Effects of speech rhythm on spoken syntax
A corpus-based study on Brazilian Portuguese and Italian

Giulia Bossaglia
Universidade Federal de Minas Gerais

This paper presents a preliminary study on the syntax/prosody interface in spoken Brazilian Portuguese and Italian, within the Language into Act Theory theoretical framework (L-AcT, Cresti 2000). According to this theory, spoken syntax has to be studied taking into consideration the way the speaker combines information along one or more information units, since each of them is understood as a syntactic/semantic island (Cresti 2014). Then, *linearized syntax* defines proper dependency relationships, performed within the same information unit, while *patterned constructions*, syntactic structures performed across more than one information unit, are to be analyzed considering the pragmatic functions of the information units involved (Cresti 2014). According to a non-discrete conception of speech rhythm (Barbosa 2000, 2006), BP can be defined as a more stress-timed language with respect to IT, and its tonal unit can host more phonological syllables than IT's, for BP's rhythm leads to strong re-syllabification phenomena. Based on data from the C-ORAL-BRASIL (Raso & Mello 2012) and Italian C-ORAL-ROM (Cresti & Moneglia 2005) spontaneous speech corpora, this paper aims at verifying if such prosodic features partake in the way the two languages display preference for *linearized syntax* or *patterned constructions*.

Keywords: spoken corpora, Brazilian Portuguese, Italian, spoken syntax, Dynamic Model of speech rhythm

1. Spontaneous speech and the Language into Act Theory: the role of prosody

The study presented in this paper is based on the Language into Act Theory (henceforth: L-AcT; see Cresti 2000; Moneglia & Raso 2014; Raso 2012), an inductive theory on spoken language built through decades of empirical work on the *LABLITA* (*Linguistic Lab of the Italian Department of Florence University*, Italy) spoken corpora (Moneglia 2011). Based on L-AcT, the C-ORAL-ROM

CHIMERA. Romance Corpora and Linguistic Studies 3.2 (2016), 265-285. ISSN 2386-2629

spoken corpora for Italian, Spanish, French and European Portuguese (Cresti & Moneglia 2005), and, later, the C-ORAL-BRASIL corpus were built (Raso & Mello 2012).

In the analysis of speech, L-AcT is innovative in taking into significant consideration its prosodic dimension, understood as the fundamental one for several reasons which will be illustrated further on in this section. The emphasis on prosody could sound kind of obvious in the analysis of speech, but it's not: for a long time, traditional studies on spoken language have based their analysis on speech *transcriptions*, which means on *written* texts and, hence, on linguistic categories derived from and adequate to written, not spoken, language.

L-AcT identifies the unity of reference for speech in the utterance, understood as the shortest linguistic unit prosodically and pragmatically interpretable as autonomous, i.e. as a speech act. Following Austin's Speech Acts Theory (Austin 1962), L-AcT assumes that the speech act is made up by three simultaneous acts:

- locutionary act, corresponding to the transmitted linguistic content;
- illocutionary act, corresponing to the action (illocution) the speaker performs through the speech act;
- perlocutionary act, corrisponding to the affective impulse that leads the speaker to perform the speech act[1].

Prosody, then, is the means by which the illocution is conveyed, as example (1) below (adapted from Raso 2012: 96) shows:

(1) *SIL: copos // copos de Urano / que tem aí // 1-5
*KAT: copos de quê //
*SIL: Urano //
*KAT: Urano //
*SIL: é // Urano // Urano // (bfamdl04, 100-107[2])
'*SIL: glasses // glasses of Uranus / that are in there //
*KAT: glasses of what //
*SIL: Uranus //

[1] In Austin's view, on the contrary, the perlocutionary act corresponds to the speaker's intended goal.

[2] Starred abbreviations stand for the speakers' names in the C-ORAL corpora; in *bfamdl* acronym *b* refers to Brazilian corpus (*i* is used to refer to the Italian one), *fam* to familiar-private context (vs. *pub* for public), *dl* to dialogue (vs. *mn* monologue, *cv* conversation). The acronym is followed by the number of the recording session, and then the number of the utterance.

*KAT: Uranus //
*SIL: yeah // Uranus // Uranus //'

It is clear that in (1) above to the same morphosyntactic content (*Urano*) correspond four different illocutions, namely, a confirmation, an expression of disbelief, and two conclusions (performed with different attitudes[3]), distinguished by their prosodic profile.

Besides conveying the illocutionary value of the utterance, according to L-AcT prosody plays an important role in the segmentation of the speech flow, through two different kinds of prosodic breaks: terminal breaks (i.e. perceived as having a terminal intonation) signal the utterance boundaries ("//" in the C-ORAL-ROM and C-ORAL-BRASIL transcriptions, see (1) above), while non-terminal breaks segment tonal units within the utterance (marked with "/").

A prosodic break does not necessarily correspond to a pause: according to L-AcT, it is defined as a "perceptively relevant prosodic variation in the speech continuum such that it causes the parsing of the continuum into discrete prosodic units" (Moneglia & Cresti 2006: 92).

Following the IPO – *Institute of Perception Research of the University of Eindhoven* model of speech perception, L-AcT assumes that the f_0 (fundamental frequency) exhibits a series of movements within the speech flow, among which some are voluntary (although unconscious) and perceptually relevant, thus carrying informational values[4] (t'Hart, Collier & Cohen 1990; Firenzuoli 2003).

The prosodic units can, therefore, convey Information Units (IUs), defined by specific prosodic profiles[5], functions and distribution within the utterance. According to the L-AcT's definition, then, the utterance needs *at least* one IU which carry the illocutionary force in order to be pragmatically interpreted as autonomous. Such a IU is the Comment, and in (1) above most of the utterances are formed by this single unit (*simple* utterances: *Urano* //), but utterances can also be formed by the Comment plus one or more IUs (*compound* utterances: see *copos* / *que tem aí* //).

---

[3] By attitude L-AcT means the *Modus on Actum*, i.e. the way the speaker performs the illocution (Bally 1932). The same illocution, actually, can be performed with a polite, arrogant, ironic, bored, etc. attitude (see Rocha & Raso 2016 on the distinction between illocution and attitude).

[4] Other parameters such as intensity, duration, and syllabic alignment of f_0 are taken into account for the prosodic analysis of speech.

[5] According to the IPO model, there are three basic typologies of prosodic profile for IUs: *root*, obligatory and non-recursive; *prefix*, optional and in some cases recursive; and *suffix*, optional, non-recursive and always following a *root* (t'Hart, Collier & Cohen 1990: 78-79).

IUs are divided into two main groups, according to their macro-functions: *textual* IUs form the text of the utterance, i.e. convey its semantic and syntactic content, or serve in order to interpret the utterance or part of it (they are: Topic, Comment, Multiple Comment, Bound Comment, Appendix of Topic/Comment, Parenthesis, Locutive Introducer); *dialogic* IUs serve in order to regulate the interaction (they are: Phatic, Allocutive, Conative, Incipit, Expressive, Discourse Conector)[6].

Units without information value are also found in speech and described by the L-AcT, as for example Time Taking units (i.e. filled pauses) and Empty units (i.e. retractings, or last units of interrupted utterances). For this paper's purposes, it is particularly worth mentioning, within this group, the Scanning units (see example (5) in section 3.2.2). These units don't bear any prosodic nucleus nor informational value and are used in order to scan the linguistic content of an IU (usually, Topic, Comment or Locutive Introducer). Actually, Scanning units correspond to different tonal parts of the same IU when the locutive content exceeds its capacity, or when it is effectively scanned by the speaker, be it for stylistic or performance reasons.

2. Premises to the study: L-AcT and spoken syntax

Before analyzing the syntax-prosody interface, two premises are necessary: the first one concerns how is spoken syntax conceived within L-AcT, and it will be presented in the following subsection.

2.1 L-AcT and spoken syntax: *linearized syntax* vs. *patterned constructions*

Since a significant amount of the linguistic content in speech is made up by strongly fragmented structures, problems arise when traditional syntactic criteria are applied to the analysis of spoken syntax.

Thus, the sentence as a unit of reference for speech has been easily discarded by scholars, among which most choose the clause or clause complexes instead (Chafe 1984; Quirk *et al.* 1985; Halliday 1985; Voghera 1992; Miller & Weinert 1998). Nonetheless, the clause as well is not a completely satisfactory unit of reference for spontaneous speech, and not only much of spoken language is made up by loose phrases, interjections or fragments, but also prosodic and

[6] For a more detailed description of the IUs according to L-AcT, see Raso (2012), Moneglia & Raso (2014).

syntactic boundaries often do not overlap (cf. the idea that intonation "is not married to syntax", Bolinger 1984: 413; see also Mithun 2009). A common approach to the "broken" surface of spoken language has been that of postulating an underlying syntactically well-formed structure, a high-level grammar in which syntactic boundaries are very clear, but then transformed and "chopped" in speech due to extralinguistic, interactional, and even performance factors.

Against this idea, Heath (1985, *apud* Miller & Weinert 1998) on Ngandi aborigenal language assumes that such a fragmented configuration of syntax can indeed be a high-level feature of languages, like Ngandi, very different from English or most European languages. Nonetheless, Miller & Weinert (1998: 356) observe that such a discrepancy between the grammars of Ngandi and English appears strongly reduced if *spoken* English is analyzed. In this way, they strongly affirm the importance of, first, avoiding the bias of a long tradition of studies on written language and, then, studying spoken language(s) as having its specific grammar and structures (see also Miller & Fernandez-Vest 2006).

Firmly against the idea of speech as a "corrupted" surface form of languages' grammars, Sornicola (1981) claims that the syntactically fragmented nature of spoken syntax is indeed speech-specific and intentionally produced in such broken-up structures by the speakers, while the extralinguistic context and the interlocutors' world knowledge make up for the strongly reduced syntactic linkage.

From similar assumptions, and from the fact that no syntactic unit could consistently cope with such a fragmentation of speech, L-AcT chooses a non-syntactic unit of reference for speech, i.e. the utterance, which is prosodically and pragmatially defined, independently from its morphosyntactic content.

L-AcT conception of spoken syntax is, instead, closely related to the way the information is packaged by the speaker: within the utterance, each single IU is conceived as a syntactic and semantic island, hence hosting true dependency relationships, while the relationship between different IUs is not syntactic, but related to their informational functions, hence, pragmatic.

Therefore, L-AcT distinguishes between *linearized syntax*, i.e. proper instances of syntactic dependency, performed within the same IU, and *patterned constructions*, i.e. syntactic structures performed across more than one IU, to be analyzed, therefore, through pragmatic/informational criteria (Cresti 2014). Examples are provided in (2) and (3):

(2) *FLA: eu acho que no saco do lixo nós já passamos //=COM= 6
‘I think that we have already passed through the trash bag (aisle) //
(bfamdl01, 506)

(3) *LUZ: porque eu acho que no mesmo concurso /=TOP= cê não pode 7
fazer duas //=COM= (bfamdl03, 181)
‘because I think that in the same competition / you can’t run for two
(positions) //’

Examples (2) and (3) are instances of complement clauses performed in *linearized* and *patterned* configuration, respectively. In (2), main predicate and complement are performed within the same IU and, therefore, a true dependency relationship is detectable. But in (3), main clause and complement are performed by the Topic/Comment information pattern: in this case, the matrix clause in Topic is not to be analyzed as the head of regency, because a relationship of pragmatic *aboutness* exists between the two IUs. Therefore, in (3) the matrix clause in Topic is serving the pragmatic function of defining the domain of application of the illocution conveyed by the Comment, in which, by the way, the complement clause is performed[7].

2.2 Linearized syntax vs. patterned constructions in spoken Brazilian Portuguese and Italian

The second premise to this paper consists in a previous study on the syntax of explicit complement clauses in spoken Brazilian Portuguese and Italian (Bossaglia 2014), whose results have inspired the research presented here in its first steps.

In Bossaglia (2014) the realization of explicit complement clauses in spoken Brazilian Portuguese and Italian is analyzed, based on two comparable subcorpora of the C-ORAL-BRASIL and Italian C-ORAL-ROM (available online in the DB-IPIC – *DataBase for Information Patterning Interlinguistic Comparison* platform[8]; see Panunzi & Mittmann 2014 and section 3.2.1 for details).

---

[7] From a pragmatic perspective, then, it is the complement, and not the matrix clause, that performs the most important function (i.e. carrying the illocution), while the matrix clause serves as background information. It goes beyond the aims of this paper to go in-depth into such a pragmatic organization of spoken syntax. For further and more detailed discussion see Cresti (2014) and Bossaglia (2014, 2015).

[8] http://lablita.dit.unifi.it/app/dbipic/

That survey on complement clauses aimed at mapping such syntactic constructions in both languages and analyzing them from a contrastive perspective with respect to the types and frequencies of complementizers, the information patterns involved, and the *ratio* of linearized syntax and patterned constructions.

BP and IT displayed a very similar realization of complement clauses in the DB-IPIC corpora: the same complementizers were detected (in order of frequency, PB/IT: *que*/*che* 'that', *se*/*se* 'whether', plus various indirect question complementizers in both languages) and showed comparable frequencies; the same information patterns appeared to be used in PB and IT (mainly Topic/Comment and chains of Multiple Comments or Bound Comments; IT data displayed also a few textually more articulated patterns than PB); linearized configurations were strongly predominant on patterned constructions in both languages.

In view of such a great, and expected, homogeneity between BP and IT in the realization of complement clauses, the only interesting deviation was found in the *ratio* of linearized syntax and patterned constructions: 83,9% *vs.* 16,1% in BP, 67,3% *vs.* 32,7% in IT (Bossaglia 2014: 48).

Our hypothesis correlates such a discrepancy to a prosodic difference between PB and IT, namely their rhythmic typology. In this paper, its preliminary experimental verification is presented.

3. Syntax-prosody interface in spoken BP and IT

3.1. Speech rhythm typologies: BP and IT

It is widely accepted that rhythm is a language-specific, psychologically relevant category. Many experimental studies focused on perception have observed that adult and newborn individuals are highly sensitive to languages' rhythm, and capable of clearly discriminating between rhythmic typologies of different and unfamiliar languages (see, for example, Bertoncini & Mehler 1981; Mehler *et al.*1996; Nazzi, Bertoncini & Mehler 1998; Ramus, Nespor & Mehler 1999).

Traditionally, natural languages have been assigned a *stress-timed* or *syllable-timed* rhythmic typology according to whether they present isochrony, i.e. periodicity, of temporal duration between, respectively, stressed syllables or syllables (Pike 1945; Abercrombie 1967). Such a dichotomy was early proven to be problematic, since natural languages, in general, don't manifest only one rhythmic typology depending on the context of the communication (Roach 1982; Dauer 1983; Bertinetto 1989). Besides, focusing exclusively on the temporal dimension for the definition of rhythmic typologies means disregarding many of the variables involved in speech rhythm.

More recently, different approaches have managed to better account for speech rhythm by analyzing it through a more complex perspective (among others: Saltzman & Munhall 1989; Cummins & Port 1998; Cummins 2002; Barbosa 2000, 2004, 2006). The inspiration for such a new perspective in the study of speech rhythm can be found in studies on human and animal motricity (e.g. Kelso, Tuller & Harris 1983; Schöner & Kelso 1988; Kelso 1995) that observed the existence of synergic systems, in which different physical components interact both one with another (functional coupling) and with the environment. According to the Dynamical Model of speech rhythm proposed by Barbosa (2000, 2004, 2006, 2007) the temporal dimension is still taken into account, of course, in the production of speech rhythm, but within a self-organized synergy whose components are (Barbosa 2006: 10-11; see also Barbosa 200, 2001, 2004, 2007):

a. interaction of high-level linguistic information with an accentual oscillator;
b. funcional coupling of two abstract oscillators, accentual and syllabic;
c. coupling of the syllabic oscillator with gestural level[9].

In this way, a less dichotomic approach to speech rhythm is proposed, since it is postulated that both the accentual and the syllabic oscillators interact in speech rhythm production in any natural language. This means that the same mechanism is responsible for both the stress-timing and syllable-timing tendencies according to a greater force of, respectively, the accentual or the syllabic oscillator in their functional coupling (Barbosa 2001: 33; Barbosa 2000, 2004, 2006).

It follows that within this perspective speech rhythm is a non-discrete category: there aren't any *tout court* stress-timed nor syllable-timed languages (as it was claimed by Abercrombie 1967: 67), and the rhythmic typology of any natural language can be described both as more stress- or syllable-timed depending on the languages to whom we compare it with[10].

---

[9] The notion of *gestural* level comes from the Articulatory Phonology paradigm (Browman & Goldstein 1988, 1989, 1993): the course of articulatory gestures is made up by the creation and distension of (various degrees of) constrictions in different parts and by different physical components of such a synergy.

[10] This holds on the condition that the comparison is made on the basis of comparable speech rates.

The hypothesis presented here springs from the assumption that Brazilian Portuguese displays a stronger stress-timing tendency with respect to Italian, and, conversely, Italian displays a stronger syllable-timing tendency with respect to Brazilian Portuguese (Barbosa 2000: 392)[11].

The consequence of such a difference in BP and IT speech rhythm is that BP's tonal unit (see section 1) can in principle host more phonetic and phonological material than IT's, since a stronger entrainment of the accentual oscillator over the syllabic one results in a sort of "temporal compression", by which less time is needed in order to pronounce a higher number of syllables (Barbosa 2001: 32-33).

Therefore, the higher *ratio* of linearized complement clauses in BP mentioned in the previous section could be due to BP stress-timing tendency with respect to IT. In the next sections, the first methodological steps in order to verify such a hypothesis are presented.

3.2 Methodology

3.2.1 *The DB-IPIC corpora*

The data presented in this study come from the two comparable minicorpora of C-ORAL-BRASIL and Italian C-ORAL-ROM, available online at the DB-IPIC platform[12]. The *Data-Base of Information Patterning Interlinguistic Comparison* was built as a rich query interface allowing contrastive studies based on the C-ORAL corpora. In this online platform, access to the Italian Informal C-ORAL-ROM corpus, a minicorpus extracted from it, and a minicorpus extracted from the Informal C-ORAL-BRASIL is possible. All DB-IPIC corpora are provided with the transcriptions of the recordings implemented with prosodic break annotation, while the audio file of each utterance is available thanks to the text-to-speech alignment. Besides the informational annotation (i.e. annotation of the IUs), Italian corpora are provided with PoS tagging as well (at the moment, PoS tagging for BP is available only in the C-ORAL-BRASIL DVD, where the whole corpus is available).

The two BP and IT DB-IPIC minicorpora are made up by 20 recorded sessions, and were built in order to be representative of their reference corpora and comparable one with another. As regards their representativeness, the proportions between monological and dialogical interactions within them is the same as within the C-ORAL-BRASIL and the Italian C-ORAL-ROM corpora,

---

[11] But, for example, BP manifests a stronger syllable-timing tendency with respect to Thai or British English (Barbosa 2000).

[12] http://lablita.dit.unifi.it/app/dbipic

i.e. one third of monologues vs. two thirds of dialogues and conversations (interactions with more than two active participants). The same holds for the proportions between Family/Private (15 recorded situations over 20 in BP, 16 over 20 in IT) and Public communicative contexts. As in their reference corpora, a high degree of diaphasic variation is detectable in the minicorpora, in order to best represent the richness of spontaneous speech situations (for more details, see Panunzi & Mittmann 2014; Panunzi & Gregori 2012).

The comparability of the architecture and size (in terms of number of words and prosodically Terminated Sequences, i.e. Utterances and Stanzas[13]) of DB-IPIC minicorpora is described in Table 1 below, adapted from Panunzi & Mittmann (2014: 140):

Table 1. Size of BP and IT DB-IPIC minicorpora

		Brazilian Portuguese			**Italian**		
Comm. context	**Int. types**	**Sessions**	**words**	**TSs**	**Sessions**	**words**	**TSs**
Family/Private	Mn	6	8635	856	6	8750	1086
	Dl	5	8360	1877	5	9306	1771
	Cv	4	6421	1407	3	5152	1283
Public	Mn	1	1616	143	2	2927	265
	Dl	2	3011	584	2	3129	555
	Cv	2	3422	645	2	8136	703
Total		**20**	**31465**	**5512**	**20**	**37355**	**5663**

Although the two minicorpora differ in the number of words (31,465 vs. 37,355), it is important to notice that in terms of number of units of reference they are perfectly comparable (5,512 vs. 5,663). Such a comparability represents a key factor for the purposes of this study.

3.2.2 *Data extraction and analysis*

In order to test our hypothesis that BP tonal unit can host more phonological syllables than IT's, and that this would be one of the factors involved in the higher rate of linearized syntax than in IT, we followed three methodological steps.

[13] The Stanza (Cresti 2010) is, like the Utterance, a prosodically terminated unit within the speech flow, but it displays a peculiar information pattern, made by a sequence of (Bound) Comments "produced by progressive adjunctions which follow the flow of thought" (Moneglia & Raso 2014: 490).

First, we extracted from the two minicorpora all the simple utterances (formed by the COM unit only) and simple scanned utterances (formed by the COM unit, performed through more than one tone unit) that contain a verbal predication[14], as it is exemplified in (4) and (5), respectively:

8 (4) *ART: qualcuno dice che ci si mette meno //=COM= (ifamdl04, 262)
‘someone says that it takes less time //’

9 (5) *LIA: ma se ti dico /=SCA= che lei /=SCA= tutte le mattine /=SCA= con uno straccetto hhh //=COM= (ifamcv01, 748)
‘but if I tell you / that she / every morning / with a rag //’

Since they are formed by a single IU, choosing simple and simple scanned utterances for our analysis was meant to use the smallest unit of reference for speech.

In both groups of utterances (simple and simple with SCA), we searched for complex syntactic structures (i.e. with a verbal head, like complex sentences, cleft sentences and so on) in both languages. Our expectation was to find a higher rate of complex syntactic structures per simple utterance (i.e. linearized complex syntactic structures[15]) in BP than in IT, in light of the the highest rate of linearization already found by Bossaglia (2014) for complement clauses.

Finally, the amount of phonological syllables per simple utterance (understood as the smallest information/prosodic unit) was calculated, expecting, once again, to find a higher rate in BP than in IT. Simple scanned utterances were excluded in this last step: although they represent the smallest unit of reference from an informational point of view, they do not from a prosodic perspective, because they are made up by more than one tonal units.

[14] This means that we excluded from our sample not only all the utterances without a verbal predication, but also utterances in which verbal predicates were used with pragmatic/interactional functions, e.g. utterances like BP: *é* //=COM= ‘yes’ (lit. ‘(it) is’) or *entendeu* //=COM= ‘did you get it?’, or IT: *vai* //=COM= ‘yes’ (lit. ‘go!’), which are very frequent in spontaneous speech.

[15] We recall here that SCA units don’t have autonomous information value, and are to be considered as different tonal parts of the same IU (i.e. Scanning unit/Comment is not an information pattern as, for example, Topic/Comment).

4. Results and data analysis

The utterances gathered according to the mentioned criteria were 1,369 for BP (1,141 simple + 228 simple with SCA) and 1,413 for IT (998 simple + 415 simple with SCA), as it is shown in Table 2 below:

Table 2. Simple and simple scanned utterances in DB-IPIC BP and IT minicorpora

Interaction typology	BP	Simple	Simple + SCA	Total	IT	Simple	Simple + SCA	Total
cv		382	85	467		440	83	523
dl		599	84	683		418	81	499
mn		160	59	219		140	251	391
Total		1141	228	1369		998	415	1413

At the outset, it seems that in IT there's much more use of simple scanned (29.4% on the total amount of utterances) utterances than in BP (16.7%), as it is shown in Figure 1 below:

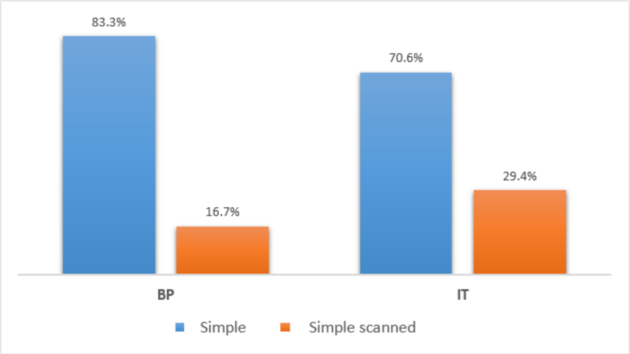

Figure 1. Simple and simple scanned utterances in BP and IT IPIC minicorpora

Such a difference turns out to be more interesting when observed within each interactional typology. In facts, the mentioned greater amount of simple utterances over simple scanned utterances is detected within conversations, dialogues and monologues in BP, but only within conversations and dialogues in IT. Actually, in monological interactions IT displays a much greater amount of simple scanned utterances (64.2%), as it can be observed in Figure 2:

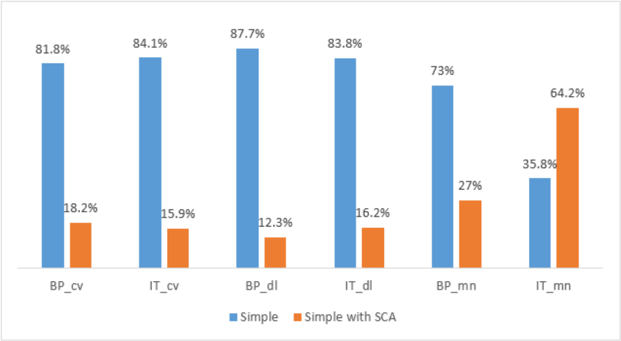

Figure 2. BP and IT simple and simple scanned utterances per interactional typology

It is worth noticing that in monological interactions BP makes a higher use of simple scanned utterances too. This fact has to do with the interactional typology itself, that is characterized by a lesser degree of actionality (there’s oly one active participant in monologues) with respect to dialogues and conversations. The counterpart of such a reduced actionality is a higher textual complexity in monologues, which usually are narratives or argumentative texts. Therefore, the fact that in IT data such a textual complexity is mainly performed through simple *scanned* utterances could already be, at the outset, an indicator of the fact that IT tonal unit presents a lesser capacity than BP one.

4.1 Complex syntactic structures per simple utterance in BP and IT

The search for complex syntactic structures within the two data samples was meant to verify if textually structured locutive content could easily be performed within a single tonal unit in both languages. The expected higher rate of complex syntactic structures per simple and simple scanned utterance in BP was verified in our data, since 16.7% (229) of BP utterances, over the total amount of them, host complex syntactic structures, while only 10.7% (151) do in IT:

Table 3. Complex syntactic structures per simple and simple scanned utterance in BP and IT

Interaction typology	**BP**	**Simple**	**Simple + SCA**	**Total**	**IT**	**Simple**	**Simple + SCA**	**Total**
Cv		50	24	74		23	27	50
Dl		99	21	120		15	10	25
mn		19	16	35		10	66	76
Total		168	61	**229**		48	103	**151**

BP data display a higher rate of complex syntactic structures per utterance not only globally, but also for each interactional typology (with the exception of monologues), as it is shown in Figure 3:

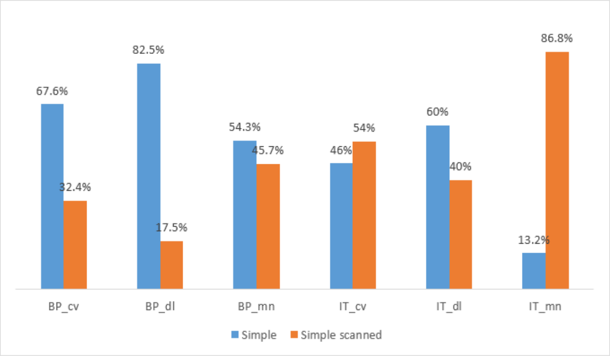

Figure 3. Simple and simple scanned utterances hosting complex syntactic structures in BP and IT

As Figure 3 shows, in monological interactions IT displays a strong preference for complex syntactic structures to be performed through simple scanned utterances, while in conversations and dialogues the percentages are quite similar (23 simple vs. 27 simple scanned utterances in cv, 15 simple vs. 10 simple scanned utterances in dl: the discrepancy in these interactional typology is not really significant).

If one looks at the data considering the relative frequency of utterances hosting complex syntactic structures with respect to the total amount of simple and simple scanned utterances separately, other observations are possible:

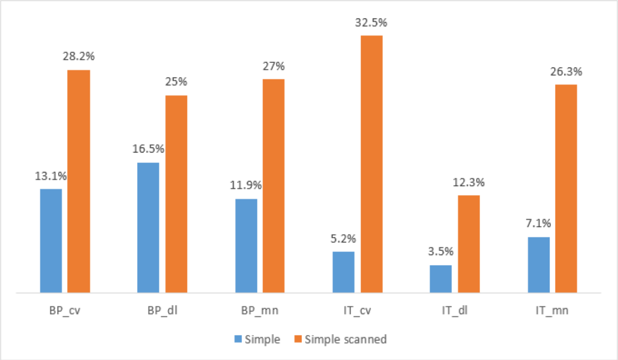

Figure 4. Simple and simple scanned utterances hosting complex sentences in BP and IT

As it is shown in the figure above, both languages display a higher relative percentage of utterances hosting complex syntactic structure within the group of scanned utterances (percentages are calculated on the total amount of simple and simple scanned utterances, respectively).

Scanned utterances' capacity is increased by the addition of one or more tonal units, hence they are at the outset capable of hosting a bigger amount of linguistic content than non-scanned ones. Therefore, a higher relative number of utterances with complex syntactic structures was expectable within the scanned utterances group in both languages.

However, it is clear that within BP simple utterances the amount of non-scanned utterances hosting complex syntactic structures is pretty higher than in IT: 13.2% vs. 5.2%, 16.5% vs. 3.5%, 11.9% vs. 7.1% in conversations, dialogues and monologues, respectively.

Within the sample of utterances hosting complex syntactic structures, the two languages display opposite trends in what concerns the privileged means of conveying such structures, i.e. non-scanned utterances for BP (168), and scanned utterances for IT (103):

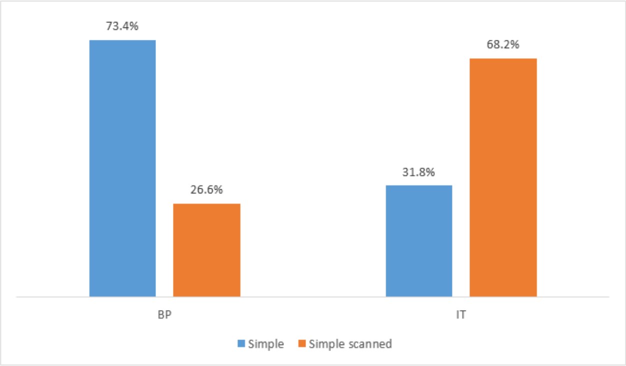

Figure 5. Typology of utterances hosting complex syntactic structures in BP and IT

However, one must have in mind that in monological interactions IT data display a much broader use of scanned utterances in conveying complex syntactic structures than in the other interactional typologies. Once again, monologues turn out to be the privileged *locus* for textual complexity.

Finally, our data showed that BP simple utterances display a greater amount of complex syntactic structures not only at a global level, but also within non-scanned and scanned utterances, as it is shown in Figure 6:

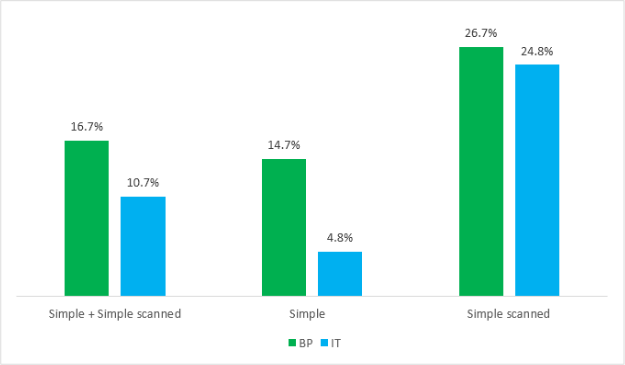

Figure 6. Complex syntactic structures in BP and IT simple (scanned) utterances

It is noteworthy that the only context in which IT displays a presence of complex syntactic structures comparable to that of BP is the group of simple scanned utterances, i.e. simple utterances whose capacity has been increased by the addition of more tonal units. All the facts illustrated in this section, therefore, seem to be in line with our hypothesis.

4.2 Phonological syllables per simple utterance in BP and IT

As we mentioned, only non-scanned utterances were used in order to calculate the phonological capacity of BP and IT simple utterances. As a first confirmation of our hypothesis, BP displayed a higher maximal number of phonological syllables per utterance than IT: 30 vs. 23, as examples (6) and (7) shown below[16].

(6) *BAL: ela não conseguiu carregar a pilha por causa que isso aqui tava em duzentos-e-vinte //=COM= (bfamdl02, 37) 10
‘she couldn’t charge the battery because this thing was set at 220 volts’

[16] Phonological syllables were counted manually. As a methodological choice, aphetisms (e.g. BP *tá* for *está* “(he/she/it) is”) were counted as whole forms (i.e. including the lost syllable in the calculation as well), with the exception of the second-person singular personal pronoun *cê* “you” (whole form *você*), which is completely grammaticalized in spoken BP (see Ferrari 2015 for details).

11 (7) *PAO: < è il ventitreesimo anniversario >[17] dell' assassinio di Aldo Moro //=COM= (ipubmn01, 117)
‘It’s the 23rd anniversary of Aldo Moro’s murder’

Besides, in each interaction typology BP displayed not only a higher maximal number of phonological syllables per utterance, but also higher mean, median and mode numbers, as it is shown in Table 4:

Table 4. Phonological syllables per utterance in BP and IT

Int. typ.	BP	Mean	Median	Mode	Max.	IT	Mean	Median	Mode	Max.
Cv		8.07	8	5	29		6.06	6	6	22
Dl		12.47	12	11	30		6.57	6	5	20
Mn		9.34	8	7	25		5.5	8	7	23

This first empirical test showed that BP tonal unit seems to be capable of hosting a higher number of phonological syllables than IT.

5. Conclusions

In this preliminary study on the prosody-syntax interface in spoken BP and IT we tried to verify the hypothesis that the difference in speech rhythm between the two languages could have consequences in the way their spoken syntax is structured. We assumed that the stronger preference for linearization BP displayed in the codification of complement clauses with respect to IT (Bossaglia 2014) could be related to the more stress-timed nature of BP rhythm: this is because a more stress-timed rhythm allows the tonal unit to host a higher number of phonological material than a more syllable-timed one.

The search for the rate of complex syntactic structures per simple (and simple scanned) utterance within the two minicorpora led to satisfactory results, since we could verify that BP actually presents a higher rate than IT. This means that within the domain of linearized syntax (i.e. a single IU, the COM) BP seems to be much more capable of performing complex syntactic structures than IT. One more verification of this fact was that IT displays a comparable ‘ability’ only in scanned utterances (see Figure 6 above), i.e. utterances whose capacity is increased by the addition of one or more tonal units. So, we can admit that BP is

[17] ‘<’ and ‘>’ symbols are used in order to mark speech overlap in the transcription.

characterized by a higher predisposition for linearized syntax than IT, probably due to its rhythmic characteristics.

Such correlation between prosodic features and syntax was confirmed also by our calculation of the number of phonological syllables the two languages present in their non-scanned simple utterances. As it was expected, BP appears to have higher mean, median and mode numbers in each interactional typology (with the only exceptions of lower mode number in conversations, and equal median and mode numbers in monologues), beside higher maximal number of phonological syllables per utterance.

Being these first results so promising, we intend to advance this study by searching for the *ratio* between phonological and phonetic syllables (viz. V-V units, see Barbosa 2000, 2006) in BP and IT simple utterances, which could represent a more reliable measure for their tonal units' capacity.

Acknowledgments

This research was possible thanks to FAPEMIG – Fundação de Amparo à Pesquisa do Estado de Minas Gerais financial support (post-doctoral grant PMPD III Program, process n. 22568).

References

Abercrombie, D. 1967. *Elements of general phonetics*, Vol. 203. Edinburgh: Edinburgh University Press.

Austin, J. L. 1962. *How to do things with words*. Oxford: Oxford University Press.

Bally, C. 1932. *Linguistique générale et linguisticque française*. Paris: Leroux.

Barbosa, P. A. 2000. Syllable-timing in Brazilian Portuguese: Uma crítica a Roy Major. *DELTA* 16(2): 369-402.

Barbosa, P. A. 2001. É Possível Integrar o Discreto e o Contínuo em um Modelo de Produção do Ritmo da Fala? *Cadernos de Estudos Lingüísticos* 40: 29-38.

Barbosa, P. A. 2004. Elementos para uma tipologia do ritmo (lingüístico) da fala à luz de um modelo de osciladores acoplados. *InCognito - Cadernos Românicos em Ciências Cognitivas* 2(1): 31-58.

Barbosa, P. A. 2006. *Incursões em torno do ritmo da fala*. Campinas: Pontes Editora.

Barbosa, P. A. 2007. Análise e modelamento dinâmicos da prosódia do português brasileiro. *Revista de Estudos da Linguagem* 15: 75-96.

Bertinetto, P. M. 1989. Reflections on the dichotomy "stress"- vs. "syllable-timing". *Revue de Phonétique Appliquée* 91-93: 99-130.

Bertoncini, J. & Mehler, J. 1981. Syllables as units in infant speech perception. *Infant Behavior and Development* 4(March): 247-260.

Bolinger, D. 1984. Intonational signals of subordination. In C. Brugman & M. Macaulay (eds), *Proceedings of the10th annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 401-413.

Bossaglia, G. 2014. Interface entre sintaxe e articulação informacional na fala espontânea: uma comparação baseada em corpus entre português e italiano. *Caligrama* 19(2): 35-60.

Bossaglia, G. 2015. Pragmatic orientation of syntax in spontaneous speech: A corpus-based comparison between Brazilian Portuguese and Italian. *CHIMERA* 2: 1-34.

Browman, C. P. & Goldstein, L. 1988. Some Notes on Syllable Structure in Articulatory Phonology. *Haskins Laboratories Status Report on Speech Research* 93/94: 85-102.

Browman, C. P. & Goldstein, L. 1989. Articulatory gestures as phonological units. *Phonology* 6: 201-251.

Browman, C. P. & Goldstein, L. 1993. Dynamics and Articulatory Phonology. *Haskins Laboratories Status Report on Speech Research* 113: 51-62.

Chafe, W. 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: The University of Chicago Press.

Cresti, E. 2000. *Corpus di italiano parlato: Introduzione*, Vol. 1. Firenze: Accademia della Crusca.

Cresti, E. 2010. La stanza: Un'unità di costruzione testuale del parlato. In *Atti del Congresso della Società di Linguistica e Filologia Italiana, SILFI 2008, Basel*. Firenze: Cesati, 713-732.

Cresti, E. 2014. Syntactic properties of spontaneous speech in the Language into Act Theory: Data on Italian complement and relative clauses. In T. Raso & H. Mello (eds), *Spoken Corpora and Linguistic Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 365-410.

Cresti, E. & Moneglia, M. (eds) 2005. *C-ORAL-ROM: integrated reference corpora for spoken romance languages.* Amsterdam/Philadelphia: John Benjamins.

Cummins, F. 2002. On synchronous speech. *Acoustic Research Letters Online* 3(1): 7-11.

Cummins, F. & Port, R. 1998. Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26(2): 145-171.

Dauer, R. M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51-62.

Ferrari, L. A. 2015. Aspectos prosódicos e sintáticos dos pronomes clíticos em português do Brasil e em verbáculo florentino. PhD diss., Universidade Federal de Minas Gerais.

Firenzuoli, V. 2003. Verso un nuovo approccio allo studio dell'intonazione a partire dai corpora di parlato: esempi di profili intonativi di valore illocutivo dell'italiano. In N. Maraschio & T. Poggi Salani (eds), *Atti del XXXIV Congresso internazionale degli studi della SLI. Firenze, 19-21 ottobre 2000*. Roma: Bulzoni, 531-546.

Halliday, M. A. K. 1985. *An introduction to functional grammar*. London: Arnold.

t'Hart, J.T., Collier, R. & Cohen, A. 1990. *A perceptual study on intonation: an experimental approach to speech melody*. Cambridge: Cambridge University Press.

Heath, J. 1985. Discourse in the field: clause structure in Ngandi. In J. Nichols, J. & A. C. Woodbury (eds), *Grammar inside and outside the clause: Some approaches to theory from the field*. Cambridge: Cambridge University Press, 89-110.

Kelso, J. A. S. 1995. *Dynamic Patterns: the self-organization of brain and behavior*. Cambridge: The MIT Press.

Kelso, J. A. S., Tuller, B. & Harris, K. S. 1983. A “dynamic pattern” perspective on the control and coordination of movement. In M. MacNeilage (ed.), *The production of speech*. New York: Springer-Verlag, 137-173.

Mehler, J., Dupoux, E., Nazzi, T. & Dehaene-Lambertz, G. 1996. Coping with linguistic diversity: The infant’s viewpoint. In J. L. Morgan & K. Demuth (eds), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates, 101-116.

Miller, J. & Fernandez-Vest, M. M. J. 2006. Spoken and written language. In G. Bernini & M. Schwartz (eds), *Pragmatic organization of discourse in the languages of Europe*. Berlin/New York: Mouton de Gruyter, 9-65.

Miller, J. E. & Weinert, R. 1998. *Spontaneous Spoken Language*. Oxford-New York: Clarendon Press.

Mithun, M. 2009. Re(e)volving complexity: Adding intonation. In T. Givón & M. Shibatani (eds), *Syntactic Complexity*. Amsterdam/Philadelphia: John Benjamins, 53-80.

Moneglia, M. 2011. Spoken corpora and pragmatics. *Revista Brasileira de Linguística Aplicada* 11(2): 479-519.

Moneglia, M. & Cresti, E. 2006. C-ORAL-ROM: Prosodic boundaries for spontaneous speech analysis. In Y. Kawaguchi, S. Zaima & T. Takagaki (eds), *Spoken Language Corpus and Linguistics Informatics*. Amsterdam/Philadelphia: John Benjamins, 89-114.

Moneglia, M. & Raso, T. 2014. Notes on Language into Act Theory (L-AcT). In T. Raso & H. Mello (eds), *Spoken Corpora and Linguistic Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 468-495.

Nazzi, T., Bertoncini, J. & Mehler, J. 1998. Language Discrimination by Newborns: Towards an Understanding of the Role of Rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24(3): 756-766.

Panunzi, A. & Gregori, L. 2012. DB-IPIC: an XML database for the representation of information structure in spoken language. In H. Mello, A. Panunzi & T. Raso (eds), *Pragmatics and Prosody: Illocution, Modality, Attitude, Information Patterning and Speech Annotation*. Firenze: Firenze University Press, 133-150.

Panunzi, A. & Mittmann, M. M. 2014. The IPIC resource and a cross-linguistic analysis of information structure. In T. Raso & H. Mello (eds), *Spoken Corpora and Linguistic Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 129-151.

Pike, K. L. 1945. *The intonation of American English*. Ann Arbor: University of Michigan Press.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. 1985. *A comprehensive grammar of the English language*. Harlow: Longman.

Ramus, F., Nespor, M. & Mehler, J. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265-292.

Raso, T. 2012. O C-ORAL-BRASIL e a Teoria da Língua em Ato. In T. Raso & H. Mello (eds), *C-ORAL-BRASIL I. Corpus de referência do português brasileiro falado informal*. Belo Horizonte: UFMG, 91-123.

Raso, T. & Mello, H. (eds) 2012. *C-ORAL-BRASIL I. Corpus de referência do português brasileiro falado informal*. Belo Horizonte: UFMG.

Roach, P. 1982. On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages. In D. Crystal (ed.), *Linguistic controversies. Essays in linguistic theory and practice in honour of F. R. Palmer*. London: Edward Arnold, 73-79.

Rocha, B. & Raso, T. 2016. The interaction between illocution and attitude and its consequences for the empirical study of illocutions. In C. Bardel & A. De Meo (eds), *Parler les langues romanes / Parlare le lingue romanze / Hablar las lenguas romances / Falando línguas românicas*. Napoli: Università di Napoli L'Orientale, 69-87.

Saltzman, E. L. & Munhall, K. G. 1989. A dynamical approach to gestural patterning in speech production. *Ecological psychology* 1(4): 333-382.

Schöner, G. & Kelso, J. A. 1988. Dynamic pattern generation in behavioral and neural systems. *Science* 239(4847): 1513-1520.

Sornicola, R. 1981. *Sul parlato*. Bologna: Il Mulino.

Voghera, M. 1992. *Sintassi e Intonazione nell'Italiano Parlato*. Bologna: Il Mulino.