Journal of Speech Sciences 7(2): 51-64. 2018.
Available at: http://revistas.iel.unicamp.br/joss

COMPLEX ILLOCUTIVE UNITS IN THE LANGUAGE INTO ACT THEORY: AN ANALYSIS OF NON-TERMINAL PROSODIC BREAKS OF BOUND COMMENTS AND LISTS[i]

SACCONE, Valentina[1*]
VIEIRA, Marcelo[2]
PANUNZI, Alessandro[3]

[1]University of Basel, University of Florence
[2]Laboratory of Empirical and Experimental Language Studies, Federal University of Minas Gerais
[3]University of Florence

Abstract: *This work presents a preliminary analysis of a prosodic description of two different structures in spoken language within the theoretical framework of the Language into Act Theory (L-AcT): (i) chains of two or more Bound Comments (COB) that do not form a compositional informative and prosodic unit; (ii) compositional Information Units formed by two or more Multiple Comments (CMM) of the List type, linked together by a conventional prosodic pattern that implements a specific meta-illocutive structure. This study aims to underline specific features of the COB units and the List-type CMM units, detecting prosodic properties of the Italian and the Brazilian Portuguese spoken languages. Through a specific script for Praat software, different parameters are automatically calculated: f_0 mean, f_0 shift, f_0 slope/variation rate, duration, spectral emphasis. Our results highlighted a common prosodic behavior in COB-units in terms of f_0 slope (rising in the stressed syllable before the boundary and falling on the unstressed one just before the boundary), and a great similarity between the two COBs and Lists.*

Keywords: prosody; prosodic breaks; Bound Comments; Lists; Language into Act Theory.

1 Introduction

This work presents a description and an analysis of prosodic breaks[1] in spontaneous spoken Italian and Brazilian Portuguese languages.

The paper starts from the points presented in Panunzi & Saccone (2018) and develops further by enhancing the prosodic analysis. It represents a first extension of the previous pilot study, with the aim of describing the formal differences between different types of non-terminal breaks co-occurring with two specific Information Units, as they are defined in the theoretical framework of Language into Act Theory (L-AcT; Cresti, 2000; Moneglia & Raso, 2014). More specifically, this work deals with the prosodic and formal features of the tone units corresponding to Bound Comments and List-type Multiple Comments, as described below, and delineates a base for future prosodic studies on this matter. For this purpose, two samples have been extracted, one from each of the two Minicorpora of Italian and Brazilian Portuguese (from DB-IPIC - Cresti & Moneglia, 2005; Panunzi & Gregori, 2012 - and C-ORAL-BRASIL corpus; Raso &Mello, 2012).

The two sample collections include 76 non-terminal prosodic units, which have been analyzed in terms of prosodic cues on both sides of the tonal breaks. To evaluate them, we used the Praat software (Boersma & Weenink, 2005) and a Praat script specially designed for this work.

Section 2 presents an introduction of the theoretical framework and deepens the nature and characteristics of the Information Unit treated in the analysis. In Section 3, we present the sample

[1] We use the term "break" to indicate the perception of discontinuity between tone units, and "boundary" for what concern unit borders.

[i] The authors conceived and discussed together all the content of this paper. However, their own contribution can be specified as follows: Valentina Saccone carried out the Italian prosodic analysis and wrote Sections 1, 3; Marcelo Vieira carried out the Brazilian Portuguese prosodic analysis, implemented the used script and wrote Sections 4, 5; Saccone and Vieira co-wrote Sections 6; Alessandro Panunzi supervised the research and wrote Section 2.

* Corresponding author: valentina.saccone@unibas.ch

collections and we introduce the prosodic parameters used for the analysis, that it is reported in detail in Sections 4 and 5. Conclusions and main results are outlined in Section 6.

2 The Language into Act Theory and the Comment Unit

2.1 General framework

The Language into Act Theory assumes that there is a systematic correspondence between pragmatic and tone units in speech. This correspondence extends on two levels in a hierarchical relation. At the higher level, terminal prosodic profiles are associated to the production of pragmatically interpretable Utterances, which correspond to the accomplishment of Speech Acts (Austin, 1962); at the lower level, tone units are associated to the segmentation of Utterances into information units. These correspondences make possible to study the pragmatic structure of spoken language based on the positive perceptual data given by the prosody (Moneglia, 2011).

In other words, L-AcT claims that prosody is the major formal component which segments the speech flow and encodes pragmatics values. As a matter of fact, it systematically signals the boundaries of each Utterance by means of a conclusive profile and the boundaries of information units within the Utterance by means of non-conclusive profiles. Moreover, the different forms of the different profiles signal the illocution values and the information values of the single units.

L-AcT assumes that, in most cases, there is a single tone/information unit carrying illocutionary force, and it corresponds to the only autonomous one, i.e. the only single tone/information unit that can be interpreted even if the others are deleted. This unit is called *Comment*, which is both necessary and sufficient for the accomplishment of the Speech Act. In this way, the definition of the information structure itself is strictly related to the fulfillment of the illocution: the nuclear unit of the Utterance is the one carrying the illocutive value.

In the framework of L-AcT analysis, each prosodic unit is labeled with respect to its informative function. Information tags divide the speech units into two main classes: Textual or Dialogic.

Textual Units bear the semantic content of an Utterance. The Comment is the only textual unit required to perform an Utterance, while the others are mainly devoted to supporting it for the accomplishment of the illocutive force. Dialogic Units are dedicated to carry out communicative functions, as taking the turn, attracting the addressee attention, or keeping the communicative channel open.

In addition to these main kinds of units, L-AcT foresees the presence of non-informative prosodic units, which mainly occur in case of disfluency or when a long information unit is dived in more prosodic units for performance reasons.[2]

2.2 Bound Comments and Multiple Comments

According to L-AcT, the Comment unit corresponds to the Utterance nucleus, since it bears the illocutive force and allows the pragmatic interpretability of the whole sequence.

As mentioned above, a prosodic conclusive sequence usually contains a single Comment unit, but it is also possible that more than one independent unit bears an illocutionary value. This happens

[2] See Moneglia and Raso (2014) for a complete list and definition of all Units, with the tags used in the information labeling.

when two of the following structures occur: a sequence of Bound Comments (COB) or a Multiple Comment (CMM) forming an illocutionary pattern (Panunzi & Mittman, 2014).

The first structure is formed by a chain of units called Bound Comments, with a homogeneous and weak illocutionary force. In this case, the conclusive prosodic profile at the end of each unit is usually not perceived,[3] and the Comments are so to speak "bound" by a continuative profile. Only the last unit of the chain brings a conclusive prosodic profile, so that it is conventionally signaled as a proper Comment unit (COM tag). The sequence of Bound Comments is functional to the realization of a unified "story", with the purpose of building an oral text. For this reason, the illocutionary strength and variation are reduced and usually limited to assertive types (Panunzi & Scarano, 2009).

Chains of COB units are typical of monologues and storytelling. They are built by progressive adjunctions, without a planned organization of the information. The sequence of Bound Comments forms a Stanza, which is conceived as another type of basic unit of Speech (and conclusive sequence), qualitatively different from the Utterance. Below are two examples of Stanza in which the progressive construction of an oral text is exemplified:[4]

(1a) *VAL: cioè /^{TMT} niente vabbè /^{PHA} si parte /^{COB} da Firenze /^{COB} eh /^{TMT} si fa i' check-in /^{COB} e si fa direttamente da [/1]^{EMP} da Firenze /^{COB} i' check-in /^{COB} eh /^{TMT} per New York //^{COM} (ifammn08_4)
*[*VAL: I mean / right well / we fly / from Florence / hm / we check-in / and directly from Florence / we check-in / hm / to New York //]*

(1b) *JOR: quer dizer /^{INT}= o brasileiro não tem a cultura /^{COB} de ler /^{COB} de &aum [/2]^{EMP} de melhorar a tecnologia /^{COB} e ter uma coisa que o satisfaça com melhor /^{SCA} &he /^{TMT} atendimento /^{SCA} a nível de audiovisual //^{COM} (bfammn06_39)
*[*JOR: I mean / Brazilian people do not have the culture / of reading / of &aum [/2] of upgrading (the) technology / and (of) having some stuff which satisfies them with a better / &he / service / in terms of audiovisual (technology) //]*

Multiple Comments (CMM) are complex information units composed by more than one Comment, each one characterized by an illocutionary force, held together by a prosodic pattern. For this reason, Multiple Comment structures are conceived as a unitary *illocutionary pattern* with a compositional structure. Each unit has its own characterization and can be, in most cases, pragmatically interpreted by itself.

CMMs are characterized by specific rhythmic and prosodic structures and their compositional structure reflects a meta-illocutionary model that need more than one information unit to be executed, producing a "rhetoric" effect (Cresti, 2000). The main models are: list, comparison, alternative, reinforcement relations, and necessary binding. In this paper we will focus on the list type.

[3] Nevertheless, there are doubtful cases in which the prosodic profile is not clearly perceived neither as terminal nor as non-terminal. We do not deal with this distinction in the paper, but further studies about continuity signals and conclusive prosodic profiles could investigate the importance of this matter.

[4] The following transcriptions of the speech are annotated with the LABLITA tagset (Cresti & Moneglia, 1997; Cresti & Moneglia, 2005; Cresti & Panunzi, 2013), which is a variant of CHAT format for speech transcription (MacWhinney, 1991). See Moneglia and Raso (2014) for abbreviations. Each sequence ends with a terminal break and is internally divided into prosodic units through non-terminal breaks. Double slash (//) is the standard sign used for terminal breaks, which characterize conclusive sequences neither interrupted (signaled with "+") nor intentionally suspended by the speaker (indicated with "…"); the question mark is used to demarcate a terminated sequence with a rising prosodic profile (interrogative, request utterances). Single slash (/) is used for non-terminal breaks. A double or single slash followed by a number, both contained in square brackets ("[/n]"), indicates retracting (i.e. false start) phenomena; *n* corresponds to the number of retracted words. Boundaries of false starts do not contribute to the informational patterning or to the semantic content of the Utterance.

The list type is a chain of two or more CMMs, usually with the same illocutionary force (e.g. assertions, suggestions, instructions, hypotheses, rhetorical questions, quotations). Here, CMM-units have specific durational and intonation patterns, which may provide the melodic and rhythmic features perceived by the listener. Additionally, the last CMM-unit of the list is defined by a conclusive prosodic profile. The following are two examples (2a, 2b) of a list in the form of a Multiple Comment:

(2a) *ROB: come l'esser cambiati /^{CMM} il bere /^{CMM} il mangiare //^{CMM} (ifammn16_38)
*[*ROB: like changing their clothes / drinks / food //]*

(2b) *EVN: tá o Aminas /^{CMM} Mauro e Filhos /^{CMM} Racing /^{CMM} Galáticos //^{CMM} (bfamcv01_183)
*[*EVN: (The teams in group A) are (the) Aminas / Mauro e Filhos / Racing / Galáticos //]*

In relation to the other types of Multiple Comments, they are usually composed by two CMMs. To briefly name their main features, in comparisons and alternatives the content of each Comment is usually complete and always semantically related to the other; the reinforcement, the most frequent type, is characterized by one semantically complete unit and one stereotyped expression in support of the previous one. Finally, the necessary binding bears two different illocutions linked by a logic relation such as causal or consecutive.

3 Samples and parameters

3.1 The Italian and the Brazilian Portuguese samples

Starting from a pilot study on a qualitative selection of Stanzas and CMM utterances, we have extended the analysis to two comparable sample collections. The recordings, as mentioned above, have been taken from the C-ORAL-ROM spoken corpus (Cresti & Moneglia, 2005; Panunzi & Gregori, 2012)[5] and the two Minicorpora of Italian and Brazilian Portuguese (from C-ORAL-BRASIL corpus; Raso & Mello, 2012). The illustrative cases reported above (1 and 2) have been selected from our samples.

After a general analysis of the COBs sets, we will focus on the comparison between COBs and CMMs of the List type. Our choice is intended to describe and differentiate chains of units (COBs and List-type CMM) that share similar features, such as illocutive monotony and a structure of subsequent adjuncts, of which only the last one bears a conclusive prosodic profile.

Utterances and stanzas have been selected from different conversations, dialogues, and monologues, both familiar and public, with high variability in terms of context and speakers (age, gender, education). The first criterion for selection was the audio quality, as we selected the ones with the greatest possible acoustic spectrogram clarity. We then selected speech turns without overlapping, high background noises or creaky voice in the areas of interest. A further selection was carried out on the base of stress position: we avoided the cases in which the last pre-boundary syllable was stressed. Moreover, we chose the parameters through which to conduct the analysis, as described below.

3.2 Parameters

[5] Freely available online at http://www.lablita.it/app/dbipic/.

In order to approach the issue of describing non-terminal breaks of COB and CMM units, we analyzed phenomena that illustrate the prosodic behaviors across the non-terminal boundaries, both left and right, regardless the information function of the unit after the break.[6]

Given that one of the functions of prosody is the demarcation of boundaries of tone units, we analyzed different instances of prosodic features with acoustic correlates such as f_0 shift, mean, variation rate and slope, and spectral emphasis (Cruttenden, 1997; Hirst & Di Cristo, 1998; Sorianello, 2006; Barbosa, 2019).

We conducted this research study implementing a Praat script specifically for this work. All the audio tracks were annotated with a TextGrid file composed by six tiers: conclusive sequence, Information Units, words, syllables, boundary (point tier), stress (point tier).

The analysis takes into account two Information Units: the one preceding and the one following the chosen non-terminal break (Unit 1 / Unit 2). The measurements concern the last syllable of Unit 1 and its last stressed syllable, and the first syllable of Unit 2 and its first stressed syllable. Clearly, it does happen that the stressed syllable is right adjacent to the boundary (at the end of Unit 1 and/or at the beginning of Unit 2). Thus, in these cases there is an overlapping of features in the same syllable position. Our choice was to avoid overlaps at the end of the Unit 1 position, but there are still cases of overlapping at the beginning of Unit 2.

More specifically, we measured what follows.

F_0 was measured in Hertz and semitones, to report shift, mean and variation rate (vr).

We smoothed f_0 using a filter bandwidth of 4 Hz. Pitch floor and ceiling of f_0 range are derived from Praat's analysis of the pitch using an adapted version of Hirst's procedure (Hirst, 2007).[7] This procedure aims to reduce the micromelodic effects on the f_0 measurement (mainly, the effects of segment articulation on f_0).

The f_0 shift indicates differences in pitch range between two adjacent intonation units, namely the differences in f_0 before and after the boundary. It was calculated as the difference between the means of the five f_0 points before and after the boundary.[8] Positive numbers correspond to upwards shifts (f_0 reset), whereas negative numbers correspond to downward shifts. To compare the results, the script provides also f_0 shift in the form of ratio related to the f_0 range of the conclusive sequence, i.e. the difference between the maximum and minimum f_0 values over the whole sequence (chain of COBs or patterns of CMMs).

The script calculates different values of f_0 mean: the f_0 mean for the syllables adjacent to the boundary and the difference between them; the f_0 mean of the stressed syllables before and after the boundary and the difference between them. Once the script gives the mean values in Hertz, it is possible to transform them in semitones using the f_0 mean of the conclusive sequence as reference – that is a variable reference; thus, values (smt) near 0 means that the f_0 value (Hz) is near the f_0 mean of the conclusive sequence.

We analyzed the f_0 slope before and after the boundary. This measurement gives two kinds of information: i. the direction of the f_0 movement (positive or negative values); ii. the variation rate (in absolute values), which indicates the speed (smt/s) of the movement regardless of the direction. In this paper, the term "f_0 variation rate" will always refer to the absolute value of f_0 slope. When discussing

[6] Our methodological choice was not to distinguish non-terminal breaks based on the next unit since the study is intended as a first step in the formalization of prosodic breaks. Further analysis may integrate such distinctions, thus taking into account possible prosodic cues determined by characteristics of specific units.

[7] To take better account of data, instead of using 1.5 * the third quantile, we have rather used 1.65 as a multiplier.

[8] The first f_0 point of each side was considered the first voiced frame (time step = 0.01) with respect to the boundary point in time. After identifying this point, the others four points were those which follow (on the right side) or precede (on the left said) the first one.

cross-boundary differences, negative values of f_0 variation rate will indicate that the pre-boundary syllable has greater values than the post-boundary syllable.

Spectral emphasis is an acoustic feature reflecting the relative intensity in the higher frequency bands. We calculated its value[9] before and after the boundary in the stressed syllables and the unstressed syllables adjacent to the boundary. Spectral emphasis is taken here as a measure related to intensity and it overcomes the variations due to the circumstances of recording.

The choice of these acoustic measurements aims to investigate and different COBs and List-type CMMs internal breaks from an acoustic point of view and underlining possible connections between different prosodic cues.

The following Sections 4 and 5 contain the analysis derived from the study of the parameters[10].

As the previous description points out, the analysis shows the prosodic behavior in a narrow and then in a wider analysis window. The first consists of the syllable just before the boundary and the adjacent just after; it allows a circumscribed analysis on the cross-boundary interval. The wider section consists of the stressed syllable before the boundary and the stressed syllable after the boundary. As already mentioned, it may occur that the syllables adjacent to the boundary coincide with the stressed ones.

4 Analysis of Italian data

The Italian sample consists of 22 non-terminal COB units and 14 non-terminal List units. The COB sample was selected so that: i. both Unit 1 and Unit 2 are COBs; ii. the stress position never corresponds to the immediately before/after- boundary position. As a result, in the COB couples it is possible to clearly differentiate the prosodic behavior of the syllables adjacent to the boundary from the prosodic behavior of the stressed syllables. Figure 1 shows an example of Italian COB.

[9] Based on the procedure described in Barbosa and Madureira (2015, pp.184-186). High frequencies were considered those above 400 Hz.

[10] The script is also measuring, when present, the durations of pauses, but we are not using this parameter in the following analysis.

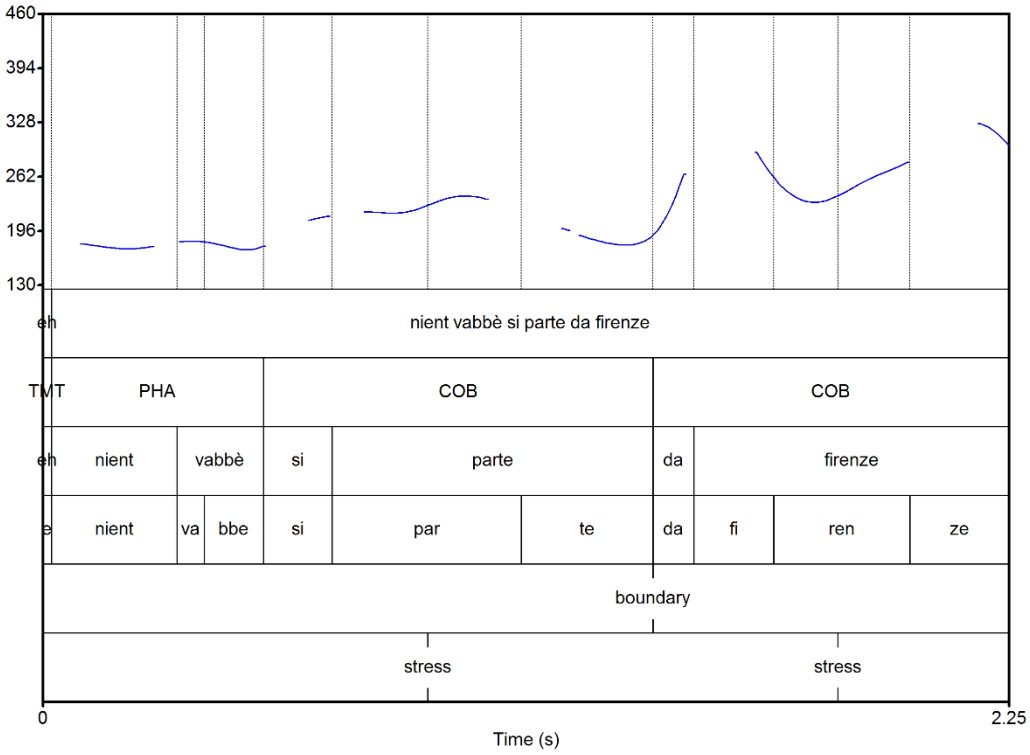

Figure 1: Example from IT_COB

On the other hand, Lists (IT_LIST) were selected so that the stress position is never overlapping with the immediately before-boundary position, while the first syllable of Unit 2 is not controlled in terms of stress position (with 5 cases of stressed syllables at the beginning of the CMM unit). Figure 2 shows an example of Italian List.

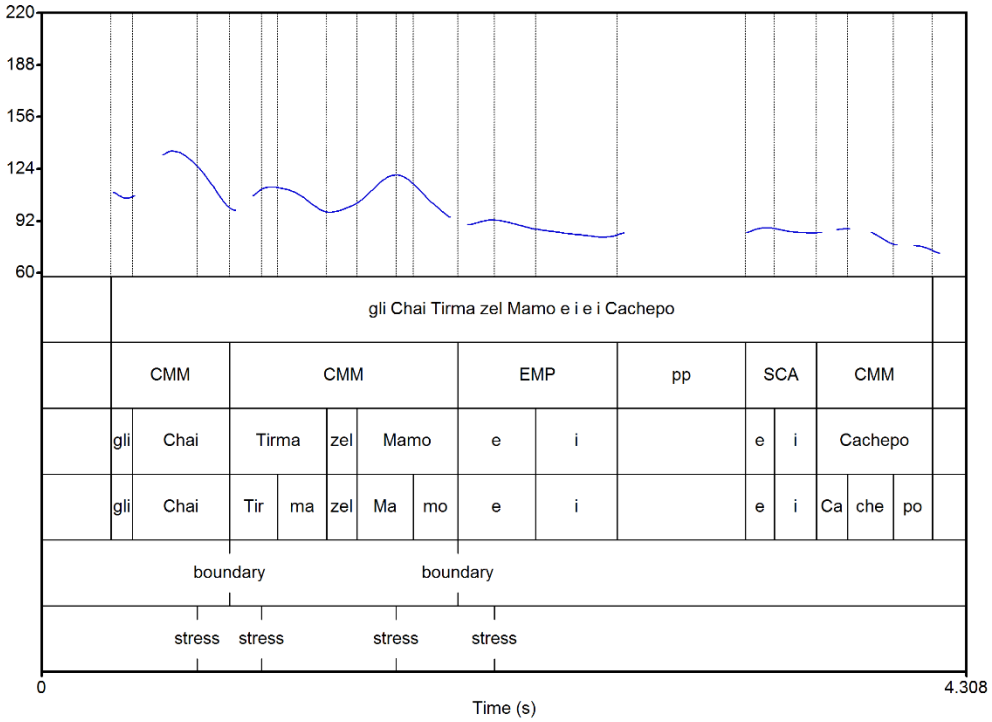

Figure 2: Example from IT_LIST

We analyzed the features of the last stressed and the last unstressed syllable before the boundary, comparing f_0 mean (ratio), f_0 slope, $f_0$vr and SE. Afterwards, we analyzed the cross-boundary contrasts (last pre-boundary unstressed syllable vs. first syllable of the following unit), comparing f_0 shift (ratio) and the differences in f_0 mean, $f_0$vr and SE. Besides, we carried out the same analysis, except for f_0 shift measurements, with the last pre-boundary and the first post-boundary stressed syllables.

At large, both IT_COB and IT_LIST seem to display an f_0 mean near the reference, rising f_0 movement on the last pre-boundary stressed syllable, and falling f_0 movements on the last pre-boundary unstressed syllable. As to the cross-boundary contrasts, we found higher f_0 mean, higher SE, and higher $f_0$vr in the pre-boundary unstressed syllables when compared to the first post-boundary syllable. On the other hand, we found lower f_0 mean and higher $f_0$vr in the pre-boundary stressed syllable compared to the post-boundary one. The f_0 shift (~ -0.09) [11] is small in both groups. See Table 1.

Table 1: Differences between IT_COB and IT_LIST, comparing the characteristics of the pre-boundary syllables

	Measure	Last stressed syllable Median (SD) Statistics	Last unstressed syllable Median (SD) Statistics
Difference between COB and List	f_0 slope	COB: 6.96 (19.2), List: 11.36 (24.5), W = 140, p = 0.67	COB: -12.64 (16.2), List: -10.29 (23.8), W = 144, p = 0.761
	$f_0$vr	COB: 16.23 (9.7), List: 19.07 (13.0), W = 128, p = 0.41	COB: 15.86 (10.1), List: 13.93 (17.8), W = 156, p = 0.96
	f_0 mean	COB: 1.31 (2.5), List: 1.37 (2.5), W = 143, p = 0.74	COB: 0.37 (2.1), List: 1.01 (2.0), W = 109, p = 0.15
	SE	COB: 4.30 (5.1), List: 4.73 (3.8), W = 148, p = 0.86	COB: 5.30 (3.4), List: 4.78 (4.5), W = 148.5, p = 0.87

To evaluate the differences between IT_COB and IT_LIST, we first assessed whether their variances are significantly different. The results of a series of Fligner-Killeen tests for each measure analyzed indicate that the variance is not different between COB and List ($p > 0.1$ for all cases).[12] Thus, we cannot say that COBs are more or less variable than Lists. To compare the medians, we also conducted a series of Mann-Whitney-Wilcoxon tests for each measure analyzed. The only significant difference between COBs and Lists is that, in COBs, the last stressed syllable before the boundary is less intense (lower SE) than the first stressed syllable after the boundary (COB: 1.64, List: -2.06, W = 227.5, $p < 0.02$). Concerning the other tests, they do not indicate any significant result ($p > 0.1$ for all cases). See Tables 1 and 2; see also Figure 3.

Table 2: Cross-boundary differences in IT_COB and IT_LIST

	Measure	Between stressed syllables Median (SD) Statistics	Last unstressed syllable *vs.* first post-boundary syllable Median (SD)

[11] Approximately, the mean of COB and List medians.

[12] Levene's tests also support this interpretation.

			Statistics
Difference between COB and List (cross-boundary)	$f_0$vr	COB: -4.88 (10.6), List: -9.41 (13.5), W = 165, p = 0.56	COB: -9.34 (16.1), List: -1.77 (15.3), W = 139, p = 0.80
	SE	COB: 1.64 (4.70), List: -2.06 (3.36), W = 227.5, p < 0.02*	COB: -1.07 (3.1), List: -0.70 (3.2), W = 122, p = 0.31
	f_0 mean	COB: -0.23 (0.30), List: -0.19 (0.31), W = 149, p = 0.89	COB: -0.15 (0.2), List: -0.11 (0.2), W = 137, p = 0.60
	f_0 shift		COB: -0.07 (0.2), List: -0.10 (0.2), W = 170, p = 0.62

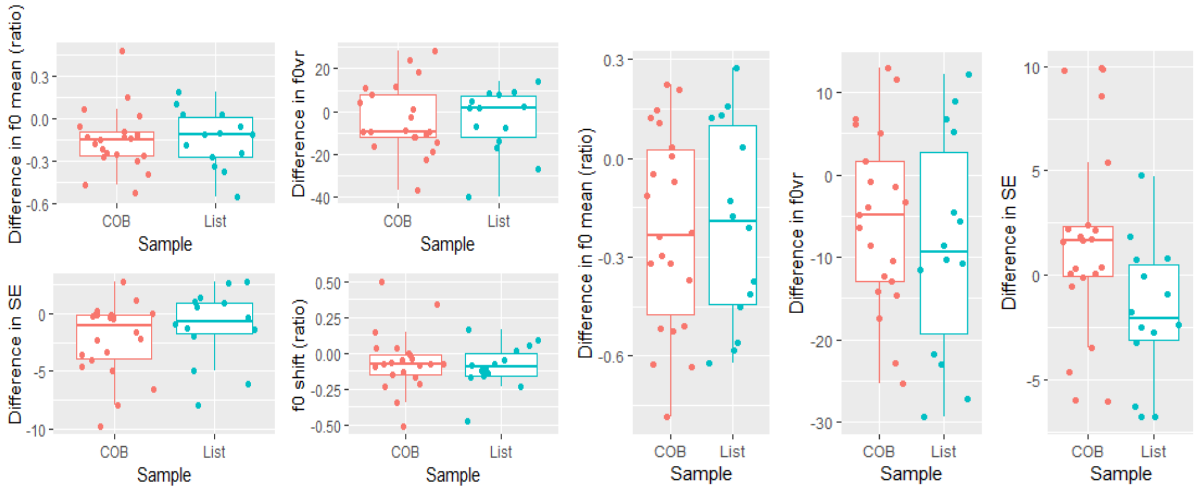

Figure 3: Boxplot displaying the cross-boundary difference in IT_COB and IT_LIST, comparing: a) the last pre-boundary unstressed syllable and the first post-boundary syllable; and b) the last stressed syllables before and the first stressed syllable after the boundary.

Finally, we split the data into two groups concerning the f_0 slope in both last stressed and last unstressed syllables before the boundary: rising f_0 contour ($f_0$vr > 0, min = 2.76 smt/s) and falling f_0 contour ($f_0$vr < 0, min= |1.03| smt/s). In this case, we only found that COBs tend to have more rising f_0 in the last stressed syllables ($\chi^2 = 2.91$, df = 1, p = 0.09), but this difference is not significant. Lists do not portray a specific pattern either ($\chi^2 = 0.29$, df = 1, p = 0.59). As would be expected, there is no significant difference between COBs and Lists ($\chi^2 = 0.10$, df = 1, p = 0.75). On the other hand, we found a high proportion of falling f_0 in the last unstressed syllable of both COBs ($\chi^2 = 8.91$, df = 1, p < 0.01) and Lists ($\chi^2 = 4.57$, df = 1, p= 0.03); but this proportion is not different between COBs and Lists ($\chi^2 = 1.15e\text{-}31$, df = 1, p > 0.9), see Table 3.

Table 3: Number of occurrences of each type of F_0 contour in IT_COB and IT_LIST

LAST STRESSED SYLLABLES	COB	List
Falling	7	6
Rising	15	8
LAST UNSTRESSED SYLLABLES	COB	List
Falling	18	11
Rising	4	3

5 Analysis of Brazilian Portuguese data

The Brazilian Portuguese sample consists of 21 non-terminal COB units (BP_COB) and 19 non-terminal List units (BP_LIST). In both groups, the last pre-boundary syllable is not stressed. However, the first syllable of the following unit is not controlled in terms of stress position. Figures 4 and 5 show an example of Brazilian COB and one of List.

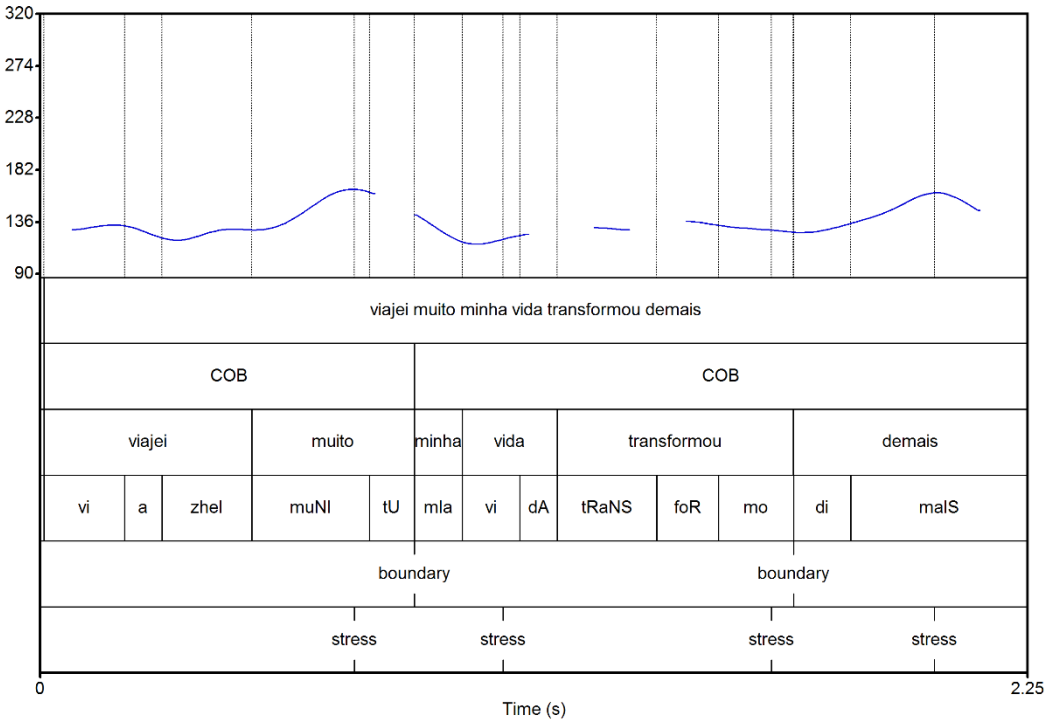

Figure 4: Example from BP_COB

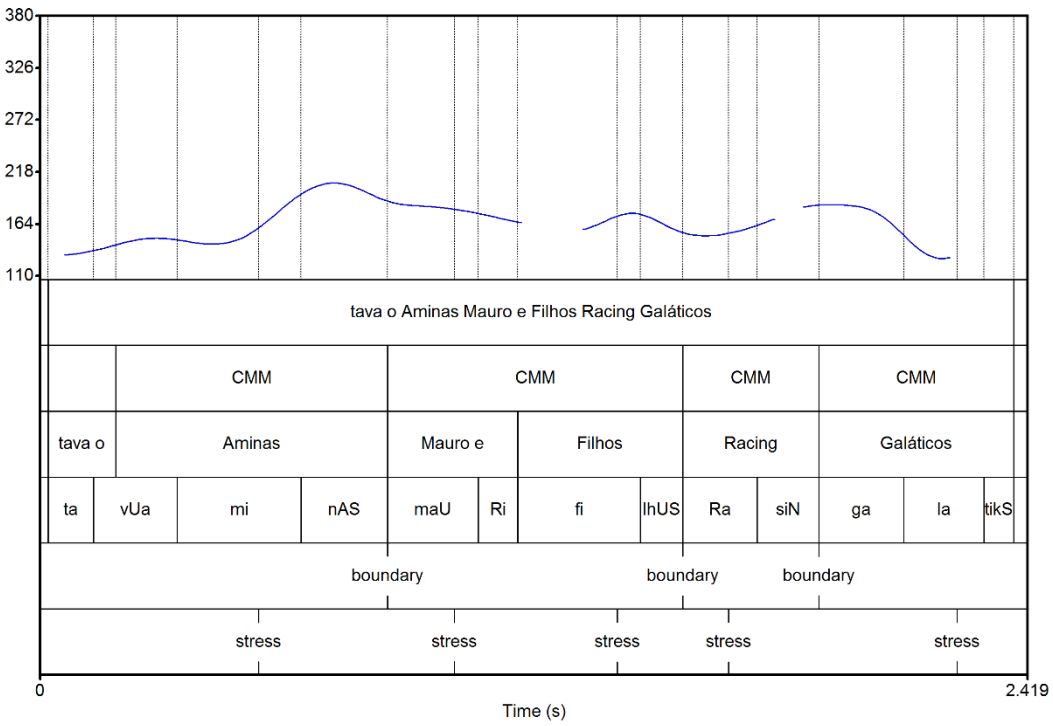

Figure 5: Example from BP_LIST

We analyzed the features of both groups in the same way as we did for the Italian samples.

In general, both BP_COB and BP_LIST seem to display, on average, an f_0 mean near the reference, rising f_0 movements on the last pre-boundary stressed syllables, and falling f_0 movements on the last pre-boundary unstressed syllables. Concerning the cross-boundary contrasts, the groups display higher f_0 mean, lower SE, and slightly higher or similar $f_0$vr in the pre-boundary unstressed syllable in comparison to the first post-boundary syllable. They also display higher $f_0$vr and lower f_0 mean in the pre-boundary stressed syllable than in the post-boundary one. Finally, it is worth mentioning that there is only a small f_0 shift (~ -0.05) in both groups. See Tables 4 and 5.

Table 4: Differences between BP_COB and BP_LIST, comparing the characteristics of the pre-boundary syllables

	Measure	Last stressed syllable Median (SD) Statistics	Last unstressed syllable Median (SD) Statistics
Difference between COBs and Lists	f_0 slope	COB: 13.17 (23.4), List: 17.27 (11.4), W = 184, p = 0.69	COB: -5.83 (19.5), List: -6.71 (12.3), W = 184, p = 0.69
	$f_0$vr	COB: 18.99 (12.0) , List: 17.27 (6.3), W = 248, p = 0.20	COB: 11.62 (14.2), List: 8.92 (6.0), W = 232, p = 0.39
	f_0 mean	COB: 0.24 (1.8), List: -0.68 (1.6), W = 256, p = 0.13	COB: 1.40 (3.5), List: 1.65 (2.3), W = 208, p = 0.83
	SE	COB: 3.68 (4.50), List: 7.46 (5.53), W = 170, p = 0.43	COB: 2.73 (2.48), List: 4.34 (2.80), W = 170, p = 0.44

As for the Italian dataset, we first assessed whether the variance in COB and List data are significantly different (Fligner-Killeen tests). The results indicate that there is no significant difference between COBs and Lists in any of the measures mentioned above ($p > 0.1$ for all cases). Thus, in BP data, COBs are not distinctly variable compared to Lists. We also conducted a series of Mann-Whitney-Wilcoxon tests for each measure analyzed. According to the results, there is no significant difference between COBs and Lists ($p > 0.1$ for all cases), thus, we cannot say that there are any distinguishable features among those analyzed here. See Tables 4 and 5; see also Figure 6.

Table 5: Cross-boundary differences in BP_COB and BP_LIST

	Measure	Between stressed syllables Median (SD) Statistics	Last unstressed syllable *vs*. first post-boundary syllable Median (SD) Statistics
Difference between COB and Lists (cross-boundary)	$f_0$vr	COB: -9.22 (15.6), List: -1.62 (10.3), W = 139, p = 0.11	COB: -1.57 (14.0) , List: 0.30 (7.8), W = 221, p = 0.47
	SE	COB: -2.67 (4.9), List: 1.34 (7.2), W = 162, p = 0.32	COB: 0.85 (2.8), List: 4.46 (6.1), W = 142, p = 0.12
	f_0 mean	COB: 0.01 (0.23), List: 0.04 (0.22), W = 149, p = 0.18	COB: -0.10 (0.3), List: -0.11 (0.3), W = 198, p = 0.98
	f_0 shift		COB: -0.05 (0.3), List: -0.05 (0.2), W = 221, p = 0.57

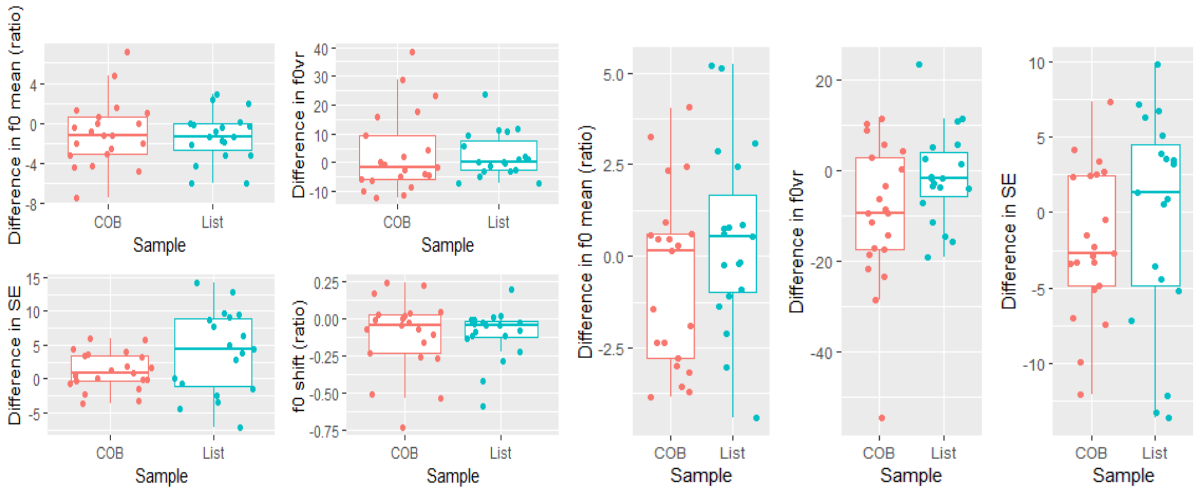

Figure 6: Boxplot displaying the cross-boundary difference in BP_COB and BP_LIST, comparing: a) the last pre-boundary unstressed syllable and the first post-boundary syllable; and b) the last stressed syllables before and the first stressed syllable after the boundary.

Finally, we compared the groups in terms of f_0 slope: rising f_0 contour ($f_0vr > 0$, min = 4.54 smt/s) and falling f_0 contour ($f_0vr < 0$, min = |1.81| smt/s). When analyzing the two groups separately, we found a higher proportion of rising f_0 in the last stressed syllables of both COBs ($\chi^2 = 8.05$, df = 1, $p < 0.01$) and Lists ($\chi^2 = 8.90$, df = 1, $p < 0.01$). Conversely, comparing the two groups (COBs and Lists), there is no significant difference between them ($\chi^2 = 4.51e\text{-}31$, df = 1, $p > 0.90$). This result confirms that COBs and Lists often have a rising contour on the stressed syllables. On the other hand, we found a high proportion of falling f_0 in the last pre-boundary unstressed syllable of COBs ($\chi^2 = 2.33$, df = 1, p= 0.13) and Lists ($\chi^2 = 4.26$, df = 1, p = 0.04), but only in the latter case (List) the difference is significant. Comparing these two groups directly, the proportion is not different between them ($\chi^2 = 0.02$, df = 1, p= 0.89). See Table 6.

Table 6: Number of occurrences of each type of F0 contour in BP_COB and BP_LIST

LAST STRESSED SYLLABLES	COB	List
Falling	4	3
Rising	17	16
LAST UNSTRESSED SYLLABLES	COB	List
Falling	14	14
Rising	7	5

6 Discussion and conclusions

This analysis was carried out in total on 76 non-terminal breaks of COBs and CMMs to investigate the prosodic features of two languages, Italian and Brazilian Portuguese, in a formal and acoustic perspective.

A relevant feature of the study is the implementation of an automatic analysis, thereby enabling accurate and detailed results. Furthermore, it deepens the observation on the stressed syllable before the boundary and it takes advantage of a statistical tests.

To sum up, the two languages in question seem to reveal quite similar patterns.

The analysis of Italian and Brazilian Portuguese COB and LIST samples shows a pronounced tendency to a falling profile in the unstressed syllable before the boundary, and a rising profile in the

stressed syllables before the boundary.

Concerning cross-boundary contrasts, the analysis showed the following trends:

i. Higher f_0 mean in the pre-boundary unstressed syllable compared to the first post-boundary syllable.
ii. Higher f_0 mean in IT and lower f0 mean in BP, comparing pre-boundary stressed syllable to the post-boundary one – this is true for both COB and List.
iii. Higher $f_0$vr in the pre-boundary region (stressed and unstressed) compared to the post-boundary one (except for BP_LIST, where the difference between the pre-boundary unstressed syllable and the first post-boundary syllables is, on average, 0.30);
iv. Small f_0 shift;
v. Lower pre-boundary than post-boundary values of SE in the unstressed syllables of Brazilian sample, whereas the Italian sample has the opposite behavior.

Thus, in general, for what concerns the comparison between Stanzas and List-CMM Utterances, the result is that they have a very similar behavior. Furthermore, they are both quite variable, and we cannot state that one group is more variable than the other.

Previous results on the pilot Italian corpus showed that one of the distinctive features of Bound Comment is non-terminal breaks with a low number of f_0 shift above 18% of the f_0 range of the whole sequence and with the absence of a rising profile before the COB break. The new study deepens the observations about f_0 contour, underlining that the continuity signal between COBs is not linked to a rising profile in the very last syllable before the boundary, but to a rising profile in the last stressed syllable of the tone unit.

Briefly, our data indicate that, by using only f_0 and SE measures, it is difficult to distinguish between COBs and Lists. Thus, rhythmic metrics and duration measures may shed light on these issues. An ongoing analysis of Brazilian data suggests that mean syllabic duration is greater in Lists than in COBs. As it would be expected, in both groups the durations of the last stressed syllables are longer than the mean syllabic duration. However, in BP_COB, 81% of these syllables are more than 1.5 times longer than the mean syllabic duration, whereas in List, only 47% meet this condition. Therefore, it seems that COBs commonly display longer last stressed syllables than Lists. This must be analyzed better using normalized duration. Besides, COB has generally more syllables than Lists in average, and this should be taken into account when analyzing rhythm.

In any case, it seems that stress position, unit length and intonation patterns interact in a more complex way to build the patterns perceived by listeners.[13] Therefore, studies with more controlled data are necessary to make these interactions clearer. For example, it could be that trends in prosodic behavior affect the whole conclusive sequence of COBs or Lists more than individual tone units. For this reason, future research could investigate the trends across the units, for example a possible trend to differently lower or rise f_0 and intensity towards the end.

The analysis of durations, such as final lengthening before the boundary or initial rush in the unit after the boundary, will be implemented in this work using normalized durations, disposable for the Brazilian Portuguese language but not yet for the Italian language. The work is now in progress following the Brazilian model (Barbosa, 2006) already used for French, British English and German among others.

REFERENCES

[13] As we have already pointed out, our dataset is formed by conclusive sequences of spontaneous spoken language that are already tagged following L-AcT Theory; that means that COB and List units have been evaluated as perceptively distinguished and different by mother tongue listeners of the two languages trained for tagging.

1. Austin JL. *How to Do Things with Words*. Oxford: Oxford University Press, 1962.
2. Barbosa, PA. *Incursões em torno do ritmo da fala*. São Paulo: Pontes, 2006
3. Barbosa PA, Madureira S. *Manual de fonética experimental: aplicações a dados do português*. São Paulo: Cortez, 2015.
4. Barbosa PA. *Prosódia*. São Paulo: Parábola Editorial, 2019.
5. Boersma P, Weenink D. *Praat: Doing phonetics by computer*. Software. Available from: http://www.praat.org/. 2005.
6. Cresti E, Moneglia M. *C-ORAL-ROM. Integrated reference corpora for spoken romance languages*. Amsterdam: John Benjamins Publishing Company, 2005.
7. Cresti E, Moneglia M. *L'intonazione e i criteri di trascrizione del parlato*. In: Bortolini U, Pizzuto E (Orgs.). *Il progetto CHILDES Italia*. Volume 2. Pisa: Il Cerro, 1997. pp. 57-90.
8. Cresti E. *Corpus di italiano parlato*. Firenze: Accademia della Crusca, 2000.
9. Cresti E, Panunzi A. *Introduzione ai corpora dell'italiano*. Bologna: Il Mulino, 2013.
10. Cruttenden A. *Intonation*. Second edition. Cambridge: Cambridge University Press, 1997.
11. Hirst D, Di Cristo A (Orgs.). *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press, 1998.
12. Hirst D. A Praat plugin for MOMEL and INTSINT with improved algorithms for modelling and coding intonation. In: *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrüken, Germany, 2007. pp. 1233-1236.
13. MacWhinney B. *The CHILDES Project. Tools for analysing talk*. Hillsdale: Lawrence Erlbaum Associates, 1991.
14. Moneglia M. Spoken corpora and pragmatics. *Revista Brasileira de Linguística Aplicada*, v. 2, n. 11, pp. 479-519, 2001.
15. Moneglia M, Raso T. Notes on the Language into Act Theory. In: Raso T, Mello H (Orgs.). *Spoken corpora and linguistic studies*. Amsterdam: John Benjamins Publishing Company, 2014. pp. 468-495.
16. Panunzi A, Gregori L. DB-IPIC. An XML database for the representation of information structure in spoken language. In: Panunzi A, Raso T, Mello H (Orgs.). *Pragmatics and Prosody. Illocution, Modality, Attitude, Information Patterning and Speech Annotation*. Firenze: Firenze University Press, 2012. pp. 133-150.
17. Panunzi A, Mittman M. The IPIC resource and a cross-linguistic analysis of information structure in Italian and Brazilian Portuguese. In: Raso T, Mello H (Orgs.). *Spoken corpora and linguistic studies*. Amsterdam: John Benjamins Publishing Company, 2014. pp. 129-151.
18. Panunzi A, Scarano A. Parlato spontaneo e testo: Analisi del racconto di vita. In: Amenta L, Paternostro G (Orgs.). *I parlanti e le loro storie.* Competenze linguistiche, strategie comunicative, livelli di analisi. Atti Del Convegno Carini-Valderice. Palermo: Centro di studi filologici e linguistici siciliani, 2009. pp. 121-132.
19. Panunzi A, Saccone V. Complex Illocutive Units in L-AcT: An Analysis of Non-Terminal Prosodic Breaks of Bound and Multiple Comments / Unidades Ilocucionárias Complexas na L-AcT: uma análise de quebras prosódicas não-terminais em Comentários Ligados e Comentários Múltiplos. *Revista de Estudos da Linguagem*, v. 4, n. 26, pp. 1647-1674, 2018.
20. Raso T, Mello H, (Orgs.). *C-ORAL-BRASIL I*: Corpus de referência do Português Brasileiro falado informal. Belo Horizonte: Editora UFMG, 2012.
21. Sorianello P. *Prosodia. Modelli e ricerca empirica*. Roma: Carocci, 2006.