0.474
middle syllable	0.032	0.005	188	6.526	6.09e-10***

[11] For reasons of space we include in the presentation only the syllable-based model. The segment-based model gave comparable results.

Table 6: statistics of the fixed factors fit in the LMM with the duration of the syllables of *allora* as the dependent variable

As shown by the table, the middle syllable has a significantly longer duration than the final syllable (the reference level), which itself is significantly different from 0. However, there is no significant difference between the duration of the first and the final syllable. The relative durations of the three syllables are visualized in Figure 5:

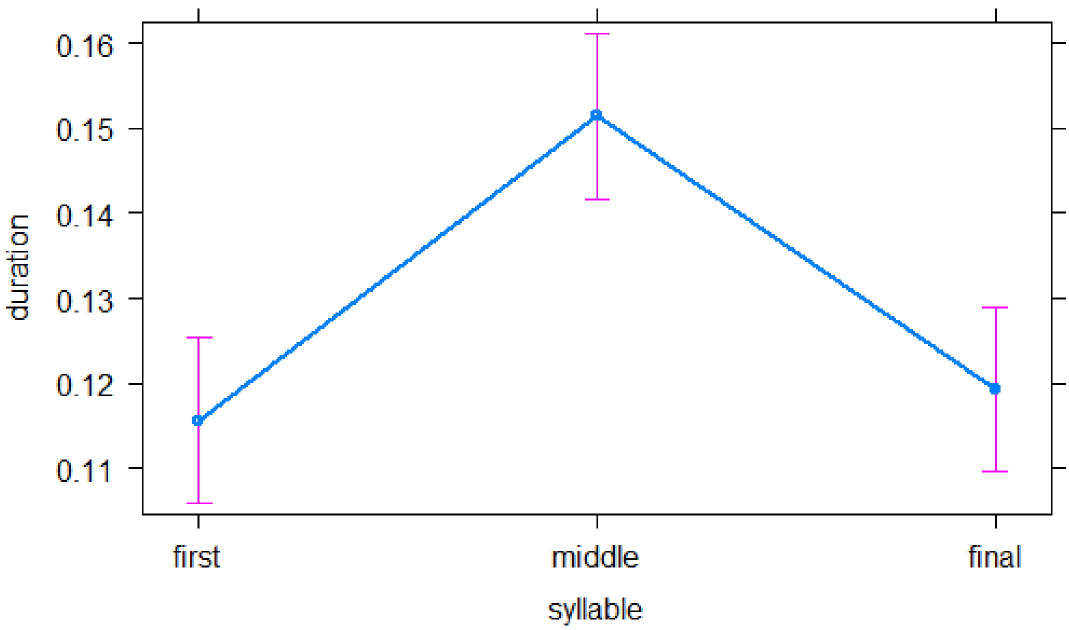

Figure 5: difference in “duration” for the “syllable” variable

Qualitative analysis indicate that strong segmental reductions are possible (ex.: [al.'lo:.ra] > [a.'lə.a], [a.'lo.ə], [a.'lə], ['lə]) yet sporadic (tot. 5 cases). The LMM indicates that the stressed/unstressed distinction remains stable in the DM, with the stressed syllable being significantly lengthened, as expected. However, no significant duration difference is found between the first and the last syllable, as mentioned, possibly due to the shortness of the examined IP. Overall, the results show that *allora* does not display systematic reduction phenomena at the segmental level and that the metrical structure of the word is preserved even at a fast speech rate.

3.2.2.2 Pitch Span

Fitting the model with PS as the dependent variable first resulted in a singular fit. The model was then refitted adding the arguments REML=FALSE and the control=lmerControl(calc.derivs=FALSE) arguments to the lmer() function to obtain interpretable results (Bates et al. 2015). These are reported in Table 7 below (variance of "turn" random effect is $1.033 * 10^9$, that of the "speaker" random effect is $5.129 * 10^2$ and the residual variance is 2.257):

Fixed Effect	Estimate	Standard Error	Degree of Freedom	t-value	p-value
Intercept (*allora*-IP)	1.690	0.239	136.510	7.064	7.52e-11***
turn length	-0.003	0.019	158.625	-0.155	0.877
conv.move:response	-0.041	0.435	45.328	-0.094	0.925
Sub-IP	1.474	0.222	126.382	6.634	8.60e-10***

Table 7: statistics of the fixed factors fit in the LMM with PS as the dependent variable

We see that the length of the turn or the difference between the Initiate and Response moves have no significant effect on PS, but there is a significant effect of "IP". The difference between the two IPs is visualized in the graph below:

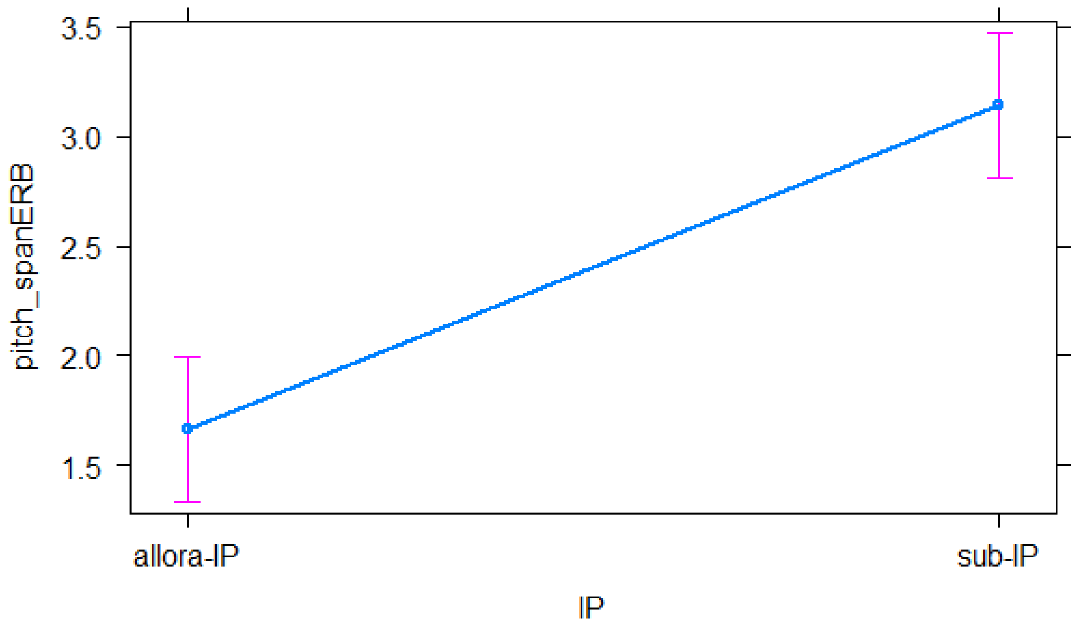

Figure 6: effect of the “IP” variable on PS

From Figure 6 it is clear that the *allora*-IP has significantly a lower pitch span than the sub-IP. In other words, the difference between the maximum and minimum f0 levels resulted in a significantly narrower range of values for the *allora*-IP compared to the sub-IP.

3.2.2.3 Pitch Height

Table 8 below reports the statistics for the fixed effects of the LMM with PH as a dependent variable (the variance of the “turn” random effect was 1.011 the variance of the “speaker” random effect was 0.524 and the residual variance was 1.660):

Fixed Effect	Estimate	Standard Error	Degree of Freedom	t-value	p-value
Intercept (*allora*-IP)	5.914	0.312	105.825	18.939	< 2e-16 ***
turn length	-0.026	0.026	79.405	-1.015	0.313
conv.move:reponse	0.346	0.631	83.789	0.549	0.585

Sub-IP	1.594	0.191	90.821	8.359	6.96e-13 ***

Table8: statistics of the fixed factors fit in the LMM with PH as the dependent variable

The table shows that, as much as for the previous measures, the results on PH are not significantly affected by the length of the turn, nor by the difference between the conversational moves (Initiate vs. Response). Figure 7 below shows that the *allora*-IP is generally produced with lower pitch values than the sub-IP:

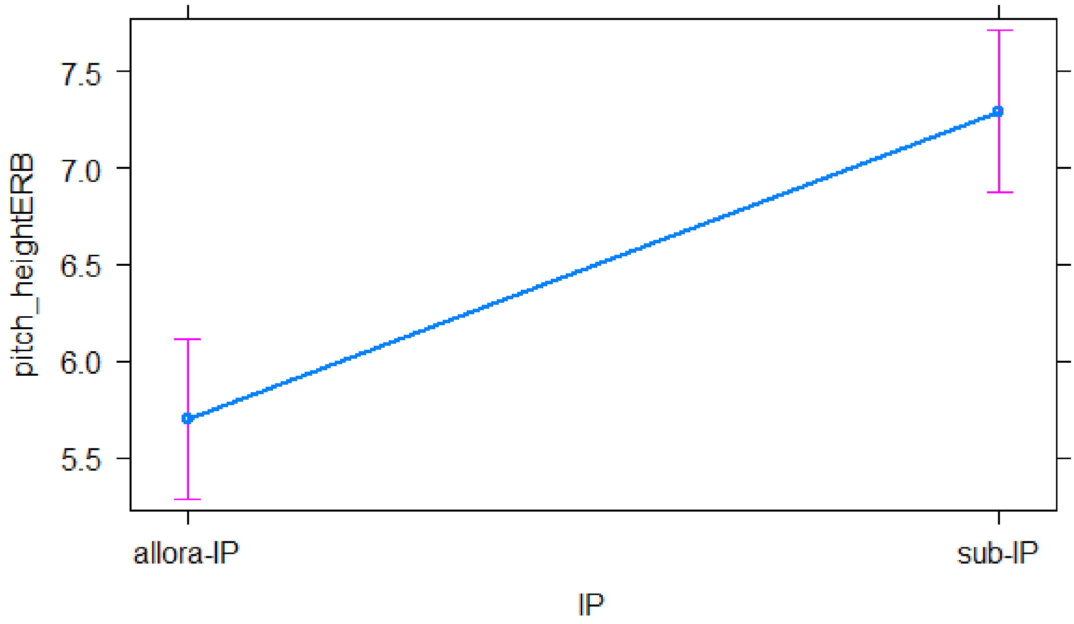

Figure 7: effect of the "IP" variable on PH

The maximum levels of f0 are significantly lower for the *allora*-IP than for the sub-IP. On average, the subtraction between the maximum pitch levels of sub-IP and the maximum pitch levels of *allora*-IP yields a positive difference of 6 semitones.[12]

[12]The semitone scale is used here (and below for the PFS measure) to express the difference between sounds, as discussed in Thomas (2011: 59). The following formula was used to convert the Hz measure into semitones: St=(log(Htz/100))/(log(2)/12) (Heylen et al. 2002: 6). The "psycho-acoustic scales", like

3.2.2.4 Pitch First Syllable (pitch reset)

The measures so far described the prosodic contours of the *allora*-IP compared to the sub-IP. To check for the prosodic relationship between the two IPs, we checked if a reset of the pitch was consistently present at the onset of the rest of the sentence following the DMs. The LMM was fit for PFS and gave the following results (the variance of the “turn” random effect was 1.185, the variance of the “speaker” random effect was 0.090 and the residual variance was 0.419):

Fixed Effect	Estimate	Standard Error	Degree of Freedom	t-value	p-value
Intercept (*allora*-IP)	4.366	0.242	95.166	18.019	<2e-16***
turn length	0.041	0.021	87.999	1.916	0.059
conv. move:reponse	0.226	0.491	76.971	0.461	0.646
Sub-IP	1.056	0.096	90.213	11.009	<2e-16***

Table 9: statistics of the fixed factors fit in the LMM with PR as the dependent variable

Table 9 indicates that the difference in f0 on the first stressed syllable of the *allora*-IP and the first stressed syllable of the sub-IP is significant. The stressed syllable of *allora* is

---

the semitones scale, “provide steps which correspond to equal perceptual intervals” (Nolan 2003: 771). If we assume that the threshold of a perceivable pitch difference is normally at 3 semitones, following Hart, Collier & Cohen (1990), this data would confirm the idea that DMs (and their prosodic features) are not a random by-effect of the processing of speech, but rather a signal given to the hearer for strategic purposes (Crible 2018). Crucially, our results show that this is true across the analyzed conversational moves. A dedicated perception study could provide further evidence on this point. For reasons of space, we leave the topic for future research.

produced at lower f0 values (4,5 semitones lower on average) than the first stressed syllable of the rest of the sentence (cf. Figure 8).

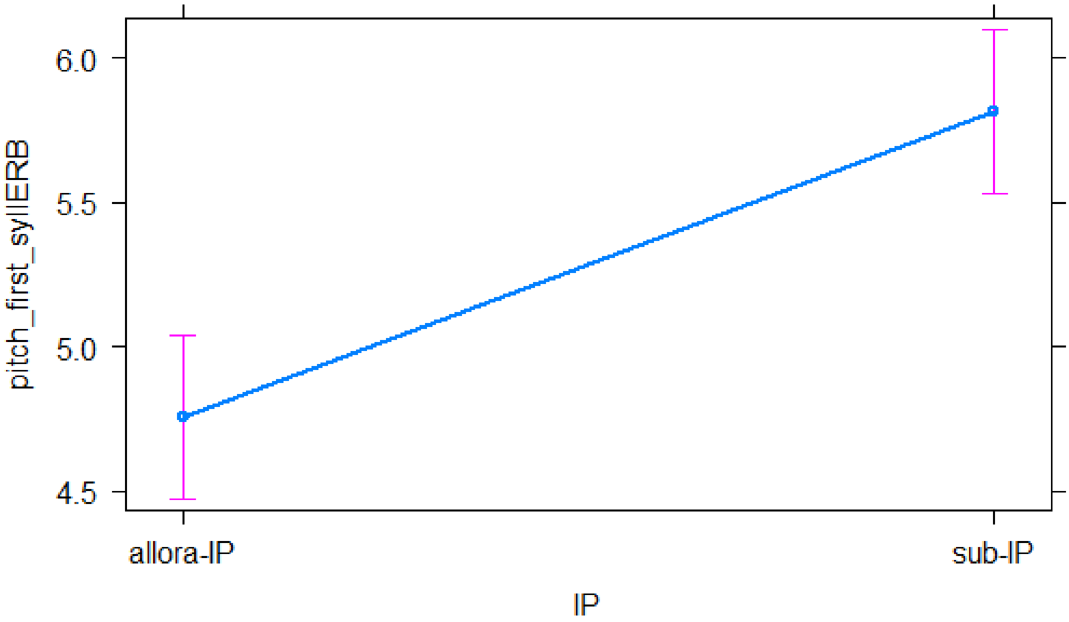

Figure 8: effect of the “IP” variable on PFS

The p-value for “turn length” in Table 8 above shows that the length of the turn in which the DM is found has an effect on PFS which can be deemed borderline significant (the p-value is very close to 0.05). Figure 9 below visualizes the effect:

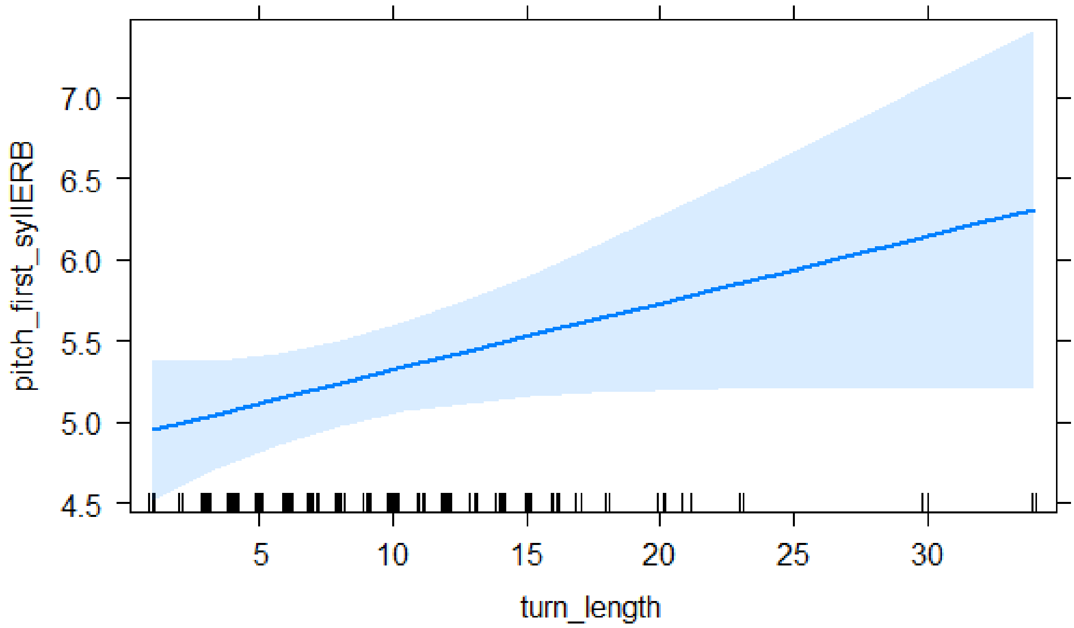

Figure 9: effect of the “turn length” variable on PFS

Since we are interested in the reset of the pitch between the *allora*-IP and the sub-IP, we finally also checked for the interaction between the “IP” on the one hand and the “turn length” and “conversational move” variables on the other, to inspect the difference between the analyzed IPs. The result of the model for the interactions are reported in Table 10 below (the variance of the “turn” random effect was 1.182, the variance of the “speaker” random effect was 0.090 and the residual variance was 0.425):

Fixed Effect	Estimate	Standard Error	Degree of Freedom	t-value	p-value
Intercept (*allora*-IP)	4.338	0.254	112.389	17.092	<2e-16
Sub-IP	1.114	0.181	88.431	6.165	2.05e-08
turn length	0.043	0.023	112.820	1.882	0.062
conv.move:reponse	0.359	0.523	97.846	0.687	0.494

IPrest:turn length	-0.004	0.016	88.221	-0.257	0.798
IPrest:conv.move response	-0.266	0.363	88.016	-0.733	0.465

Table 10: statistics of the fixed factors fit in the LMM with the interaction between “IP” on the one hand and “turn length” and “conversational move” on the other

As indicated in the table, none of the interactions is significant. We can therefore conclude that a reset of the pitch is present in our data (as evidenced by our previous analysis) but it does not depend on the length of the turn in which the IPs occur nor on the conversational move.

3.3 Summary of the results

As shown so far, the prosodic analysis confirmed the hypotheses made in Section 2. The turn-initial *allora* displays a higher speech rate and a narrower range of pitch values compared to the rest of the sentence. However, we do not observe systematic reduction at the segmental level. The DM is normally uttered at a lower level of pitch and an upward reset is visible at the onset of the following IP. The main result derivable from the previous sections is that all the measures display a significant difference between the two IPs. The variable “turn length” had a borderline significant effect on PFS, but crucially not on the reset of the pitch (Ladd 1988), whereas the variable “conversational move” never showed a significant effect. This indicates that the discursive status of the DM is overall not influenced by the conversational move that is being realized by the speaker. The results are further discussed in the next section.

4. Discussion

Our data confirmed the hypothesis made in the introduction: the fact that the turn-initial *allora* behaves as an appendage to their host sentence, and that this property is conveyed by prosodic means. The first prediction regarded the duration of the DMs compared to the rest of the turn. The *allora*-IP was consistently produced at a faster pace than the sub-IP, even though this did not imply a systematic segmental reduction of syllables elision. On the one hand, the high frequency of *allora* in Italian spontaneous speech (Bazzanella et al. 2007) would make us expect a stronger reduction (Frank-Job 2005). On the other hand, the DMs considered in our analysis were always in the initial position and followed by a silent pause: both these factors probably played a role in yielding a fully-realized DM (Bell et al. 2003; Schubotz, Oostdijk & Ernestus 2015). In any case, the faster SR that the *allora*-IP exhibits compared to the rest shows once again the “independent” behavior of the DM from the rest of the turn. In the same vein, Schubotz et al. (2015:376) propose that a higher reduction on DMs as opposed to their lexical counterparts could result from the looser relationship that the formers entertain with the surrounding prosodic and syntactic context (see also Local 2003: 326). If on one side our results are in line with these analyses, on the other side a systemic comparison of *allora* as a DM with *allora* used as an adverb could provide further insights on the prosodic (and syntactic) relationship of DMs with respect to the surrounding discourse. For reasons of space, we leave this point as an open issue for future research.

The second and third predictions concerned the flattened f0 profile of the marker and the presence of a pitch reset following it. As we have shown, the turn-initial *allora* introducing a conversational move displays a global falling contour. This confirms the observation by Wennerstrom (2001), cf. Section 2, according to whom DMs display L* pitch accent when they are external to the propositional content of the host sentence. Crucially though, the author refers to other possible pitch accents (H* or L+H*) in the cases in which a marker does participate in the construction of the information structure. Some variability in the f0 contour of DMs is also found in other works (see Leemann & Siebenhaar 2006; Raso & Vieira 2016). In fact, DMs do not *need* to be lower than the rest of the sentence. We propose (following Wennerstrom 2001), that the speaker can "decide" how integrated the marker is, for informational or discursive needs. More data on turn-integrated DMs would be necessary to test this hypothesis.

The constant presence of a pitch reset after the DM provides further insights about the discursive relationship between *allora* and the following IP. It has been proved that an upward shift of f0 –i.e., a reset of the pitch after its physiological declination (Ladd 1988)– can individuate the boundaries of discourse (Swerts 1997; Nakajima & Allen 1993), mark the alternation of turns (Fuchs et al. 2016; DeLooze et al. 2014), or signal the beginning of something new: a new topic (Wichmann 2000), a new speech paragraph (Wiklund 2014) or a new action (Couper-Kuhlen 2001, 2004). In our data, the upward shift of the pitch at the onset of the sub-IPs, (across the types of conversational move) would then indicate that the *allora*-IP and the sub-IP constitute two distinct discourse-planning units within the turn, in the sense of Krivokapic (2012) (see below).

These results are of particular interest for a better understanding of how turns are planned and organized -on the prosodic as well as discourse-pragmatic and syntactic levels- by the speakers in the interaction. As mentioned in section 2, the turn is traditionally taken as the unit of analysis by many scholars of the Phonetics of Talk-in-Interaction (PTI) approach, which investigates the role of prosody in interaction, within the CA tradition[13] (cf. Zellers and

[13] More precisely, the turn-constructional unit (TCU) is taken as the minimal unit of analysis in CA (see Zellers & Post 2012).

Post 2012). Couper-Kuhlen (2015: 85) writes that "viewed from the perspective of interaction, prosodic phenomena can be thought of as furnishing a format or design for turns at talk." We argue that this is precisely what the prosodic properties of the *allora*-IP (as opposed to the sub-IP) did in our data. In the spirit of Krivokapic (2007, 2012) we suggest that the two IPs analyzed in this work (*allora*-IP and sub-IP), being uttered at different levels within the speaker's range, constitute two distinct units in the turn's planning. In particular, we have shown that the DM (being lowered in the speaker's range) is peripheric to the overall prosodic planning of the turn. The isolated, low DM at the beginning of the turn hints to various levels of programming within the same turn and demonstrates that the initial portion of the turn already provides information to the hearer in this respect. In fact, the turn-initial position has received great attention within the PTI approach, as part of the investigation over the organization of the intonation units into larger declination units: "when there are several intonation units in a declination unit, they have slightly different shapes, depending on their relative position in the larger structure. The position of a single intonation unit within the larger unit is detectable in its final pitch, but also – importantly – in its initial pitch. It is the way intonation units begin that forms one of the new territories for exploration beyond the intonation phrase" (Couper-Kuhlen 2015: 87).

As we have shown, the DM *allora* is in fact outside the general declination trend.[14] Thus, our data corroborate the idea, advanced in previous works, that speakers in interaction exploit pitch range variations, producing higher or lower turn onsets, or placing their sentences in a lower or higher level of their range. Within the PTI framework, it has been shown that the speaker manipulates the level of his/her range to introduce a new (discourse) topic or to delimit

---

[14] As a matter of fact, the presence of turn-initial elements outside the high onset was already detected (though not directly addressed) in some of the above mentioned works. In Couper-Kuhlen (2001:38–39), for example, filled pauses (like "uh") precede the high onset, as much as our *allora*.

reported speech or parentheticals (Couper-Kuhlen 2015). This fundamental finding on pitch range variation emerges also within the other main approach to prosodic analysis mentioned in Section 2, namely the formal tradition of the Autosegmental-Metrical phonology (AM). Within AM framework, it has been shown indeed that sentence-level pitch manipulations can be used to mark the syntactic and –broadly speaking– the semantic relationship between the different IPs that compose the sentence (Féry & Truckenbrodt 2005; Author 2013; Author & Author 2020). Crucially, most of the works within the AM have focused their attention on the right side of the sentence, to investigate the syntactic recursion and the related prosodic properties at the end of the turn planning. The novelty of our contribution resides on the focus on the side in which the planning begins, i.e. the left side of the turn, which has been less explored in the AM tradition. As much as the syntactically recursive side has been shown to display lowering of the pitch (as a physiological consequence of declination; Ladd 1988), so the left side can host elements like DMs, being lower in the speaker's range and having an introductory function. From a syntactic point of view, this is in line with the tradition of studies mentioned in Section 2, ascribing elements like DMs to the Speech Act Layer in the Left Periphery: our data suggest that *allora* is indeed a unit outside the layer expressing the propositional content of the sentence. Despite being exterior to the main illocution, though, these units function as discursive and syntactic junctures: they are at the same time outside the core of the turn, and inside the speaker's turn planning. Crucially, this property is conveyed through prosodic means (see Brazil 1995, 1997 and Haegeman 2009 for similar insights).

To conclude, we suggest that a unitary explanation, putting together not only the formal and functional approaches to prosody but also the two margins of the turn's principal unit, is possible. We have shown that the phonetic parameters are used in conversation to mark the periphery and the "center" of a turn, and the relationship between the units that compose it.

While previous works within the AM approach have shown that speakers manipulate the pitch level on the recursive side to convey syntactic dependencies, our contribution demonstrates –in line with previous findings with the PTI tradition– that the variation of pitch levels are employed at the beginning of the turn, to convey discursive-pragmatic relations to the listener.

5. Conclusions

In this work we have analyzed the prosodic characteristic of an Italian DM, *allora*, occurring in turn-initial position in spontaneous speech as a means to take the floor and introduce a new conversational move. The analysis has shown that, when followed by a pause, *allora* is produced with a lower pitch accent and a faster speech rate than the rest of the sentence, and a pitch reset is present between the two. We conclude that the DM prosodic properties mark it as belonging to the discursive *and* syntactic periphery of the sentence. From a conversational perspective, this indicates that speakers can manipulate the levels of their pitch range to convey not only syntactic but also discursive and pragmatic relationships within their dialogical turns.

We believe that a highlight of this work is the combination of different theoretical perspectives. As we have shown, by analyzing the same linguistic object via various approaches, it is possible to get interesting insights regarding the different modules of grammar. In this way, not only one can provide a better answer to complex phenomena, but also obtain wide-ranging results that can inspire future work in different research fields. In fact, even if our data provided significant evidence of the peripheral status of *allora* as a DM, some issues remain open for further investigation. For example, it would be useful to examine cases in which no silent break is present between the DM and the rest of the sentence, and cases in which *allora* is used as a temporal adverb, and/or in other positions of the sentence. Lastly, our analysis could be reproduced with other DMs, and with more DMs occurring together in clusters. A broader analysis with different types of markers would help us understand how this sketchy class of elements interacts with prosody in the planning and production of natural speech.

References

Aijmer, Karin. 2002. *English discourse particles: evidence from a corpus*. Amsterdam: John Benjamins Publishing Company.

Anderson, Anne H., Miles Bader, Ellen Gurman Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, et al. 1991. The Hcrc Map Task Corpus. *Language and Speech* 34(4). 351–366. https://doi.org/10.1177/002383099103400404.

Andorno, Cecilia. 2007. Apprendere il lessico: elaborazione di segnali discorsivi (sì, no, così). In Marina Chini, Paola Desideri, M. Elena Favilla & Gabriele Pallotti (eds.), *Imparare una lingua: recenti sviluppi teorici e proposte applicative*, 95–122. Perugia: Guerra.

Author. 2013.

Author & Author. 2020.

Author & Author. 2019.

Bates, Douglas, Martin Machler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). 1–48. doi: 10.18637/jss.v067.i01.

Bayer, Josef, Roland Hinterhölzl & Andreas Trotzke. 2015. *Discourse-oriented syntax*. (Ed.) Josef Bayer, Roland Hinterhölzl & Andreas Trotzke. John Benjamins Publishing Company.

Bazzanella, Carla. 1995. I segnali discorsivi. In Lorenzo Renzi, Giampaolo Salvi & Anna Cardinaletti (eds.), *Grande Grammatica Italiana di Consultazione*, 1339–1357. Il Mulino.

Bazzanella, Carla. 2006. Discourse Markers in Italian: towards a “compositional” meaning. In Kerstin Fischer (ed.), *Approaches to Discourse Particles*, 504–524. Elsevier Ltd.

Bazzanella, Carla, Cristina Bosco, Alessandro Garcea, Barbara Gili Fivela, Johanna Miecznikowski & Francesca Tini Brunozzi. 2007. Italian allora, French alors: Functions, convergences and divergences. *Catalan Journal of Linguistics* 6. 9–30. https://doi.org/10.5565/rev/catjl.122.

Beckman, Mary E. & Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309. https://doi.org/10.1017/s095267570000066x.

Bell, Alan, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Michelle Gregory & Daniel Gildea. 2003. Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America* 113(2). 1001–1024. https://doi.org/10.1121/1.1534836.

Bennett, Ryan & Emily Elfner. 2019. The syntax-prosody interface. *Annual Review of Linguistics* 5. 151–171.

Bertinetto, Pier Marco. 1981. *Strutture prosodiche dell'italiano: accento, quantité, sillaba, giun-tura, fondamenti metrici*. Firenze: Accademia della Crusca.

Bertinetto, Pier Marco & Michele Loporcaro. 2005. The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association* 35(2). 132–151. https://doi.org/10.1017/S0025100305002148.

Bocci, Giuliano. 2013. *The Syntax Prosody Interface: A cartographic perspective with evidence from Italian*. John Benjamins Publishing Company.

Boersma, Paul & David Weenink. 2021. Praat: doing phonetics by computer [Computer Program]. 6.1.51. http://www.praat.org/.

Bolden, Galina B. 2009. Implementing incipient actions: The discourse marker “so” in English conversation. *Journal of Pragmatics* 41(5). 974–998. https://doi.org/10.1016/j.pragma.2008.10.004.

Brazil, David. 1995. *A grammar of speech*. Oxford University Press.

Brazil, David. 1997. *The communicative value of intonation in English book*. Cambridge University Press.

Brazil, David, Malcom Coulthard & Catherine Johns. 1980. *Discourse intonation and language teaching*. London: Longman.

Büring, Daniel. 2016. *Intonation and meaning*. Oxford University Press.

Cardinaletti, Anna. 2015. Italian verb-based discourse particles in a comparative perspective. *Discourse-oriented syntax* 71–91.

Carletta, Jean, Amy Isard, Stephen Isard, Jacqueline Kowtko, Alison Newlands, Gwyneth Doherty-Sneddon & Anne Anderson. 1996. *HCRC dialogue structure coding manual*. Human Communication Research Centre. https://www.lancaster.ac.uk/fass/projects/eagles/maptask.htm.

Cerrato, Loredana & Mariapaola D’Imperio. 2003. Duration and tonal characteristics of short expressions in Italian. In *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, 1213–1216.

Chow, Ivan. 2005. Prosodic cues for syntactically-motivated junctures. In *Ninth European Conference on Speech Communication and Technology*, 2373–2376.

Coniglio, Marco & Iulia Zegrean. 2012. Splitting up force. In Lobke Aelbrecht, Liliane Haegeman & Rachel Nye (eds.), *Main clause phenomena: New horizons*, vol. 190, 229–

256. John Benjamins Publishing Company.

Corver, Norbert. 2015. Interjections as structured root expressions. In Marc van Oostendorp & Henk van Riemsdijk (eds.), *Representing Structure in Phonology and Syntax*, vol. 124, 41–83. De Gruyter Mouton. https://doi.org/10.1515/9781501502224-003.

Couper-Kuhlen, Elizabeth. 2001. Interactional prosody: High onsets in reason-for-the-call turns. *Language in Society*. Cambridge University Press 30(1). 29–53. https://doi.org/10.1017/S0047404501001026.

Couper-Kuhlen, Elizabeth. 2004. Prosody and sequence organization in English conversation. In Elizabeth Couper-Kuhlen & Cecilia E. Ford (eds.), *Sound patterns in interaction*, 335–376. John Benjamins Publishing Company.

Couper-Kuhlen, Elizabeth. 2014. Prosody as dialogic interaction. In Dagmar Berth-Weingarten & Beatrice Szczepek Reed (eds.), *Prosodie und Phonetik in der Interaktion—Prosody and phonetics in interaction*, 221–251. Verlag für Gesprächsforschung.

Couper-Kuhlen, Elizabeth. 2015. Intonation and discourse analysis. *The Handbook of Discourse Analysis* 82–104.

D'Imperio, Mariapaola & Sam Rosenthall. 1999. Phonetics and phonology of main stress in Italian. *Phonology* 16(1). 1–28. https://doi.org/10.1017/S0952675799003681.

Degand, Liesbeth & Jacqueline Evers-Vermeul. 2015. Grammaticalization or pragmaticalization of discourse markers?: More than a terminological issue. *Journal of Historical Pragmatics* 16(1). 59–85. https://doi.org/10.1075/jhp.16.1.03deg.

Dehé, Nicole & Anne Wichmann. 2010. Sentence-initial I think (that) and I believe (that):

Prosodic evidence for use as main clause, comment clause and discourse marker. *Studies in language* 34(1). 36–74. https://doi.org/10.1075/sl.34.1.02deh.

De Looze, Céline, Irena Yanushevskaya, John Kane & Ailbhe Ní Chasaide. 2014. Pitch range declination and reset in turn-taking organisation. *Proceedings of the International Conference on Speech Prosody* 1100–1109. https://doi.org/10.21437/speechprosody.2014-209.

De Looze, Céline, Irena Yanushevskaya, Andy Murphy, Eoghan O'Connor & Christer Gobl. 2015. Pitch declination and reset as a function of utterance duration in conversational speech data. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* 3071–3075.

Didirková, Ivana, George Christodoulides & Anne Catherine Simon. 2018. The prosody of discourse markers alors and et in French: A speech production study. In *Proceedings of the International Conference on Speech Prosody*, vol. 2018-June, 503–507. https://doi.org/10.21437/SpeechProsody.2018-102.

Elordieta, Gorka. 2008. An Overview of Theories of the Syntax-Phonology Interface. *Journal of Basque Linguistics and Philology* 42(1). 209–286. http://www.ehu.es/ojs/index.php/ASJU/article/view/2309.

Elordieta, Gorka & Pilar Prieto. 2013. *Prosody and Meaning*. (Ed.) Gorka Elordieta & Pilar Prieto. De Gruyter Mouton.

Féry, Caroline & Hubert Truckenbrodt. 2005. Sisterhood and tonal scaling. *Studia Linguistica* 59(2–3). 223–243. https://doi.org/10.1111/j.1467-9582.2005.00127.x.

Frank-Job, Barbara. 2005. A dynamic-interactional approach to discourse markers. In Kerstin

Fischer (ed.), *Approaches to Discourse Particles*, 395–413. Elsevier Ltd.

Frascarelli, Mara. 1999. The prosody of Focus in Italian (and the syntax-phonology interface). *Probus* 11. 209–238.

Frascarelli, Mara. 2008. The fine structure of the Topic field. In Cécil De Cat & Katherine Demuth (eds.), *The Bantu-Romance Connection*, 261–292. John Benjamins Publishing Company.

Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 31(7). 931–952. https://doi.org/10.1016/s0378-2166(98)00101-5.

Fuchs, Susanne, Leibniz Zentrum, Allgemeine Sprachwissenschaft & Uwe Reichel. 2016. F0 declination and speech planning in face to face dialogues. In *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung*, 145–152. https://doi.org/10.13140/RG.2.1.4909.0320.

García García, Marta. 2021. Turn-Initial Discourse Markers in L2 Spanish Conversations: Insights from Conversation Analysis. *Corpus Pragmatics*. Springer International Publishing 5(1). 37–61. https://doi.org/10.1007/s41701-019-00075-8. https://doi.org/10.1007/s41701-019-00075-8.

Ghezzi, Chiara & Piera Molinelli. 2020. Connectives and cyclicity: From the Latin temporal phrase illa hora to the Italian discourse marker allora. *Journal of Historical Pragmatics* 21(2). 208–235. https://doi.org/10.1075/jhp.00042.ghe.

Haegeman, Liliane. 2009. Parenthetical adverbials: the radical orphanage approach. In Benjamin Shaer, Philippa Cook, Frey Werner & Cluadia Maineborn (eds.), *Dislocated elements in discourse : syntactic, semantic and pragmatic perspectives*, 331–347.

Haegeman, Liliane & Virginia Hill. 2014. The Syntacticization of Discourse. *Syntax and its Limits* (1). 370–390. https://doi.org/10.1093/acprof:oso/9780199683239.003.0018.

Heritage, John. 2013. Turn-initial position and some of its occupants. *Journal of Pragmatics*. Elsevier B.V. 57. 331–337. https://doi.org/10.1016/j.pragma.2013.08.025.

Heylen, L., Floris L. Wuyts, F. Mertens, M. De Bodt & P. H. Van de Heyning. 2002. Normative voice range profiles of male and female professional voice users. *Journal of Voice* 16(1). 1–7. https://doi.org/10.1016/S0892-1997(02)00065-6.

Hill, Virginia. 2007. Vocatives and the pragmatics-syntax interface. *Lingua* 117(12). 2077–2105. https://doi.org/10.1016/j.lingua.2007.01.002.

Hirschberg, Julia & Diane J. Litman. 1993. Empirical Studies on the Disambiguation of Cue Phrases. *Computational Linguistics* 19(3). 501–530. https://doi.org/10.7916/D8PV6TGB.

Krämer, Martin. 2009. *The phonology of Italian*. Oxford University Press.

Krivokapic, Jelena. 2007. Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of phonetics* 35(2). 162–179.

Krivokapic, Jelena. 2012. Prosodic planning in speech production. *Speech Planning and Dynamics* 157–190.

Ladd, D. Robert. 1988. Declination “reset” and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84(2). 530–544. https://doi.org/10.1121/1.396830.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Leemann, Adrian & Beat Siebenhaar. 2006. Prosodic features of spontaneous utterance-initial phrases in Bernese and Valais Swiss-German. In *Symposium A Quarterly Journal In Modern Foreign Literatures*, 127–142. Tapei, Taiwan.

Lehmann, Christian. 1985. Grammaticalization: Synchronic Variation and Diachronic Change. *Lingua e Stile* 20. 303–318.

Local, John. 2003. Variable domains and variable relevance: Interpreting phonetic exponents. *Journal of Phonetics* 31(3–4). 321–339. https://doi.org/10.1016/S0095-4470(03)00045-7.

Local, John & Gareth Walker. 2004. Abrupt-joins as a resource for the production of multi-unit, multi-action turns. *Journal of Pragmatics* 36(8). 1375–1403. https://doi.org/10.1016/j.pragma.2004.04.006.

Maschler, Yael. 2009. *Metalanguage in interaction: Hebrew discourse markers*. Amsterdam: John Benjamins Publishing Company.

Miecznikowski, Johanna, Barbara Gili Fivela & Carla Bazzanella. 2008. Words in Context: Agreeing and Disagreeing with allora. (Ed.) Giovanni Gobber, Sibilla Cantarini, Sara Cigada, Maria Cristina Gatti & Sara Gilardoni. *Proceedings of the IADA Workshop Word Meaning in Argumentative Dialogue* 1(16). 205–218.

Molinelli, Piera. 2018. Different Sensitivity to Variation and Change: Italian Pragmatic Markers dai vs. Discourse Marker allora. In Salvador Pons Bordería & Oscar Lourdena Lamas (eds.), *Beyond Grammaticalization and Discourse Markers*, 271–303. Leiden: Brill.

Nakajima, Shin & James F Allen. 1993. A study on prosody and discourse structure in cooperative dialogues. *Phonetica* 50(3). 197–210.

Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Nolan, Francis. 2003. Intonational equivalence: an experimental evaluation of pitch scales. In *Proceedings of the 15th international congress of phonetic sciences*, 2–5.

Ogden, Richard. 2006. Phonetics and social action in agreements and disagreements. *Journal of Pragmatics* 38(10). 1752–1775. https://doi.org/10.1016/j.pragma.2005.04.011.

Orrico, Riccardo & Mariapaola D'Imperio. 2020. Tonal specification of speaker commitment in Salerno Italian wh-questions. *Speech Prosody* 361–365. https://doi.org/10.21437/SPEECHPROSODY.2020-74.

Petrone, Caterina, Mariapaola D'Imperio, Leonardo Lancia & Susanne Fuchs. 2014. The interplay between prosodic phrasing and accentual prominence on articulatory lengthening in Italian. In *Proceedings of the International Conference on Speech Prosody*, 192–196. https://doi.org/10.21437/speechprosody.2014-26.

Pierrehumbert, Janet. 1979. The perception of fundamental frequency declination. *Journal of the Acoustical Society of America* 66(2). 363–369. https://doi.org/10.1121/1.383670.

Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*. MIT press.

Prieto, Pilar & Paolo Roseano. 2021. The encoding of epistemic operations in two Romance languages: The interplay between intonation and discourse markers. *Journal of Pragmatics*. Elsevier Ltd 172. 146–163. https://doi.org/10.1016/j.pragma.2020.11.008.

QSR International. 2018. Nvivo qualitative data analysis software [Computer Program]. (released in March 2018). https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home.

R Core Team. 2021. R Studio [Computer Program]. PBC, Boston, MA: RStudio: Integrated Development for R. RStudio. http://www.rstudio.com/.

Raso, Tommaso & Marcelo Vieira. 2016. A description of Dialogic Units/Discourse Markers in spontaneous speech corpora based on phonetic parameters. *CHIMERA: Romance Corpora and Linguistic Studies* 3(2). 221–249.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. *Elements of grammar*. Springer, Dordrecht 281–337. https://doi.org/10.1007/978-94-011-5420-8_7.

Rizzi, Luigi. 2004. On the cartography of syntactic structures. In Luigi Rizzi (ed.), *The Structure of CP and IP*, vol. 2, 3–15. Oxford University Press.

Roberts, Ian & Anna Roussou. 1999. A formal approach to "grammaticalization." *Linguistics* 37(6). 1011–1041. https://doi.org/10.1515/ling.37.6.1011.

Savino, Michelina. 2011. The intonation of backchannel tokens in Italian collaborative dialogues. In *Language and Technology Conference*, 28–39. Cham: Springer. https://doi.org/10.1007/978-3-319-08958-4_3.

Savy, Renata. 2005. Specifiche per la trascrizione ortografica annotata dei testi. In Federico Albano Leoni & Rosa Giordano (eds.), *Italianoparlato. Analisi di un dialogo*. Naples: Liguori Editore. http://www.clips.unina.it/it/documenti/11_specifiche_trascrizione_ortografica.pdf.

Schegloff, Emanuel A. 1996. Turn organization: one intersection of grammar and interaction. *Interaction and Grammar* 13(52). https://doi.org/10.1017/cbo9780511620874.002.

Schegloff, Emanuel A. & Gene H. Lerner. 2009. Beginning to respond: Well-prefaced responses to wh-questions. *Research on Language and Social Interaction* 42(2). 91–

115. https://doi.org/10.1080/08351810902864511.

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Schubotz, Louise, Nelleke Oostdijk & Mirjam Ernestus. 2015. Y ' know vs you know: What phonetic reduction can tell us about pragmatic function. In S Lestrade, P.J.F. Swart & L. Hogeweg (eds.), *Addenda: artikelen voor Ad Foolen*, 361–380. Radbpud University.

Selkirk, Elisabeth. 1984. *Phonology and syntax: the relationship between sound and structure*. Cambridge, MA: MIT press.

Singman, Henrik, Ben Bolker, Jake Westfall & Frederik Aust. 2016. Afex: Analysis of Factorial Experiments. R package version 0.16-1. https://cran.r-project.org/package=afex.

Sluijter, Agaath MC & Jacques MB Terken. 1993. Beyond sentence prosody: Paragraph intonation in Dutch. *Phonetica* 50(3). 180–188.

Sobrero, Alberto & Immacolata Tempesta. 2006. Definizione delle caratteristiche generali del corpus: informatori, località. *Relazione Tecnica CLIPS/W0-a1/CGP/001*. http://www.clips.unina.it/it/documenti/1_scelta_informatori_e_localita.pdf.

Speas, Peggy & Carol L. Tenny. 2003. Configurational properties of point of view roles. *Asymmetry in Grammar1* 315–344. https://doi.org/10.1075/la.57.15spe.

Swerts, Marc. 1997. Prosodic features at discourse boundaries of different strength. *The Journal of the Acoustical Society of America* 101(1). 514–521. https://doi.org/10.1121/1.418114.

The MathWorks Inc. 2021. Matlab [Computer Program]. Natick, Massachusetts: version 9.11.0.1837725 (R2021b). https://nl.mathworks.com/products/matlab.html.

Thomas, Erik. 2011. *Sociophonetics: an introduction*. Basingstoke: Palgrave Macmillan.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30(2). 219–255.

Walker, Gareth. 2012. Phonetics and Prosody in Conversation. In J. Sidnell & T. Stivers (eds.), *The Handbook of Conversation Analysis. Blackwell Handbooks in Linguistics*, 455–474. Chichster: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118325001.ch22.

Watson, Duane & Edward Gibson. 2004. The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* 19(6). 713–755. https://doi.org/10.1080/01690960444000070.

Wennerstrom, Ann. 2001. *The music of everyday speech: Prosody and discourse analysis*. Oxford University Press.

Wichmann, Anne. 2007. Sound patterns in interaction: Cross-linguistic studies from conversation. *Language in Society* 36(2). 259–263. https://doi.org/10.1017/s0047404507210139.

Wiklund, Mari. 2014. The realization of pitch reset in Finnish print interpreting data. *Text and Talk* 34(4). 491–520. https://doi.org/10.1515/text-2014-0013.

Zellers, Margaret & Brechtje Post. 2012. Combining Formal and Functional Approaches to Topic Structure. *Language and Speech* 55(1). 119–139. https://doi.org/10.1177/0023830911428872.

Zimmermann, Malte. 2008. Discourse particles in the left periphery. In Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (eds.), *Dislocated Elements in*

Discourse: Syntactic, Semantic, and Pragmatic Perspectives, 200–231. London: Routledge. https://doi.org/10.4324/9780203929247.

Zimmermann, Malte. 2011. Discourse particles. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics - Sentence and Information Structure*, 2011–2038. https://doi.org/10.1515/9783110589863-015.

Elisa De Cristofaro – first and corresponding author

Is PhD student of Linguistics in the Department of Translation, Interpreting and Communication at Ghent University. She completed her MA in Linguistics and Cognitive Studies at the University of Siena (2018), and her bachelor in Italian Literature at the University of Rome La Sapienza (2015).

Her research interests are theoretical linguistics (particularly syntax, the syntax-prosody interface and the syntax-discourse interface) and second language acquisition, with a special focus on the Italian language. At Ghent university, she is currently working on a project on discourse markers in Italian as a second and heritage language.

Claudia Crocco – co-author

is Associate professor of Italian linguistics at Ghent University. Claudia completed her Ph.D. in General linguistics (2004) and her MA in Classical languages (1999) at the University of Naples Federico II. Before joining Ghent University, she was a post-doctoral researcher at the Universities of Antwerp (2004) and Salerno (2005), and a visiting professor of phonetics at the University of Macerata (2004).

Her research interests range from grammar to sociolinguistics. Current research topics include intonational phonetics and phonology of Italian, syntax-prosody interface, and the (re)standardization of Italian.

Linda Badan – co-author

is Associate Professor of Linguistics at Ghent University. She obtained her PhD at University of Padova and different research fellowships as the Chiang Ching-kuo Foundation for International Scholarly Exchange (Taiwan) and the International Fellowships for Experienced Researchers Fernand Braudel-IFER (École Normale Supérieure, Paris). She has been post-doc fellow at City University of Hong Kong, Leiden University, École des Hautes Études in Science Sociales (Paris).

Her main research interests include theoretical linguistics at the interface between syntax and prosody and syntax and pragmatics, comparative syntax between Romance and Sinitic languages and the acquisition of Italian as second language.

Koen Plevoets – co-author

Is Data Scientist and Tutor in statistical methodology in the Department of Translation, Interpreting and Communication, and in the Department of Applied mathematics, computer science and statistics at Ghent University. He is among the coordinators of the inter-university training network FLAMES. He has obtained his doctorate in Germanic Languages at KU Leuven (2008). He has completed a master in in Artificial Intelligence (2002) and a master Statistics (2003), both at KU Leuven.

His research interests are the visualization of complex data and the analysis of categorical data. He makes use of the open-source programming languages Python and R.