

Phrasal prominences do not need pitch movements: postfocal phrasal heads in Italian

Giuliano Bocci¹, Cinzia Avesani²

¹Università di Bologna, Italy, ² ISTC, CNR , Italy

giuliano.bocci@unibo.it, cinzia.avesani@pd.istc.cnr.it

Abstract

Informationally *Given* phrases following an instance of focus are generally realized in a compressed pitch range and are assumed to lack prosodic prominences above the word level. In this paper, we address the question of the metrical representation of postfocal constituents in Tuscan Italian. The results of a production experiment show that, despite their being realized with a low and flat F0 contour, postfocal constituents are not extrametrical, but are phrased and assigned metrical prominences of phrasal level. The impact of our results on the prosodic representation of Italian is discussed.

Index Terms: Prosody, Focus, prosodic hierarchy, *Givenness*.

1. Introduction

When the Focus phrase precedes the elements in its scope, the latter are phonologically subordinated to the focus phrase and this results in an inversion of the default prominence pattern. In these contexts, post-focal elements included in the scope of Focus are cross-linguistically characterized by a compressed pitch contour. In English, where post-focal elements are realized with a flat and low contour, it is generally assumed that they are not associated with intonational prominences.

To account for the effect of focus on the prosodic structure, Féry and Samek-Lodovici [5] and Selkirk [11] assume two distinct rules/constraints. The Contrastive Focus Prominence Rule, proposed by Truckembrodt [14], states that a focus phrase must be more prominent than the elements in its scope; “destress/deaccent *Given*”, modeled after Swartzchild [10], states that an element which is informationally *Given* must be non-prominent. It is this second constraint that accounts for the fact that post-focal and *Given* elements in Germanic languages lack metrical and intonational prominences above the word level.

With regards to Romance languages, it has been proposed in the literature on syntax-prosody interface that post-focal elements fail to be assigned metrical and intonational prominences above the word-level, and, in this sense, are extrametrical [15,13]. In these proposals, Italian Spanish and Catalan are assumed to have a rigid prosodic template, where the main prominence is invariably assigned to the rightmost metrical head in the utterance since all the prosodic constituents above the word must be right headed (*Rightmostness*). Extrametricality of postfocal constituents thus becomes the formal device through which the alignment of the Focus phrase with the main prominence is achieved, in compliance with *Rightmostness*.

However, this view appears quite problematic in light of the empirical research on Portuguese [6] and on southern varieties of Italian [8], where post-focal elements, despite being *Given* and within the scope of Focus, are systematically associated with compressed pitch accents (PAs). These findings show that post-focal and *Given* constituents are not invisible to intonation. Among Italian varieties, a puzzling

case is Tuscan where the pitch contour does not show any evidence of postfocal PAs: the pitch contour is invariably low and flat, as observed in English. However, it has been proposed that Tuscan Italian does not substantially differ from the other Italian varieties. In Bocci [4] it is postulated that post-focal elements in Tuscan cannot be intonationally unspecified, but are rather associated with a L* PA whose distribution is driven by the occurrence of a focus phrase. In this paper, we assume this analysis, although no much hinges on it for the current discussion.

Although many researchers have reported that post-focal constituents are realized with compressed PAs in many varieties of Italian, it is not clear which kind of metrical heads in the prosodic hierarchy they associate with and, actually, no attempt has been made to identify the metrical representation of post-focal constituents. In this paper, we address this issue and we show that post-focal constituents in Tuscan Italian, despite their being *Given* and realized with a low and flat pitch contour, are phrased and assigned phrase-level prominences by virtue of default phrasing and stress assignment rules.

Several interesting issues are tied to the metrical representation of post-focal constituents. In fact, if post-focal constituents are assigned phrasal metrical prominences, we have to discard the idea that Italian is characterized by a rigid prosodic template which invariably obeys *Rightmostness*, since the focus phrase is not aligned with the rightmost phrasal prominence. Finally, if post-focal constituents in Tuscan Italian are endowed with genuine phrasal prominences which are not realized via pitch movements, we have to reject the common view that the heads of the higher layers in the prosodic hierarchy are always expressed by means of F0 movement, as concluded in [3] with regard to English.

2. Our experiment

To investigate the prosodic properties of post-focal constituents, we carried out an experiment on read speech Tuscan Italian. We varied the pragmatic properties and syntactic structures of superficially quasi-identical sentences in order to manipulate the prosodic constituency and the distribution of the prominences at the higher level of the prosodic constituency. Consider the set of stimuli exemplified in Tab. 1. Our aim was to compare the properties of the infinitive – the target word – across the following conditions.

Condition A(ccented): the infinitive verb occurs in sentences with the basic word order (SVO) in Broad Focus (BF) and with Informational and Contrastive Focus (IF and CF) on the last word. We hypothesized that, independently of the focus structure, the verb would be phrased along with its complement into a single phonological phrase (φ), whose head is assigned to the complement (see Tab. 2). Although the verb in this prosodic configuration tends to be unaccented, Welby [16] shows that it can be optionally associated with a PA.

In Condition P(ost-focal), the verb occurs in post-focal position, as the preverbal subject was elicited with a contrastive focus interpretation which takes scope over the

whole sentence. We expected that the background of the sentence would be realized with a flat and low F0 contour, and that the verb would be phrased along with its complement, exactly as in Cond. A. The infinitive is thus supposed not to receive a degree of prominence higher than the word-level one. In fact, if the postfocal elements were phrased into headed prosodic constituents, the head of such constituents would be assigned to the object, rather than to the infinitive. If postfocal elements could not receive metrical prominences above the word level, the verb would not qualify *a fortiori* as a head of phonological or intonational phrase (ι) (see Tab. 2).

Condition H includes two types of sentences in which the verb occurs postfocally, being the first constituent elicited with a contrastive focus interpretation, as in condition P. In H1, the last element is the direct object preceded by a resumptive clitic on the verb. The occurrence of the clitic was expected to force speakers to interpret the object as a Right-Dislocated (RDed) topic, since clitic doubling of objects only occurs in Italian with dislocated topics. In H2, the focussed object is fronted and the last word is the subject. Given the properties of postverbal subjects in Italian and the pragmatic conditions of elicitation, the subject was expected to be interpreted as a RDed topic, although no visible morphosyntactic marker of dislocation occurs when the dislocated element is a subject.

The rationale and the hypotheses underlying the experiment are the following. By assuming that right dislocated topics are phrased into independent prosodic constituents, speculatively intonational phrases (ι), the last word in H was expected to be phrased into an independent ι. Accordingly, the insertion of a prosodic boundary before the topic is supposed to induce preboundary lengthening on what precedes it, i.e. the infinitive. In such a configuration, the durations of the final syllable and vowel of the infinitive verb in the condition H should be longer than in conditions A and P, where the verb is phrased along with its complement. If that were the case, the restructuring verb and the infinitive should occur between two prosodic boundaries: the ι boundary preceding the topic, and the boundary following the focus phrase. Actually, it is commonly assumed that a preverbal focus phrase in Italian is followed by a prosodic boundary, although it is disputed if it is a boundary of φ- or ι-level [see 4]. For the sake of argument, we assume that the focus phrase is followed by a φ-boundary, but not much hinges on this specific assumption for the current discussion. In either case, the verbs in Cond. H are expected to form an independent prosodic constituent by being wrapped between boundaries. Hence, we can formulate two alternative hypotheses with regard to the prosodic representation of the sentences in condition H.

Table 1. A set of sentences of our corpus. The infinitive (in boldface) occurs in three conditions: A=(potentially) pitch accented, P=postfocal, H= metrical head.

Condition	Sentences
A	1 [Germanico vorrebbe invitare Pierangela] _{BF}
	2 Germanico vorrebbe invitare [Pierangela] _{IF}
	3 Germanico vorrebbe invitare [PIERANGELA] _{CF} (Germanico would like to invite Pierangela)
P	[Germanico] _{CF} vorrebbe invitare Pierangela (Germanico would like to invite Pierangela)
H	1 [Germanico] _{CF} la vorrebbe invitare [Pierangela] _{RD} (Germanico her-would like to invite Pierangela)
	2 [I Germanico] _{CF} vorrebbe invitare [Pierangela] _{RD} (The Germanicos would like to invite Pierangela)

If postfocal and *Given* elements cannot bear phrasal prominences by virtue of *Destress given*, the prosodic constituent in which the verb occurs should lack its head and

be “enclitic” on the prosodic constituent formed by the focus phrase. Conversely, if default stress assignment rules apply to postfocal and *Given* elements, the prosodic constituent (φ) including the restructuring verb and the infinitive should be endowed with its (φ-) head, and this head should be assigned to the rightmost element in the constituent, i.e. the infinitive. If this is the case, regardless the pitch contour, the stressed syllable of the infinitive is expected to be more prominent in condition H where it carries a phrasal prominence than in both conditions A and P.

Table 2. Prosodic structures across conditions. The prominences marked by “+” are not directly tested in the experiment, but assumed in light of our conclusions.

Condition	Prosodic structure
A BF or late IF/CF	{ * } _v [*] _i (*) _φ (*) _φ Germanico vorrebbe invitare Pierangela
P initial CF	{ * } _v [*] _i (*) _φ (* +) _φ [Germanico] _{CF} vorrebbe invitare Pierangela
H initial CF, RDed	{ * } _v [*] _i [+] _i (*) _φ (*) _φ (+) _φ [Germanico] _{CF} la vorrebbe invitare [Pierangela] _{RD} [I Germanico] _{CF} vorrebbe invitare [Pierangela] _{RD}

2.1. Method

10 speakers of Tuscan Italian (2 from Florence, 8 from Siena) read 12 short scripts written as dialogues (4-6 dialogical turns) in pseudo-randomized order. The scripts were designed for eliciting the desired semantic and pragmatic properties of the target sentences (one per script). Except full stops and question marks, no other punctuation marks have been used in the scripts in order to avoid spurious phrasing effects. Subjects were warned that punctuation marks had been mistakenly left out in the stories’ printouts. Subjects were recorded individually in a quiet room by means of a head-mounted microphone (Shure Beta 53). Each subject produced 4-5 repetitions of 12 target sentences (6 types of sentences x 2 lists). The first list is reported in Tab 1. Out of the collected data, we analyzed 3-4 repetitions for each target sentence. The target sentences have been manually segmented into phonemes and ToBI transcribed.

A number of acoustic measures were computed on the target word, in both lists, ending in [-‘tare]. For the final syllable [re], we computed the duration of the vowel and the syllable; for the stressed syllable [‘ta], besides vowel duration we analyzed other acoustic properties known to correlate with prominence: formant structure (F1 and F2) and spectral emphasis. Two measures related to spectral emphasis were calculated at vowel midpoint: i) spectral tilt (A1*-A2*), computed as the difference in dB between the strongest harmonic of F1 and the strongest harmonic of F2, corrected with the formula proposed in [7]; ii) spectral balance [see 12], calculated as the difference in dB between the frequency bands B1 (0-500 Hz) B2 (500-1000 Hz) and B3 (1-2 kHz).

All the data were analyzed by using linear mixed models including subject and items as random effects [9]. As for continuous variables, the p-values were estimated using 10000 Markov Chain Monte Carlo simulations. The analyses were also replicated by means of Repeated Measures Anovas.

2.2. Results

2.2.1. Validating condition A.

Before comparing the conditions A, H, and P, it is relevant to ascertain whether all the three types of sentences included in the Cond. A were actually realized with the metrical structure reported in Tab. 2. Are narrow focus phrases preceded by a metrical boundary signaling the focus size? If this were the case, the infinitive in narrow focus sentences would qualify as a phrasal head. According to our findings, however, all the different focus structures in Cond. A were realized with same prosodic structure sketched in Tab 2., since neither focus size nor focus type affect phrasing. No preboundary lengthening effect was found on the final syllable of the infinitive. A linear mixed model (reference group=BF) indicated that the duration of the final syllable in BF (mean=83 ms.) does not significantly differ either from CF ($\beta=-3$, $t=1.866$, $pMCMC>.05$) or IF ($\beta=-1$, $t=-0.593$, $pMCMC>.05$).

However, the focus structure of the sentence has a marginal impact on the association of prenuclear PAs with the infinitive. In our data, the probability for the infinitive to be PAed is 0.41 when it occurs in BF, 0.59 when it is *Given* and precedes a narrow CF phrase, and 0.29 when it is *Given* and precedes a narrow IF phrase. A mixed logit model indicated that the comparison between BF (=reference group, $\beta=.423$) and CF was significant ($\beta=.806$, $z=2.014$, $p=.044$), while BF vs. IF was not ($\beta=-.451$, $z=1.107$, $p=.268$). It seems to us quite plausible to conclude that the distribution of prenuclear PAs is highly optional, as argued by [W], and that the higher probability of prenuclear PAs in CF is to be tracked back to a more expressive style of speech. In fact, it cannot result from differences in phrasing, nor be accounted for in terms of *Deaccent Given* (being the infinitive non-*Given* only in BF). Overall, we found that when PAed, the infinitive in Cond. A was associated with a H* PA 48% of the time, with H+L* 41%, L+H* 11%.

Moreover, the occurrence of a prenuclear PA on the infinitive in Cond. A, correlates with longer durations of the stressed vowel (and syllable). When PAed, the stressed vowel is on average 12 ms longer than when unPAed (unPAed=ref. group $\beta=84$, PAed $\beta=12$, $t=-5.063$, $pMCMC<0.0001$). For the following comparisons across Cond. A, H, and P, we take into consideration only the cases in which the infinitive was associated with a PA, i.e. the cases in which the stressed vowel and syllable are longer. This shrinks the total number of the sentences taken into account for further analyses to 303.

2.2.2. Phrasing of postfocal constituents.

If phrasing applies to postfocal material, the RDeD topic in Cond. H should be preceded by a τ -boundary. This is expected to cause preboundary lengthening at the right edge of the infinitive. This prediction is actually born out. The final vowel of the infinitive in Cond. A and P – where the infinitive is phrased along with its complement – result to be respectively 12 ms. and 16 ms. shorter on average than in Cond. H (H=reference group $\beta=77.167$; H vs. A $\beta=-12.3$ $t=-3.611$ $pMCMC=.0007$; H vs. P $\beta=-16.3$ $t=-4.824$ $pMCMC<.0001$). Analogous results are found with regard to the final syllable of the infinitive (H $\beta=99.079$; H vs. A $\beta=-15.502$ $t=-4.199$ $pMCMC=.0001$; H vs. P $\beta=-17.355$ $t=-4.771$ $pMCMC<.0001$)

2.2.3. Postfocal phrasal heads

Since there is evidence that the occurrence of a RDeD topic in Cond. H triggers the insertion of an τ -boundary after the infinitive, the two alternative hypotheses on the degree of

prominence associated with the infinitive become relevant. If default prosodic rules regularly apply to postfocal and *Given* elements, the infinitive in Cond. H is the head of the resulting prosodic constituent, and should bear a higher degree of prominence than in A or in P. On the contrary, if the assignment of phrasal heads to postfocal and *Given* elements is prevented, the infinitive in Cond. H should be assigned the same prominence than in A and P.

Pitch contour. As expected, the postfocal region of the sentences in Cond. H and P was realized with a flat and low F0. In both cases, no visible obtusation in the F0 contour that could be clearly interpreted as a high, rising or falling PA was found. Consider Fig. 1. To quantify the amount of pitch movement on the stressed syllable of the infinitive across conditions, we compared the standard deviations of the F0 trajectory on the central part of the stressed vowel. In Cond. A, i.e. when a PA is realized on the vowel, the standard deviation values result to be twice as higher than in Cond. H and P (A=reference group $\beta=.686$; H vs. A $\beta=-0.311$ $t=-4.966$ $pMCMC<.0001$; A vs. P $\beta=-.365$ $t=-4.971$ $pMCMC<.0001$). Moreover, at vowel midpoint, the F0 values (in st.) in Cond. A are significantly higher than in Cond. P and H (A=reference group $\beta=6.827$; A vs. H $\beta=-2.853$ $t=-6.733$ $pMCMC<.0001$; A vs. P $\beta=-2.205$ $t=-5.204$ $pMCMC<.0001$).

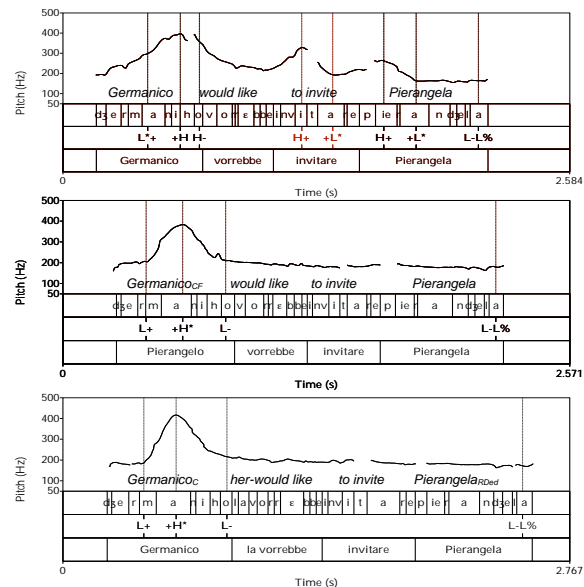


Figure 1. Examples of F0 contours in Condition A (top), Condition P (middle) and Condition H (bottom).

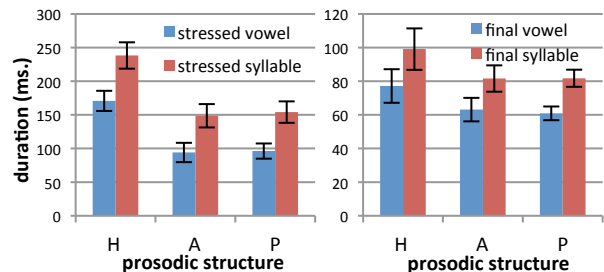


Figure 2. Mean durations (averaged over subjects) of stressed syllables and vowels (left), and of final syllables and vowels of the target word (right). Bars=C.I. 95%.

Durations. Unlike F0 values, durations straightforwardly indicate that the stressed syllable and the stressed vowel of the infinitive are more prominent in Cond. H than in Cond. A and

P as expected if the infinitive qualifies as a phrasal head in Cond. H. The stressed vowel in Cond. H ($\beta=170$) results to be on average 75 ms. (i.e. 79%) longer than in Cond. A ($\beta=75.399$ $t=-14.02$ $pMCMC<.0001$) and P ($\beta=74.552$ $t=-14.06$ $pMCMC<.0001$), while the stressed syllable in Cond. H ($\beta=238$) results to be 88 ms. (i.e. 58%) longer than in Cond. A ($\beta=87.648$ $t=-15.69$ $pMCMC<.0001$) and 84 ms. (i.e. 55%) longer than in Cond. P ($\beta=84.117$ $t=-15.55$ $pMCMC<.0001$) (see Fig. 2).

F1 and F2. The stressed vowel [a] of the infinitive, in all the stimuli, is characterized by more extreme formant values in cond. H. F1 values, measured in Hz at midpoint, result to be significantly higher in Cond. H ($\beta=799$) than in Cond. A ($\beta=70$ $t=-6.687$ $pMCMC<.0001$) and P ($\beta=103$ $t=-6.687$ $pMCMC<.0001$). Furthermore, F2 values, measured in Hz at midpoint, are significantly lower in Cond. H ($\beta=1474$) than in Cond. A ($\beta=181$ $t=9.93$ $pMCMC<.0001$) and P ($\beta=170$ $t=9.45$ $pMCMC<.0001$). Accordingly, also the difference between F1 and F2 is significantly lower in Cond. H ($\beta=799$; H vs. A $\beta=251$ $t=10.15$ $pMCMC<.0001$; H vs. P $\beta=273$ $t=11.19$ $pMCMC<.0001$). Analogous results are found for the values of F1 and F2 measured at the overall intensity peak, at the maximum point of F1 and at the minimum point of F2.

Spectral emphasis. We found a clear effect of spectral tilt on the stressed vowel in condition H. In this condition ($\beta=5.149$), the value of A1*-A2* results to be significantly lower (3dB) than in Cond. A ($\beta=3.047$ $t=4.143$ $pMCMC<.0001$) and 2.2 dB lower than in P ($\beta=3$ $t=3.087$ $pMCMC=.0031$). As for spectral balance, however, the measure of B1-B2 in Cond. H ($\beta=-7.556$) does not significantly differ from Cond. A ($\beta=.594$ $t=1.118$ $pMCMC>.05$), while it is significantly lower than in cond. P ($\beta=2.2$ $t=4.384$ $pMCMC<.0001$). Analogous results are obtained for B1-B3. Notably, the fact that Cond. H and A do not significantly differ in spectral balance values is somehow expected. Since B1 encompasses F0, the measures of spectral balance are affected by the values of F0, and accordingly the comparison between H and A is biased, being the words in the latter case PAed. However, that bias does not apply to the comparison between H and P and consequently the result should be taken as reliable and meaningful.

3. Discussion and conclusions

Our findings show that phrasing in Italian applies to postfocal elements, despite their being *Given* and within the scope of focus. The occurrence of a RDeD topic (Cond. H) calls for the insertion of an τ -boundary on its left-hand and the target word, preceding the topic, undergoes pre-boundary lengthening. We conclude that default stress-assignment rules apply on postfocal constituents, and that the target word in Cond. H qualifies as a phrasal head, despite its being *Given* and within the scope of focus. In our view, this can be elegantly accounted for by assuming that the phonological constraint “Destress Given” is not active in Italian.

Although no F0 movement signals the target word as more prominent in Cond. H than in Cond. A and P, all the other discussed correlates of prominence indicate that the stressed syllable and the stressed vowel are more prominent in H than when the target word is phrased along with its complement, either in prenuclear (as in A) or postfocal context (as in P). In particular, the stressed syllable undergoes a remarkable duration increase which cannot simply be imputed to the deceleration of the articulators in front of a boundary as: i) the duration increase on the stressed syllable is much higher than the one observed on the final syllable; ii) the stressed vowel in Cond. H is characterized by more extreme formant trajectories and spectral emphasis, which index hyperarticulation and

increase of the articulatory effort. In light of these results, we conclude that phrase-level accents in Tuscan Italian do not necessarily imply visible F0 movement [cfr. 1], differently from what has been claimed for English [3].

Moreover, we conclude that the prosodic system of (Tuscan) Italian cannot be characterized by a rigid prosodic template in which the prominences above the word level are invariantly assigned in compliance with Rightmostness. Since postfocal material is not extrametrical, structures with early focus such as in Cond. H and P violate *Rightmostness*, being the main prominence not aligned with the rightmost subordinated prominence in the structure.

According to our analysis, the insertion of a τ -boundary before the topic in cond. H is ruled by the discourse-related properties characterizing topics or by the syntactic configuration involved in right dislocation. However, we claim that the prominence assigned to the infinitive in Cond. H does not stem from specific discourse-related properties of the infinitive itself, but that it is assigned only by virtue of default prominence assignment rules, which require every prosodic constituent to be headed. In fact, despite our results are somehow reminiscent of what is observed in nested foci configurations in English [2], the infinitive in Cond. H does not qualify semantically as a second occurrence of focus. As a consequence, we conclude that phrasal prominences at the higher levels of the prosodic hierarchy can be inserted only by virtue of prosodic requirements, and this conclusion strengthens the crucial role of phonology in prosody.

4. References

- [1] Avesani, C., Vayra, M. and Zmarich, “On the articulatory bases of prominence in Italian”, Proc. XVI ICPhS, II: 981-984, 2007.
- [2] Beaver, D., Clark, B., Flemming, E., Jaeger, T. and M. Wolters, “When Semantics meets Phonetics”, Language 83, 2, 2007.
- [3] Beckman, M. and Edwards, J., “Articulatory evidence for differentiating stress categories”, in P.A. Keating P.A. [Ed], Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form, 7-3, Cambridge University Press, 1994.
- [4] Bocci, G., The syntax-prosody interface from a cartographic perspective: evidence from Italian, John Benjamins, forthcoming
- [5] Féry, C. and Samek-Lodovici, V., “Focus Projection and Prosodic Prominence in Nested Foci”, Language, 82:131-150, 2006.
- [6] Frota, S., Prosody and focus in European Portuguese: phonological phrasing and intonation, Garland Pub., 2000.
- [7] Fulop, S., Ethelbert, K. and Ladefoged, P., “An acoustic study of the tongue root contrast in Degema vowels”, Phonetica, 55:80-98
- [8] Grice, M., D’Imperio, M. P., Savino, M. and Avesani, C., “Towards a strategy for ToBI labelling varieties of Italian”, in S.-A. Jun [Ed], Prosodic Typology: The Phonology of Intonation and Phrasing, 362-389, Oxford University Press, 2005.
- [9] Pinheiro, J.C. and Bates, D. M., “Mixed Models in S and S-PLUS”, Statistics and Computers, Springer, New-York, 2000
- [10] Schwarzschild, R., “Givenness, Avoid F, and other constraints on the placement of accent”, Natural Language Semantics, 7:141-177, 1999.
- [11] Selkirk, E., “Contrastive focus, Givenness and the unmarked status of ‘Discourse-new’”, Interdisciplinary Studies on Information Structure, 125-145, 2007.
- [12] Sluijter, A. M. C. and van Heuven, V. J., “Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch”, Phonetica, 52:71-89, 1995.
- [13] Szendrői, K., “Stress-focus correspondence in Italian”, *Proceedings of Going Romance 2000*, 2002.
- [14] Truckenbrodt, H., Phonological phrases. Their relation to syntax, focus, and prominence, PhD dissert., MIT, 1995.
- [15] Vallduví, E., The Informational Component, Garland, 2002.
- [16] Welby, P., “Effects of pitch accent type and status on focus projection”, Language and Speech, 46, 1: 53-81, 2003.