DOI: 10.17469/O2104AISV000009

MARIA CRISTINA PINELLI, CINZIA AVESANI, CECILIA POLETTO

Is it prosody that settles the syntactic issue? An analysis of italian cleft sentences[1]

The present study aims at investigating the prosodic realization of Italian cleft sentences, in order to provide some new cues for their still debated syntactic interpretation. A monoclausal approach to the analysis of cleft sentences (a.o. Meinunger, 1998) parallels them to left focalization constructions, while a biclausal approach (a.o. Belletti, 2008) considers them composed of a main copular clause and an embedded pseudo-relative clause. A systematic comparison between cleft sentences and left focalizations – carried out through an experimental study and an analysis of pitch accent distribution, scaling, and prosodic phrasing – leads to conclude that their prosodic realization is very similar, thus suggesting that a monoclausal analysis for cleft sentences is supported by prosodic data.

Keywords: Cleft sentences, Left focalization, Prosody-syntax interface.

1. *A syntactic puzzle*

Cleft sentences have been subject of discussion among syntacticians since Jespersen's (1927) work on English, and no unitary theory has yet been developed. One of the reasons for this is that several different languages display different kinds of cleft structures, which often have slightly different syntactic/semantic/discourse properties (Hartmann, Veenstra, 2013).

However, most scholars agree on the assumption that cleft sentences at least can be used with a highlighting function: given the canonical clause in (1), the corresponding cleft clause in (1b) marks focus on the clefted constituent and treats the relative clause as presupposed backgrounded material.

(1) a. Ho visto Andrea in cucina
See-PST.1SG Andrea in the kitchen
"I saw Andrea in the kitchen"

b. È Andrea che ho visto in cucina
be-PRS.3SG Andrea that see-PST.1SG in the kitchen
"It is Andrea that I saw in the kitchen"

In line with this analysis of clefts' information structure, the cleft sentence in (1b) can easily be compared with another syntactic structure that operates the same fo-

[1] Authorship note: this study has been jointly designed by the three authors. Main responsibility for this paper is divided as follows: §1: Pinelli, Poletto; §2: Avesani; §3: Pinelli; §4: Pinelli, Avesani; §5: Avesani.

cus-background partition through a similar word order change: left focalization. Example (2) shows that the focalized constituent (in capitals) of left focalization structures has the same informational status and nearly the same position of the clefted constituent in (1b):

(2) ANDREA ho visto in cucina
Andrea see-PST.1SG in the kitchen
"ANDREA I saw in the kitchen"

These considerations have led scholars to hypothesize that cleft sentences and left focalizations can also have similarities at a deeper level in the syntactic structure.

The most widespread theories on cleft sentences consider them biclausal structures, composed of a main copular clause and an embedded (pseudo-)relative clause (Declerck, 1988; Den Dikken, 2013), as showed in (3a). However, some other authors like Meinunger (1998) and Frascarelli, Ramaglia (2013) proposed a monoclausal analysis of cleft sentences based on the similarities with left focalizations mentioned above (3b).

(3) a. [È Andrea]$_{\text{COPULAR}}$ [che ho visto in cucina]$_{\text{(PSEUDO-)RELATIVE}}$
b. [È Andrea che ho visto in cucina]
"It is Andrea that I saw in the chichen"

In Frascarelli & Ramaglia's view, the free relative clause (*che ho visto in cucina)* is directly merged in the left periphery of the main clause, in a TopP projection (Familiarity Topic), while the clefted constituent moves from its base position to a FocP projection in the same left periphery. This analysis accounts for the similarities with focalizations, since the clefted constituent undergoes the same A'-movement of the focalized constituent and ends up in the same FocP projection in the left periphery, giving as a result a monoclausal structure (4).

(4) [$_{GP}$[$_{IP}$ t'$_{DP}$ è [$_{SC}$ t$_{DP}$ t$_{NP}$]] [$_{\textbf{FocP}}$ [$_{NP}$ Andrea] [$_{FamP}$ [$_{DP}$[$_{SC}$ [$_{NP}$*pro*] [$_{CP}$ che *pro* ho visto]]] t$_{IP}$]]][2]

A third proposal for the analysis of cleft sentences was made by Belletti (2008). The author claims that in Italian there are two different types of cleft sentences, corrective/contrastive clefts and new information clefts, which have different syntactic structures[3]. Corrective/contrastive clefts are used when the clefted constituent is

[2] GP = Ground Phrase; IP = Inflectional Phrase; DP = Determiner Phrase; SC = Small Clause; NP = Nominal Phrase; FocP = Focus Phrase; FamP = Familiarity Topic Phrase; CP = Complementizer Phrase.

[3] Actually, other non-corrective/contrastive clefts can be found in the literature on Italian clefts: they are the so-called non-prototypical temporal cleft sentences – an example of which is reported in (i) – first analysed by Benincà (1978).

i) Sono due ore che ti aspetto
be-PRS-3PL two hours that you-ACC wait-PRS-1SG
"it is two hours that I'm waiting for you"

Whether they can be considered fully-fledged new information clefts is not clear yet. We aim to further investigate this topic from a syntax-prosody-discourse interface perspective.

explicitly opposed to an antecedent in the preceding context or in the common ground: only in this type of sentences the clefted constituent can be a NP, a PP, an AdvP or an AdjP, and if it is an NP it can function both as a subject or as an object[4]. The syntactic structure of corrective/contrastive clefts is said to be still biclausal – with a main copular clause and an embedded pseudo-relative clause – but slightly different from the one in (3a). In fact, the clefted constituent moves to the left periphery of the lexical verb – that is, of the relative clause – and does not reach the copular clause, as the schematic representation in (5) shows.

(5) a. $[\text{È}]_{\text{COPULAR}}$ [Andrea che ho visto in cucina]$_{\text{(PSEUDO-)RELATIVE}}$

b. $[_{\text{CP}} \ldots [_{\text{TP}} \ldots [_{\text{vP}}$ *be* $[_{\text{CP/}\mathbf{FocPcorr}}$ Andrea$_1$ $[_{\text{FinP}}$ che $[_{\text{TP}}$ ho visto t$_1$ ...

The three structures in (3a), (4) and (5) represent the starting point for this prosodic study. It is important to note that in all three structures the clefted constituent ends up to be hosted in a FocP projection, i.e. it is treated as a focus from a syntactic and information-structural point of view.

The similarity between corrective focalizations and clefts, which has been detected by the supporters of the monoclausal analysis, intuitively seems to hold for their prosodic realization as well, but a systematic comparison of these two structures has not been carried out yet: investigating the prosodic phrasing of clefts and left focalizations, as well as their pitch accent selection, can add new elements for the comprehension of their underlying syntactic structures from a prosody-syntax interface perspective.

2. *The prosody of corrective/contrastive Focus in Italian*

In the literature on syntax-prosody interface it has been proposed that when a focus phrase is moved to the left periphery of the sentence – as it is the case for corrective/contrastive Focus – it is followed by an intonational phrase (*IP*) boundary (Nespor, Guasti, 2002): it separates the phrase with the main intonational prominence from the postfocal material, which is argued to be extrametrical (i.e. not included in the "core IP" as proposed by Szendröi, 2001) or to be right-dislocated (Samek-Lodovici, 2006). Frascarelli (2000) proposes instead that an *IP* boundary follows a contrastive focus phrase only if it is not adjacent to the verb, while an intermediate/phonological phrase occurs in the other cases.

Contrary to the claims of the first authors and more in line with Frascarelli, in two production experiments Bocci, Avesani (2005) and Bocci (2013) show that an initial focus is followed by the right boundary of an intermediate/phonological phrase (*ip*). The evidence resides in the acoustic lengthening of the segments preceding the *ip* right boundary: the final syllable and vowel of a fronted focus phrase were not significantly longer than the corresponding elements at the end

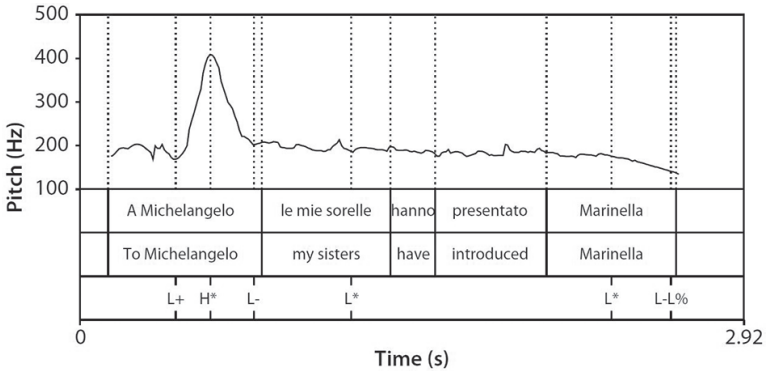
[4] In new information clefts the clefted constituent can only be a subject NP, which means that in Italian new information clefts can only be subject clefts (Belletti, 2008).

of a preverbal subject in broad focus (Bocci, Avesani, 2005) and were shorter than the corresponding syllable and vowel at the end of contrastive and partial topics (Bocci, 2013). They conclude that an utterance divided in left peripheral focus and background, as the one presented in Fig. 1, is metrically phrased as in (6), in which the focus phrase is mapped into an *ip* which occurs in the same *IP* along with the rest of the clause.

The focus phrase attracts the main prominence of the *IP*, leading to an Intonational Phrase whose head is *not* assigned to its rightmost element. The backgrounded material in postfocal position is intonationally realized as a low and flat pitch contour but, despite the apparent absence of intonational cues, it cannot be considered as extrametrical (i.e., dephrased and deaccented). Bocci, Avesani (2011), Bocci (2013) and Bocci, Avesani (2015) show that the postfocal material is prosodically "visible": it is phrased, and L*[5] pitch accents are associated with the metrical head of any postfocal phrase.

(6) $[[[\text{A MICHELANGELO}]_{ip}\,[\text{le mie sorelle hanno presentato Marinella}]_{ip}\,]_{IP}]_{U}$
"To MICHELANGELO my sisters presented Marinella"

Figure 1- *From Bocci (2013:145). Left focalization: "A MICHELANGELO le mie sorelle hanno presentato Marinella" (To Michelangelo my sisters introduced Marinella). U = utterance, IP = Intonational Phrase, ip = intermediate Phrase*



3. *Hypotheses and predictions*

The syntax-prosody mapping rules as formulated by Selkirk (e.g., Selkirk 2005), require that a syntactic clause is mapped into a prosodic constituent of *IP* level. According to this view, if corrective/contrastive cleft sentences are biclausal, we expect that each clause is phrased in a separate Intonational Phrase, i.e. that an *IP* boundary occurs between the main copular clause and the pseudo-relative clause.

[5] In Bocci analysis, L* pitch accents are non fully-fledged accents: they are inserted in the metrical structure only in fulfilment of phonological requirements, and have the function to mark the right side of the focus phrase.

Specifically, in Declerck (1988) and Den Dikken (2013) syntactic analysis (3a) we expect that an *IP* boundary occurs *after* the focused element, while in Belletti (2008) analysis (5) we expect that the *IP* boundary occurs after the copula and *before* the focused element. Conversely, if the corrective/contrastive clefts are monoclausal (Meinunger, 1998 and Frascarelli, Ramaglia, 2013), we expect cleft sentences to be prosodically phrased in one *IP* and, assuming Bocci's analysis for sentences with a left peripheral contrastive/corrective focus, that a prosodic boundary of *ip* level occurs after the focused element. No boundary is expected to occur between the copula and the focused element, as they are both wrapped in the same Intermediate/Phonological Phrase.

Accordingly, if corrective/contrastive clefts are biclausal as in (3a), we expect that the prosodic boundary after the focused element in the copular phrase will be stronger than in left-focalized sentences. Specifically, that the boundary will be cued by a pre-boundary lengthening longer than in focused sentences and by the presence of a pause. If corrective/contrastive clefts are biclausal and have the syntactic structure proposed in (5), we expect the copula "è" to be endowed with a nuclear pitch accent and to be followed by a pause, differently from left-focalized sentences where we expect the copula to have at most a prenuclear accent and no following pause.

Moreover, we expect different prosodic phrasings for biclausal clefts and left-focalized sentences independently of their structural position, i.e. whether or not they are embedded in a superordinate main sentence.

In order to test these predictions, we run a production experiment on minimal pairs of cleft sentences and left focalizations in which we examined their prosodic phrasing and tonal structure with an analysis cast within the framework of the Autosegmental-Metrical Theory of Intonation (e.g. Beckman, Pierrehumbert, 1986; Ladd, 1996).

The paper is structured as follows: after a Method section (§4) in which we present the corpus chosen, the speakers and the prosodic measurements, in section 5 we report the results on phrasing (§5.1) and on type and distribution of pitch accents across the sentence pairs (§5.2). In §5.3 we examine in detail the alignment and scaling properties of the most widely used focal pitch accent in both cleft and focus sentences. In section 6 we discuss how the prosodic analysis impinges on the syntactic interpretation of cleft sentences.

4. *Method*

In order to systematically compare corrective cleft sentences and left focalizations, a corpus of minimal pairs was created, taking into account three syntactic conditions. The first set of minimal pairs includes 8 main clauses with a singular clefted/focalized constituent; the second set includes 4 main clauses with plural clefted/focalized constituents, while the third set includes 4 embedded cleft sentences and focalizations with a singular clefted/focalized constituent. For all three conditions,

subject and object clefts and focalizations were equally distributed within each set of minimal pairs.

In the clefted/focalized constituents, target words are stressed on the penultimate or on the antepenultimate syllable (paroxytones *vs* proparoxytones) and the stressed syllable can be open or close (CV *vs* CVC). All segments of the target words are sonorants, in order to minimize microprosodic effects. The corpus included indeed 5 paroxytone words with an open stressed syllable, and 2 proparoxytone words, 1 with an open stressed syllable and 1 with a closed stressed syllable.

Each target sentence has been inserted in a short conversational context, with the aim of suggesting the desired interpretation and making the reading as natural as possible; the resulting short paragraphs have been pseudo-randomized and presented to the subjects as Power Point slides. 40% of similar extra texts have been added to the experimental set, in order to serve as fillers.

Four female speakers of the Italian variety of Rome, aged 20-28, were asked to read out the small texts three times. They have been recorded with a Shure WH20QTR microphone and a Zoom H2 digital recorder. Two out of three repetitions were segmented and analysed. Out of 144 sentences, was 144, 7 were discarded because reading errors had occurred.

The resulting 137 target sentences (61 focalizations, 76 corrective clefts) have been then extracted from the contexts, segmented and transcribed at phone and syllable level in Praat, and ToBI transcribed. The transcriptions have been carried out separately by the first two authors, and the diverging cases have been discussed untill an agreement was reached[6].

In order to find out whether Italian corrective/contrastive cleft sentences and left focalizations have the same prosodic realization, we measured and compared the following parameters: i) duration of the last vowel[7] of the clefted/focalized constituents; ii) distribution of focal and postfocal pitch accents types in the two structures; iii) alignment of the tonal targets of focal pitch accents with the stressed syllables of the clefted/focalized constituent, obtained by calculating the latency of L and H targets from the stressed syllable onset; iv) scaling of the L and H targets of the focal pitch accents in the clefted/focalized constituents (absolute height, Δraising and Δfalling, in Hz).

Measures of pre-boundary lengthening (i) should inform us about presence and level of prosodic phrasing between the clefted/focal constituents and the rest of the sentence; distribution of focal and postfocal pitch accents (ii) should reveal if the

[6] Out of 137 sentences, in 11 cases the two authors diverged in the trascription of the focal pitch accent (agreement: 80%). An agreed upon transcription for those cases was reached after a discussion of each one. As for prosodic phrasing, a 100% agreement was reached.

[7] Even if desirable, in analyzing the prosodic phrasing we did not take into consideration the duration of the syllable, as the last unstressed syllable of the focal word varies across sentences both in structure (CV, CjV, V) and in segmental composition. The rather limited size of the corpus would not balance out the significant variation in the duration of the syllable as induced by the presence/absence of an onset consonant and by its class (nasal, liquid, stop). The variation induced by the quality of the vowel has been taken care of by setting the target words as a random effect ("Item") in the statistical analyses.

intonation contours are comparable in the two structures; alignment and scaling measures of the focal pitch accents (iii, iv) should provide phonetic evidence for the phonological categorization of the pitch accents.

5. *Results*

5.1 Phrasing

In order to verify if a prosodic boundary occurs after the focal word in the sentence pairs (an example is shown in (7a, b)), we first inspected the F0 contours and then compared the duration of the last unstressed vowel of the focal word in focus and cleft sentences. The focal word may occur in the main clause (7a,b) or in an embedded position (7c, d):

(7) a. $[\text{MARINA}]_{\text{FOC}}$ regala gioielli di valore
b. $[\text{è MARINA}]_{\text{COPULAR}}$ $[\text{che regala gioielli di valore}]_{\text{(PSEUDO-)RELATIVE}}$
c. Ho sentito dire che $[\text{MARINA}]_{\text{FOC}}$ regala gioielli di valore
d. Ho sentito dire che $[\text{è MARINA}]_{\text{COP.}}$ $[\text{che regala gioielli di valore}]_{\text{P.-RELATIVE}}$

Out of 137 sentences analysed, a pause occurred after the focused word only in 7 cases (1 main and 2 embedded left-focus sentences; 2 main and 2 embedded clefts). The pause was easily identified in the set of left-focalized sentences (n=3, average duration = 123.6 ms; $\sigma = 20.8$); while in cleft sentences, as the focused elements is followed by the complementizer "che" ([ke]), the pause, if present, cannot be easily distinguished from the closure duration of the velar stop. Therefore, we calculated the average velar closure duration of the pre-focal "che" in the declarative main clause of embedded clefts (e.g.: 7d) and set that duration as the threshold for distinguishing a pure closure from a closure-plus-pause.

On average, the [k] closure duration of the prefocal "che" is 33.67 ms (n =14; $\sigma = 5.86$); durations of post-focal [k] closures above that threshold were considered the combined outcome of the stop closure and a pause. Four cases exceeded that value, with an average duration of closure-plus-pause of 122.5 ms ($\sigma = 26.5$).

Given that a pause was detected after the focused element only in 5% of the left-focalized sentences (3 cases out of 61) and in 5% of the corrective/contrastive clefts (4 cases out of 76), we excluded the systematic presence of an IP boundary after the focal accent in both sentence types.

In order to ascertain the presence of a boundary of a lower hierarchical level, we fit a Generalized Linear Model (JMP platform) to the duration of the last unstressed vowel with "Type of Sentence" (Focus *vs* Cleft) and "Sentence Position" (main *vs* embedded clause), as fixed factors and "Subject" and "Item" as random effects. Our corpus does not include sentences in which the target word occurs also as a subject in a broad focus sentence, condition that would allow a direct comparison with cases of absence of boundary after the target word. Therefore, we choose to compare the final vowel of the focal word with the final vowel of the post-focal following verb that we can safely assume is phrase-internal, as a verb is usually

phrased with its complement in a same prosodic constituent (see Table 1). As a consequence, a third fixed factor "Position in ip" (ip-internal vs ip-final) and the interaction "Type of Sentence* Position in ip" were included in the model.

Table 1 - *Boundary vowels and phrase-internal vowels compared for final lengthening in main and embedded clauses. Target vowels are bold and underlined.*

	ip-final	*ip-internal*
Main	È Debor**a**	che ved**o** bene in un'azienda a Milano
Embedded	Ho sentito dire che è Marin**a**	che regal**a** gioielli di valore

Results show that the factors "Type of Sentence" and "Sentence Position" are not significant, indicating that the unstressed final vowel of the focal word is not significantly different in focus and cleft sentences, nor is it different if the focal word occurs in a main or in an embedded clause. Instead, "Position in ip" is a significant factor, indicating that the duration of the unstressed final vowel is significantly longer when the word is focal (mean: 127.79 ms) compared to when it is post-focal and ip-internal (mean: 49,30 ms), $F(1,9.73) = 24.063$, $p = .0007$. The interaction between "Type of Sentence" and "Position in ip" is non-significant.

We can infer from these data that both the clefted phrase and the focus phrase are followed by a boundary of *ip* level and consequently we can conclude that the presence of an *ip* boundary points to a monoclausal interpretation of the cleft sentence, in which the copular clause and the pseudo-relative clause are wrapped in two *ips* included in the same *IP*.

However, such a phrasing would not rule out the biclausal interpretation proposed by Belletti, as it shows the presence of an *ip* boundary on the right of the focal word but it does not exclude a boundary on its left. That is, it does not exclude a phrasing in which the copula is mapped into an autonomous *IP* and the rest of the sentence is mapped into two *ips* included in the same *IP* as in (8):

(8) $[[[\text{è}]_{ip}]_{IP}[[\text{MARINA}]_{ip}\,[\text{che regala gioielli di valore}]_{ip}]_{IP}]_{U}$

The prosodic analysis shows that no pause follows the copula in any of the cleft sentences, thus excluding the presence of an *IP* boundary between the copula and the focal word. Moreover, in the vast majority of cases, the copula is not realized with any pitch accent, thus resulting included with the following focal NP in the same intermediate phrase. In the few cases in which a pitch accent is associated with the copula (10% of the total cases), its height is lower than the following focal accent, resulting in a prenuclear PA prosodically subordinated to focal accent within the same *ip*.

Summarizing, results on final lengthening and pitch accent distribution indicate that clefted and focalized constituents are phrased similarly, and that the boundary right-flanking the focus word is of *ip* level. Since there is no evidence that the copula is independently phrased nor headed, the data suggest for cleft sentences the phrasing in (10) that points to a monoclausal interpretation:

(9) $[[[\text{È MARINA}]_{ip}\ [\text{che regala gioielli di valore}]_{ip}]_{IP}]_{U}$

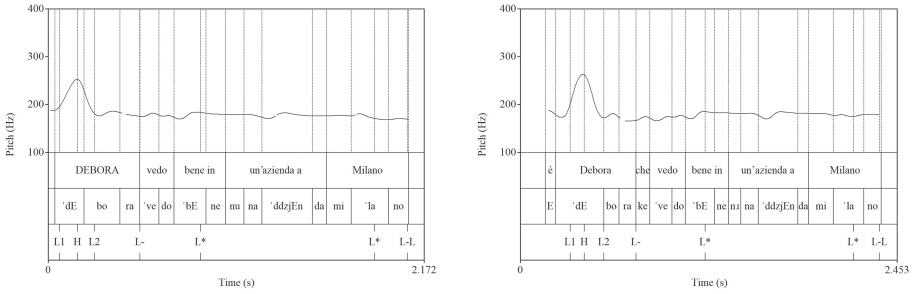
5.2 Clefts and left-focalized sentences: types of pitch accents and their distribution

All cleft sentences and left focalizations collected for this study have been clearly realized as focus-background structures, with a first prosodic constituent bearing a focal pitch accent, and a second constituent realized with a low and flat F0 contour.

In both sentence types the pitch contour starts at a medium-low pitch level and continues roughly flat until the onset of the stressed syllable is reached. There, the pitch sharply rises to a peak in the accented syllable and sharply falls at the end of it. From the offset of the accented syllable on, the F0 contour is slightly falling or flat till the end of the focalized word and continues on a low level until the end of the sentence (Fig. 2). The same pattern is found when the clefted and focalized constituents are embedded within the main clause and no intervening intermediate phrase boundary separates the two clauses (Fig. 6, top). If the main clause is wrapped into an *ip* marked by L- at its right boundary, at the start of the embedded clause the F0 contour continues low and flat until the onset of the accented syllable (Fig. 6, bottom).

In the postfocal constituent, the great majority of prominences that can be found in the corpus are of the L* type (120 out of 137) comparably to what reported by Bocci, Avesani (2006) for left focalization in Tuscan Italian; in cleft sentences, L* amounts to 87% of the postfocal accents and in left-focalized sentences to 96%. In all other few cases (17 out of 137), a compressed !H+L* pitch accent can be detected, comparable to the post-focal pitch accent in the southern varieties of Italian (D'Imperio, 2000; Grice, D'Imperio, Savino & Avesani, 2005) and in European Portuguese (Frota, 2000). The distribution of such compressed post-focal accent is not uniform across speakers but appears to be speaker-specific: one speaker (CC) is responsible for most cases of compressed !H+L* pitch accents (11/32); another speaker produces 4/32 cases and the remaining two speakers one case each.

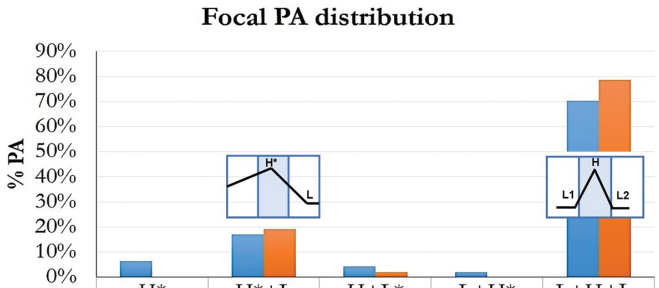
Figure 2 - *Left: Corrective left focalization: "DEBORA vedo bene in un'azienda a Milano" (DEBORA I can imagine in a company in Milan); right: Corrective cleft sentence "è DEBORA che vedo bene in un'azienda a Milano" (it is Debora that I can imagine in a company in Milan)*



The pitch accent selection for the focal constituent is more variable, but no significant difference in distribution between clefts and focalizations can be observed. As showed in Fig. 3, the most widespread pitch accent in the corpus is a rising-falling accent both in clefts (70%) and focalizations (79%). It is phonetically characterized by a sharp rise aligned at the onset of the accented syllable and by a sharp fall that ends at the offset of the same syllable or slightly later, soon after the onset of the postonic syllable (see Fig. 3). We label this pitch accent LHL and will discuss it further in §4.3.

17% of remaining cleft sentences and 19% of remaining focalizations bear an H*+L pitch accent: in contrast with the preceding one, this early falling pitch accent is preceded by a plateau or by a gradual rising movement which spreads from the sentence beginning and reaches a peak in the focal syllable, as shown by the contour in Fig. 4. Rare realizations of H*, H+L* and L+H* pitch accents are found, with a slightly higher frequency in cleft sentences than in focalizations. Fig. 3 shows percentages of occurrence of the focal pitch accents in corrective/contrastive clefts and left-focalized sentences.

Figure 3 - *Distribution of focal pitch accents in corrective clefts and focalizations. Schematic representation of the most widespread pitch accents in the F0 contours: LHL and H*+L*



	H*	H*+L	H+L*	L+H*	L+H+L
CleftC	6%	17%	4%	2%	70%
FocC	0%	19%	2%	0%	79%

In a subset of the corpus (see Appendix "Main clause: plural", 27 items) the clefted/focalized constituent is a complex NP in which two NPs are coordinated as subjects of the clause, as in the example in (8):

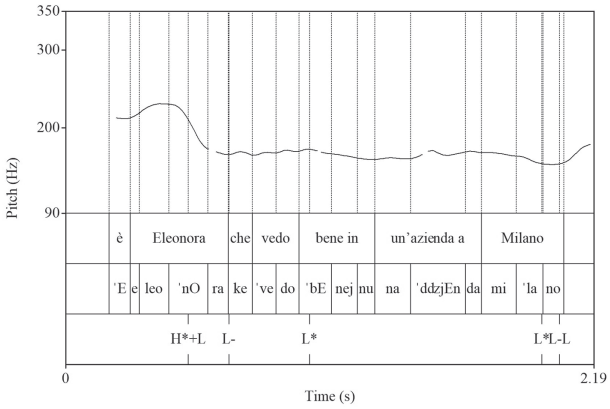
(8) Sono Andrea ed Angelo che vivranno due anni a Londra
It is Andrea and Angelo that will be living in London for two years

In 81% of those cases, a prefocal pitch accent has been realized on the stressed syllable of the first NP. The prefocal pitch accent type is variable, with a great majority of L+H* (59%) and some occurrences of H* (27%) and LHL (14%). The variability of prefocal pitch accents is predicted by the AM theory: their presence is not compulsory, and they do not convey any relevant pragmatic information – as opposed to focal pitch accents. Moreover, all but 2 prefocal pitch accents have been realized with a lower pitch span than the focal pitch

accent of the clause, which in all other cases reaches the highest F_0 of the whole contour.

Overall, the most frequent focal pitch accent in cleft sentences and in early focussed sentences has a rising-falling pattern in which three tonal targets LHL can be detected, all of which appear to align with the stressed syllable.

Figure 4 - *Corrective cleft with a H*+L pitch accent: "è Eleonora che vedo bene in un'azienda a Milano" (it is Eleonora that I can imagine in a company in Milan)*



5.3 Alignment and scaling of the LHL focal pitch accent

To evaluate whether the focal accent displays the same phonetic features in clefts an in left focalizations we analysed the alignment and the scaling properties of each tonal target in both sentence types.

Each tonal target was manually tagged by visual inspection of the F0 curve. The location of the H tone was defined as the point in time where the rise reaches the F0 peak and the peak was automatically detected by Praat as the F0 maximum within the pitch accent. No cases of high plateau are present in the set of the LHL focal accents. The particular nature of our corpus, in which each target sentence has one focal accent which is followed by a long stretch of low and flat pitch contour and generally preceded by no prenuclear pitch accent[8], highly facilitates the identification of the low turning points that precede and follow the F0 peak in the focal accent (see Figures 2 and 4). The first low target L_1 was identified as the last F0 minimum right before the rise (the "elbow" or inflection point of the pitch curve). The second low target L_2 was identified as the first F0 minimum after the peak from which the F0 curve continues low and flat till the end of the utterance. Cases of proparoxytonic focal words (e.g. "Debora", Figure 2 right) are particularly clear in showing that the minimum of the fall is reached at the offest of the accented syllable (or soon after it), and that the postonic syllables lay on the baseline of the contour as the rest of

[8] Prenuclear accents occur only in the subset of sentences in which the focal word is a plural NP.

the sentence. As the speakers of the Roman variety of Italian we analysed are highly consistent in the production of the focal accents, we felt no need to use a line-fitting procedure to automatically calculate the position of the L tones (e.g. D'Imperio, 2000; Frota, 2002; Welby, 2006), in line with Lickley, Schepman, Ladd (2005).

5.3.1 Alignment

For each L and H target in the tritonal L_1HL_2 accent of clefted/focalized constituent we performed the following measurements: i) distance in ms (expressed as a percentage of syllable duration) of the beginning of the rising movement relative to syllable onset (latency L_1-syll); ii) distance of the F0 peak relative to syllable onset (latency H-syll); iii) distance of the end of the falling movement relative to syllable onset (latency L_2-syll).

As different alignments could result from different syllabic and accentual structures (for a review see Prieto, 2011; D'Imperio, 2012), the results for proparoxytone and paroxytone focal words have been kept separate, as well as those for open and close stressed syllables[9]. In Table 2 latency percentages are reported for the alignment of every target of the L_1HL_2 pitch accent, by syllable type (open *vs* close) and word accentual structure (proparoxytones *vs* paroxytones).

Results show that the high target H is always aligned in the middle of the stressed syllable, both in clefted and in focalized constituents, independently of syllable type and word accentual structure. The timing of H is roughly central in paroxytones with open syllable (Focus: 46%, Cleft: 51%); it is slightly later in proparoxytones with open syllable (slightly more than 60%); while it is the earliest in proparoxytones with closed syllable (about 45%).

Table 2 - *Alignment of the three targets of the LHL pitch accents with respect to the stress syllable onset, organized by syllabic and accentual structure*

word	*#items*	*type*	*L1*		*H**	*L2*	
syllable			*pretonic*	*tonic*	*tonic*	*tonic*	*postonic*
*CV.'**CV**.CV*	35	Focus	91%		51%		22%
	23	Clefts	91%		46%	99%	
*'**CV**.CV.CV*	7	Focus		21%	62%		22%
	7	Clefts		31%	65%		21%
*'**CVC**.CV.CV*	7	Focus		0%	42%	86%	
	11	Clefts		2%	48%	93%	

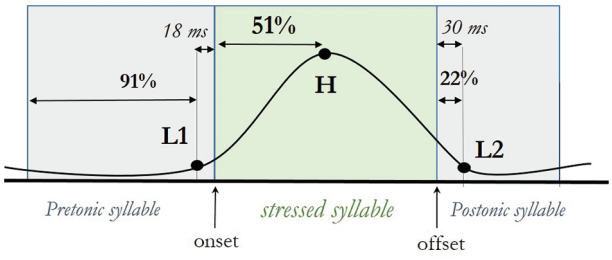
As for the low targets, their alignment appears to depend on the position of the stressed syllable in the word and on the syllable structure. The leading low tone L_1 alignes slightly before the syllable onset in paroxytones (at 91% of the pretonic

[9] Note that, as already stated in §3, the number of items per category is unbalanced, since the CV.'CV.CV words represent the great majority of the corpus.

syllable), both in clefted and in focalized constituents; while it aligns at stressed syllable onset or slightly later in proparoxytones, in a comparable number of cases in focalized and clefted constituents: earlier if the syllable is open, (respectively, 0-2%), later if the syllable is closed (respectively, 21-31%). The trailing tone L_2 aligns with the postonic syllable if the stressed syllable is open, both in focalized paroxytones and proparoxytones (22%); while it aligns differently in clefted constituents: in proparoxtitones, on the postonic syllable (21%) and in paroxytones at the end of the tonic syllable (99%).

Considering the whole LHL pattern, results show that the three tonal targets stably align with the stressed syllable, with small variations induced by the syllabic structure and by the position of the accent in the word. When the trisyllabic word is stressed initially and the syllable is closed, the rising movement is tightly aligned with the syllable onset and the falling movement completed before the syllable offset. When the trisyllabic word has the same accentual structure but the syllable is open, the rising and falling movements are shifted forward in the syllable, with the rising movent starting within the stressed syllable and the falling movement ending in the following unstressed syllable, roughly with the same percentages. If the stressed syllable is word-medial in the trysillable and it is separated by the following *ip* boundary by only one untressed syllable, the LHL pattern is shifted back in the syllable, with the rising movement starting at the end of the preceding unstressed syllable (91%) and the falling movement ending at the offset of the stressed syllable (99%), but only for the clefted constituent. The leftward push of the tonal targets relative to segments in the proximity of a phrase boundary is compatible with what observed in the literature (e.g., Mücke, Hermes, 2007). In the focus cases, the fall takes longer to complete and reaches the low target soon after the onset of the following unstressed syllable (22%, see Fig. 5).

Figure 5 - *Schematic representation of the alignment of the LHL targets in paroxytone words with open syllable in focalizations*



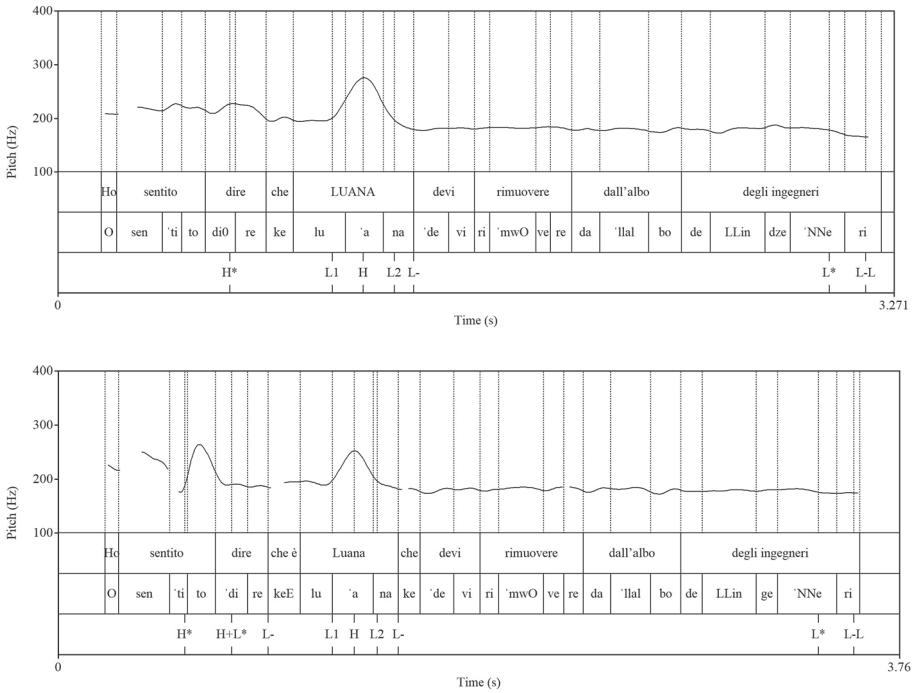
FOCUS – PAROXYTONE – OPEN SYLLABLE

Overall, results show that the tonal targets of the focal pitch accent maintain the same alignment both in cleft and left-focalixed sentences.

5.3.2 Scaling

The focal pitch accent in focalizations and cleft sentences appears to have a symmetrical shape, as can be appreciated in the contours of the following figure.

Figure 6 - *top: Corrective left focalization: "Ho sentito dire che LUANA devi rimuovere dall'albo degli ingegneri" ("I heard that LUANA you should remove from the Register of Engineers"); bottom: Corrective cleft sentence "Ho sentito dire che è LUANA che devi rimuovere dall'albo degli ingegneri" ("I heard that it is Luana that you should remove from the Register of Engineers")*



The height of the peak is comparable in both sentence types. A Generalized Linear Model (JMP platform) applied to the height (Hz) of the H target with "Sentence Type" (focus *vs* cleft), "Sentence Position" (main *vs* embedded), and their interaction ("Sentence Type*Sentence Position") set as fixed factors and "Subject" and "Item" set as random effects indicates that the peak height is not statistically different in clefted and focalized words, and that it does not vary according to of the position of the focalized word in the sentence (mean height: 262.8 Hz, $\sigma = 26.5$).

The low targets of the focal pitch accent, L_1 and L_2, are scaled in the low pitch range of the speakers, with a difference in height compared to the following low phrase accent L- and low boundary tone L% that is compatible with their lying on the declining baseline of the F0 contour[10].

[10] A Generalized Linear Model applied to the F0 of the low targets with "L Type" (L1, L2, L-, L%),

Table 3 - *Mean F0 in Hz (and standard deviation) of the low tones in focus and cleft sentence, in main and embedded position*

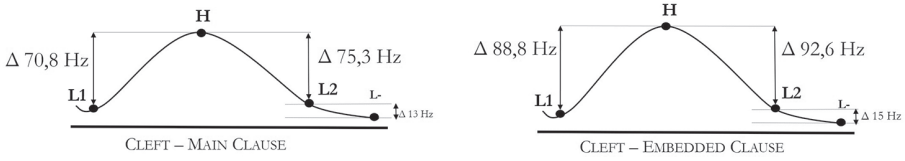
	L1	*L2*	*L-*	*L%*
Main	187.4 (σ 26.4)	179.2 (σ 14.7)	165.4 (σ 12.6)	159.1 (σ 12.9)
Embedded	182.6 (σ 18.3)	175.8 (σ 18.5)	166.1 (σ 9.8)	156.9 (σ 13.6)

In order to verify the symmetry of the focal accent and whether it is kept constant in both sentence types, a Generalized Linear Model was fit on the pitch span of the the rising and in the falling movement as dependent measure, i.e. on the difference in Hz between H and L_1 (Δ raising) and H and L_2 (Δ falling). "Type of Movement" (rising *vs* falling), "Sentence Type" (focus *vs* cleft), "Sentence Position" (main *vs* embedded), interaction "Type of Movement*Sentence" position were set as fixed factors; "Subject" and "Item" were included as random effects. Results show that the pitch span of the raising movement is not significantly different from the falling movement, while Sentence Type and Sentence Position are significant (Sentence Type: $F(1,172.5)= 73.257$ $p <0.0001$; Sentence Position: $F(1,172.3)= 13.595$ $p =0.0003$). The span of the pitch accent is higher in focalized constituents relative to clefted ones (LSM-Hz focus: 86.5 vs cleft: 78.4) and it is higher when it occurs in embedded relative to main clauses (LSM-Hz embedded: 88.9 *vs* main: 76.1; see Table 4). The interaction between type of pitch movement and its position within the sentence is not significant. A summary of the results is presented in Fig. 7 and 8.

Table 4 - *Δ raising and Δ falling values in cleft and left focalized sentences*

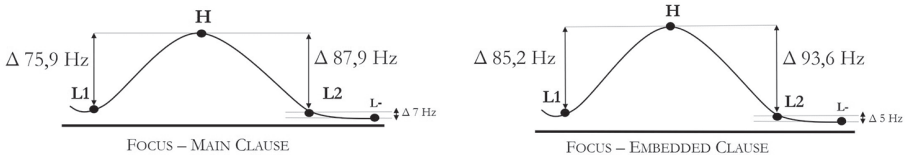
		Δ raising	*Standard Error*	*Δ falling*	*Standard Error*
Clefts	Main	71 Hz	6.235	75 Hz	6.235
	Embedded	89 Hz	12.665	93 Hz	12.665
Focus	Main	76 Hz	6.143	88 Hz	6.143
	Embedded	85 Hz	9.249	94 Hz	9.249

Figure 7 - *Scaling in main and embedded clauses: cleft*



"Sentence Type" (focus vs cleft), set as fixed factors and "Subjects" and "Item" set as random effects, shows that "Type of L target" is the only significant factor ($F(3,353)= 61.892$ $p< 0.001$). A Tuckey HSD post-hoc test indicates that each low target is significantly different from all other targets, with a pitch height that steadily decreases from L_1 to L%.

Figure 8 - *Scaling in main and embedded clauses: Focus*



Summarizing, the results presented so far clearly show that the characteristics of the focal pitch accents of clefts and focalizations do not differ: their distribution, alignment and scaling are very similar in all syntactic conditions and for all syllabic structures considered.

6. *Discussion and conclusion*

At the start of this paper we asked whether an analysis of the prosody of cleft sentences would help in disentangling the syntactic issue related to their internal constituency, namely whether they can be considered monoclausal or biclausal structures. Both positions are represented in the literature, with some scholars considering them composed of a copular clause followed by a pseudo-relative clause (see (3a)), while other considering them as monoclausal structures (see (3b)) on the basis of their similarity with left-focalized sentences. Crucially, the similarity is established not only on an informational ground, but also on a prosodic ground: both clefts and left focalizations are sentences partitioned in a focus/background structure and share what appears to be a common prosodic structure, in which the clefted and focalized constituents attract the main prominence of the intonation contour and the following backgrounded material is prosodically subordinated.

Prompted by such reported similarity, our analysis centred on the prosodic properties of a set of minimal pairs of corrective clefts and corrective left focalizations, by examining their phrasing and their accentual structure. Results of the comparative analysis reveal that in both sentence types the focalized/clefted constituent is wrapped in an autonomous prosodic phrase that separates it from the rest of the sentence. The nature of such prosodic constituent is indicated by tonal and metrical evidence: we showed that in both cases i) a L target marks the right boundary of the target constituent, ii) no pause occurs after it, iii) the final vowel of the focalized/clefted word has a comparable duration and it is longer than in a constituent internal position. All prosodic cues concur to indicate the metrical nature of such constituent: in both clefts and left focalizations, the prosodic boundary right-flanking the clefted and focalized constituent is the boundary of an intermediate phrase. The metrical structure of cleft sentences is therefore equivalent to that of sentences in which a corrective focus appears in the left periphery (Bocci, 2013). Namely, the sentence is prosodically phrased in one intonational phrase divided in two intermediate phrases coextensive respectively with the focalized/clefted constituent and with the background. Moreover, no

evidence is found that could indicate a phrasing of the copular clause in an autonomous prosodic constituent: the copula "è" is not associated with any nuclear pitch accent and it is not followed by a pause. Therefore, we argued that the copula is phrased with the clefted constituent in the same *ip* and that the third possible syntactic structure proposed in the literature (see (5)) can be safely ruled out.

That same phrasing holds true also if the linear and structural position of the cleft or left-focalized sentence is manipulated. We increased the syntactic complexity of the sentence by embedding the cleft/focalized sentence in a main one as in (10), and we created the condition in which the prosodic realization of the sentence could be changed: the clefted or the focalized word could become nuclear in a longer prosodic phrase, inclusive of the main clause (Fig. 6, top) or it could remain nuclear in the clefted/focalized constituent but necessarily shifted rightward in the utterance if a prosodic boundary is inserted after the main clause (Fig. 6, bottom).

(10) Ho sentito dire che (è) LUANA (che) devi rimuovere dall'albo degli ingegneri
I heard that it is LUANA you should remove from the Register of Engineers

Results show that the syntactic inclusion in a sovraordinate structure does not change the prosodic phrasing: an *ip* boundary is inserted after the clefted/focalized phrase, while the copula itself has no pitch accent nor is it followed by any boundary that separates it from the clefted noun. Therefore, we can conclude that all metrical evidence supports a monoclausal analysis for cleft sentences.

Clefted and left-focalized sentences are equivalent also on intonational (melodic) ground. They share the same accentual structure: a focal accent on the clefted or left-focalized constituent and postfocal L* accents on the backgrounded material. If, in a minory of cases, a compressed !H+L* pitch accent is used instead, it is equally distributed across sentence types. Clefted and left-focalized sentences also share type and phonetic properties of the most used focal accent: LHL, in which both the rising and the falling movements are tightly aligned with the accented syllable. We did not take any stance on the phonological categorization of such an accent, as we think a more thorough investigation of the intonational system of the Rome variety of Italian is needed before we can define it as truely tritonal pitch accent[11].

In conclusion, our study has shown that corrective clefts and left focalizations share a common prosodic realization in terms of phrasing, accent placement and accent type, and that a detailed analysis of their phrasing suggests a monoclausal interpretation of clefts' syntactic structure. More generally, our study confirms that a thorough prosodic analysis can help to disentangle syntac-

[11] Note that the pitch accent of contrastive/corrective focus in the Rome variety of Italian has been categorized as H*+L in Gili Fivela et al. (2015). No tritonal pitch accents are used in describing the intonational system of the Rome variety of Italian in other previous works (Sardelli , 2006; Sardelli, Marotta, 2009; Giordano, 2006). Contrastive accents with three tonal targets have been attested also in Pisa Italian (Gili Fivela, 2002), and the LHL sequence was categorized as [L]H*+L, i.e. as a bitonal falling accent preceded by a L tone deemed to be a structural property of the peak accent.

tic issues, adding to our comprehension of underlying syntactic structures from a prosody-syntax interface perspective.

Acknowledments

Our special thanks go to one anonimous reviewer for his precious comments and suggestions.

Appendix

Corrective clefts vs corrective focalizations

Main clause, singular

È Andrea che rimane due anni a Londra *it is Andrea that stays two years in London*	ANDREA rimane due anni a Londra *ANDREA stays two years in London*
È Angelo che rimane due anni a Londra *it is Angelo that stays two years in London*	ANGELO rimane due anni a Londra *ANGELO stays two years in London*
È Eleonora che vedo bene in un'azienda a Milano *it is Eleonora that I can imagine in a company in Milan*	ELEONORA vedo bene in un'azienda a Milano *ELEONORA I can imagine in a company in Milan*
È Debora che vedo bene in un'azienda a Milano *it is Debora that I can imagine in a company in Milan*	DEBORA vedo bene in un'azienda a Milano *DEBORA I can imagine in a company in Milan*

Main clause, plural

Sono Andrea ed Angelo che vivranno due anni a Londra *it is Andrea and Angelo that will be living in London for two years*	ANDREA ED ANGELO vivranno due anni a Londra *ANDREA AND ANGELO will be living in London for two years*
Sono Marianna e Valeria che vedo bene in un'azienda a Milano *it is Marianna and Valeria that I can imagine in a company in Milan*	MARIANNA E VALERIA vedo bene in un'azienda a Milano *MARIANNA AND VALERIA I can imagine in a company in Milan*

Embedded clause, singular

Ho sentito dire che è Luana che devi rimuovere dall'albo degli ingegneri *I heard that it is Luana that you should remove from the Register of Engineers*	Ho sentito dire che LUANA devi rimuovere dall'albo degli ingegneri *I heard that LUANA you should remove from the Register of Engineers*
Ho sentito dire che è Marina che regala gioielli di valore *I heard that it is Marina that gives valuable jewelry as a gift*	Ho sentito dire che MARINA regala gioielli di valore *I heard that MARINA gives valuable jewelry as a gift*

Bibliography

Beckman, M.E., Pierrehumbert, J.B. (1986). Intonational structure in Japanese and English. In *Phonology*, 3, 255-309.

Belletti, A. (2004). Aspects of the low IP area. In Rizzi, L. (Ed.), *The structure of CP and IP. The Cartography of Syntactic Structures*, Vol. 2, New York: OUP, 16-51.

Belletti, A. (2008). *Structures and Strategies*. New York: Routledge.

Benincà, P. (1978). Sono tre ore che ti aspetto. In *Rivista di Grammatica Generativa*, 3 (2), 231-245.

Bocci, G. (2013). *The Syntax-Prosody Interface: A Cartographic Perspective with Evidence from Italian*. Linguistik Aktuell 204, Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bocci, G., Avesani, C. (2005). Focus Contrastivo nella periferia sinistra: un accento ma non solo un accento. In Savy, R., Crocco, C. (Eds.), *Analisi Prosodica. Teorie, modelli e sistemi di annotazione. Atti del secondo convegno AISV- Associazione Italiana di Scienze della Voce*, 1-30.

Bocci, G., Avesani, C. (2011). Phrasal prominences do not need pitch movements post-focal phrasal heads in Italian. In Cosi, P., De Mori, R., Di Fabbrizio, G. & Pieraccini, R. *Proceedings of Interspeech 2011*, Firenze, 27-31 August 2011, International Speech Communication Association, 1357-1360.

Bocci, G., Avesani, C. (2015) Can the metrical structure of Italian motivate focus fronting?. In Shlonsky, U. (Ed.), *Beyond Functional Sequence. The cartography of syntactic structures*, vol 10. Oxford: Oxford University Press, 23-41.

Declerck, R. (1988). *Studies on Copular Sentences, Clefts, and Pseudo-Clefts*. Leuven & Dordrecht: Leuven University Press & Foris.

Den Dikken, M. (2013). Predication and specification in the syntax of cleft sentences. In Hartmann, K., Veenstra, T. (Eds.), *Cleft Structures*. Linguistik Aktuell 208, Amsterdam/Philadelphia: John Benjamins Publishing Company, 35-70.

D'Imperio, M.P. (2000). *The Role of Perception in Tonal Targets and Their Alignment*. PhD dissertation, Ohio State University.

D'Imperio, M.P. (2012). Tonal alignment. In Cohn, A.C., Fougeron, C., Huffman, M.K. (Eds.), *The Oxford Handbook of Laboratory Phonology*, Oxford-New York: Oxford University Press, 275-287.

Frascarelli, M. (2000). *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*. Dordrecht: Kluwer.

Frascarelli, M., Ramaglia, F. (2013). (Pseudo) Clefts at the Syntax-Prosody-Discourse Interface. In Hartmann, K., Veenstra, T. (Eds.), *Cleft Structures*. Linguistik Aktuell 208, Amsterdam / Philadelphia: John Benjamins Publishing, 97-138.

Frota, S. (2000). *Prosody and Focus in European Portuguese: Phonological Phrasing and Intonation*. New York NY: Garland.

Frota, S. (2002). Tonal association and target alignment in European Portuguese nuclear falls. In Gussenhoven, C., Warner N. (Eds.), *Papers in Laboratory Phonology VII*, Berlin: Mouton de Gruyter, 387-418.

Gili Fivela, B. (2002). Tonal alignment in two Pisa Italian peak accents. In Bell, B., Marlien, I. (Eds.), *Proceedings of the 1st International Conference on Speech Prosody*. Aix-en-Provence: Laboratoire Parole et Langage, 339-342.

Gili Fivela, B., Avesani, C., Barone, M., Bocci, G., Crocco, C., D'Imperio, M., Giordano, R., Marotta, G., Savino & M., Sorianello, P. (2015). Varieties of Italian and their intonational phonology. In Frota, S., Prieto, P. (Eds.), *Intonation in Romance.* Oxford: Oxford University Press, 140-197.

Giordano, R. (2006). The intonation of polar questions in two central varieties of Italian. In Hoffmann, R., Mixdorff, H. (Eds.), *Proceedings of the Third International Conference on Speech Prosody (Dresden)*. CD-ROM: PS8-10-155.

Haegeman, L., Meinunger, A., Vercauteren, A. (2013). The syntax of it-clefts and the left periphery of the clause. In Shlonsky, U. (Ed.), *Beyond functional sequences: The cartography of syntactic structures*, vol. 10. Oxford: Oxford University Press, 73-90.

Hartmann, K., Veenstra, T. (2013). *Cleft Structures*. Linguistik Aktuell/Linguistics Today 208, Amsterdam / Philadelphia: John Benjamins Publishing Company.

Jespersen, O. (1927). *A Modern English Grammar on Historical Principles*. Heidelberg: Winter.

Ladd, D.R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Lickley, R., Schepman, A. & Ladd, R. (2005), "Alignment of "Phrase Accent" lows in Dutch falling-rising questions: Theoretical and methodological implications". In *Language and Speech*, 2005, 48 (*2*), 157-183.

Meinunger, A. (1998). A monoclausal structure for (pseudo-)cleft sentences. In Tamanji, P.N., Kusumoto, K. (Eds.), *Proceedings of NELS* 28. *GLSA,* University of *Massachusetts,* 283-298.

Mücke, D. Hermes, A. (2007). Phrase Boundaries and Peak Alignment: An Acoustic and Articulatory Study. *Proceedings 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany.

Nespor, M., Guasti, M.T. (2002). Focus-stress alignment and its consequences for acquisition. In *Lingue e linguaggio*, 1 (1), 79-196.

Samek-Lodovici, V. (2006). When right dislocation meets the left-periphery. A unified analysis of Italian non-final focus. In *Lingua,* 116 (12), 836-873.

Sardelli, E. (2006). *Prosodiatopia: alcuni parametri acustici per il riconoscimento del parlante.* PhD dissertation, Università di Pisa.

Sardelli, E., Marotta, G. (2009). Prosodiatopia: parametri prosodici per un modello di riconoscimento diatopico. In Ferrari, G., Benatti, R., Mosca, M. (Eds.), *Linguistica e modelli tecnologici di ricerca: Atti del XL Congresso della SLI.* Rome: Bulzoni, 411-435.

Selkirk, E. (2005). Comments on Intonational Phrasing in English. In Frota, S., Vigario, M. (Eds.), *Prosodies.* Berlin: Mouton de Gruyter, 11-58.

Szendröi, K. (2001). *Focus and the Syntax-Phonology Interface*. PhD dissertation, University College, London.

Welby, P. (2006). French intonational structure: Evidence from tonal alignment. In *Journal of Phonetics*, 34(*3*), 343-371.