

SHORT INFORMATION UNITS: A CORPUS-BASED PROSODIC STUDY ON THE LEXEME "ASSIM" IN BRAZILIAN PORTUGUESE

RASO, Tommaso^{1*}

SANTOS, Saulo²

¹Federal University of Minas Gerais

²Federal University of Minas Gerais

Abstract: *Our main goal is to show that short information units (one phonological word immediately preceded and followed by a prosodic boundary, at least one of which with non-terminal value) can be classified only on the basis of their formal prosodic characteristics. That is to say that lexicon and syntax may vary with respect to the information function, while what formally marks the informational function is the regularity of the prosodic profile of the lexical item. We also maintain that an information unit corresponds to an intonation unit (except in one specific circumstance). We use just one lexeme for the analysis, the Brazilian Portuguese ASSIM, and extract all the occurrences where this lexeme is found in a dedicated prosodic unit in the C-ORAL-BRASIL corpus. According to our analysis, this lexeme can fulfill at least five different information units, i.e. it can fulfill at least five different linguistic functions, recognizable by their prosodic regularities.*

Keywords: information structure; short information units; prosody; corpus.

Resumo: *Nosso objetivo principal é mostrar que as unidades informacionais curtas (uma palavra fonológica imediatamente precedida e seguida por fronteira prosódica, das quais pelo menos uma não terminal) podem ser classificadas somente com base em suas características prosódicas. Isso significa que o léxico e a sintaxe podem variar com relação à função informacional, enquanto o que marca formalmente a função informacional é a regularidade do perfil prosódico do item lexical. Também defendemos que a unidade informacional corresponde a uma unidade entonacional (com exceção de uma circunstância específica). Para a análise usamos apenas um lexema, ASSIM do português brasileiro, e extraímos todas as ocorrências em que esse lexema aparece em uma unidade prosódica dedicada dentro do corpus C-ORAL-BRASIL. Segundo a nossa análise, o lexema pode preencher pelo menos cinco funções informacionais diferentes, ou seja, pode preencher pelo menos cinco funções linguísticas diferentes, reconhecíveis pelo perfil prosódico.*

Palavras chave: estrutura informacional; unidades informacionais curtas; prosódia; corpus.

*Corresponding author: tommaso.raso@gmail.com

1 Introduction

This research aims at taking a first step towards an in-depth understanding of the relation between functions and prosodic forms regarding what we call *short information units* (SIUs). In this work SIUs are defined as *information units (IUs) performed by just one phonological word and encapsulated in a single intonation unit, i.e. preceded and followed by a prosodic boundary, at least one of which necessarily being of non-terminal nature.*

We use data from the C-ORAL-BRASIL corpus (Raso and Mello, 2012) and concentrate our attention on just one very frequent lexeme of Brazilian Portuguese (BP), the lexeme ASSIM. This lexeme, which could be in principle translated as “this way” or “so”, has become very vague in some of its uses and has been increasingly used with many different functions. Therefore, we deemed it a good candidate for an exploratory study intended to establish a direct bridge between an information function and its prosodic form. In fact, we maintain that the main features that convey an information function are prosodic in nature and that the lexicon, as well as the syntactic structure, plays a very secondary role. With the expression *information unit* we refer to the units that make up the information structure of the utterance; more precisely, we refer to the functional value that, from a pragmatic point of view, each intonation unit may establish with the other intonation units in the structure of the utterance. This perspective has already been adopted in the analysis of different information units within the paradigm of the Language into Act Theory (L-Act – Cresti, 2000; Moneglia and Raso, 2014; Cavalcante, 2020), and it will be used for the first time here to investigate SIUs.

The main principle underpinning this work is that, in speech – and especially in spontaneous speech data –, information functions are marked primarily by prosody and not by syntax and lexicon. While different syntactical constructions and lexical items may vary almost freely within speech chunks conveying the same information function, different prosodic realizations can completely change the informational statuses of speech chunks made up of the very same syntactic constructions and lexical items. In other words, there is no strong relationship between lexicon and syntax on the one hand and information functions on the other. There may be some correlations between them, but these correlations are not the main feature that convey the informational function.

We will argue that resorting to syntax or the lexicon for features that characterize an information function can lead to confounding different IUs with one another and to categorizing in a wrong way the functional distinctions. Obviously, it is always possible to categorize the data in a plausible way, but this does not mean that this is the way in which they are categorized in natural speech. In fact, different functional categorizations have been proposed in the literature, usually using the lexicon and/or syntax as the main formal cue that conveys or marks the function. Our perspective is that we can only consider a certain categorization to be correct or useful to understanding natural speech if we can find a coherent prosodic correlate in spontaneous speech data.

The prosodic aspects allowing the understanding of information structure in speech are twofold: prosody allows us to segment speech, and thus enables us to individualize units that constitute the informational object to be studied; after this step, prosody guides us to observe (and perceive) the specific informational value of the unit.

All the audio files of the examples presented in this work can be found at www.c-oral-brasil.org > multimedia > RASO, T., SANTOS, S. (2020).

2 Speech segmentation and theoretical framework

2.1 Terminated sequences as a reference unit for speech

It is a general consensus that speech is segmented into intonation units that encapsulate a certain number of words (at least one, but not so many) in the same prosodic envelope, which establishes some contrast with the words encapsulated in different prosodic envelopes (see Barth-Weingarten, 2016; Barbosa and Raso, 2018; Izre’el et al., 2020b, for a survey and a discussion of the relation between segmentation and functional consequences in speech; see Izre’el et al., 2020a for different proposals that basically follow the same principle).

L-AcT assumes that there is an isomorphic relationship between intonation unit and IU. Each intonation unit carries an informational value. The only exception to this are the so-called *scanning* units (SCAs; see the cases in the example for stanza at the end of this section¹). SCAs are intonation units that pertain to an IU formed by two or more (but not so many) intonation units. They can be recognized (i) by their neutral prosodic profile, since the functional profile is always positioned on the last intonation unit of the IU; and (ii) by their being syntactically compositional with the other intonation units forming the same IU. Syntactic compositionality is a property of the IU, and IUs establish functional relations with each other by means of prosodic features; syntactic compositionality does not play a necessary role (Cresti, 2014). SCAs are motivated by either voluntary or, more frequently, non-voluntary reasons: the former occurring when we scan an IU in order to be more clear or emphatic; the latter, when the segmental content is too long to fit into just one intonation unit or when there is some problem due to performance. As expected, SCAs are more common in lower diastatic varieties, in monologues, and in media speech.

Except for SCA, speech is segmented into intonation units that carry IUs. In the speech flow, we can recognize intonation unit boundaries with terminal or non-terminal value. Terminal boundaries mark terminated sequences, while non-terminal boundaries mark intonation units, which (again, except for SCA) form IUs inside the same terminated sequence.

There are two types of terminated sequences: *utterances* and *stanzas* (see examples at the end of this section). Utterances are formed by one pattern of IUs, whose nucleus is the illocutionary unit called *Comment* (COM). The illocutionary unit is the only one that is necessary and sufficient to build an utterance, since it carries the action that the sequence communicates (Austin, 1962). Sometimes, two or more (but not many more) patterned illocutions create what is called a *Multiple Comment* (CMM), like in rhetorical holistic illocutionary sequences such as *Comparison*, *Reinforcement*, *List*, and other patterns. These are conventionalized as a holistic pattern that requires more than one illocutionary unit to be performed. On the contrary, stanzas (Cresti, 2009) are formed by two or more (sometimes many more) sub-patterns, each of which featuring a mandatory illocutionary unit, called *Bound Comment* (COB), and optional non-illocutionary units; each sub-pattern ends with a non-terminal boundary, signaling that, despite the accomplishment of the illocution, the speaker has not finished his sequence. Stanzas are very frequent in monologues and less interactive exchanges, and they are typically characterized by weakened illocutions that lose part of their actional potential, in order to express a flow of thought (Chafe, 1994; Du Bois, 1993; Mithun,

¹ The examples of utterance, stanza, and all the IUs explained in section 2 are extracted from a minicorpus built upon texts of the Santa Barbara Corpus for Spoken American English (Du Bois et al., 2000-2005), resegmented in accordance with the L-AcT criteria (Cavalcante et al., 2018; <http://www.c-oral-brasil.org/> > corpora; see also Cavalcante and Ramos, 2016).

2014 and 2020. For different traditions that share many theoretical and methodological aspects with L-AcT, see also Blanche-Benveniste, 2002, many chapters of Raso and Mello, 2014, and of Izre'el et al., 2020a). It is, for instance, what happens in narrations, descriptions or argumentations. The sub-patterns of a stanza are juxtaposed one after the other, as they are not patterned. They constitute a speech structure different from the utterance.

We exemplify below a case of utterance (Sound file 1) with a pattern formed by three IUs, respectively AUX, TOP, and COM, and a case of stanza² (Sound file 2). Sound file 1.a exemplifies that the COM is interpretable in isolation.

🔊) Sound file 1 – 056_apubdl01
*RAN: [56] but/=AUX= if you can 't get him /=TOP= **just go on** //COM=

🔊) Sound file 2 – 004_afammn01
*LYN: [4] that's another thing too /=COB= is I kinda had a general idea of [1]=SCA= of kinda how to do it just watching him /=COB= or [1]=EMP= **and watching** /=SCA= **people come to our place and doing it and stuff** /=COB= **you know** /=AUX= and /=AUX= I don't know then /=COB= down there /=TOP= &he /=EMP= it's mandatory /=COB= you have to [3]=EMP= &he /=EMP= to graduate /=TOP= **you know** /=AUX= or well /=AUX= to /=SCA= get the degree /=TOP= you know /=AUX= you have to take this class //COM=

In the latter example (a rather big stanza), it is possible to remark some interesting aspects of the information structure: firstly, the fact that we have six juxtaposed sub-patterns (five COBs, followed by a continuity prosodic signal and the final COM) that form the so-called stanza. The first sub-pattern is a simple one, built up just by the illocutionary unit. The second one is also formed only by the illocutionary IU, which this time is scanned into two intonation units. The third one also presents an illocution scanned into two intonation units, followed by a DM (AUX). The fourth one begins with a DM and is followed by another illocution. The fifth one features a pattern TOP-COB with a disfluency in the middle. The last one is more complex: after some disfluencies, it features a TOP, two DMs, a second scanned TOP, followed by another DM and by the terminal illocution.

2.2 Different IUs

Besides the illocutionary unit (COM, CMM, or COB, which are mandatory units to form an utterance or a stanza sub-pattern), other optional units may be present in the utterance or in a stanza sub-pattern. All these units are defined in functional, prosodic and distributional terms. These units can be of two types: textual units or dialogic units. The first type constitutes the semantic text of terminated sequences, while the second type (corresponding to what in other frameworks are called *discourse markers*) regulates the interaction and are directed to the interlocutor. In the remainder of this section, we outline these units, giving more details about those that are more important for studying SIUs, namely, the parenthetical, the locutive introducer, the appendix, and some dialogic units.

² Each example is preceded by an abbreviation that indicates the number of the utterance, separated by underscore to the abbreviation of the file in the corpus. The first letter indicates the language (a=American English; b=Brazilian Portuguese); the second part of the abbreviation indicates whether the example is extracted from the private-familiar (fam) or the public (pub) context; the rest of the abbreviation indicates the interactional typology (dl=dialogue; cv=conversation; mn=monologue) and the number of the text in each section of the corpus.

(i) Topic (TOP). It is defined, in terms of pragmatic aboutness, as the cognitive domain for the illocution (Firenzuoli, 2003; Mittmann, 2012; Cavalcante, 2015 and 2020; Raso et al., 2017). Prosodically, TOP has a functional focus on its last stressed syllable and the first post-stressed (if any)³. Distributionally, it always occurs to the left of COM. It is the most important among the optional IUs. Some examples of TOP can be seen in the examples of utterance and stanza, given at the end of section 2.1, and in the example of APT, at the end of point (ii), just below. Another example is the following:

🔊 Sound file 3 – 113_afamdl03
*ANE: [113] **those cashews** /=TOP= I just kept eating em //COM=

As the reader can observe, for our definition of TOP it does not make any difference whether the compositionality with the other units is maintained (as in the other examples of TOP) or not (as in the example above). Since the relation between TOP and the illocutionary unit is signaled by prosody, and since the syntactic compositionality (except in SCA) is interrupted by the prosodic boundary, from a functional point of view there is no need for a distinction between these two kinds of TOP. They all establish a relation of pragmatic aboutness with the COM, providing the cognitive domain of reference for the illocution.

(ii) Appendix of Topic (APT) and Appendix of Comment (APC). These two units serve as textual integrations to TOP and COM (see examples at the end of the paragraph), and are always on the right side of the main units (TOP or COM). While APC always has flat or falling f0 profile and low intensity, ATP may have a falling or a TOP-like profile without functional focus. APC is frequently composed of just one phonological word, and is, therefore, especially interesting for our purpose. Because of its prosodic characteristics, and since it frequently appears at the end of the utterance, it often features a change in voice quality, mainly to creaky or breathy qualities. These two units are rare, especially APT. We will say more about APC in section 3.

🔊 Sound file 4 – 057_afammn02
*ALN: [57] well /=AUX= before that /=TOP= he took us /=SCA= around the city
/=COB= still got his card /=COM= **somewhere** //APC=

In the example above, which is a small stanza, the first sub-pattern shows a DM at the beginning, a TOP and a scanned illocution, while the second sub-pattern features another illocution and an APC. Listening just to the COM (Sound file 4.a) it is easy to perceive its interpretability in isolation, without any need of the APC.

🔊 Sound file 5 – 053_afamdl02
*PAM: [53] **the things I know most** /=TOP= **about life and death** /=APT= come
from [/1]=SCA= from /=SCA= my grandmother //COM=

This example, in its turn, displays an APT, which just integrates the text of the TOP without carrying a prosodic nucleus like TOP.

³ See Raso et al., 2017, and Cavalcante, 2020, for a more detailed description of the three forms of TOP we found in English and four analyzed Romance languages.

(iii) Parenthetical (PAR). This unit may be very long or short, and the short PAR is one of the very objects of this work. PARs have a metalinguistic function in that they express a commentary that instructs the addressee on how to interpret the rest of the utterance, or part of it, and they may modalize some part of the utterance. Prosodically, PAR exhibits the following features: overall flat f0 profile on a different level with respect to the rest of the utterance, usually a lower one, but sometimes a higher one; frequently, long parentheticals frequently display a transition phase from the f0 level of the previous unit and to the following one; other prosodic features are: higher articulation rate and lower intensity. The flat f0 profile may display non-functional movements, below the glissando threshold. PAR can occupy almost any position, including interrupting another textual unit (mainly COM and TOP), except for the initial position. It is not clear yet whether long and short PARs (see the examples at the end of the paragraph) can be dealt with as the same information unit. This issue, however, is out of the scope of this work. We deem it possible that long PARs may be performed on a communicative level different with respect to that of the host utterance and may be articulated in more than one sub-units. An example of long PAR is provided below:

🔊) Sound file 6 – 013_afamnm04
 *LAJ: [13] and I know my brother /=TOP= **who lives in Dallas** /=PAR= &he
 /=TMT= **the one that played football at Perdue** /=PAR= he [1]=EMP= he knew
 it //COM=\$

Besides the TOP-COM pattern, which is perfectly interpretable even if we cut out whatever occurs between them (Sound file 6.a), this utterance features two long parentheticals, separated and followed by short disfluencies. It is easy to note the metalinguistic function of the PAR, aimed at providing information that helps the interpretation of the rest of the utterance or part of it.

🔊) Sound file 7 – 207_apubmn01
 *LAJ: [207] [...] last time you guys were here /=TOP= there was /=i_COB=
maybe /=PAR= ten /=COB= [...]

The example above displays a short PAR, embedded within the illocutionary unit⁴, which is modalized. The COB would function perfectly without the PAR, as shown by Sound file 7.a.

(iv) Locutive Introducer (INT). This unit has the function of marking that what follows does not refer to the *hic et nunc* of the utterance level (see example at the end of the paragraph). In fact, INT introduces meta-illocutions, mainly reported speech, but also emblematic exemplifications, and other meta-illocutions that share their reference either with a time, space or people different from the ones of the utterance or with any time, space or people (i.e. they are neutralized in terms of time, space or people). INT can also introduce lists, signaling that what follows pertains to a unified domain in contrast to the rest of the utterance. Prosodically, INT has a very recognizable profile. It may be a long or a very short stretch of speech, but it is always expressed by a falling final f0 profile, with a very high articulation rate and/or a strong syllabic reduction, which gives the impression of a very high speech rate. At the same time, the meta-illocution (especially when it comes to reported speech) is performed at a higher f0 level,

⁴ In this example, i-COB (or interrupted COB) and COB form a single unit, which is momentarily interrupted by PAR.

with longer duration and higher intensity. This produces a sharp contrast between INT and the introduced meta-illocution.

- 🔊) Sound file 8 – 011-012_afamnn03
* LEE: [11] she said /=INT= &sh [1]=EMP= they were sitting there watching TV /=CMM= and the next thing she know there 's Mike /=SCA= on [1] =EMP= on TV //CMM=\$
*LEE: [12] **and she goes** /=INT= oh //COM_r=\$

Utterance 12, preceded by another utterance just to provide the reader with some context, shows a typical example of INT followed by reported speech. It is very clear how the INT is poorly articulated and how the prosodic contrast with the meta-illocution it introduces is evident. However, utterance 11 also features an example of INT, in this case not followed by reported speech.

(v) **Dialogic units (AUX).** These units correspond to what, in other frameworks, are called *discourse markers* or *pragmatic markers* (and even other terms). Functionally, they are addressed to the interlocutor in order to regulate the communication, with various sub-functions (many of them are exemplified, with the generic tag AUX, in the previous examples). Cresti (2000) individualizes five different subtypes of AUX; Frosali (2008) adds a sixth one with a cohesive function to Cresti's set; Raso (2014) analyzes the six AUX in order to account for the function of discourse markers in terms of IUs; Raso and Vieira (2016) propose a prosodic description for three of the subtypes identified by Cresti and Frosali, while considering that the other three subtypes need further research and may be divided into more than three different functions, given that they do not seem prosodically coherent; Gobbo (2019) statistically validates the three units described in Raso and Vieira (2016); Raso and Ferrari (forthcoming) refine the previous prosodic description, as well as study the lexical makeup and frequencies of these IUs using comparable corpora of Italian and BP. So far, three different types of AUX have been satisfactorily described on the prosodic level: *Conative* (CNT), *Allocutive* (ALL), and *Incipit* (INP). In this work we propose a fourth AUX, which has already been observed (but not analyzed) associated with several lexical items in Gobbo (2019). This fourth unit, however, needs to be better accounted for on the functional and prosodic levels. We will return to this in section 5.2.

We now proceed to review the literature on the main IUs found in our data, namely parentheticals, appendices, locutive introducers, and different types of dialogic units/discourse markers.

3 The literature on short parentheticals, discourse markers, and other SIUs

3.1 Short parentheticals

The literature is somewhat convergent with respect to the general function of parentheticals. They constitute a specialized communicative behavior that enables the speaker to comment on his own speech (Schneider, 2014). This behavior results from the speaker's assessment of the interlocutor's contextual knowledge and processing skills (Kavalova, 2007). Parentheticals allow the speaker to go beyond the domain of linearity by creating a disruption whose main purpose is to add information in a position that is maximally relevant to the addressee (Kavalova, 2007; Schneider, 2014). Evaluating the contextual knowledge of the interlocutor, the speaker plans the utterance so that the necessary references can be established as soon as

possible. This allows the interlocutor to make anticipatory hypotheses about the general meaning of the utterance. The speaker obtains optimum relevance at a minimum processing cost by inserting additional information exactly where it is most useful. Kavalova (2007) points out that the increased syntactic processing cost of parentheticals is offset by its semantic contribution in optimal distribution. Schneider (2014) also points out that, despite violating the maxim of mode – by inserting information in a place that implies more costs –, parentheticals satisfy the maxim of relevance by providing information in the most relevant place. For L-AcT, PAR offers instructions on how to interpret any part of the utterance (i.e. it may have backward, forward, or global scope) and hence has a metalinguistic/metacommunicative function (Cresti, 2000; Firenzuoli and Tucci, 2003; Tucci, 2004 and 2009). PAR, and especially the short PAR, often modalizes the utterance (Santos, 2020; Schneider, 2007; Tucci, 2009).

Parentheticals have been studied from a syntactic and lexical perspective at least since the 1950's (see Urmson, 1952, for parenthetical verbs, for instance). The bulk of the research assumes syntax as the departing point to examine whether and/or how parentheticals are mapped onto prosody (Dehé, 2014; Paixão and Serra, 2018). Syntactic parentheticals are generally defined as structures whose syntax is not governed by the host structure (Dehé, 2014; Espinal, 1991; Haegeman, 1989, for some syntactic accounts). In other words, syntactic parentheticals are structures that fail to pass constituency tests and that display some syntactic properties that lie beyond the scope of this work.

The idea that parentheticals are mapped onto prosody and that their statuses are signaled by prosodic features dates back to the 1980's. Dealing with syntactic issues of adverbial parentheticals, Haegeman (1989) puts forth that the speaker may provide prosodic cues from which the interlocutor infers the parenthetical status of a given constituent and its non-syntactic relationship with the matrix sentence. This approach was influenced by the Relevance Theory (Sperber and Wilson, 1986). Bolinger (1989), without the constraints of formal syntax, observes that parentheticals are generally characterized by a higher speech rate, marginal pauses (or boundaries), and pitch lower than that of the host sentence. Within L-AcT, Firenzuoli and Tucci (2003) observe, by examining spontaneous speech in Italian, that PAR is marked by an overall flat f_0 contour that is mostly on a lower (but sometimes higher) level with respect to the host utterance, showing lower intensity levels, and a higher articulation rate.

Although the idea that parentheticals are prosodically signaled in speech is not new, a considerable amount of descriptive research is still carried out without taking prosody into consideration. The conclusions may seem obvious. Descriptions often result in an incoherent set of properties and functions, or a *motley crew*, as Dehé (2014) describes it. For instance, Schneider (2007) analyzes a large set of reduced parenthetical clauses, which were extracted from various speech corpora according to criteria requiring that, among other things, the clauses must a) bear a parenthetical verb; b) have no overt syntactic link to the host structure; and c) be placed in medial or final position within the host structure. Although the author does acknowledge that parentheticals may exhibit a certain number of prosodic features, no prosodic criterion is adopted to select the data. We deem it possible that many clauses in the sample may have functions other than those of short parentheticals, as we try to show in this work. For instance, fr. *tu vois* “you see”, fr. *vous savez* “you know” or sp. *ya ves* “you see” are good candidates to function as interactional regulators rather than proper parenthetical mitigators. We argue that the function these speech chunks take on will depend first and foremost on their prosodic realization.

A second example can be drawn from Kaltenböck, Heine and Kuteva's (2011) typological effort. The authors put forth an umbrella term to accommodate a large number of extra-clausal constituents that often go embedded in utterances. The *theticals*, as the authors call

them, are: a) syntactically independent; b) prosodically detached from the host utterance (i.e. prosodically phrased); c) non-appositive; d) tendentially moveable within the utterance; and e) similar to any other piece of language, with which internal syntax of theticals share the same syntactic principles. The functions assumed by theticals may vary greatly. For instance, they may (i) mark reported speech and evidentiality, (ii) offer comments and hedging (modalization) to what is being said, and even (iii) induce the interlocutor to take some action within the scope of the interactional relationship. Although the authors do acknowledge a special status to prosody – in that they assume theticals are prosodically detached –, they do not consider that prosody may be interwoven with more specific functions performed by theticals. Once again, we argue that many of the elements categorized as thetical may be better classified if the role played by prosody, which we consider to be central, is acknowledged. Within the L-AcT framework, depending on the prosodic realization, function (i) is perfectly compatible with INT, (ii) with PAR, and (iii) with one of the dialogic units, all of which being marked by prosodic regularities.

For L-AcT, PAR may take on no matter what syntactic form. Indeed, it may even display an overt mark of subordination to its host structure. If a unit is clearly delimited by prosodic boundaries (i.e. is prosodically phrased) and if it displays the typical prosodic profile and function of a parenthetical, it is a PAR unit. This does not mean that all syntactically independent units inserted into the utterance must necessarily bear the prosodic features and function displayed by parentheticals. It is perfectly possible that such a condition be met by other information units. Indeed, longer excursus (i.e. longer detailed explanations that divert from the main topic of the conversation) are often observed in speech without displaying the typical prosodic cues of parentheticals. However, those stretches of language will not be interpreted by the addressee as actual parentheticals.

It is worthwhile to mention that L-AcT aims to account for how informational statuses are signaled within the domain of terminated spoken sequences, which are highly marked by pragmatic rather than syntactic linkage. For L-AcT, the formal cue guiding this process is prosody. Although there might be some peril of circularity in defining parentheticals (or any other information unit) by its prosodic form, as some authors claim, this risk is cancelled out when formal regularities are considered together with function, i.e. when both prosodic form and function are coherently examined in spontaneous speech.

3.2 Locutive Introducers and Appendices

These two units are not frequently acknowledged in the literature on information units. Within the L-AcT framework, they are clearly distinguished and studied (Cresti, 2000 and 2003; Tucci, 2006; Raso and Ulisses, 2008; Oliveira, 2012, for the Appendix; and Cresti, 2000; Giani, 2003; 2004; Corsi, 2009; Maia Rocha, 2011; Maia Rocha and Raso, 2011, for the Locutive Introducer).

Although APC has already been outlined in section 2.2, it merits further consideration. Some frameworks treat APC as a postponed Topic (Lombardi Vallauri, 2009) or as an anti-topic or afterthought (Gordon, 2008; Averintseva-Klisch, 2008; Shimoyama *et al.*, 2015, among others). L-AcT maintains that the notion of postponed Topic cannot correspond to any IU if we define TOP as the cognitive domain to which the illocution makes reference. In fact, it is not only on its prosodic form that APC differ from TOP. If this was the only reason, the difference in prosodic profile could be considered a consequence of the position.

In fact, APC differs from TOP for cognitive reasons. Let us compare a sequence like *Have you already BOUGHT* /=COM= *the house* //APC= with the sequence *The house* /=TOP= *have you already BOUGHT (it)* //COM=. If *the house* is performed after the COM unit and

with a flat or falling profile and low intensity, it cannot instantiate any cognitive domain to which the illocution makes reference, since the unit can refer only to given referents. We can only utter the unit with this profile if the referent (in this case, what the interlocutor should have bought) is already clearly given, in the discourse or in the context. We cannot use this kind of unit to establish a new domain, as it can only be used to recall – and with little informational gain – something that is already known. On the contrary, TOP is frequently used to change the cognitive domain of the illocution in discourse. This is a function commonly carried out by TOP, and it can typically be observed in narrations, when something in the setting may be changed by means of a TOP unit informing the interlocutor that the next illocution must be interpreted with respect to a cognitive domain other than that of the previous illocutions.

We can recover a referent after the illocution has been performed, in order to make it clear that the illocution must be interpreted with respect to a new referent, one that has not been mentioned yet. But, in this case, we would need a clearly different prosodic profile, and hence we would utter a new illocution, such as in *have you already BOUGHT / the HOUSE //*. However, this case is quite different (both prosodically and functionally) from the COM/APC pattern.

Cresti (2000 and 2014) claims that TOP has a semantic modality independent of that of COM, while APC always shares it with COM. That means that APC is just a textual integration of COM. In fact, APC is usually very short and functions as a repetition or reminder, or as a filler that concludes the utterance. Besides, the syntactic correlates of TOP and APC are quite different (see Raso and Ulisses, 2008).

In sum, what APC adds, informationally speaking, is a textual integration that usually includes (i) repetitions of an expression of the utterance or of the previous one; (ii) fulfillment of the content of the utterance, marking its conclusion with a filler (this would be our case with the lexeme ASSIM); (iii) delayed information, that textually complete the semantic content of the COM (Tucci, 2006; Oliveira, 2012).

3.3 Discourse markers

Discourse Markers (DM), or Dialogic Units (DU) in the L-AcT terminology, have become an important object of study during the last 30 years, at least since Schiffrin (1987). Nevertheless, with few exceptions, DMs have been always studied from a lexical perspective, without taking into account prosody in all of its aspects. Although several frameworks have been proposed, the following questions remain unresolved: (i) how to identify a DM, i.e. how can we describe the conditions that lead a lexical item to perform the function of a DM, given that they can perform functions that are clearly different? (ii) Once we have identified a DM, how can we understand its specific sub-function (take the turn, mark contrast, push the interlocutor, or any other function)?

Of course, if we want to treat DM as a real linguistic category, we need to discriminate the formal principles that allow us to answer these two questions. Once again, our proposal is that these formal principles can only be offered by prosody, not by the lexicon. We may find some lexical correlations, but they are very vague, since the same lexical item can fulfill different functions, and, even when a lexical item behaves as DM, it may carry out different sub-functions.

What we propose is that DMs (DUs) should be dealt with as information units, hosted in a dedicated intonation unit and marked by specific prosodic features: first, their prosody does not allow a pragmatic interpretation in isolation, which distinguishes them from illocutions; secondly, their prosodic isolation marks the non-compositionality with respect to the rest of the utterance. This allows us to distinguish cases such as the following:

- (a) *You see my father’s house.*
- (b) *You see, my father’s house.*

In (a) *you see* is syntactically and semantically compositional with the rest of the utterance, while in (b) it is not; in this last case, depending on other prosodic features, *you see* can either have an illocutionary value or function as a DM (DU). For more in-depth considerations about DMs in L-AcT, see Raso (2014), Raso and Vieira (2016), and Raso and Ferrari (forthcoming).

Once we have observed that a lexical item behaves as a DM, we still need to distinguish its possible sub-functions. As we have already said, Raso and Vieira (2016) satisfactorily describe three of the DUs proposed by Cresti (2000) and Frosali (2008). These descriptions were statistically validated by Gobbo (2019). These three DUs are the *Conative*, the *Allocutive* and the *Incipit*. The other DUs proposed by Cresti and Frosali seem to need further research, and they should probably be analyzed as more than three sub-functions, each of which marked by specific prosodic features.

In analyzing the item ASSIM, we found two different DMs. The first one, the *Conative*, is frequent, well known and described. Its function is to push the utterance toward an actional solution, or, when put at the end of the utterance, to confirm that what was said was meant to give an actional solution to the utterance. It is characterized by a falling f₀ movement beginning on the stressed syllable. When there is enough segmental material before the accented syllable, a rising movement below the glissando threshold precedes the falling movement, which seems to create appropriate conditions for the falling functional movement, which usually exhibits a high f₀ variation rate. The duration of the *Conative* is short, and its intensity is relatively high (see example in section 5.1).

In Gobbo’s research, a new and recurrent prosodic profile that seemed to convey a specific function was found but could not be dealt with at the time. This profile was lexically realized by expressions such as BP *entendeu* “did you get it”, *viu* “did you see” or *hein* (an interjection having basically the same function of *entendeu* and *viu*). Searching now for the different informational values fulfilled by the lexical item ASSIM, we found again the same profile, apparently with same function. The profile is characterized by a rising f₀ movement, which seems to convey the function of highlighting and calling the attention of the addressee to what has just been said. Of course, we need a better description of the prosodic features of this unit, also taking into account intensity and duration and comparing the tokens of ASSIM with other lexical items that seem to have the same function. In this work, we simply propose that this new unit, provisionally called AUX/ASC (see section 5.2), may be another sub-function of DMs and we show a prototypical example in our results section.

4 The C-ORAL-BRASIL corpus and the selection of the lexeme ASSIM

4.1 The corpus

As aforementioned, we based this study on data from the C-ORAL-BRASIL I corpus (Raso and Mello, 2012; Mello, 2014). The corpus, together with several statistics, measurements, and its companion book can be freely downloaded at www.c-oral-brasil.org.

C-ORAL-BRASIL is a third-generation spontaneous speech corpus, with text-to-speech alignment, prosodic segmentation, and a good to outstanding acoustic quality, thanks to the high-quality wireless equipment used in its recording sessions. It documents the informal speech

of the state of Minas Gerais, with a focus on the variety spoken in the metropolitan area of the state's capital, Belo Horizonte. For a description of the corpus, see Raso (2012a; 2012b). Here, we provide just some general information, useful to understanding the main characteristics of the data used in this work.

C-ORAL-BRASIL I features 139 texts of approximately 1,500 words each, adding up to 208,130 words. The texts are divided according to diaphasic categories: firstly, into family/private (circa 75%) and public (circa 25%) contexts; in each of these two contexts, one third of the texts are monologues, one third are dialogues and one third are conversations (that is, dialogues with more than two main participants). The corpus is diastatically balanced (sex, age, and school level), but its main goal is to be diaphasically diverse: within each interactional type (monologue, dialogue or conversation), the largest possible variety of situations is documented. Consequently, just two texts in the corpus represent interviews and chats, unlike what usually happens in spoken corpora, since these are obviously the easiest situations to record. On the contrary, C-ORAL-BRASIL documents a large variety of situations, including many in which people move around (i.e., not only static situations), for instance, people doing exercises at the gym, people travelling by car, a waiter serving people at a party, a family dinner, people cleaning the house, people playing soccer, people playing table games, an engineer and a construction worker at a work site, a broker showing an apartment to clients, people shopping at the supermarket and other stores, and many other situations. The reason for this is to elicit, in natural context, the highest possible variability of speech acts and spoken information structures. The corpus contains speech by 362 speakers, two thirds of which being the main participants of the interactions, recorded with clip-on microphones. Understandably, a corpus like this features many unforeseen speakers in the recordings, usually for short periods of time.

The corpus transcriptions are prosodically parsed into terminated sequences, marked by a double-slash symbol (//), and non-terminal intonation units, marked by a single-slash symbol (/). Different tags indicate retractions ([/n]), with a single slash followed by the number of retracted words within brackets, interrupted utterances (+), time-takings (&he), which are more commonly known as filled pauses, and interrupted words, which are preceded by an ampersand (&). The kappa coefficient k (Fleiss, 1971) expressing the inter-rater agreement among the annotators who reviewed the segmentation of the corpus is $k = 0.86$.

4.2 The lexeme ASSIM and the data sampling

We chose the lexeme ASSIM because it is very frequent in BP and because it is the most frequent among those lexemes which, according to the original segmentation of the corpus, constitute alone an entire intonation unit. Therefore, we thought it would be a good choice for a first comparison of different functions marked by prosodic regularities.

This lexeme can, however, be used in different ways: (i) as a clear modifier, like in BP *falei assim* “I said this / I was like”, *assim sendo* “being so” or *grande assim* “this big”. In cases like these, ASSIM is always compositional with its accompanying word and cannot be a candidate to fulfill a dedicated intonation unit and to constitute an IU by itself; (ii) as a connector indicating consequentiality, being translatable as *therefore* or *thus*, like in the following example:

🔊) Sound file 9 – 032-034_bnatpr01
 *ANT: [32] e essa parábola / se a gente analisar bem / ela reflete a nossa vida /
 atual // \$

*ANT: [33] este pastor / a esposa / e Jane / nós podemos **assim** dizer que o pastor é Deus / podemos dizer / que a esposa que faleceu / é o nosso Senhor Jesus Cristo / que deu a vida por mim e por você // \$

*ANT: [34] e também / podemos **assim** dizer / que Jane / representa cada um de nós // \$

*ANT: [32] *and this parable / if we analyze it well / it reflects our life / nowadays* // \$

*ANT: 33] *this pastor / the wife / and Jane / we can **therefore** say that the pastor is God / we can say / that the wife who passed away / is our Lord Jesus Christ / who gave his life for me and for you* // \$

*ANT: [34] *and also / we can **thus** say / that Jane / represents each of us* // \$

(iii) as a way to modalize, to attenuate or to accomplish some similar function, similarly to the example in section 5.4; and (iv) as a discourse marker, signaling something directly to the interlocutor, and not as a contribution to the semantic text of the utterance, as in the examples in sections 5.1 and 5.2. Clearly, only the cases (ii) to (iv) are candidates to form an IU.

The lexeme appears in the corpus 1,366 times, and for 310 times it appears in a dedicated prosodic unit. We left aside the tokens in which the lexeme was part of a larger intonation unit – i.e. when it is compositional – and examined only the tokens in which it was immediately preceded and followed by a prosodic boundary, except for the few tokens (5) in which both boundaries were terminal, since this automatically marks an illocutionary use of ASSIM, thus falling outside the scope of this work. In the following example, ASSIM appears three times, always as the illocutionary nucleus, even if only once in a dedicated unit. It is straightforward to perceive that all cases are interpretable in isolation, each time with a different illocutionary value.

») Sound file 10 – 191-196_bfamcv04
*BRU: [191] *cê pode fazer assim* //
*BRU: [192] *que isso é <similar>* //
*HEL: [193] *<tá>* //
*HEL: [194] *e assim* //
*BRU: [195] *não* //
*BRU: [196] **assim** //
*BRU: [191] *you can do it this way* //
*BRU: [192] *'cause this is <similar>* //
*HEL: [193] *<ok>* //
*HEL: [194] *and this way* //
*BRU: [195] *no* //
*BRU: [196] **this way** //

The total number of ASSIM in dedicated units of which at least one of the boundaries was non-terminal, according to the corpus annotation, is 305.

Our first analysis aimed at checking the actual presence of the boundaries. We noticed that annotators transcribe many neither perceivable nor phonetically noticeable boundaries around this word. Sometimes, just one boundary (before or after the word) is perceivable, and the insertion of the other boundary seems to rely rather on non-acoustic factors. This process led to the elimination of 168 tokens, where one of the boundaries was not attested at a more careful observation by three annotators. This is a very important finding since it indicates that, even when no boundary is present, annotators still are induced by syntactic or pragmatic factors to isolate ASSIM in dedicated intonation unit. This phenomenon is so frequent that it deserves

further consideration, even if it is outside the scope of this work. We will come back to it in section 6.

We decided to check the segmentation error rate by comparing the compositional and non-compositional interpretations of ASSIM. Therefore, among the 1,056 tokens not annotated between two boundaries, we randomly extracted 200 tokens. We listened to these 200 tokens and observed that 166 resulted compositional and 30 non-compositional, while 4 could be interpreted either way. By checking the segmentation of the 166 compositional ASSIM, we found 7 errors, i.e. less than 5%; the error rate among the non-compositional was over 13% (4 tokens). This means that, in principle, the chance of prosodic annotation error is almost three times higher when the interpretation is non-compositional.

We excluded from the analysis, due to unreliable segmentation, 168 tokens out of 305 that were originally annotated in dedicated units, which is a much higher error rate than that found among the 30 non-compositional tokens. This may be due at least to three reasons: (i) 30 non-compositional tokens may not be enough to give a reliable percentage of errors; (ii) when reviewing our sample, we wanted to be very conservative, and, therefore, we eliminated all the tokens exhibiting at least one unclear boundary; and (iii) we can expect that the error rate for non-compositional tokens to be higher when the annotator marks both boundaries, thus isolating ASSIM in one prosodic unit, when compared to the tokens in which no more than one boundary was identified, as usually happened in the sample containing 200 tokens of ASSIM. In any case, it is clear that the non-compositional interpretation of the lexeme leads to a much higher segmentation error rate.

As for the distribution of the tokens that do not constitute an IU, we observed that the compositional ones were positioned in 6 tokens immediately after a boundary (4 terminal and 2 non-terminal), 20 were in the middle of the unit, without any boundary around them, 100 were immediately before a non-terminal boundary and 39 before a terminal one. As for the non-compositional tokens, no one was positioned at the beginning of a unit, 6 were in the middle of the unit and 18 at the end (17 followed by a non-terminal boundary and only 1 by a terminal boundary). What can be observed is a strong preference, in both cases, for the end of the unit, be it terminal or non-terminal.

A second analysis focused on the acoustic quality of the audio files. If for any reason the acoustic quality did not allow a reliable analysis with Praat (Boersma and Weenink, 2019), the file was eliminated from our sample. This happened in 12 tokens, due to background noise, in 13 tokens due to creaky voice, in 15 tokens due to devoicing, in 12 tokens due to overlapping, and in 5 tokens due to other quality issues.

We still found 3 more tokens in which ASSIM had an illocutionary value and 2 tokens in which the word was not correctly transcribed. After this phase, only 75 occurrences of ASSIM were saved for the functional-prosodic analysis. We excluded not only the cases in which one of the boundaries was clearly absent, but also those cases in which they were, according to our perception, very unreliable. We did not check all the 1,056 occurrences that were not surrounded by boundaries, but only 200 of them, which were randomly extracted. Among these, we found only 3 tokens to add to our sample, since one boundary was missed. The prosodic/informational analysis was, therefore, performed only on those tokens in which the presence of both boundaries seemed absolutely clear. Table 1 below presents the summary of the revision of the sample.

Table 1: Summary of the revision

| # | Discard reason | # tokens |
|---|----------------|----------|
|---|----------------|----------|

| | | |
|----------------------------------|-----------------------------------|------------|
| 1 | Absence of prosodic boundary(ies) | 168 |
| 2 | Quality issues | |
| 2.1 | Background noise | 12 |
| 2.2 | Creaky voice | 13 |
| 2.3 | Devoicing | 15 |
| 2.4 | Other quality issue | 5 |
| 2.5 | Overlapping discourse | 12 |
| 3 | Illocutionary unit | 3 |
| 3 | Transcription error | 2 |
| A. Total discarded tokens | | 230 |
| B. Total kept tokens | | 75 |
| Total tokens (A + B) | | 305 |

There are advantages and disadvantages in choosing only one very frequent lexeme to pursue the goal of this work, which is to identify different prosodic profiles that convey different information functions. The advantages are that (i) we can easily show how lexical content does not say much about informational value, since we can extract from the same lexeme a small set of prosodic profiles that realize different functions; (ii) we do not have to deal with micro-prosodic effects that change according to the segmental material, since it does not vary substantially. The main micro-prosodic effect we observed in this lexeme is the small fall at the beginning of the second syllable, due to the effect of the unvoiced consonant [s]. The disadvantages arise mainly from the very vague meaning that this lexeme acquires when it is not used as a clear modifier or a clear connector. In the other cases, the vagueness of its meaning does not help the functional interpretation, which, once more, shows how the interpretation of the pragmatic function, without any clear formal correlate, may easily lead to different categorizations, all of which in principle being acceptable.

5 The prosodic forms associated with the different functions

We start from the IUs that seem easier to be functionally identified, and then move on to those that present more problems. We will also examine some isolated cases that may suggest the existence of different functions, which should obviously be confirmed with other lexical items. Finally, we will outline the prosodic forms and the distributions of each IU found in the sample.

5.1 Conative (CNT)

CNT makes up 33.3% of the sample (25 out of the 75 tokens). Distributionally, CNT was found mainly in the absolute beginning of the sequence but also in the middle of sub-patterns and more rarely in the absolute final position (a table containing the frequencies of all IUs found in the sample arranged by position is provided in section 7). Other studies (mainly Raso and Ferrari, forthcoming) support the conclusion that the preferred position for CNT is the initial one. The figures below provide contextualized examples (Figure 1, in particular, shows a token of CNT). Over the waveform in the first layer of Figure 1, f0 contours are represented by blue lines, and the direction of their movements is emphasized by the red arrows. Since we dealt with a spontaneous speech corpus, it is worth mentioning that the f0 signal may present errors in some portions of the audio due to background noise, voice quality (especially falsettos, creaky

and breathy voice), besides some unavoidable coarticulation effects. However, all audio files effectively included in the sample do present reliable signals in the portions of interest, i.e. the dedicated intonation unit containing ASSIM. The second layer of the figure, from top to bottom, contains the syllabic segmentation and the broad phonetic transcription in ASCII characters in accordance with the standards required by the SG Detector script (Barbosa, 2006). The third layer delimits the unit of interest, in this case, CNT. The fourth and the fifth layers display the transcriptions in BP and English, respectively. It is worthwhile to say that the translation should not be taken literally, since, as we said, the same lexical item can perform different information functions; this also being true for languages other than BR. As said earlier, all the audio files of the examples presented in this work can be found at www.c-oral-brasil.org > multimedia > RASO, T., SANTOS, S. (2020).

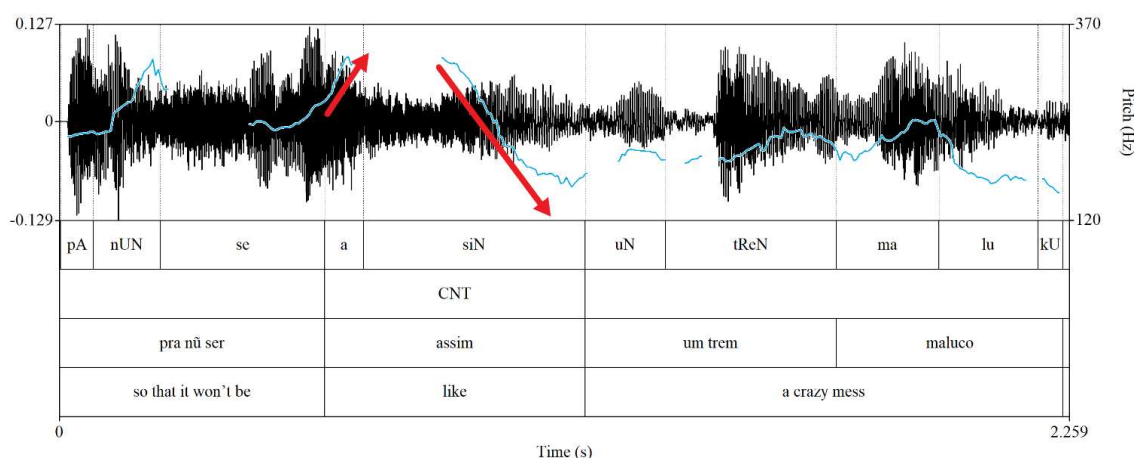


Figure 1: Example of a CNT unit in context.

The transcription and the translation of the whole utterance in Figure 1 can be found below. Note that this utterance is momentarily interrupted by another speaker's utterance, not shown below for the sake of brevity.

- 🔊) Sound file 11 – 139-143_bpubcv03
 *FER: [139] <por isso que cês têm que ser as intermediárias> disse aí /
 *FER: [143] / pra nũ ser / **assim** / um trem / maluco //
 *FER: [139] <that's why you have to be the mediators> of that /
 *FER: [143] / so that it won't be / **like** / a crazy mess //

The prosodic form of CNT is in line with descriptions by other works (Raso and Vieira, 2016; Raso and Ferrari, forthcoming). It is characterized by a falling f_0 movement aligned with the stressed syllable, preceded by a rising movement below the glissando threshold when there is enough segmental material before the stressed vowel. Its intensity is lower than that of the illocution in the same utterance, but it is still rather high; its duration is clearly short with respect to the illocutionary unit. Its f_0 variation rate is usually high, but it seems to vary (as well as its intensity) depending on the attitude with which the unit is performed. By attitude we mean the concrete way in which an abstract informational schema is performed (Mello and Raso, 2011). The attitudinal variation is particularly clear for illocution, but it is true for any information unit. For instance, an order (that can be defined as an abstract actional schema), may be concretely performed with several attitudes (polite, aggressive, seductive, urgent, etc.).

This also happens with IUs other than illocution. Figure 2 below displays the form of the CNT unit in question set off from its context.

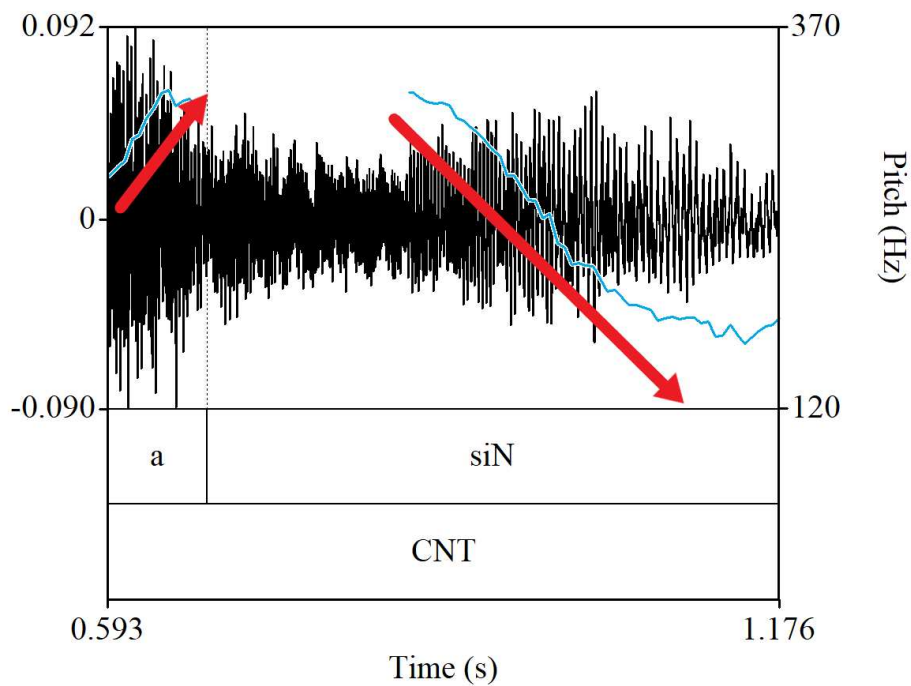


Figure 2: Example of a CNT unit detached from its context.

5.2 A new dialogic unit

This new dialogic unit makes up approximately 13% of the sample (10 out of the 75 tokens). Distributionally, it was found most often in medial positions but also at the end of sub-patterns and terminated sequences. Figure 3 below displays an example of this new dialogic unit in context. It is provisionally identified by the tag AUX/ASC.

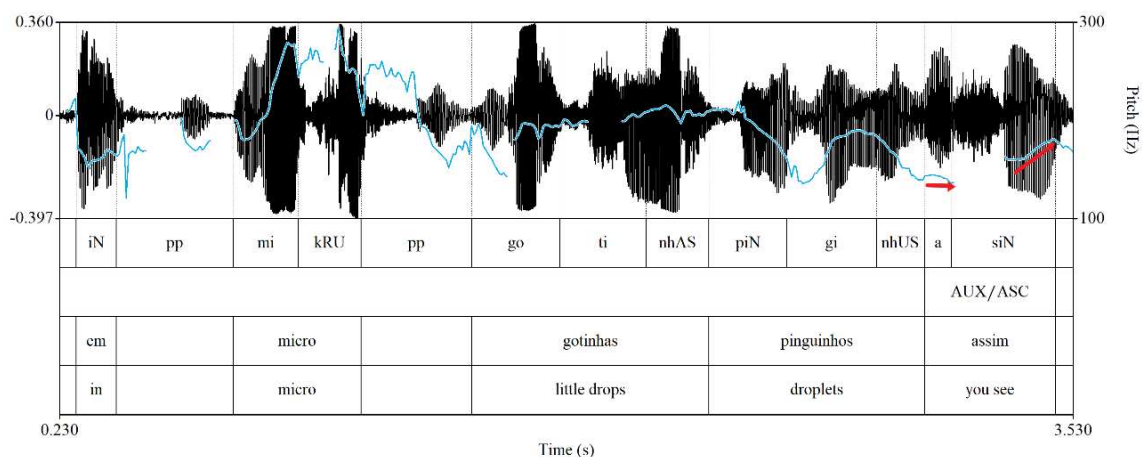


Figure 3: Example of the new dialogic unit in context.

The transcriptions and the translation of the whole utterance can be found below.

🔊 Sound file 12 – 105_bfamdl09
 *LUC: [105] em / micro / gotinhas / pinguinhos / **assim** / de cores / e [1] e [1] e
 ele só pinta desse jeito //
 *LUC: [105] in / micro / little drops / droplets / **you see** / of colors / and [1] and
 [1] and he only paints this way //

The prosodic form of the new dialogic unit is characterized by a rising f0 movement aligned to the stressed syllable. As mentioned before, this unit needs a better description, considering different lexical items, but we think that the profile and the function of these tokens of ASSIM meet those of other items found by Gobbo (2019) while he was conducting his study, whose goal was only the statistic validation of other DMs (namely Conative, Allocutive, and Incipit). Figure 4 below displays the form of the unit set off from its context.

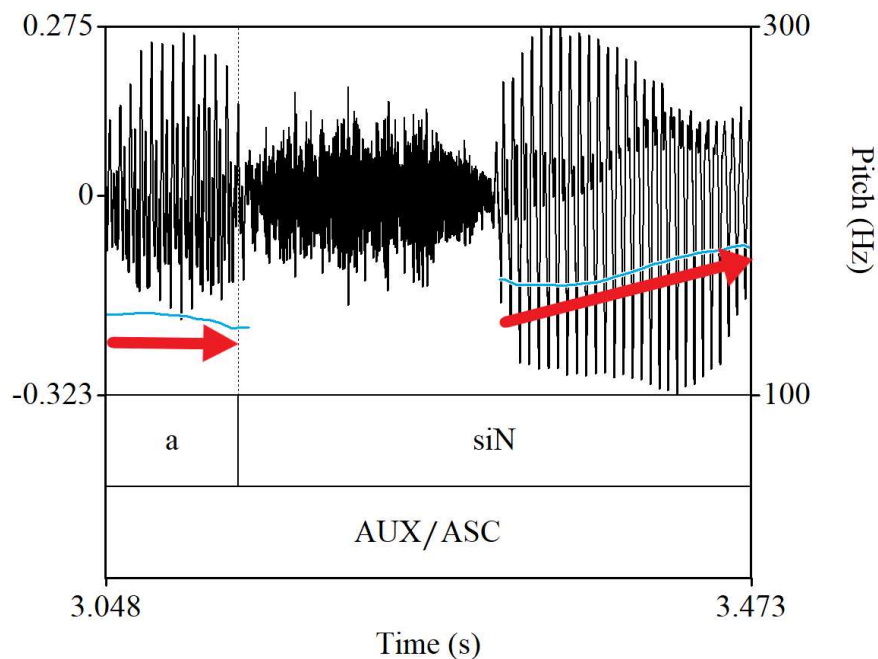


Figure 4: Example of the new dialogic unit detached from its context.

5.3 Locutive Introducer (INT)

INT makes up 8% of the sample (6 out of the 75 tokens). Distributionally, it was found mostly in initial position both of sub-patterns and sequences. In medial position, it was always placed before the introduced meta-illocution. Figure 5 displays an example of INT in context.

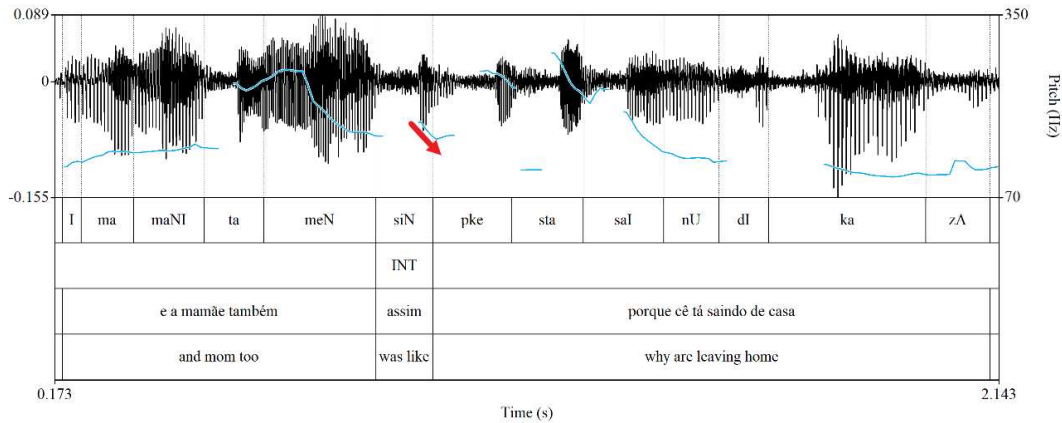


Figure 5: Example of an INT unit in context.

The transcripts and the translation of the whole utterance can be found below.

🔊 Sound file 13 – 040_bfammn33
 *ADR: [40] e mamãe também / **assim** / o que cê tá saindo de casa //
 *ADR: [40] and mom too / **was like** / why are you leaving home //

The prosodic form of INT is in line with other works (Giani, 2003 and 2004; Corsi, 2009; Maia Rocha, 2011; Maia Rocha and Raso, 2011). It is characterized by a falling f_0 profile at the end of the unit, with a high or very high f_0 variation rate. Its intensity is low and its duration short or very short, often with a reduced phonetic form. Figure 6, below, displays the form of INT detached from its context.

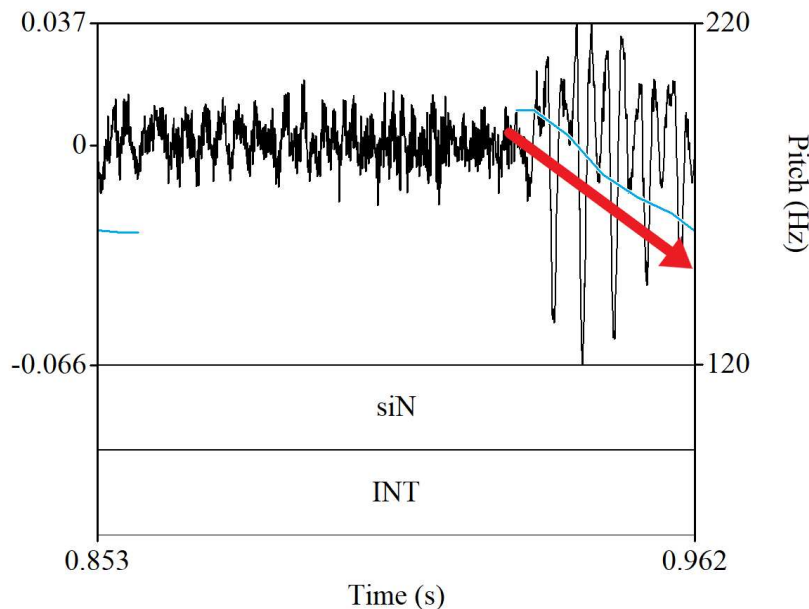


Figure 6: Example of an INT unit detached from its context.

5.4 Short Parenthetical (PAR)

Short PAR is the second most frequent IU of the sample after CNT (but the difference is of just 1 token; therefore, we can say that PAR and CNT are the most frequent IUs filled with ASSIM). It makes up 32 percent of the sample (24 out of 75 tokens). Distributionally, it exhibited a strong preference for medial position, but it was also found in final positions. Short PAR was found neither at the absolute beginning of the sequence nor at the beginning of a sub-pattern, an expected outcome considering the distributional constraints of this unit. This is coherent with the fact that PAR cannot occur in initial position (Cresti, 2000; Tucci, 2004 and 2009) because of its functions: it expresses a commentary on (part of) an unfolding utterance. Figure 7 displays an example of short PAR in context.

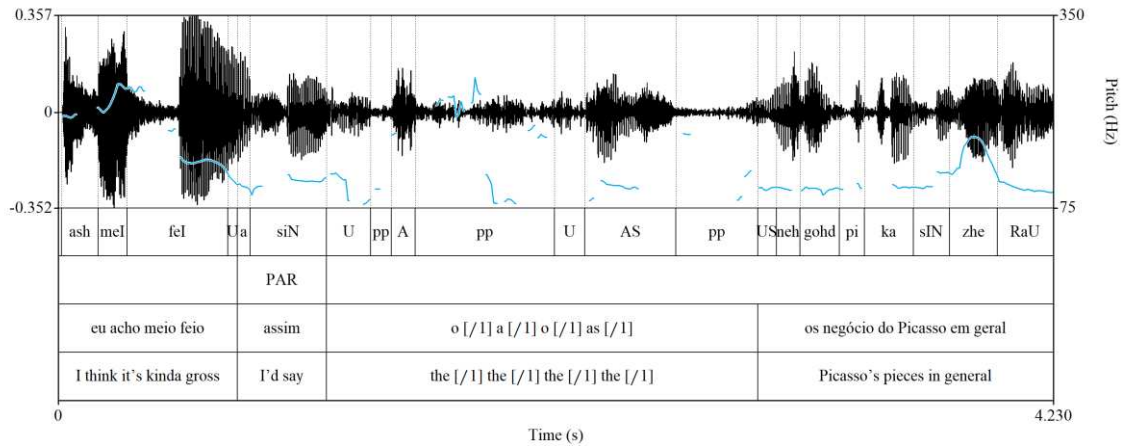


Figure 7: Example of a short PAR unit in context.

The transcriptions and the translation of the whole utterance can be found below.

🔊 Sound file 14 – 344_bfamdl09

*LUC: [344] eu acho meio feio / **assim** / o [/ 1] a [/ 1] o [/ 1] as [/ 1] os negócio do Picasso em geral / assim / mas / como importância / tem que ser considerado //

*LUC: [344] I think it's kinda gross / **I'd say** / the [/ 1] the [/ 1] the [/ 1] the [/ 1] Picasso's pieces in general / you see / but / for his importance / he has to be considered //

The prosodic form of short PAR is in line with other works (Firenzuoli and Tucci, 2003). It is characterized by a flat f0 profile that tends to be lower or higher than the level of the surrounding units. This is more evident in longer PARs, probably due to the fact that short PAR does not afford enough time for a clear switch of level. Its articulation rate is higher than that of the units that surround it, but not so high as the rate of INT. Intensity is lower than that of the surrounding units. Figure 8 below displays a short PAR unit set off from its context.

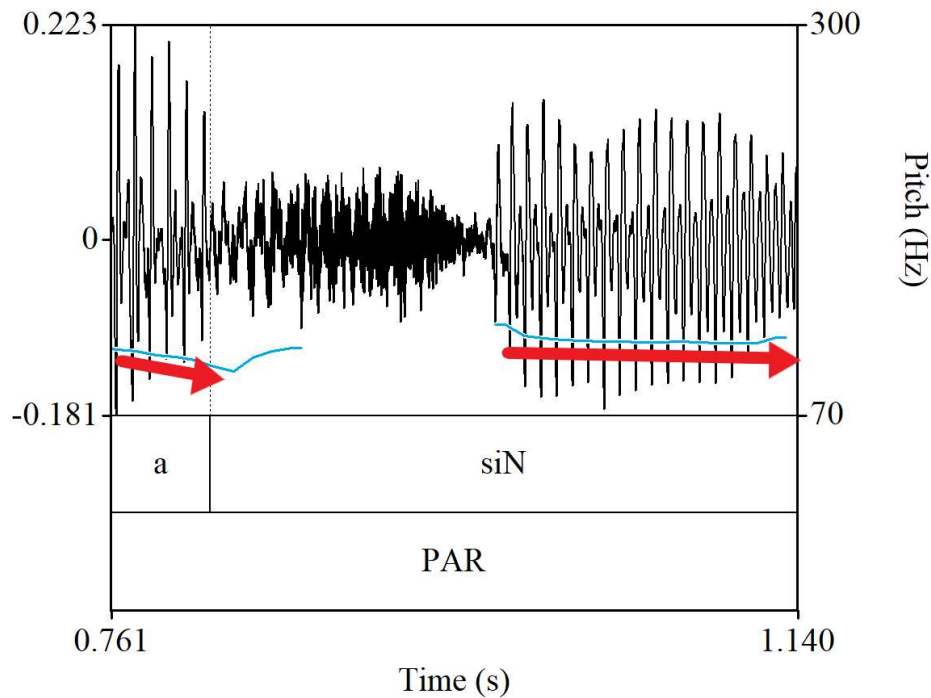


Figure 8: Example of short PAR detached from its context.

5.5 Appendix of Comment (APC)

APC makes up approximately 9% of the sample (7 out of the 75 tokens). Distributionally, it was found mostly at the end of sub-patterns and terminated sequences. When it was found in the middle of a sub-pattern, it always followed the illocutionary unit in a semifinal position, which means that it was followed by just one DU. Figure 9 displays an example of short APC in context.

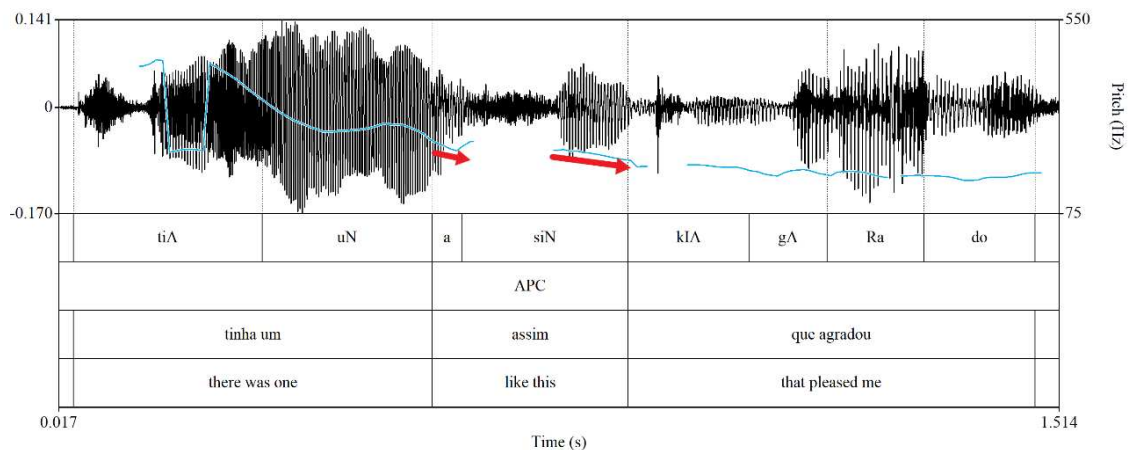


Figure 9: Example of an APC unit in context.

The transcriptions and the translation of the whole utterance can be found below.

🔊 Sound file 15 – 333_bfamdl26
 *MBA: [333] tinha um / **assim** / que agradou //

*MBA: [333] there was one / **like this** / that pleased me //

The prosodic form of APC is in line with other works (Cresti, 2000; Raso and Ulisses, 2008; Oliveira, 2012). It is characterized by a flat or falling f_0 profile, with low intensity. Figure 10, below, displays the form of APC detached from its context.

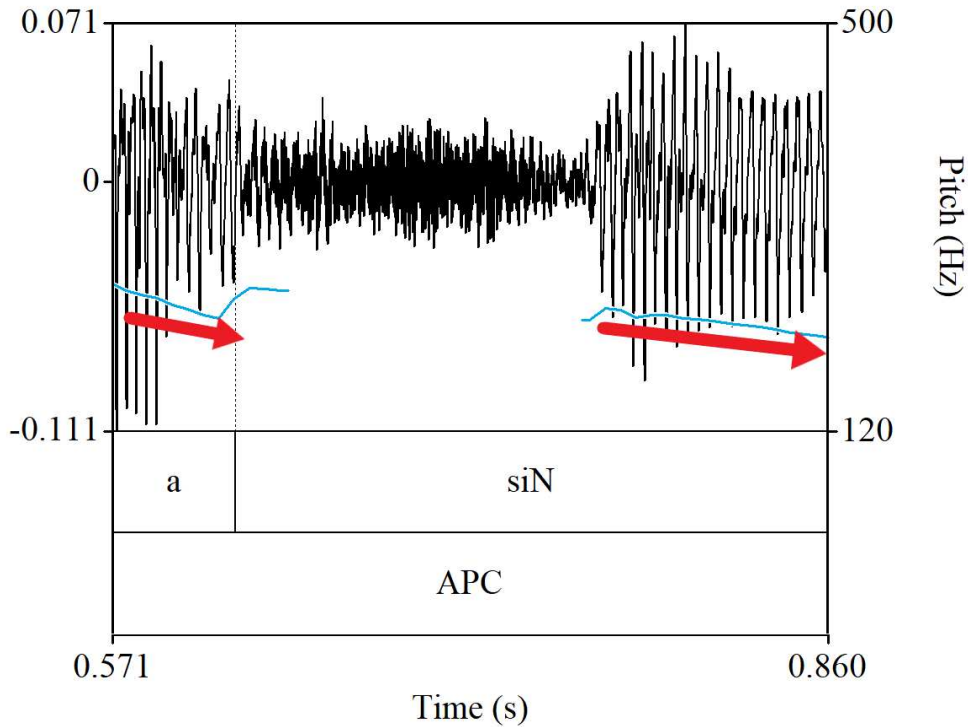


Figure 10: Example of APC detached from its context.

5.6. Other isolated cases

5.6.1 List of Topic

Lists of Topic (TPL) are a rather rare phenomenon. It is a sequence of prosodic units that together functions as one single cognitive domain for the interpretation of the illocution. In this respect it is like a regular Topic. However, it is neither composed of just one intonation unit, nor is it like multiple Topics that constitute more than one identification domain for the same illocution, as in the following case:

Yesterday /^{TOP} *John* /^{TOP} *at his house* /^{TOP} *everything got wrong* //^{COM}

In this case, there are three different identification domains: one temporal, one personal, and one local, each of which individually establishing a cognitive and prosodic pattern with the COM and carrying a TOP profile of its own. Conversely, TPL establishes a unique identification domain, made up of three semantically and prosodically compositional units; the functional prosodic profile of the TPL is carried by just one prosodic unit of the list, which happens to be the last one in the example below. In this example, ASSIM is equivalent to *etcetera*.

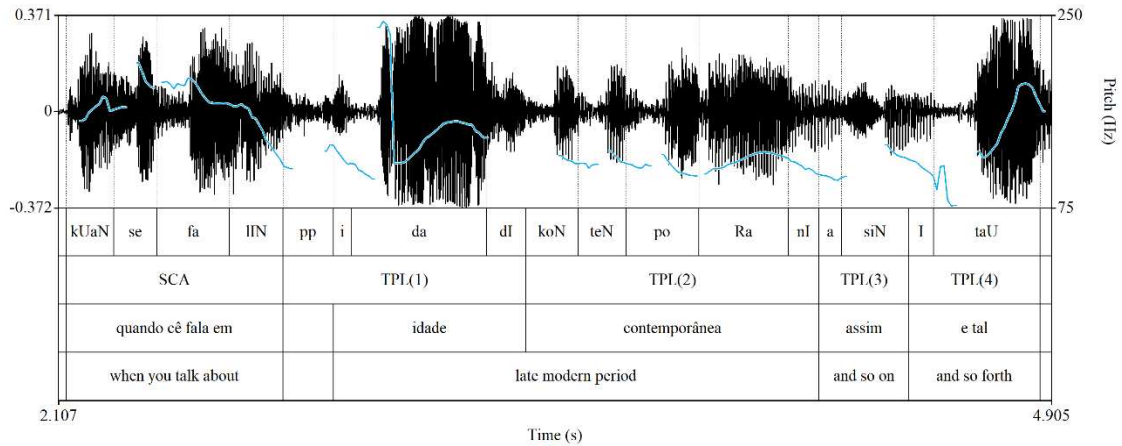


Figure 11: Example of ASSIM in a List of Topics.

🔊) Sound file 16 – 279_bfamdl09
 *LUC: [279] &he / quando cê fala em / idade / contemporânea / **assim** / e tal +
 *LUC: [279] &he / when you talk about / late modern period / **and so on** / and so forth +

5.6.2 Tokens linked to disfluency phenomena

We deal in this subsection with a few cases that just have in common the fact that they are accompanied by disfluencies. This does not mean all of them have the same function, and we cannot say much about them, given that only 3 tokens were found. One of them, which was the third one to be found, comes from the 200-token sample (dealt with in subsection 4.2), but it actually seems to be linked to the time-taking token (&he) shown in Figure 14.

The first token was found in initial position after an interrupted utterance. A pause is clearly noticeable before the lexeme ASSIM. Prosodically, it features falling f0 movements on both of its syllables. It is interesting to note that the movement of the second syllable begins higher than the end of the first one. The lengthening of the last syllable is also clearly perceivable, and it may be related to a need for some extra time to process a new program.

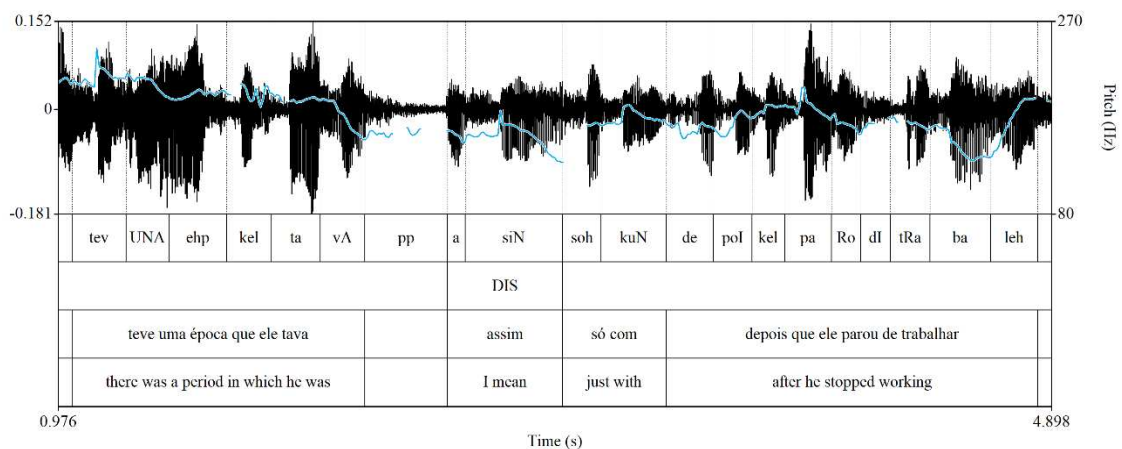


Figure 12: First disfluency-related ASSIM.

- 🔊 Sound file 17 – 189_bfamcv13
 *JON: [189] <meu vô / meu vô teve uma época> que ele tava / **assim** / só com [2]
 depois que e' parou de trabalhar / essas pessoa que vive em trecho aí / eas bebe
 demais //
 *JON: [189] <my grandpa / my grandpa there was a time when> he was / **I mean**
 / just with [2] after he stopped working / those people who are always on the road
 / they drink too much //

The second token appears in medial position and clearly functions as a transition between a retracted word and its substitute. Prosodically, it also features falling movements on both of its syllables, and again the movement of the second syllable begins higher than the end of the first one, with an even larger range difference, which may be due to the preceding and following prosodic context. No lengthening is perceivable.

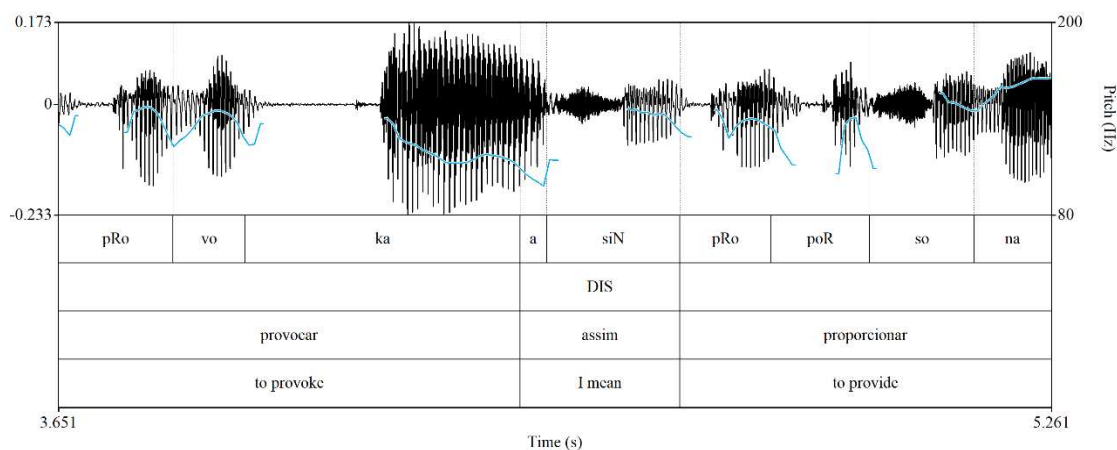


Figure 13: Second disfluency-related ASSIM.

- 🔊 Sound file 18 – 076_bpubmn13
 *JAN: [76] é justamente / a Casa Açoriana tem como idéia provocar [1] &he /
assim / proporcionar um espaço de vivência / onde / as pessoas possam retomar as
 suas raízes culturais / não é / saber que elas vieram [2] nós viemos de algum lugar
 //
 *JAN: [76] precisely / the Azorean House aims to provoke [1] &he / **I mean** / to
 provide a living space / where / people can go back to their cultural roots / you see
 / to know that they came [2] we came from somewhere //

Finally, the third token was found in medial position, but preceded only by a DU. It is preceded and followed by pause and linked inside the same intonation unit to a time-taking token (also called filled pause). Here, the need for some extra processing time is evident, too. The prosodic profile is similar to those of the first two tokens.

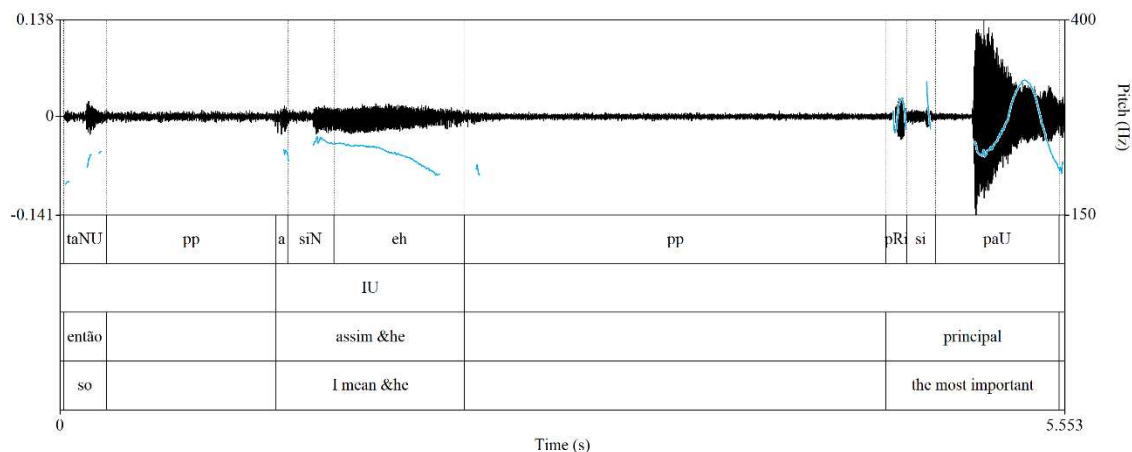


Figure 14: Third disfluency-related ASSIM.

Sound file 19 – 019_bfamnm08

*HMB: [19] então / **assim &he** / nts / principal / é a pessoa falar que que ela quer //

*HMB: [19] so / **I mean &he** / nts / the most important thing / is to say what we want //

These three tokens seem very interesting because of their functional and prosodic similarities. However, we cannot assume that they constitute a unique information unit based on the few tokens that we have found and without the support of other studies considering different lexemes.

6. When ASSIM does not occur in a dedicated intonation unit

By checking the sample used to study the different information functions of ASSIM, we observed many cases (168 out of the 305 tokens) in which at least one of the boundaries did not seem to be perceptually or phonetically justifiable, despite the fact that the general inter-rater agreement among annotators of the C-ORAL-BRASIL I corpus is very high ($k = 0.86$) according to the agreement test using Fleiss’s Kappa (1971) reported in Mello *et al.* (2012).

In order to better account for this phenomenon, we analyzed an additional sample of 200 tokens that were not annotated in dedicated units (Subsection 4.2). We draw the reader’s attention to the fact that this lexeme is very frequent in the corpus (1,366 hits or 65 occurrences per 10k words). If we exclude other functional words, the negation (bp. *não*, “no” or “not”), and bp. *ai* “so” and bp. *né* “huh” or “isn’t it” – two typical discourse markers –, the only words that have a frequency as high as that of ASSIM are the locative adverbs bp. *lá* “there” and bp. *aqui* “here” (for further information on the frequency of lexemes, we refer the reader to the Appendix of the C-ORAL-BRASIL I corpus). The results of this analysis are summed up in the second column of Table 2 (Sample 2), which is organized as follows:

a) Part A refers to tokens with full lexical meaning; Part B refers to tokens with weakened meaning, functioning as modalizers, attenuators, and other non-canonical values;

b) the first column specifies the positions in which the tokens may occur within terminated sequences and intonation units;

c) the second column refers to the so-called Sample 1. Sample 1 is comprised of the tokens that had been originally annotated in dedicated intonational units, but which were excluded from our analysis after the sample revision, since one or both of their boundaries were not attested;

c) the third column presents the data of the so-called Sample 2. Sample 2 is comprised of tokens from the randomized sample of ASSIM originally annotated in non-dedicated intonational unit (as presented in Subsection 4.2), except for the 3 tokens that, after our revision, were identified as actually being in dedicated units. Therefore, only 197 tokens are tallied in Table 1;

d) the fourth column provides row-wise totals. Column-wise subtotals (Parts A and B) and totals are also provided.

Table 2: Position of ASSIM in non-dedicated IUs by sample and compositionality

| A. Full lexical meaning tokens⁵ | | | |
|--|-----------------|-----------------|--------------------|
| Position | Sample 1 | Sample 2 | Total (1+2) |
| Beginning of terminated sequence | 1 | 7 | 8 |
| Beginning of intonation unit | 9 | 2 | 11 |
| Middle of intonation unit | 2 | 22 | 24 |
| End of intonation unit | 30 | 100 | 130 |
| End of terminated sequence | 2 | 39 | 41 |
| Subtotal (A) | 44 | 170 | 214 |
| B. Weakened lexical meaning tokens (modalizers, conclusion markers, and other non-canonical uses) | | | |
| Beginning of terminated sequence | 1 | 0 | 1 |
| Beginning of intonation unit | 23 | 0 | 23 |
| Middle of intonation unit | 30 | 5 | 35 |
| End of intonation unit | 46 | 17 | 63 |
| End of terminated sequence | 24 | 5 | 29 |
| Subtotal (B) | 124 | 27 | 151 |
| Total (A+B) | 168 | 197 | 365 |

Based on the segmentation error rate of Sample 2 (1.5%) and on the transcription error rate (0.4%) observed in all items checked in this work, we can estimate that approximately 1,210 items are not in dedicated intonation units. This estimation is summarized in Table 2.

⁵ Among the tokens considered as having full lexical meaning are included 4 items that can be interpreted in both ways, with full or weakened lexical meaning. We decided to include them in the group with full lexical meaning tokens, because this goes against our hypothesis. Therefore, we adopted a conservative procedure.

Table 3: Estimation of tokens not in dedicated units

| A. Tokens checked in this work | | | |
|---|-----------------|----------------------------------|---|
| | Total tokens | Attested in dedicated unit | Not attested in dedicated unit |
| Sample 1 | 302 | 132 | 70 |
| Sample 2 | 200 | 3 | 197 |
| Illocutionary items | 5 | 5 | 0 |
| Attested tokens in dedicated units (I) | | | 367 |
| B. Tokens not checked in this work | | | |
| Total not checked | | | 859 |
| Transcription error rate (0.4%) | | | -3 |
| Boundary annotation error rate (1.5%) | | | -13 |
| Estimated to be in dedicated units (II) | | | 843 |
| Estimation of tokens not in dedicated units (I + II) | | | 1210 |

In this section, we are especially interested in observing the behavior of ASSIM when it is used with weakened lexical meaning and is not placed in a dedicated unit (Part B of Table 2). Firstly, we examined the items from the 200-token sample (13.7% of the 200-token sample). No item was found at the very beginning of the intonation unit. 18.25% occurs in medial position. 22 tokens (81.5%) were found at the end of the intonation unit, either at the end of a sub-pattern (22 tokens) or at the end of the terminated sequence (5 tokens).

If we consider the original 302-token sample⁶ that annotators had originally transcribed in dedicated units and look for those tokens of each at least one boundary was excluded, we obtain 168 tokens (55.63% of the original annotation). 124 of them (73.81%) do not exhibit full lexical meaning, of which 24 were found at the very beginning of the intonation unit (just one was found at the beginning of the terminated sequence). 30 tokens were found in the middle a of the intonation unit (this meaning that both boundaries were excluded). Finally, 70 tokens were found at the very end of the intonation unit (24 tokens at the end of the terminated sequence). Therefore, 56.45% of these 124 tokens seem just to mark the end of the IU. Another interesting observation about these 124 tokens is that 19.35% of them appear in initial position. This was not observed in the 200-token sample.

We also observed an intriguing characteristic in the tokens used with weakened lexical meaning. Put together, the two samples confirm a strong tendency for ASSIM to occur not only in final position but also, even if less frequently, in initial position. In both positions, these tokens frequently behave as a sort of morpheme that marks the boundary either at the end or at

⁶ 3 tokens from the 200-token sample were eventually added to this sample, add up to the 305 tokens described in subsection 4.2.

the beginning of the IU. We believe the lexeme is undergoing a grammaticalization process, which, in face of the distribution observed, may be more advanced at the end of the IU.

There is more to support this hypothesis. A considerable number of the tokens that appear in initial or final position of intonation units are in a cliticized form, having lost the stress on the second syllable. It seems that this cliticization process is more frequent when ASSIM appears at the beginning of the unit. In any case, this process and, more in general, what is presented in this section need to be studied more closely, which cannot be done here, since this analysis does not constitute the main goal of this paper.

The following examples illustrate cliticized tokens of ASSIM. The first and the second ones in initial position and the second one in final position. We advise the reader not to take the translations of ASSIM to English literally since even in BP it is difficult to retrieve any semantic value from the cliticized forms, in spite of hearing the audios.

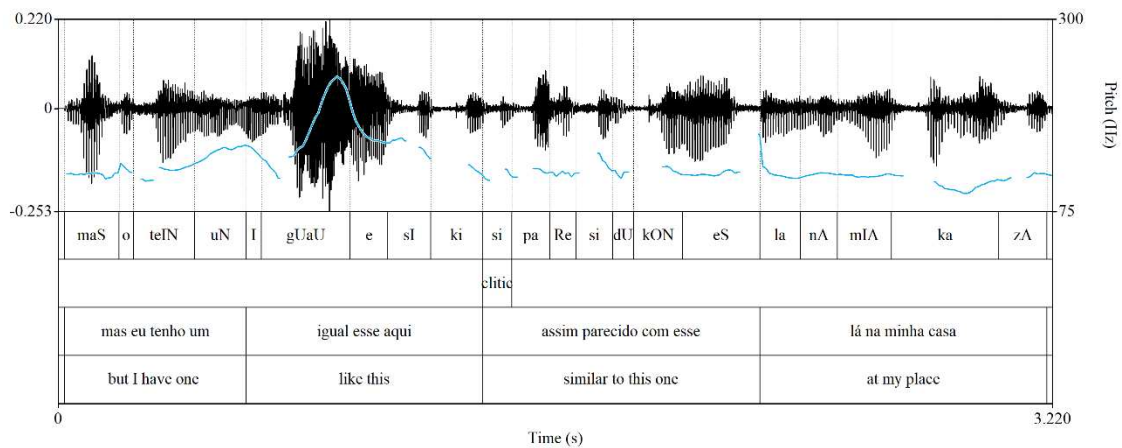


Figure 15: Example of cliticized ASSIM in initial position.

🔊) Sound file 20 – 060_bfammn33
 *ADR: [60] mas eu tenho um / igual esse aqui / **assim** parecido com esse / lá na minha casa / só com coisas da Madonna / que ele me deu de presente //
 *ADR: [60] but I have one / like this one here / **somehow (cliticized item)** similar to this one / at my place / only with Madonna stuff / that he gave me as a gift //

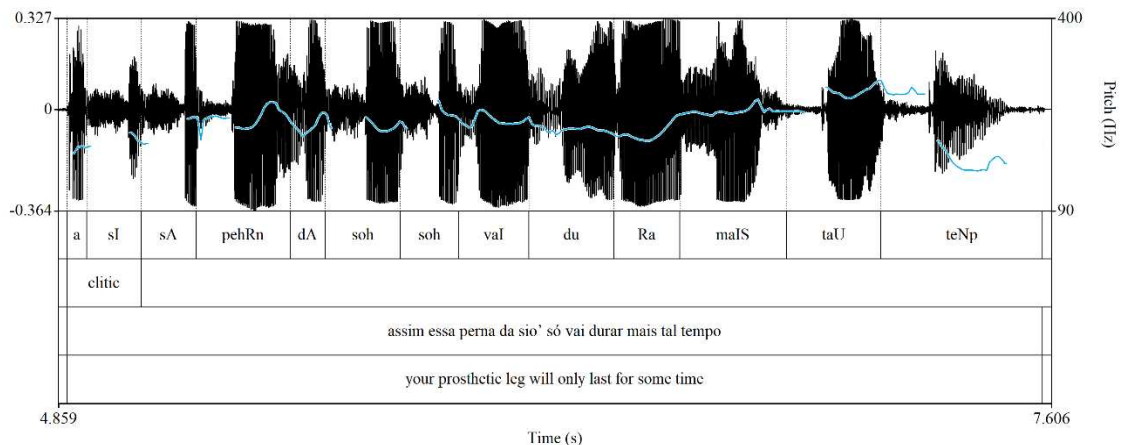


Figure 16: Example of cliticized ASSIM in initial position.

- 🔊) Sound file 21 – 156_bfamdl25
 *LIA: [156] então es vão consertando a perna dela / consertando a perna / e fala assim / o' / **assim** essa perna da sio' só vai durar mais tal tempo //
 *LIA: [156] so they'll keep fixing her prosthetic leg / fixing the leg / and they say / like (*cliticized item*) your prosthetic leg will only last for some time //

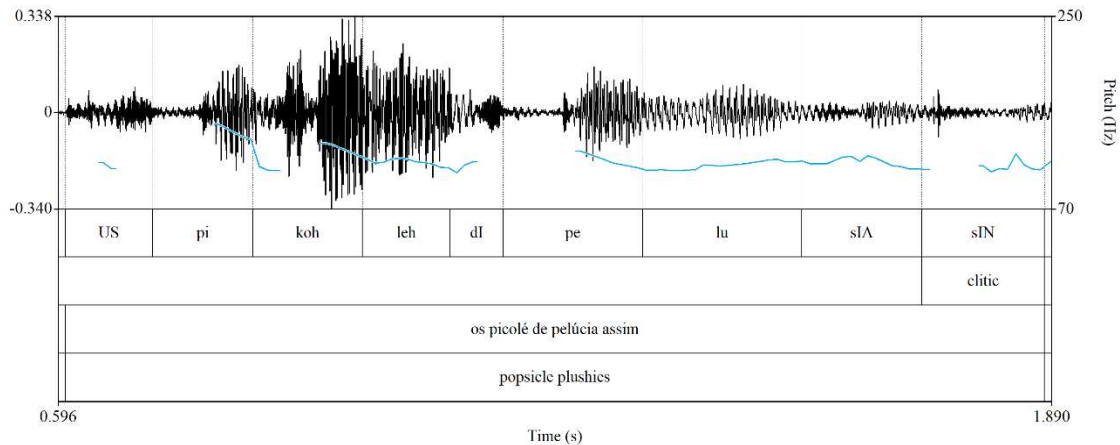


Figure 17: Example of cliticized ASSIM in initial position.

- 🔊) Sound file 22 – 673_bfamdl09
 *LUC: [673] os picolé de pelúcia **assim** //
 *LUC: [673] popsicle plushies (*cliticized item*) //

Finally, it is worth mentioning that the data in Table 2 clearly supports the hypothesis put forth in Subsection 4.2. The hypothesis is that annotators tend to incorrectly insert ASSIM in prosodically dedicated units when the item is used with weakened lexical meaning (i.e. less compositionally). It is worth repeating that Sample 1 is comprised of items in prosodic units which had at least one of their originally annotated boundaries called into question, and that Sample 2 is comprised of items that had not been originally annotated between boundaries. The two samples display exactly the opposite balancing with respect to the frequency of items used with full lexical meaning (Part A) and weakened lexical meaning (Part B). While Sample 1 has only 44 tokens (26.2%) used with full lexical meaning and 124 tokens (74.8%) used with weakened lexical meaning, Sample 2 contains 170 tokens (86.3%) with full lexical meaning and only 27 with weakened lexical meaning (13.7%).

7. Summary

This section constitutes a summary of the most relevant aspects discussed in this paper, particularly those related to:

- i) the five prototypical intonational forms associated with the five information functions that have been the focus of our analysis, namely, CNT, the new dialogic unit, which is characterized by a rising f0 movement, INT, PAR, and APC, plus a sixth possible form associated with the three tokens treated in subsection 5.6, in which ASSIM is accompanied by disfluency phenomena (Figure 18);

ii) the information displayed in Table 3 relative to the frequencies of each information unit in our sample and their distribution within terminated sequences (Table 3).

The five IUs feature clearly distinct intonation forms, except for PAR and APC when compared to each other. This represents a problem only when PAR is positioned after an illocutionary unit, since APC necessarily occurs in this very position. The functions of these two IUs are, in principle, different, but it is not easy to clearly distinguish them based on as vague a lexeme as ASSIM. Therefore, what we have said about APC and PAR tokens occurring after illocutionary units should be taken with a grain of salt, as a more accurate distinction will only become possible once studies considering different and less vague lexical items have been conducted.

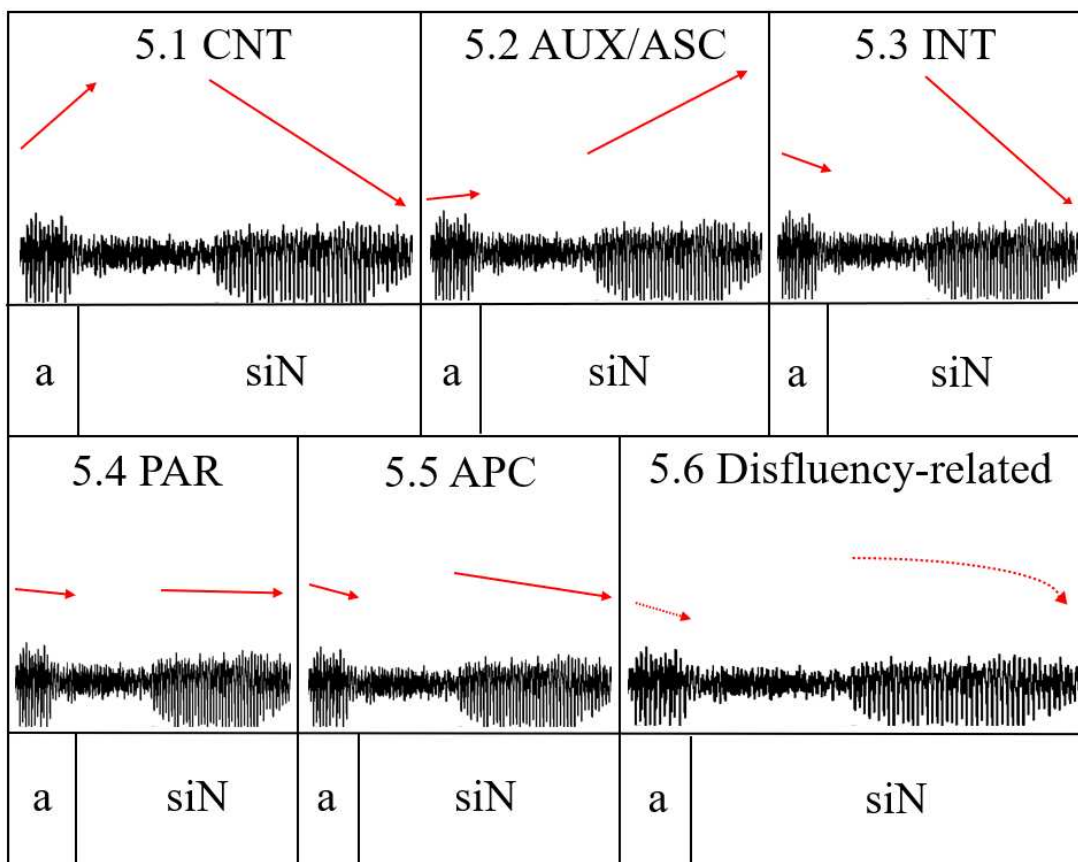


Figure 18: Summary of the prosodic forms and functions associated with.

With respect to duration and intensity, we only have a few comments to make. CNT tends to have high intensity and short duration. The rising dialogic unit (AUX/ASC) tends to show low intensity and short duration. INT features very low intensity and tends to be very short, frequently showing a reduced phonetic realization. PAR tends to have lower intensity and shorter duration with respect to its surrounding units. APC tends to have much lower intensity with respect to its preceding unit (which is very frequently an illocutionary unit). Finally, the disfluency-related units are characterized by very long duration. Of course, one must also

consider the position of these different units. For instance, APC tends to occur at the end of the utterance or sub-pattern, a position that contributes for a longer duration.

Table 3 provides a complete picture of the different IUs occurring in our sample. It shows that ASSIM can fulfill many information functions with different frequencies and distributions. CNT and PAR are the most frequent functions, but their distributional patterns are quite different, since PAR cannot appear in initial position, which is the most frequent position for CNT. The rising dialogic unit (AUX/ASC) is also frequent in the sample, and, as already mentioned, it must be better accounted for by comparing what we found for ASSIM to other lexemes potentially performing the same function. Given its function, it is not surprising that AUX/ASC was not found in initial position. APC occurs always in final position, except for 2 tokens which are followed by dialogic units. It is worthwhile to note that we found a token of TPL, which is a rather rare IU. This suggests that the lexeme ASSIM may also perform other functions not found in our sample.

Table 3: Frequency and distribution of each IU

| Position within the terminated sequence | CNT | AUX/ASC | INT | PAR | APC | DIS | TPL | Row totals |
|--|------------|----------------|------------|------------|------------|------------|------------|-------------------|
| Initial | 13 | 0 | 3 | 0 | 0 | 0 | 0 | 16 |
| Sub-pattern | Initial | 0 | 0 | 1 | 0 | 0 | 1 | 2 |
| | Medial | 11 | 7 | 2 | 17 | 2 | 1 | 41 |
| | Final | 0 | 1 | 0 | 5 | 1 | 0 | 7 |
| Final | 1 | 2 | 0 | 2 | 4 | 0 | 0 | 9 |
| Column totals | 25 | 10 | 6 | 24 | 7 | 2 | 1 | 75 |

8 Final remarks

Our main goal in this paper was to show how lexicon and syntax constitute weak correlates when we want to define the informational status of a stretch of speech. Prosody, on the other hand, provides the necessary formal features for the informational status to be conveyed. It allows us not only to identify the sequence of words that constitutes an IU, thanks to speech segmentation, but also to interpret the specific function of an IU by means of intonational, durational and intensity parameters. We carried out our analysis by confronting SIUs and using ASSIM, one very frequent lexeme extracted from a spontaneous speech corpus.

Of course, it will be important to confirm our proposal for SIUs using different lexical items. At least two IUs we have identified require a clearer description. On the one hand, the vagueness of the lexeme ASSIM makes it highly frequent and, therefore, especially appropriate for a first analysis; on the other hand, this very vagueness also represents a problem when it comes to the interpretation of the functional value. Since it is way too vague, the semantic content of ASSIM does allow for a sharp functional distinction when prosodic and distributional correlates do not set a clear direction. This is what happens when we need to distinguish between APC and PAR in post-illocutionary position. Furthermore, the so-called rising dialogic unit (AUX/ASC) must be confirmed and better accounted for by studying different lexemes that display the same function and prosodic features.

Examining the tokens of ASSIM that had originally been annotated in dedicated intonation units, we also chanced upon a very intriguing problem: many of them proved to be overannotated with respect to prosodic boundaries. This mostly occurred when the lexeme

ASSIM was being used with a somewhat weakened lexical meaning, functioning as a very vague modalizer, an attenuator or with other values, sometimes being ambiguous or even completely desemantized. We decided, therefore, to look into a sample of tokens that had not been annotated in dedicated unit. We observed that the vast majority of the tokens with weakened lexical meaning were found at the very end of an IU and, less frequently, at the very beginning of an IU. We also found tokens of ASSIM in these peripheral positions, and especially in initial position, tending to lose stress and becoming a clitic. We proposed that in such cases this lexeme is acquiring the function of morphologically marking the boundary of an IU. As already said, analyzing these cases was not our main goal and this hypothesis need to be checked.

In conclusion, it will be important to carry out in-depth analyses of SIUs by looking at different lexemes. Likewise, all the prosodic analyses will have to be confirmed using more sophisticated techniques, taking automatic phonetic measurements, as well as other statistic techniques such as data normalization. This was an exploratory study that, we believe, offers interesting hypotheses to be tested.

The authors of this paper are grateful to Fapemig for financing the research.

REFERENCES

1. Austin J L. *How to do things with words*. Cambridge: Cambridge University Press, 1962.
2. Averintseva-Klisch M. *German right dislocation and afterthought in discourse*. Amsterdam-Philadelphia: John Benjamins, 2008.
3. Barbosa P. *Incursões em torno do ritmo da fala*. São Paulo: Fapesp, 2006.
4. Barbosa P, Raso T. *A segmentação da fala espontânea: aspectos prosódicos, funcionais e aplicações para a tecnologia*. In: Revista estudos da linguagem, Belo Horizonte, 2018.
5. Boersma P, Weenink D. *Praat: doing phonetics by computer [Computer program]*. Version 6.1.08, retrieved 5 December 2019. Available at <<http://www.praat.org/>>.
6. Barth-Weingarten D. *Intonation units revisited*. Cesura in talk-in-interaction. Amsterdam: John Benjamins, 2016.
7. Blanche-Benveniste C. *Macro-syntaxe et micro-syntaxe : les dispositifs de la rection verbale*. Macro-syntaxe et macro-sémantique. Actes du colloque international d'Århus (17-19 mai 2001), Andersen, H. L., Nølke, H. (eds.). Bern: Peter Lang, 2002: 95-118.
8. Bolinger D. *Intonation and its uses: melody on grammar and discourse*. Stanford: Stanford University Press, 1989.
9. Cavalcante F. *The topic unit in spontaneous American English: a corpus-based study*. Master Thesis. Federal University of Minas Gerais, Belo Horizonte, 2016.
10. Cavalcante F. *The information unit of topic: a crosslinguistic, statistical study based on spontaneous speech corpora*. PhD Thesis. Federal University of Minas Gerais, Belo Horizonte, 2020.
11. Cavalcante F, Ramos A. *The American English spontaneous speech minicorpus*. Architecture and comparability. In: *CHIMERA: Romance Corpora and Linguistic Studies*. 3.2, 2016: 99-124.
12. Chafe W. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago, IL: The University of Chicago Press, 1994.
13. Corsi G. *L'introduttore locutivo: una ricerca corpus-based di italiano parlato informale*. Tesi di Laurea triennale. Facoltà di Lettere e Filosofia, Università degli Studi di Firenze, Firenze, 2009.
14. Cresti E. *Corpus di Italiano parlato*. Firenze: Accademia della Crusca, 2000.
15. Cresti E. *La Stanza: un'unità di costruzione testuale del parlato*. In: Ferrari, A. (Ed.), *Sintassi storica e sincronica dell'italiano. Subordinazione, coordinazione, giustapposizione*. Atti del X Congresso della Società Internazionale di Linguistica e Filologia Italiana, vol. 2. Firenze: Firenze University Press, 2009: 713–732.

16. Cresti E. *Syntactic properties of spontaneous speech in the Language into Act Theory: Data on Italian complements and relative clauses*. In: Raso T, Mello H (Eds.), *Spoken corpora and linguistic studies*. Amsterdam: John Benjamins, 2014: 365-410.
17. Dehé N. *Parentheticals in Spoken English: The Syntax-Prosody Relation*. Cambridge: Cambridge University Press, 2014.
18. Du Bois J, Schuetze-Coburn S, Cumming S, Paolino, D. *Outline of Discourse Transcription*. In: Edwards J, Lampert M (Eds). *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1993.
19. Du Bois J, Chafe W, Meyer C, Thompson S, Englebretson R, Martey N. *Santa Barbara corpus of spoken American English*, Parts 1-4. Philadelphia: Linguistic Data Consortium, 2000-2005.
20. Espinal M. *The representation of disjunct constituents*. In: *Language*, Vol. 67, 1991: 726-62.
21. Firenzuoli V. *Le forme intonative di valore illocutivo dell'Italiano parlato: Analisi sperimentale di un corpus di parlato spontaneo*. PhD dissertation. LABLITA, University of Florence. Florence, 2003.
22. Firenzuoli V, Tucci I. *L'unità informativa di inciso: correlati intonativi*. In: *Atti delle XIII giornate del Gruppo di Fonetica Sperimentale (GFS), ETS, Pisa, 2003:185-192*.
23. Fleiss J. *Measuring nominal scale agreement among many raters*. In: *Psychological Bulletin*, 76(5), 1971:378-382.
24. Frosali F. *Le unità di informazione di ausilio dialogico: valori percentuali, caratteri intonativi, lessicali e morfo-sintattici in un corpus di italiano parlato (C-ORAL-ROM)*. In: Cresti E. (Ed.). *Prospettive nello studio del lessico italiano*. Firenze: Firenze University Press, 2008:417-424.
25. Giani D. *Le discours directe rapporté dans l'italien parlé et écrit*. In: Scarano A (Org.). *Macrosyntaxe et Pragmatique: l'analyse de l'oral*. Roma: Bulzoni, 2003:203-213.
26. Giani D. *Una strategia di costruzione del testo parlato: l'introduttore locutivo*. In: Leoni F (Org.). *Il parlato italiano. Atti del congresso*. Napoli: D'Auria, 2004:84-97.
27. Gobbo O. *Marcadores discursivos em uma perspectiva informacional: análise prosódica e estatística*. Master Thesis. Federal University of Minas Gerais. Belo Horizonte, 2019.
28. Gordon B. *Antitopic and afterthought Givenness and grammar in Ponca and Omaha*. In: *Kansas Working Papers in Linguistics*, Vol. 30, 2008:74-88.
29. Haegeman L. *Parenthetical adverbials: The radical orphanage approach*. In: Chiba S, Shuki A, Ogawa A, Fuiwara Y, Yamada N, Koma O, Yagi T (Eds). In: *Aspects of Modern Linguistics: Papers Presented to Masatomo Ukaji on his 60th Birthday*. Tokyo: Kaitakushi, 1988:232-254.
30. Izre'el Sh, Mello H, Panunzi A, Raso T (Eds.). *In Search of a Basic Unit of Spoken Language: A Corpus-driven Approach*. Amsterdam-Philadelphia: John Benjamins, 2020a.
31. Izre'el Sh, Mello H, Panunzi A, Raso T. *In search of a basic unit of spoken language: Segmenting speech*. In: Izre'el Sh, Mello H, Panunzi A, Raso T (Eds.). *In Search of a Basic Unit of Spoken Language: A Corpus-driven Approach*. Amsterdam-Philadelphia: John Benjamins, 2020b.
32. Kaltenböck G, Heine B, Kuteva T. *On thetical grammar*. *Studies in Language*, vol. 35 (4), 2011:852-897.
33. Kavalova Y. *And-parenthetical clauses*. In: Dehé N, Kavalova Y. *Parentheticals*. In: *Linguistics Today*, vol. 106. Amsterdam-Philadelphia: John Benjamins, 2007.
34. Lombardi Vallauri E. *“Appendix” or “Postposed Topic”: where does the difference lie?* In: Mereu L (Org.). *Information structure and its interfaces*. Berlin: De Gruyter Mouton: 2009:387-411
35. Maia Rocha B. *A unidade informacional de introdutor locutivo no português brasileiro: uma análise baseada em corpus*. Master Thesis. Federal University of Minas Gerais. Belo Horizonte, 2011.
36. Maia Rocha B, Raso T. *A unidade informacional de Introdutor Locutivo no português do Brasil: uma primeira descrição baseada em corpus*. In: *Domínios da Lingu@gem*, Vol 5, No 1, 1º Sem, 2011:327-343.

37. Mello H. *Methodological issues for spontaneous speech corpora compilation*. In: Raso T, Mello H (Ed.). *Spoken corpora and linguistics studies*. Amsterdam-Philadelphia: John Benjamins, 2014:411-467.
38. Mello H, Raso T. *Illocution, Modality, Attitude: Different Names for Different Categories*. In: Mello H, Panunzi A, Raso T (Eds.). *Pragmatics and Prosody. Illocution, Modality, Attitude, Information Patterning and Speech Annotation*. Firenze: University Press, 2011:1-18.
39. Mithun M. *Syntactic and prosodic structures: Segmentation, integration, and in between*. In: Raso T, Mello H (Eds.). *Spoken corpora and linguistics studies*. Amsterdam-Philadelphia: John Benjamins, 2014:297-330.
40. Mithun M. *Prosody and the organization of information in Central Pomo*. In: Izre'el Sh, Mello H, Panunzi A; Raso T (Eds.). *In Search of a Basic Unit of Spoken Language: A Corpus-driven Approach*. Amsterdam-Philadelphia: John Benjamins, 2020.
41. Mittmann M. *O corpus C-ORAL-BRASIL e a análise da fala informal: Um novo olhar sobre o Tópico no Português do Brasil*. PhD dissertation. Federal University of Minas Gerais. Belo Horizonte, 2012.
42. Moneglia M, Raso T. *Notes on the Language into Act Theory*. In: Raso T, Mello H (Eds.). *Spoken corpora and linguistics studies*. Amsterdam-Philadelphia: John Benjamins, 2014:468-494.
43. Paixão V, Serra C. *Fraseamento prosódico de estruturas parentéticas em dados de leitura no português do Rio de Janeiro*. In: *Working Papers em Linguística*, 19 (2), 2018:113-135.
44. Raso T. *O corpus C-ORAL-BRASIL*. In: Raso T, Mello H. *C-ORAL-BRASIL: corpus de referência do português brasileiro falado informal*. I. Belo Horizonte: UFMG, 2012a.
45. Raso T. *O C-ORAL-BRASIL e a Teoria da Língua em Ato*. In: Raso T, Mello H. *C-ORAL-BRASIL: corpus de referência do português brasileiro falado informal*. I. Belo Horizonte: UFMG, 2012b.
46. Raso T, Cavalcante F, Mittmann M. *Prosodic forms of the Topic information unit in a cross-linguistic perspective: a first survey*. In: SLI-GSCP International Conference. Proceedings. 13-15 June, 2016. Anna de Meo & Francesca Maria Dovetto. Rome: Aracne editrice, 2017.
47. Raso T, Ferrari L. *Uso dei Segnali Discorsivi in corpora di parlato spontaneo italiano e brasiliano*. In: Ferroni R, Birello M (Orgs). *La competenza discorsiva a lezione di lingua straniera*. Roma: Aracne. Forthcoming.
48. Raso T, Mello H. *C-ORAL-BRASIL: corpus de referência do português brasileiro falado informal*. I. Belo Horizonte: UFMG, 2012.
49. Raso T, Mello H (Eds.). *Spoken corpora and linguistics studies*. Amsterdam-Philadelphia: John Benjamins, 2014.
50. Raso T, Viera M. *A description of Dialogic Units/Discourse Markers in spontaneous speech corpora based on phonetic parameters*. In: CHIMERA: Romance Corpora and Linguistic Studies, Vol. 3, No 2, 2016.
51. Santos S. *Parentético e outras unidades informacionais curtas: uma distinção através da análise das diferentes realizações prosódicas do lexema assim*. Master Thesis. Federal University of Minas Gerais. Belo Horizonte, 2020.
52. Shimoyama J, Drummond A, Schwarz B, Wagner M. *Movement and ellipsis in afterthought right dislocation, sluicing and fragments*. In: Venerte S (Ed.). 2015 Annual Conference of the Canadian Linguistic Association. Proceedings. Available at: <http://cla-acl.ca/wp-content/uploads/Shimoyama_Drummond_Schwarz_Wagner-2015.pdf>.
53. Schneider S. *Reduced Parenthetical Clauses as Mitigators: A corpus study of spoken French, Italian and Spanish*. Amsterdam-Philadelphia: John Benjamins, 2007.
54. Schneider S. *Parenthesis: fundamental features, meanings, discourse functions and ellipsis*. In: Marlies K, Ott D, De Vries M (Eds.). *Parenthesis and ellipsis: cross-linguistic and theoretical perspectives*. Berlin: De Gruyter Mouton, 2014:277-300.
55. Sperber D, Wilson D. *Relevance: communication and cognition*. Oxford: Basil Blackwell, 1986.

Short information units: a corpus-based prosodic study on the lexeme “assim” in Brazilian Portuguese

56. Tucci I. *L'inciso: caratteristiche morfosintattiche e intonative in un corpus di riferimento*. In: Albano F, Cutugno F, Pettorino M, Savy R (Orgs.). *Il parlato italiano. Atti del Convegno Nazionale*, D'Auria Editore, Napoli, 2004:1-14.
57. Tucci I. *Obiter dictum: La funzione informativa delle unità parentetiche*. In: Pettorino M, Giannini A, Dovetto F (Eds.). *La Comunicazione parlata 3. Atti del GSCP, vol. 3*. Napoli: Università l'Orientale Press, 2009:635–654.
58. Urmson J. *Parenthetical verbs*. In: *Mind*, Vol 61, 1952:480–496.