

MELLO, Heliana; SILVA, Luis Filipe Lima e. Focus and information patterning: refining terminology and distinguishing categories in a spoken corpus. *ReVEL*, edição especial n. 10, 2015. [www.revel.inf.br].

FOCUS AND INFORMATION PATTERNING: REFINING TERMINOLOGY AND DISTINGUISHING CATEGORIES IN A SPOKEN CORPUS

Heliana Mello^{1,2}

Luis Filipe Lima e Silva^{1,3}

hmello@ufmg.br

luisf.1397@gmail.com

RESUMO: Comumente atribui-se a inepção do estudo da estrutura informacional da fala às noções de *point de départ* e *but du discours* de Weill (1844) e sujeito psicológico e predicado psicológico (PAUL, 1880; GABELEENZ, 1891). Posteriormente, a literatura linguística, desde a proposta de articulação da frase em tema e rema no contexto da Escola de Praga (cf. AMMANN, 1928; MATHESIUS, 1929), tem apresentado diferentes possibilidades terminológicas para categorias linguísticas geralmente consideradas semelhantes em teorias distintas. As noções de tema e rema mantiveram-se em alguns quadros teóricos e foram transpostas para outros sob os rótulos de tópico e comentário (HOCKETT, 1958; CHAFE, 1970; GUNDEL, 1977); pressuposição e foco (CHOMSKY, 1971; JACKENDOFF, 1972), bem como para ainda outras abordagens como dado e novo (HALLIDAY, 1976), e tópico e foco (LAMBRECHT, 1994). Assim, fala-se crucialmente de categorias que se aproximam, mas que não são idênticas, no contexto de teorias de estruturação informacional, estruturação funcional da sentença, articulação tópico-foco, dentre outras possibilidades. Tais categorias são por vezes reconhecidas em níveis analíticos distintos, por exemplo, quando se reconhece a noção de foco expresso através de elementos prosódicos, morfológicos, lexicais ou sintáticos, para além da organização mais tradicional da estrutura dado-novo, que encontraria no componente novo, o elemento foco. Neste artigo, adotando a Teoria da Língua em Ato (CRESTI, 2000) como arcabouço teórico e dados do C-ORAL-BRASIL (RASO & MELLO, 2012) – um corpus de fala espontânea do português brasileiro – propomos que a noção de foco prosódico é exclusiva a unidades informacionais textuais, nomeadamente o tópico e o comentário e tem correlatos funcionais no reconhecimento destas, enquanto que os fenômenos de focalização e de ênfase possuem uma natureza diversa, qual seja, a de destacar elementos linguísticos, e se aplica em qualquer unidade informacional através de recursos linguísticos distintos.

Palavras-chave: Foco; Focalização; Ênfase; Teoria da Língua em Ato; Fala Espontânea; Padronização Informacional.

¹ Universidade Federal de Minas Gerais - UFMG

² H. Mello is grateful for research grants from CNPq and FAPEMIG.

³ L. F. L. Silva is grateful for a graduate scholarship grant from CAPES.

INTRODUCTION

In this paper we discuss the notion of focus as it is referred to in the Language into Act Theory (henceforth, LAcT), as proposed by Cresti (2000), and situate it within the theoretical premises that place it within textual information units (topic and comment) while differentiating it from the notion of comment, as assumed by other theories (cf. CHOMSKY 1971, JACKENDOFF 1972, WARD 1985, PRINCE 1986, VALLDUVI 1990). Additionally, we distinguish the notions associated to focus, focalization and emphasis which can be confused in the literature. We support our claims through the analysis of examples extracted from the C-ORAL-BRASIL corpus (RASO & MELLO, 2012). The motivation for the discussion put forth in this paper stems from the fact that different theories use the same terminology to refer to phenomena which may be very dissimilar, while distinct terms might be used to refer to approximate notions. Therefore we hope to clarify both terminology and notions adopted in the linguistic literature associated to focus.

The paper is organized around the following topics: the development of the notion of focus; information patterning and focus; focalization and emphasis, revisiting focus.

1. INFORMATION ORGANIZATION: SOME BACKGROUND

Notions referring to observable differences between the syntactic and the information organization of linguistic expression have been referred to in the literature for well over a century. References stemming from discourse level phenomena are made to Weill (1844) who coined the terms *point de départ* and *but du discours* as well as to German linguists Gabelentz (1868) and Paul (1880) who identified the notions of psychological subject and predicate, the very notions that later would be called in the Praguian tradition and its heirs as theme and rheme. This latter terminology has evolved within different 20th century schools of linguistics, which might associate syntactic/phonological and semantic/pragmatic functions in their analytical proposals, namely generative and functionalist approaches such as Chomsky (1971), Givón (1983), Gundel (1974/ 89), Halliday (1967), Jackendoff (1972), Kuno (1972, 1976), Kuroda (1965, 1972), and Reinhart (1981), among others.

Analogically speaking, the well-established syntactic notions of subject and predicate would be paralleled in information structuring by a bipartite organization that can be referred

to by multiple terms such as theme and rheme (AMMANN 1911-1912, 1928; MATHESIUS 1928, FIRBAS 1966, 1972, HALLIDAY 1967), logical subject and predicate (CHAO,1968), presupposition and focus (CHOMSKY 1971, JACKENDOFF 1972), topic and comment (GUNDEL 1974, 1978), open proposition and focus (WARD 1985, PRINCE 1986), and ground and focus (VALLDUVI 1990). What seems to unite such different terminology and the notions they stand for is the perception that an utterance joins information which is, in some sense, new to presupposed knowledge. Under this general (and quite imprecise) view, the new information in some way would need to semantically stand out from what is already known to a hearer, hence the term “focus” finds its way into information structuring. Focus (or comment or rheme) would be the information unit that carries new information in an utterance and is, therefore, cognitively salient. As will be seen in section 2, LAcT proposes the dissociation between the notion of focus and the introduction of new information. For LAcT, focus is actually a semantic phenomenon that licenses both topic and comment by marking their prosodic cusp; therefore, it does not coincide with the notion represented by comment in other theoretical frameworks, as will be seen below. In the LAcT view, thus, focus is a prosodically marked semantic property recognizable within the major textual units that make up an utterance, namely, topic and comment.

2. INFORMATION PATTERNING AND FOCUS

In this section we will discuss information packaging into information units, presenting the view of LAcT about major textual information units, i.e., topic and comment, and differentiating these pragmatic notions from the semantic notion conveyed by focus.

Additionally, we introduce focus from the LAcT perspective, that is, a semantic notion bound by pragmatics, as will be shown.

2.1 INFORMATION PATTERNING IN THE LANGUAGE INTO ACT THEORY

The Language into Act Theory is anchored in the tradition inaugurated by Austin’s Speech Act Theory (1962); however it developed based on corpus-based and experimental research carried out on large spontaneous speech corpora over the past thirty years at the

LABLITA Laboratory at the University of Florence⁴. Cresti (2000), following Austin, proposes the utterance as the basic unit for the analysis of speech; as such it is pragmatically defined and corresponds to a speech act. In LAcT the illocutionary force is carried by one specific information unit, namely, the comment. The comment can be stringed with other optional information units, therefore fulfilling various pragmatic functions. Information units are combined into information patterns; these in turn will constitute utterances. Therefore, an utterance may be simple if its information pattern is made up of a single mandatory unit, the comment, or compound, if made up of two or more information units which would combine into a complex information pattern. An information pattern will have a prosodic interface whose compounding units conform to prosodic patterns, as indicated by ‘t Hart et al. (1990).

The following schema, proposed by Moneglia & Raso (2014:479) represents the correspondence between prosodic and information unit types as found in L-AcT:

Prosodic Pattern			Information Pattern	
root		→	Comment	
			Tag: COM	
(prefix)	(suffix)	→	(Topic)	(Appendix)
			Tag: TOP	Tag: APC
	(introducer)	→		(Locutive Introducer)
				Tag: INT
	(parenthetical)	→		(Parenthesis)
				Tag: PAR
(incipit)	(phatic)	→	(Incipit)	(Phatic)
			Tag: INC	Tag: PHA

Table 1 Correspondence Prosodic Pattern / Information Pattern (apud Moneglia & Raso, 2014:479)

The segmentation of an utterance into information units is based on perceived prosodic breaks. These breaks can be non-terminal, thereof marking the boundaries of units within the utterance, or terminal when corresponding to the final marking in an utterance. The utterance has therefore a prosodic interface. On prima facie grounds, each information unit corresponds to a prosodic unit. There can be more complex configurations, which shall not concern us in this paper (cf. MONEGLIA & RASO, 2014).

The information pattern of an utterance is determined by its illocutionary activation, and has a prosodic interface which constitutes a prosodic pattern. Different information units

⁴ <http://lablita.dit.unifi.it/>

will have dedicated prosodic profiles as indicated by Cresti (2000), which in combination will exhibit a prosodic pattern which interfaces with an information pattern.

Compound utterances may feature two kinds of information units (IU) in addition to the mandatory comment: textual IUs and dialogic IUs. Textual IUs build the very text of an utterance, carrying its semantic interface, whereas dialogic IUs are directed to interlocutors and have the function of regulating the on-going interaction.⁵ Each IU is identified through three parameters: its function, its prosodic profile and its distribution with respect to the comment.

Besides the comment unit, the topic - also a textual unit - is relevant for the discussion raised in this paper. The topic is considered to be the most important textual unit besides the comment. Its function is to realize the field of application for the illocutionary force carried by the comment. What this means is that the cognitive reference for the illocutionary force of the comment is the semantic content of the topic unit – this pragmatic relationship amounts neither to semantic nor syntactic compositionality. In (i) below, the topic provides the field of pragmatic application for the act accomplished by the comment, however there is neither propositional nor syntactic compositionality:

(i) [bfamd103]: porque quando cê chega num lugar que cê se sente em casa/=TOP= cê sabe imediatamente //COM=

‘because when you arrive at a place you feel at home/=TOP= you immediately know it//COM=’

Topic and comment are also related through specific prosodic configurations. This comes forth through the identification of topic as a prosodic prefix unit and comment as a prosodic root unit, following the IPO School prosodic model described in ‘t Hart et al. (1990) and later developed at the LABLITA Laboratory in relation to information units by Firenzuoli (2003) and adopted by LAcT (cf. Table 1).

According to Moneglia and Raso (2014:479):

The *root* and *prefix* prosodic units, corresponding to the comment and topic information units, have a *nucleus* which contains the prosodic cues necessary for the identification of the unit information function. Besides the nucleus (frequently realized in one or two syllables), the unit can be composed of other syllables that allow the lexical expression of the semantic and syntactic content. Syllables that do not pertain to the nucleus can be positioned before it (*preparation*), after it (*tail*), or between the different syllables that form it (*binding*).

⁵ For a thorough description of the state of the art of LAcT, see Moneglia & Raso (2014).

As indicated in section 1, the comment information unit would be the one which is referred to as focus, a semantic notion, in other theoretical frameworks. In LAcT, focus refers to prosodic features, identified as a functional prominence, found solely in the comment and the topic units, as will be explained in 2.2. Accordingly, an utterance will have a single focus if it is simple, or two or more, if it is compounded by a comment one or multiple topics.

2.2 FOCUS IN THE LANGUAGE INTO ACT THEORY

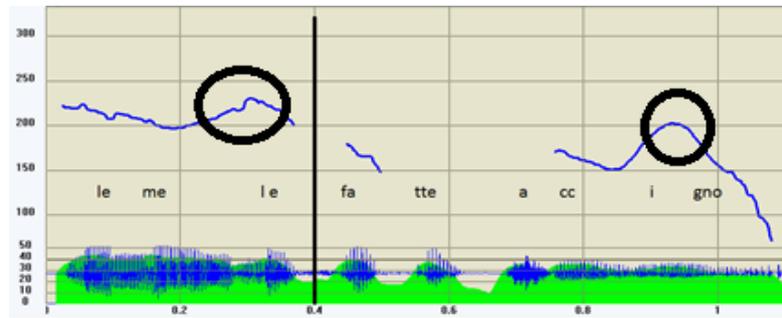
Differently from other theories, LAcT does not base information structuring on context, therefore not recognizing the identity between comment (focus in other frameworks) as a novel semantic input, nor topic as constituted by deriving contextual semantic aboutness on the comment. On the contrary, LAcT views the basic analytic unit for speech, the utterance, as a tripartite notion as per the Speech Act Theory, whereby the locutory act accomplishes an illocutionary act, motivated by an affective pulsion, the perlocutionary act. Therefore, as far as the information pattern of an utterance is concerned, it relates to the pragmatic level, and LAcT indicates the comment as its nucleus, as it carries the illocutionary force. Focus, on the other hand, remains as a semantic notion, associated to the locutory act, constituted by segmental and prosodic features, in both the topic and comment units.

Among the textual information units, topic and comment are the most tightly bound, as they are configured through a pragmatic aboutness relation, i.e., the topic realizes the cognitive domain for the application of the illocutionary force carried by the comment and distributionally is always placed to its left (cf. (i) in section 2.1). Quantitative research into information patterning has shown that the topic-comment pattern is the most significantly found in compound utterances, as reported by Cresti (2012:47). The locutory material contained in topic and comment will peak in functional prosodic prominences (their nuclei), recognized as focus by LAcT. Therefore, both topic and comment will exhibit focus, noted as a semantic concept, but realized prosodically. As observed by Cresti (2012:72), focus prosodic prominence is characterized by several parameters, the most relevant of which are: “(a) pitch with perceptually relevant f_0 movement (rising-falling, or rising) or a strong modulation movement; (b) duration with the lengthening of the syllables (plus a high intensity value).”

Topic and comment focus prominence can be observed in example (i) in Italian, portrayed in Graph 1. The topic-comment information pattern is represented through its f0 curve and corresponding foci circled:

(i)) *VER: le mele/=TOP= fatte a cigno//=COM=

‘(for what regards) the apples, (the right shape should be) like a swan’



Graph 1: Topic-Comment pattern and respective foci

Cresti (2012:71) defines focus in LAcT as: “A focus signals the apex of a semantic domain which develops a Topic or a Comment information unit.” Therefore, Cresti recognizes T-Focus as the semantic nucleus of the topic and C-Focus as its counterpart within the comment. T-Focus and C-Focus differ prosodically and functionally from other types of prominences found in IUs. As already stated, T-Focus and C-Focus perform, respectively, the functional roles of identifying the field of application and the specification of an illocutionary act; whereas other types of prominence may mark language specific melody, end of a phonetic group and, relevantly for this paper, focalization and emphasis, as explored in section 3.

3. FOCALIZATION AND EMPHASIS IN BRAZILIAN PORTUGUESE

Having discussed what focus means in LAcT in 2.2, the goal of this section is to characterize prominence strategies, namely focalization and emphasis, in Brazilian Portuguese (BP), through analysis of data extracted from the spontaneous speech corpus C-ORAL-BRASIL (RASO & MELLO, 2012) in order to characterize them. What we call focalization and emphasis are clearly distinct phenomena from focus as defined in 2.2, and carry particular functions not associated to the licensing of textual information units.

We propose focalization to encode two semantic meanings - either exhaustiveness or contrastiveness - through highlighting. The mechanisms which speakers use to focalize an

item vary in nature and can be syntactic (3.1.1), morphological (3.1.2) or prosodic (3.1.3). Therefore, speakers do not focalize arbitrarily, i.e., they are aware of which item they want to focalize, when and how. The context that induces a speaker to focalize a phrase can be either a linguistic action (cf. section 3.1.3, example 13) or a non-linguistic one (cf. 3.1.3, example 12). On the other hand, we consider that emphasis encodes various semantic meanings different from those two encompassed by focalization. Although emphasis can be expressed by the same linguistic operations as those employed by focalization, namely, syntactic, morphological and prosodic ones, its context of occurrence induces the speaker to highlight an item encoding other types of semantic meanings majorly related to reinforcement through prosodic prominence, therefore distinct from exhaustiveness and contrastiveness, as will be shown.

The methodological steps we used to extract and analyze data in order to demonstrate the differences and the characterization of focalization and emphasis are listed below:

(i) Five texts randomly chosen from the C-ORAL-BRASIL. Three of these texts come from the informal part of the corpus: bfamcv02, bfamcv03 and bfamcv04. They are familiar/private conversations. The other two texts were taken from the formal part of the corpus still in compilation, therefore unpublished as of yet. They are two telephonic dialogues.

(ii) We listened to the audio signals while reading the transcriptions through the aligned sound-text files provided in the corpus by using the Winpitch software (MARTIN, 2011). All the utterances in which any kind of focalization or emphasis occurred were selected. The identification of focalization and emphasis was done through perceptual means of prosodic prominence, syntactic and morphological analysis.

(iii) The utterances selected in step (ii), were extracted from the corpus through the Winpitch software.

(iv) The selected data were acoustically analyzed through the Praat software (BOERSMA & WEENINK, 2011). The acoustic analysis consisted of measuring mean f_0 , maximum f_0 , mean intensity, maximum intensity and duration of phonetic syllables in all utterances (Cf. CHEN, 2012). In order to harvest the acoustic information from the utterances, we used a script developed by Katherine Crosswhite⁶, adapted for our purposes by Marcelo

⁶ To contact the developer of the script write to <crosswhi@ling.rochester.edu>.

Vieira⁷. Duration, intensity and f0 are important parameters found in the literature about both identification and measuring of focus.

(v) We additionally smoothed the pitch curve through Praat commands in order to check the relation between the alignment of the syllables with the pitch curve, noticing moreover the relation between pitch curve-syllable alignment with the parameters mentioned in (iv).

(vi) We maintained pitch range for male voices within 80-200 Hz and 120-350 Hz for female voices. These are the common ranges for male and female voices (cf. TITZE, 1994). However, we found some octave errors in the data. Therefore, we changed the pitch range of male voice to 80-400 Hz and 120-700 Hz for female voice in octave error data. The pitch range of each datum will be shown in the analysis in sections 3.1 and 3.2.⁸

3.1 FOCALIZATION

There are at least three linguistic devices which speakers use to focalize an item in an utterance in BP. They include prosodic, syntactic and morphological strategies, conveying specific semantic properties. Here we examine each of them considering acoustic features, syntactic structures and morphological elements, observing the pragmatic context of enunciation of each one.

Focalization plays a role in highlighting an item in an utterance in speech, drawing the attention of the interlocutor to what is being uttered. The item highlighted by this property is named here Focalization focus or F-focus, in contraposition to T-focus (Topic-focus) and C-focus (Comment-focus) proposed by Cresti (2012) to identify the nucleus portion of respectively a topic unit and of a comment unit (cf. section 2.2). F-focus is not expressed necessarily by prosodic prominence, although in many cases it is. We assume that speakers do not focalize an item arbitrarily. What this presupposes is that focalization has a function in oral discourse. The element highlighted or the F-focus of the utterance carries a specific semantic property conveyed by the act of focalization. There are two semantic properties that F-focus can convey: exhaustiveness and contrastiveness.

⁷ To contact the developer of the adapted version of Crosswhite script for the purposes of this paper write to <marceloletrasi@gmail.com>.

⁸ We thank Marcelo Vieira and Bruno Rocha for helping us with some Praat commands and especially Marcelo Vieira for adapting Crosswhite Praat script for the purposes of our analysis presented in this paper and for revising a first syllable segmentation of our data.

Exhaustiveness is the property of identification of a variable in a set of variables, placing the one chosen prominently in relation to the others. On the other hand, contrastiveness is the property of dissimilarity of a variable over another one. These two semantic properties can be formalized in generic terms as [x and only x], in the case of exhaustiveness and [x and not y], in the case of contrastiveness. We illustrate these two properties with examples below. Consider the following stretch:

- (1) a. Speaker A: What did John give to Mary?
b. Speaker B: It was [a book] that John gave to Mary.

This sentence belongs to a group of syntactic constructions known as cleft sentences (cf. RESENES, 2009, 2014). This is an it-cleft sentence which will be discussed through corpus data in the 3.1.1. The constituent between the copula and the complementizer is the item highlighted or the focus of it-cleft sentences. In example (1), exhaustiveness determines that in a set of things that John could have given to Mary (e.g., a flower, a pencil, a note, a car etc) he gave a book and only a book. Therefore, in a set of possible gifts that John could have given to Mary [x, y, z, w], he picked out one and only that one [x and only x] to give to Mary. At first, this would be what the semantic property of an F-focus in an it-cleft would express in this context.

Now, consider (2) below:

- (2) a. Speaker A: John bought a Lamborghini last week.
b. Speaker B: No, John bought [a Maserati].

Speaker A makes an assumption of a state of affairs, namely, that somebody bought something. Then, speaker B rectifies what speaker A has said, by contrasting his affirmation over the previous one uttered by the first speaker. This is a case of contrastive F-focus. The item highlighted has the property to state that, as a matter of fact, John bought a Maserati, and not a Lamborghini [x and not y].

Taking into consideration the notions of exhaustiveness and contrastiveness, in the next sections we examine the occurrence of focalization and we identify F-focus in BP spontaneous speech.

3.1.1 SYNTACTIC FOCALIZATION

In current literature about focus, authors describe a group of syntactic constructions which highlights a phrase of the sentence. These syntactic constructions are grouped under the name of cleft sentences⁹. They are described as a bi-clausal structure, namely, two clauses which are derived from one simple sentence. In (3) there is a simple sentence from which the sentence in (4) would be derived by a clefting operation:

(3) John kissed Mary.

(4) It was John that kissed Mary.

This assumption becomes clearer if it is considered that there was a movement of the NP in the subordinate clause to the matrix clause. The subscript symbol in (5) indicates a coindexation between the NP that was moved from the subordinate clause to the matrix one:

(5) It was John_i that ____i kissed Mary.

(5) above is an it-cleft sentence and it is composed of a matrix clause which contains a cleft pronoun (it), a copula (the verb to be), a clefted constituent (an XP, in this example an NP), and a subordinate clause which contains a complementizer (in this example, that), and a clefted clause¹⁰, featuring a verb and its internal argument (as seen in the example above). Other kinds of cleft sentences in English can be seen below:

(6) Pseudo-cleft: What John gave to Mary was a book.

(7) Reversed pseudo-cleft: A book was what John gave to Mary.

⁹ In what concerns other focal syntactic constructions besides cleft ones, Ross states that “I don’t have the foggiest idea of how many ways there are to build a syntactic structure whose rhetorical effect is that of picking out some element(s) in it for highlighting purposes. I have the general impression that it would be not hard to find more than one or two hundred of these, but I may be guesstimating far too low” (Ross, w/d date, ms). Although there are plenty of syntactic constructions which would serve the purpose of highlighting an element of the sentence, the cleft ones seem to be the most known or maybe the only ones, as far as we know, presented in the literature.

¹⁰ The clefted clause can be composed by a verb and its internal argument as in (5), but it can be composed by the verb with its external argument as well. In this case, the NP moved is the internal argument of the verb, as can be seen below:

a. The brunette girl kissed John.

b. It was John that the brunette girl kissed.

c. It was John_i that the brunette girl kissed ____i

Pseudo-cleft sentences are made up by a wh-item (what, in the example (6)), a subject clause, a copula (verb to be) and a XP (a NP, in (6))¹¹. In reversed pseudo-clefts, XP is fronted to the beginning of the sentence.

We present two examples of cleft sentences from C-ORAL-BRASIL, an it-cleft and a pseudo-cleft. In the excerpt in (8), four people play *Imagem e Ação* – a mimic game whose objective is for players to guess what an acting player is trying to demonstrate through mimics. LUC asks BRU if it is HEL’s playing turn to guess the next word. BRU answers affirmatively to LUC. Then, BRU makes mimics for HEL to try to guess the word selected. HEL tries several times until she gets the right word, which was *rolha*. Then, BRU uses an it-cleft sentence to inform that it was then HEL’S turn to take up the mimics to proceed with the game. In this context, the pronoun *ocê* is focalized through an it-cleft structure, receiving the semantic property of exhaustiveness, namely, *now it’s you and only you who must perform the mimics*. Therefore, *ocê* ‘you’ is the F-focus of utterance 453:

(8) bfamcv04

*LUC: [430] a [1] ea [1] ea ela tem que adivinhar //

*BRU: [431] ela vai adivinhar //

*HEL: [432] hhh barril // [433] alambique // [434] garrafa //

*BRU: [435] uhn //

*HEL: [436] gargalo // [437] tampa // [438] &he / tampinha //

*LUC: [439] hhh <tampinha> //

*HEL: [440] <&abri> + [441] &he / bico // [442] &he / &he / saca-rolhas // [443] champanhe // [444] hhh garrafa // [445] &garra + [446] rolha //

*BRU: [447] uhn // [448] yes hhh // [449] yyyy //

*HEL: [450] <ai / que foda> //

*BRU: [451] <yyyy / né> //

*HEL: [452] <Nossa> //

*BRU: [453] **agora é ocê que faz** //

“now it is you who goes //”

*LUC: [454] pera aí //

¹¹ Real pseudo-cleft sentences must have a specificational reading rather than a copular predicational one. These two readings are illustrated below:

a. What John gave to Mary was a book.

b. What John gave to Mary was amazing.

In sentence (a) there is a specificational reading in which the NP *a book* was highlighted by the semantic feature that F-focus expresses, discussed in section 3.1. In sentence (b), *amazing* is just an attribute to the thing that John gave to Mary, receiving, therefore, a copular predicational reading. It can be noted also that the thing that John gave to Mary is not really expressed in sentence (b).

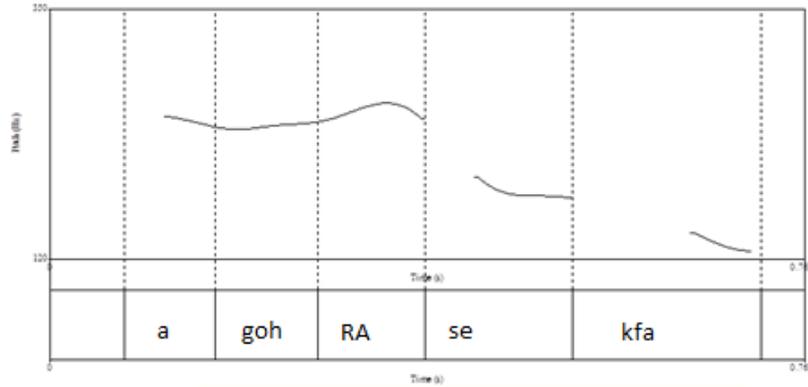


Figure 1: Utterance 453 pitch curve (bfamcv04)

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
a	247	251	91.463	65.783	68.976
goh	242	246	103.681	68.900	72.898
RA	256	263	108.109	63.778	69.116
se	197	244	149.399	58.637	62.975
kfa	238	277	191.058	45.994	51.219

Table 1: Utterance 453 acoustic measures (bfamcv04)¹²

A first observation about (8) above is the fact that the phonetic realization of speech is commonly different from what we perceive. The copula of the it-clefted sentence above is not really produced and the complementizer plus the verb are performed in just one syllable. The pitch curve begins with a high onset in the first two syllables of the adverb *agora* and begins to fall in the last syllable until the end of the utterance. Note that the F-focus of this utterance does not carry prosodic prominence, although it is the second longest syllable in duration. The pitch range used in this example was 120-350 Hz.

In the next excerpt four people are playing snooker. TON performs a move and he tries to play at CAR's playing turn. Then, using a pseudo-cleft sentence, CAR warns TON that he, CAR, would play in that moment, and not TON. In this case, CAR uses a pseudo-cleft sentence to focalize that it was his turn to play, and not TON's. Thereby, the semantic property that the word *eu* 'I' expresses is contrastiveness, namely, *I and not you will play now*. Therefore, the F-focus of utterance 175 applies on *eu*.

(9) bfamcv03

*REN: [165] joguei atrás das suas bola //

*CAR: [166] então é que o Damião [2] o [1] o [1] o [1] o [1] o &Tonin [2] o Renato então //

*TON: [167] é // [168] foi atrás das outras //

*CAR: [169] pôs / ué // [170] pôs <não> //

*CEL: [171] <uhn> //

*CAR: [172] lá na frente //

¹² Acoustic measure tables in this paper have the following abbreviations: Max f0 = Maximum f0, Mean Int = Mean Intensity, Max Int = Maximum Intensity. In all examples shown, duration is given in milliseconds (ms), f0 is given in Hertz (Hz) and Intensity is given in decibels (dB).

*TON: [173] dá pra &ma [3] dá pra jogar ela aqui / ela vem na frente da quatro o' //

*CAR: [174] não // [175] **quem vai jogar agora sou eu** / Toninho //
 “no// who is playing now is me / Toninho //”
 [176] cê nũ faz besteira não / porque // [177] cê nũ entendeu / cê nũ //

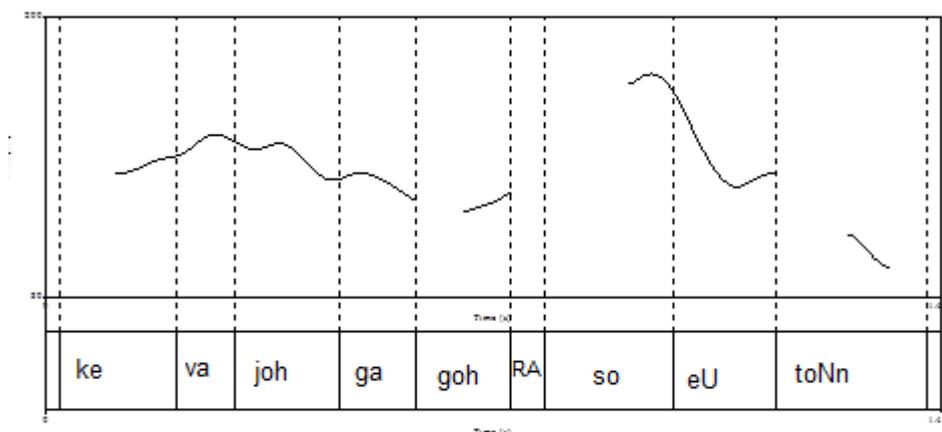


Figure 2: Utterance 175 pitch curve (bfamcv03)

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
ke	136	140	182.460	55.449	59.211
va	146	150	93.425	57.998	61.371
joh	140	146	165.094	58.236	62.672
ga	130	133	122.856	62.457	64.080
goh	120	126	149.130	60.151	63.523
RA	133	139	54.363	54.432	60.324
so	160	175	203.107	60.240	66.564
eU	139	168	164.195	63.076	66.651
toNn	135	135	237.406	51.894	58.838

Table 2: Utterance 175 acoustic measures (bfamcv03)

In Fig. 2 above it can be noted that there is an external sandhi between the last syllable of the verb *jogar* ‘play’ and the first one of the adverb *agora* ‘now’. There is a prosodic prominence in the copula, which presents the highest values for f0, intensity and duration. It is also noteworthy that the F-focus of this pseudo-cleft does not receive prosodic prominence. A downward profile can be verified at some point in the copula until at the first phone of the F-focus. The pitch range used in this example was 80-200 Hz.

3.1.2 MORPHOLOGICAL FOCALIZATION

Morphological focalization is realized through a word in an utterance that expresses exhaustiveness or contrastiveness. In the data analyzed for this paper, we found that in some of its occurrences the adverb *mesmo* in BP can be a focalizer. In order to exemplify such a

case, let us take a previously presented excerpt from bfamcv03 in which four people are playing snooker. TON tells Crossinho he should hit the ball in a certain way, in order to make a specific move. TON conveys the move Crossinho should perform using the deictic *aqui* ‘here’, meaning the place Crossinho should hit the ball. This deictic is focalized by the adverb *mesmo* ‘right, exactly’ which follows it. The F-focus in utterances 187 and 188 is the word *aqui* and it carries the semantic property of exhaustiveness, namely, *hit here and only here*.

(10) bfamcv03

*TON: [179] ah / aqui nũ dá não // [180] vai ter que ser <aqui> //

*CEL: [181] <puxa> esse três cá pra cima / sô //

*TON: [182] não // [183] que "puxa <cá pra cima"> //

*CAR: [184] <não> // [185] <quem> joga agora sou eu / Toninho // [186] cê nũ entendeu //

*TON: [187] **aqui mesmo / &Cel //**

“right here / &Cel //”

[188] **aqui mesmo / &he / ô [1] ô [1] ô [1] ô [1] ô / Crossinho //**

“right here / &he / hei [1] hei [1] hei [1] hei [1] hei / Crossinho //”

*CEL: [189] tem que pensar muito não / Renato // [190] isso aqui ou cê joga pra cá +

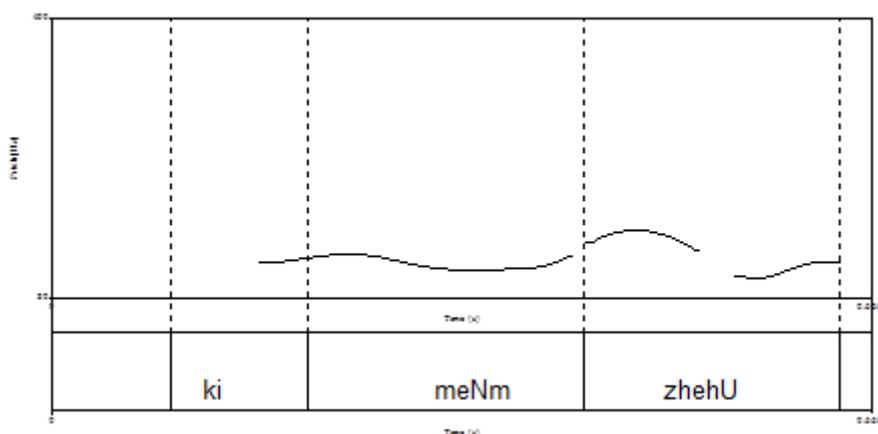


Figure 3: Utterance 187 pitch curve (bfamcv03)

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
ki	244	251	109.232	57.002	62.085
meNm	232	259	223.754	65.231	68.784
zhehU	189	255	206.096	57.492	61.256

Table 3: Utterance 187 acoustic measures (bfamcv03)

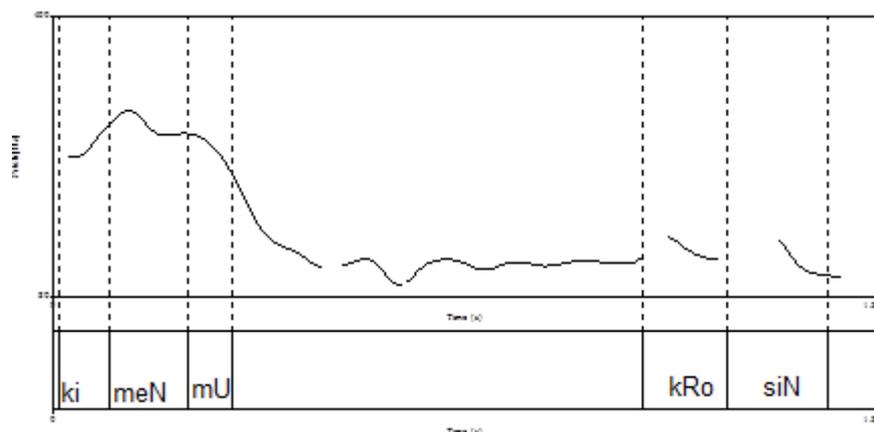


Figure 4: Utterance 188 pitch curve (bfamcv03)

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
ki	253	276	110.629	58.458	60.825
meN	275	292	175.126	61.803	64.032
mU	249	265	96.270	58.789	59.336
kRo	133	147	182.629	51.639	55.667
siN	126	144	224.354	48.066	51.843

Table 4: Utterance 188 acoustic measures (bfamcv03)

Utterance 187 presents a modulated pitch curve with a little ascending movement in the last syllable. Performed in just one syllable, the focalizing adverb *mesmo* is the longest in duration in the utterance. The F-focus in this utterance does not receive prosodic prominence. On the other hand, the F-focus in utterance 188, the deictic *aqui*, begins with a high onset f0 showing an ascending profile up to some point in the following syllable, which begins to descend due to the retracting that the speaker will perform next. The pitch range used in two examples above was 80-400 Hz.

Adverbs usually have several placement positions in BP, but in this case it would not be possible to have inverted *mesmo aqui*¹³. Thereby, the adverb *mesmo* as a focalizer in this case has a fixed position, namely, the one after the element it focalizes. It is possible to depict intuitively the focalization function of the adverb *mesmo* if we compare the pair of sentences below:

- (i) Joga aqui, Crossinho ‘Strike here, Crossinho’
- (ii) Joga aqui mesmo, Crossinho ‘Strike right here, Crossinho’

¹³ Note that depending on the position of the adverb *mesmo*, it can convey different meanings in sentences:
a. Mesmo o João foi à festa ‘Even John went to the party’
b. O João mesmo foi à festa John ‘It is the case that John went to the party’
c. O João foi à festa mesmo ‘John really went to the party’

Comparing the two sentences above, we can note that the adverb *mesmo* is not necessary to convey the place meaning the sentence carries; however, if it is used, it highlights the previous element.

3.1.3 PROSODIC FOCALIZATION

Prosodic focalization exclusively conveys contrastiveness, in pragmatic contexts contrasting a negation to or a correction of some propositional content, or adding informative content to create contrast to an utterance or a previous extra-linguistic action. Contrastiveness can be observed in example (11), in which three friends talk about the wedding of the daughter of one of them. The women discuss who should give a sofa as a gift to the bride. TER says that a sofa could be given to the bride with the pitching in of all family members. RUT says that a sofa shouldn't be cheap. Then TER focalizes the word *caro* 'expensive'. The F-focus is performed through a prosodic prominence in the adjective *caro* and carries the semantic property of contrastiveness, namely, *a sofa is expensive, and not cheap*.

(11) bfamcv02

*TER: [135] e / aí / ela tá assim / mãe // [136] só falta só o sofá // [137] <aí eu> falei assim então <hhh / os> [/1] os +

*RUT: [138] <Nossa> //

*JAE: [139] <ocê vai dar> //

*TER: [140] não //

*TER: [141] o sofá vão vez / que às vezes a Rute distribui com ela / Tonita / <Ção> //

*RUT: [142] <a família> toda // \$

*TER: [143] a família toda e dá um sofá <hhh> //

*RUT: [144] <é pobre mas &so> [/2] mas <sofã nũ deve ser> barato não //

*TER: [145] <mas é caro> // [146] **é caro** //

“it is expensive //”

*RUT: [147] sofá é caro //

*TER: [148] é //

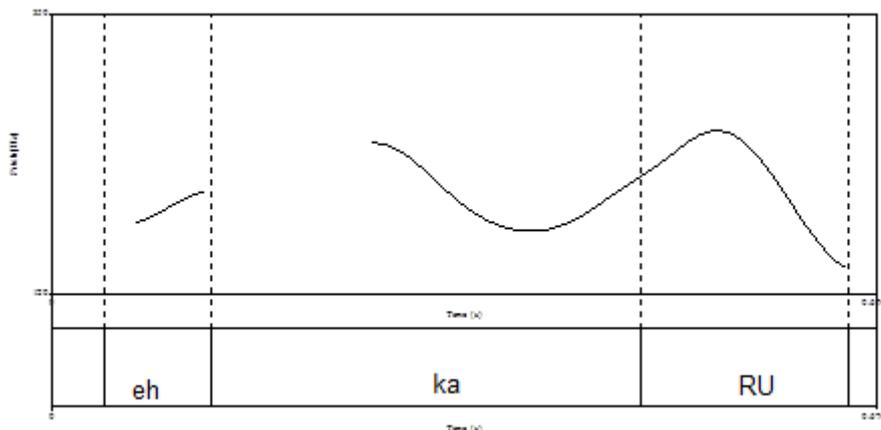


Figure 5: Utterance 146 pitch curve (bfamev02)

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
eh	193	207	88.092	55.104	56.782
ka	208	244	351.180	61.686	65.287
RU	216	254	168.866	60.125	62.616

Table 5 – Utterance 146 acoustic measures (bfamev02)

The F-focus in (11) above shows an ascendent profile that begins in the previous syllable, the verb *é*. The tonic syllable in *caro* is the longest in duration and the post-tonic receives prosodic prominence showing the highest f0 and intensity values in the utterance. The pitch range used in the example (11) was 120-350 Hz.

Prosodic focalization can also express contrastiveness without overt mention to any linguistic reference to a previous utterance. F-focus is possible in such a scenario due to inferences provided by the pragmatic context, as will be discussed for (12) below. In (12) friends are playing snooker and at a certain point of the game, CEL warns TON (who does not speak in the excerpt) to hit ball number five. Another player, REN, goes for the hit, instead of TON, and hits ball 8. Although REN never said anything about hitting ball 8, this is inferred from the context through CEL's reprimand performed in utterances 64 and 65, in which he focalizes the NP *o cinco* 'five'. This is the F-focus of the utterance and it carries the semantic property of contrastiveness, namely, *hit ball number five and not number eight*.

(12) bfamev03

*CEL: [58] <agora tem que matar o cinco> //

*CAR: [59] *cê erra <também / Toninho / aí ficava o seis> +*

*CEL: [60] <ô Toninho / *cê tem que matar o cinco / sô*> //

*REN: [61] *agora *cê mata o onze* aí // [62] *nũ dá não* //*

*CEL: [63] **não / sô / o cinco / Renato** // [64] *oito* // [65] *que isso* //

“no / pal / five / Renato // eight // what the heck //”

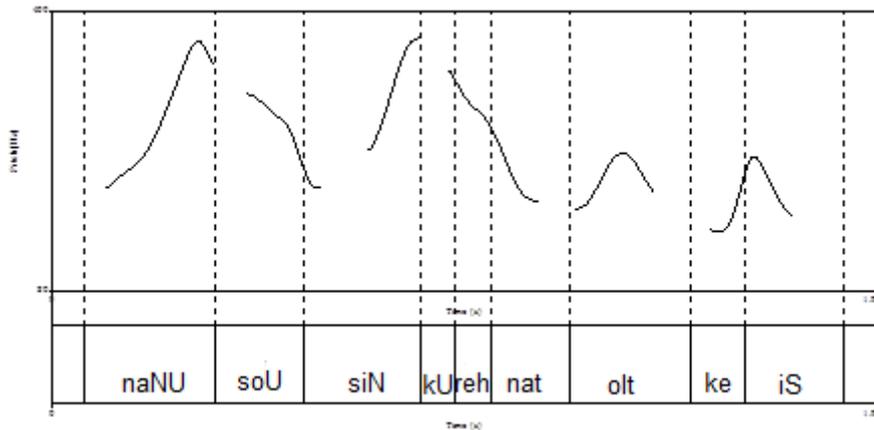


Figure 6: – Utterance 63 pitch curve (bfamcv03)

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
naNU	280	365	306.726	78.933	84.120
soU	299	336	210.773	76.158	82.532
siN	291	369	273.874	61.155	67.335
kU	345	366	83.409	55.822	61.592
reh	292	320	83.493	72.666	74.709
nat	199	268	186.157	67.662	72.702
oIt	198	237	288.333	68.389	74.523
ke	164	218	128.852	65.182	68.707
iS	203	233	230.243	60.313	64.796

Table 6: Utterance 63 acoustic measures (bfamcv03)

In order to focalize the NP *o cinco*, the speaker begins the utterance with a negation *não* ‘no’, carrying a high pitch onset. It begins to descend at the end of this syllable up to the allocutive *sô* ‘pal’, which forms an external sandhi with the determinant *o* ‘the’, clustering up in one syllable. The F-focus of this utterance receives prosodic prominence, showing a very high onset pitch of maximum f0 value of 369 Hz, the highest in the utterance. Interestingly, the pitch curve of the F-focus begins to descend in the post-tonic syllable *co* throughout the next two syllables of the allocutive *Renato*, which is the same pattern observable in the first two syllables of the utterance. The pitch range used in this example was 80-400 Hz.

Focalization through prosody can also convey the semantic property of exhaustiveness. The context can be the identification of a discursive referent, generally, concrete, signaling a property of that referent to be exclusive, namely, x and only x. In example (13) the context is the same as in (11). TER has some doubt about whether she should invite certain people to her daughter’s wedding. She asks RUT for an opinion on the issue at hand. RUT tells TER she could send the wedding invitation to anybody, namely, the

wedding invitation and only the wedding invitation could be sent to whomever; whereas the best men and maid of honor invitations should be well thought about.

(13) bfamcv02

*TER: [329] aí cê acha que eu devo convidar o Guilherme // [330] todo mundo / <né / Rute> //

*RUT: [331] <com certeza / Terezinha> //

*TER: [332] <e / Zé> Levi //

*RUT: [333] <não> //

*TER: [334] <todo> mundo //

*RUT: [335] **convite de casamento cê pode mandar pa todo mundo //**

“a wedding invitation you can send to anybody //”

[336] agora / <padrim> / aí que ocê tem que / ver //

*TER: [337] <uhn> //

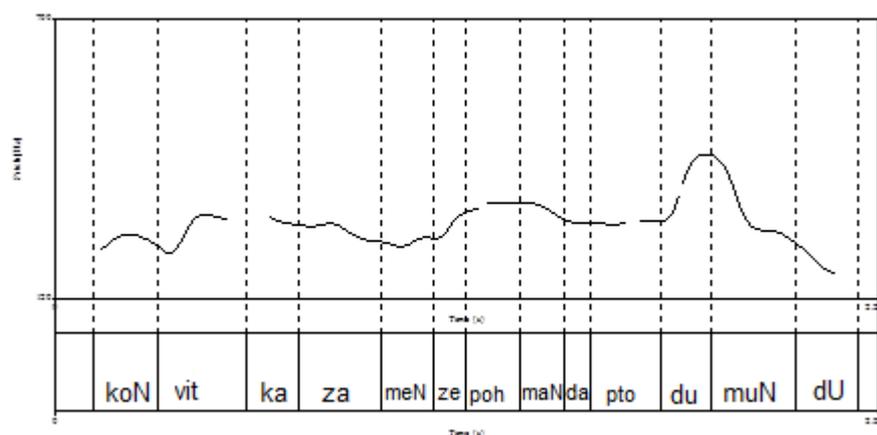


Figure 7: Utterance 335 pitch curve (bfamcv02)

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
koN	240	253	179.924	62.992	66.512
vit	267	295	248.965	61.006	65.433
ka	282	287	142.080	61.137	65.168
za	260	275	231.760	68.974	73.299
meN	237	248	146.919	62.772	64.975
ze	270	298	91.049	58.761	63.045
poh	313	318	150.202	64.263	68.082
maN	307	318	120.874	62.950	66.607
da	278	284	74.636	65.176	67.258
pto	279	285	196.441	56.044	60.407
du	366	419	140.987	54.010	56.787
muN	302	417	232.994	60.897	64.045
dU	201	235	175.934	55.175	60.772

Table 7: Utterance 335 acoustic measures (bfamcv02)

In (13) the F-focus *convite de casamento* ‘wedding invitation’ curve shows a slightly high modulated profile, except in the second syllable. It can be noted that there are also some phonetic reductions in the syllables of the orthographic written transcription. There is an ascending movement in the post-tonic syllable of *todo* ‘every’ which reaches a peak in the next syllable and begins to descend up to the last syllable of the utterance. The first two

syllables of this ascending movement show high f0 values. We verified that this movement and these high f0 values stem from the fact that the speaker uses a falsetto at the end of the utterance. The pitch range used in this example was 120-700 Hz.

Having explored the devices found in corpus data for focalization, we now move to emphasis in the next section, using the same methodology.

3.2 EMPHASIS

In section 3.1 we discussed focalization in BP analyzing spontaneous speech corpus data to illustrate this phenomenon. We showed that focalization highlights an item in an utterance and proposed that this item be named F-focus. The semantic nature of F-focus is conditioned by the pragmatic environment in which it occurs for its expression of two properties: exhaustiveness and contrastiveness. By the same token, we propose that the difference between focalization and emphasis lies in the fact that an item highlighted by emphasis does not carry any F-focus semantic property. Therefore, an emphatic constituent expresses other semantic meanings because its context of occurrence does not require exhaustiveness or contrastiveness. Generally, emphasis is associated with reinforcement through prosodic prominence. We will show that speakers use the same linguistic devices employed in focalization to emphasize an item. However, what emphasis expresses in semantic terms has a diverse nature, as will be shown in the following subsections through corpus data.

3.2.1 SYNTACTIC EMPHASIS

In BP there is a syntactic construction which conveys emphatic highlighting. This construction can be expressed as *um(a) x de (um/a) y* ‘an x of y’. There is a list of nouns that can fill the position of x which indicate scalar measurement, either augmentative or subtractive, for example, *baita, puta, lasca, sepa, titiquinha, trenzico, bostinha* etc.^{14, 15} Position y can be filled by NPs. some examples illustrating this construction are:

¹⁴ This construction might have other highlighted, evaluative meanings, depending on the meaning of x. More often than not, in **these** evaluative cases, x will be derogatory. For example: *A porra da burocracia me mata*, literally: the fuck of the bureaucracy me kills ‘the fucken bureaucracy kills me’.

- (i) A seleção mexicana tem um baita de um time ‘The Mexican team is great’.
- (ii) Tive uma puta de uma febre na semana passada ‘I had a high fever last week’.
- (iii) O João tem um lasca de um braço ‘João has a very strong arm’.
- (iv) Parou uma sepa de uma carreta em frente à minha garagem ‘A huge truck parked in front of my garage’.
- (v) Uma titiquinha de uma cidade que não dá para encher o Maracanã ‘Such a small town that all its population fits in a stadium’.

Depending on the separate meanings of *x* and *y* and their combined phraseological meaning, the construction stands for different meanings, such as: in (i) it means a great team, in (ii) it means a high fever, in (iii) it means a strong arm, in (iv) it means a very big truck – all have augmentative meaning. On the other hand, (v) means a very small town with a subtractive meaning. Therefore through this construction it is possible to express scalarity in quality, intensity, thickness, length or size, for example. All concepts, in some way, are related to measurement.

Example (14) brings an example of the *um(a) x de um(a) y* construction. The excerpt is taken from a phone conversation which integrates the formal part of C-ORAL-BRASIL not yet published. (14) is a phone conversation between a father and his daughter. MAR tells THA, his daughter, that he is at a relative’s enjoying a *bacalhoadada* ‘cod stew’ – a typical Portuguese dish. After some chatting with her father, THA says she might visit the relative’s when they cook some steak. Then MAR uses the syntactic construction *um(a) x de um(a) y* to emphasize the wonderfulness of the *bacalhoadada* his daughter missed having.

(14) Phone dialogue 1

- *MAR: ah / eu tô aqui na casa da tia Célia //
- *THA: cê tá aí / fazendo o quê //
- *MAR: ah / eu vim comer aqui hoje uma bacalhoadada que ela fez //
- *THA: uhn / que delícia // que horas cês vão voltar //
- *MAR: ah / lá pras / nove horas / dez / nove nove-e-meia / nós já devemos tá aí //
- *THA: ah / então tá bão //
- *MAR: teve massa // cê perdeu / cê nũ [2] cê nũ gosta de comer / bacalhoadada / né //
- *THA: nũ é que eu nũ goste / papis //
- *MAR: uhn //
- *THA: é porque é complicado / né / <comer bacalhoadada> //
- *MAR: <pois é> // podia ter vindo //
- *THA: <mas outro dia eu vou> / quando ela fizer um / filé mignon //
- *MAR: <xxx> // perdeu //
- *THA: uhn //
- *MAR: **perdeu uma big duma bacalhoadada** / muito saborosa com azeitona preta / &he / regada a um vinho branco / né / <fazer o quê> //

¹⁵ These constructions are known and documented in the dialect spoken in the state of Minas Gerais, we are not aware of what their status is in other Brazilian dialects.

“You missed a wonderful cod stew / very tasty with black olives / &he / served with White wine / yeah / < what can I do about it> //”

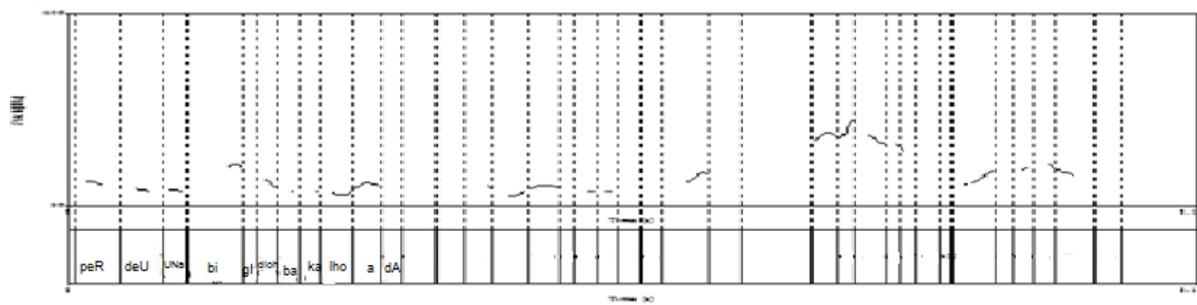


Figure 8: Pitch curve of MAR's utterance in a phone dialogue

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
peR	116	120	342.486	57.708	62.560
deU	111	113	293.811	57.096	59.434
UNa	112	134	174.184	60.554	64.436
bi	229	302	404.434	56.681	63.190
gI	231	274	99.784	53.546	54.628
dIoh	139	192	148.817	61.984	65.270
ba	107	111	144.436	60.449	66.421
ka	104	104	160.433	58.635	63.952
lho	102	107	239.698	57.853	61.545
a	115	118	192.874	67.159	68.590
dA	114	115	156.018	54.754	58.063

Table 8 – Acoustic measures of MAR's utterance in a phone dialogue

Analyzing phonetic material from telephonic data is a great challenge; therefore we would like to state that the phonetic analysis proposed here should be seen as a tendency and not an attested profile. For the acoustic measurements in table 8 we extracted the values for the syllables in the emphatic construction *um(a) x de um(a) y*. It can be noted that the pitch curve increases considerably in the tonic syllable of the word *big*, pronounced by most Brazilian speakers as two syllables, namely, *bigi* as was the case in the example. This syllable presents the longest duration and the highest f0 value. The pitch range used in this example was 80-400 Hz.

3.2.2 MORPHOLOGICAL EMPHASIS

Morphological emphasis is expressed by the addition of a word to highlight an element of an utterance. This word can be reduplicated, as shown in (15). (15) is a phone

recording of a conversation between a mother (IVA) and her daughter (CAM). The topic of the conversation is money left by IVA's husband to pay for her other daughter's college tuition. CAM says that her sister has taken the money and IVA asks how much it was. CAM answers it was four hundred *reais* and says the tuition is about eight hundred *reais*. IVA asserts that her other daughter should pay what she has got and negotiate the rest. IVA uses the word *lá* two times, one to highlight the verb *negociar* 'negotiate' and the other to highlight the complement *com eles* 'with them'.

(15) Phone dialogue 2

*CAM: já pegou o Gabriel / e já levou o dinheiro lá que cê pediu po / meu pai deixar aqui pra ela //

*IVA: quanto que seu pai deixou aí //

*CAM: quatrocentos //

*IVA: tá //

*CAM: &pare [1] eu acho que o curso dela é oitocentos mais ou menos //

*IVA: é / lá [1] ea passa o resto lá / uai / ea vai pagando / uai //

*CAM: é &la [1] meu pai deu quatrocentos /

*IVA: **ea negocia lá com eles lá //**

“she negotiates *lá* with them *lá*”

*CAM: / os outros restante <quatrocentos> //

*IVA: <é> / joga mais pra frente / né //

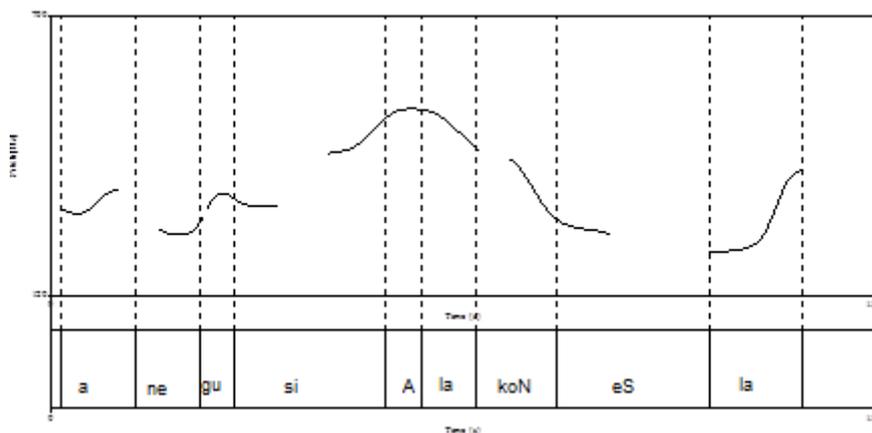


Figure 9 – Pitch curve of IVA's utterance in phone dialogue 2

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
a	308	332	175.662	70.475	73.155
ne	264	312	148.706	61.228	66.302
gu	315	332	80.323	63.081	66.015
si	377	485	355.993	54.637	58.680
A	502	506	85.640	62.499	64.622
la	474	505	127.100	66.604	69.763
koN	370	425	192.443	52.663	56.553
eS	240	277	360.539	52.087	57.534
la	265	380	218.298	62.151	66.141

Table 9 – Acoustic measures of IVA's utterance in phone dialogue 2

The acoustic analysis shows a very high ascending profile for the curve in the third syllable of the verb *negociar*, reaching its peak in the last syllable of this verb, following the next syllable of the *lá*. Then, the curve begins to descend up to the syllable of the personal pronoun *elas*, performed as *es*. It can be noted that the speaker performs this utterance in a falsetto voice. The pitch range used was 120-700 Hz.

In the analyzed example, morphological emphasis conveyed by the particle *lá* is supported by prosodic emphasis as well. It should also be noted that particle *lá* which in many uses is the location deictic form for *there* does not have this meaning in the example explored.

3.2.3 PROSODIC EMPHASIS

Emphasis can be conveyed through a prosodic prominence alone as in (16) below. The context of (16) is the same presented in (13). Three friends are talking about the wedding of the daughter of one of them. RUT says that she doesn't want to be invited to be one of the maids of honor because she feels shy to “parade” at the church. She mentions that in other similar occasion she failed to show up. In order to emphasize this fact, she says *eu não fui não* ‘I didn't show up’ changing her quality of voice, which becomes more high pitched, highlighting the verb *fui*. The semantic effect it conveys seems to be assurance about the information presented and persuasiveness.

(16) bfamcv02

*RUT: [75] nũ me convida pa ser &pa [1] madrinha não / hein //
 *TER: [76] <não> //
 *RUT: [77] <já vou> avisando com antecedência // [78] <morro de vergonha> //
 *TER: [79] <escuta só que que a> Dani tava <falando> //
 *RUT: [80] <morro de> vergonha de <desfilar> //
 *TER: [81] <psiu> //
 *RUT: [82] Nossa <Senhora> //
 *TER: [83] <hein> // [84] é porque é da família lá dele / né // [85] <o'> //
 *RUT: [86] <é> //
 *TER: [87] a mãe da &Fa [2] da +
 *JAE: [88] mas é lógico que ea vai pôr ocês / uai //
 *RUT: [89] não // [90] pra quê // [91] eu nũ quero não // [92] Deus me livre //
 *JAE: [93] ocê / <Tonita / yyyy> //
 *RUT: [94] <morro de vergonha // [95] cê nũ viu que o> da Tiane eu nem <entrei> na igreja //
 *JAE: [96] <é> //
 *RUT: [97] **eu nũ fui não / minha filha / lá não** // [98] morro de vergonha //
 “I didn't show up / girl / there!”

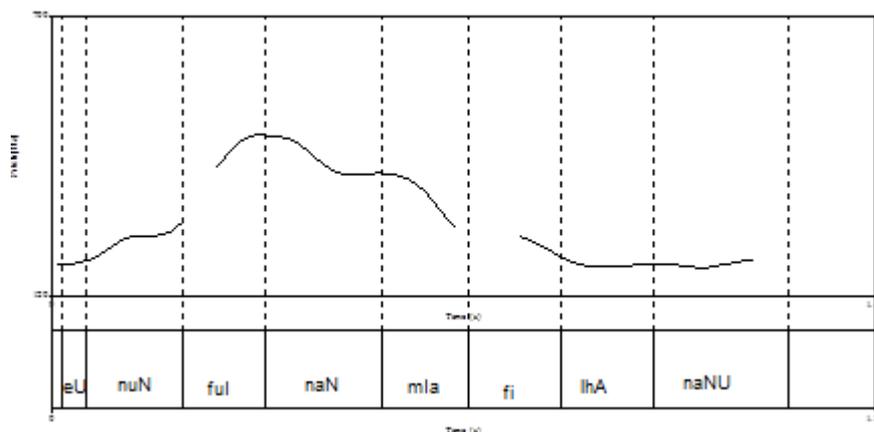


Figure 10 – Utterance 97 pitch curve (bfamev02)

Syllable	Mean f0	Max f0	Duration	Mean Int	Max Int
eU	188	193	40.307	58.995	59.178
nuN	235	283	160.304	56.181	59.584
ful	391	453	138.008	55.925	59.458
naN	403	452	195.112	57.882	62.448
mla	323	374	143.555	51.606	54.333
fi	239	259	152.469	49.685	53.791
lhA	185	201	153.834	55.911	57.624
naNU	185	195	224.787	53.430	56.685

Table 10 – Utterance 97 acoustic measures (bfamev02)

In Figure 10 above a rising-falling movement of the curve can be noted, beginning exactly in the first syllable of the utterance and concluding in the last one. The highest values of f0 reach its peak in the last two syllables of the VP *nũ fui não*. The longest syllable is the final *não* of the double negation. This syllable presents also the highest intensity value. Possibly this corroborates some studies that state that double negation in BP is used to emphasize the utterance. In this example, the speaker also uses a falsetto voice. The pitch range used in this example was 120-700 Hz.

In the next section we provide our final remarks about the two phenomena discussed in the sections above.

3.3 SOME REMARKS ON THE DIFFERENCES BETWEEN FOCALIZATION AND EMPHASIS

In the sections 3.1 and 3.2 above we discussed the characterization of focalization (F-Focus) and emphasis, taking spoken corpus data to show their acoustic properties and

semantic effects. We argued that (i) the linguistic or extra-linguistic context in which the F-focus occurs enables it to perform the semantic effect of exhaustiveness or contrastiveness; (ii) the motivation for the use of emphasis, conditioned by the context in which it occurs, causes it to convey several semantic effects, depending also on what the speaker wants to express, for example, certainty about information presented, persuasiveness, assurance of quality, reinforcement, etc; (iii) prosodic focalization and prosodic emphasis may have the same prosodic profile, namely, a prosodic prominence to draw the listener's attention by highlighting an item in an utterance; however, what differentiates one from the other lies both in semantics, which reveals differences between properties and effects of meaning, and in pragmatics, which is determined by the utterance context.

4. REVISITING FOCUS

The term focus, as has been show in this paper, may refer to different phenomena in distinct theoretical frameworks. On the other hand, information categories that are approximate might be referred to through different labels (cf. comment, rheme, focus). Working within the LAcT framework we support the distinction between comment as an information unit and focus as a semantic property, characterized by prosodic parameters, that licenses the textual information units topic and comment as their nucleus.

The commonality among different theoretical views in what concerns focus is its prominence feature. We have, however, shown that prominence, be it prosodic, syntactic or morphological, might have very distinct semantic functions. In doing so, we have initially made a distinction between T-Focus and C-Focus on the one hand, and topicalization or T-Focus and emphasis on the other.

At this point, these different notions might be joined within a general label FOCUS, which stands for subspecified prominence performing a semantic function. Within FOCUS, different functional properties are conveyed, namely the nuclei of topic and comment, contrastive focus, exhaustive focus and emphasis. As such, FOCUS encompasses the following subcategories, joined solely by the prominence effect: T-Focus (topic focus), C-Focus (comment focus), F-Focus (focalization performing contrastiveness and exhaustiveness) and E-Focus (emphasis pertaining to augmentation, intensification, salience, etc).

Our proposal is able to account for different linguistic levels and devices in the performance of prominence and its specific functions, at the same time that distinguishes

among them. Naturally, further research into spoken language and the manifestation of prominence is needed in order for this proposal to be refined.

REFERENCES

1. AMMANN, Hermann. Die Stellungstypen des lateinischen attributiven Adjectivums und ihre Bedeutung für die Psychologie der Wortstellung auf Grund von Ciceros Briefen an Atticus untersucht, *Indogermanische Forschungen*, 29, 1911/1912.
2. AMMANN, Hermann. *Die menschliche Rede: sprachphilosophische Untersuchungen*. Teil 2: Der Satz. Schauenburg: Lahr i.B., 1928.
3. AUSTIN, John. *How to Do Things with Words*. Oxford: Oxford University Press, 1962.
4. BIRNER, Betty. J.; WARD, Gregory. *Information status and noncanonical word order in English*. Amsterdam: John Benjamins, 1998.
5. BLAND, Susan R.K. Topic/comment sentences in English. *Cornell Working Papers in Linguistics*, v. 2, 1980.
6. BOERSMA, Paul; WEENINK, David. *Praat*. 2011. Available at: <<http://www.fon.hum.uva.nl/praat/>>. Accessed on: 23/07/2015.
7. BOLINGER, Dwight. Contrastive accent and contrastive stress. *Language*, v. 37, 1961.
8. BOSCH, Peter.; van der SANDT, Rob. (eds.). *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge: Cambridge University Press, 1999.
9. BÜRING, Daniel. 1999. Topic. In: BOSCH, Peter.; van der SANDT, Rob. (eds.). *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge: Cambridge University Press, 1999.
10. CHAFE, Wallace. L. *Discourse, consciousness, and time*. Chicago: Chicago University Press, 1994.
11. CHAO, Yuen Ren. *A grammar of spoken Chinese*. Berkeley: University of California Press, 1968.
12. CHEN, Aoji. The prosodic investigation of information structure. In: KRIFKA, Manfred; MUSAN, Renate. *The Expression of Information Structure*. Berlin: De Gruyter Mouton, 2012.
13. CHOMSKY, Noam. *Aspects of the theory of syntax*. Cambridge: MIT Press, 1965.
14. CHOMSKY, Noam. Deep structure, surface structure and semantic interpretation. In: STEINBERG, Danny D.; JAKOBOVITS, Leon A. (eds.). *Semantics: An Interdisciplinary*

Reader in Linguistics, Philosophy and Psychology. Cambridge: Cambridge University Press, 1971.

15. CRESTI, Emanuela. *Corpus di Italiano Parlato*. vols. I-II, CD-ROM. Firenze: Accademia della Crusca, 2000.

16. CRESTI, Emanuela. The Definition of Focus in Language into Act Theory (LAcT). In: MELLO, Heliana; PANUNZI, Alessandro; RASO, Tommaso. *Pragmatics and Prosody: Illocution, Modality, Information Patterning and Speech Annotation*. Firenze: Firenze University Press, 2012.

17. CULICOVER, Peter; McNALLY, Louise. (eds.). *The limits of syntax*. New York: Academic Press, 1997.

18. ERTESCHIK-SHIR, Nomi. *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press, 1997.

19. FIRBAS, Jan. On defining the theme in functional sentence analysis. In: VACHEK, Josef (ed.). *Travaux linguistiques de Prague, Volume 1: L'École de Prague d'aujourd'hui*. University of Alabama Press, 1966.

20. FIRBAS, Jan. On the interplay of prosodic and non-prosodic means of functional sentence perspective (A theoretical note on the teaching of English intonation). In: FRIED, Vilem. (ed.). *The Prague School of Linguistics and Language Teaching*. London: Oxford University Press, 1972.

21. FRETHEIM, Thorstein. Pragmatics and intonation. In: VERSCHUEREN, Jef; BERTUCCELLI-PAPI, Marcella (eds.). *The Pragmatic Perspective*. Amsterdam: John Benjamins, 1987.

22. FRETHEIM, Thorstein. Grammatically underdetermined theme-rheme articulation. *ROLIG*, n. 49, Roskilde University Center, Denmark, 1992a.

FRETHEIM, Thorstein. Themehood, rhemehood and Norwegian focus structure. *Folia Linguistica* XXVI/1-2, 18, 1992b.

23. GABELENTZ, Georg von der. Ideen zur einer vergleichenden Syntax: Wort-und Satzstellung. *Zeitschrift für Völkerpsychologie und Sprachwissenschaft*, 6, 1868.

24. GIVÓN, Talmy. *Topic Continuity in Discourse: A Quantitative Cross Language Study*. Amsterdam: John Benjamins, 1983.

25. GUNDEL, Jeanette K. *The Role of Topic and Comment in Linguistic Theory*. Ph.D. Dissertation, University of Texas at Austin, 1974.

26. GUNDEL, Jeanette K. Stress, pronominalization and the given-new distinction. *Working Papers in Linguistics*, University of Hawaii, 10.2, 1978.

27. GUNDEL, Jeanette K. Zero NP-anaphora in Russian: a case of topic-prominence. In: *Proceedings from the 16th Meeting of the Chicago Linguistic Society*. Parasession on Anaphora, vol. 10, n. 2, 1980.
28. GUNDEL, Jeanette K. Shared knowledge and topicality. *Journal of Pragmatics*, 9, 1985.
29. GUNDEL, Jeanette K. Universals of topic-comment structure. In: HAMMOND, Michael; MORAVCSIK, Edith A.; WIRTH, Jessica (eds.). *Studies in syntactic typology*. Amsterdam: John Benjamins, 1988.
30. GUNDEL, Jeanette K. Topic, focus and the grammar pragmatics interface. In: ALEXANDER, Jim HAN, Na-Rae; MINNICK, Michelle (eds.). *Proceedings of the 23rd Annual Penn Linguistics Colloquium*. Penn Working Papers in Linguistics, vol. 6.1, 1999b.
31. GUNDEL, Jeanette K.; HEDBERG, Nancy; ZACHARSKI, Ron. Cognitive status and the form of referring expressions in discourse. *Language*, 69, 1993.
32. GUNDEL, Jeanette K. On different kinds of focus. In: BOSCH, Peter.; van der SANDT, Rob. (eds.). *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge: Cambridge University Press, 1999.
33. HALLIDAY, Michael A.K. Notes on transitivity and theme in English. Part II. *Journal of Linguistics*, 3, 1967.
34. JACKENDOFF, Ray. *Semantic interpretation in generative grammar*. Cambridge: MIT Press, 1972.
35. KISS, Katalin É. Identificational focus versus information focus. *Language*, 74, 1998.
36. KUNO, Susumo. Functional sentence perspective. *Linguistic Inquiry*, 3.3, 1972.
37. KURODA, Sige-Yuki. *Generative grammatical studies in the Japanese language*. Ph.D Dissertation, Massachusetts Institute of Technology, 1965.
38. KURODA, Sige-Yuki. The categorical and thethetic judgement: evidence from Japanese syntax. *Foundations of Language*, 9, 1972.
39. LADD, D. Robert. *The structure of intonational meaning: evidence from English*. Bloomington: Indiana University Press, 1978.
40. LAKOFF, George. On generative semantics. In: STEINBERG, Danny D.; JAKOBOVITS, Leon A. (eds.). *Semantics: An Interdisciplinary Reader in Linguistics, Philosophy and Psychology*. Cambridge: Cambridge University Press, 1971.
41. LAMBRECHT, Knud. *Information structure and sentence form: topic, focus and the mental representation of discourse referents*. Cambridge: Cambridge University Press, 1994.
42. LI, Charles N.; THOMPSON, Sandra A. Subject and topic: a new typology of language. In: LI, Charles N. (ed.). *Subject and topic*. New York: Academic Press, 1967.

43. MARTIN, Philippe. *WinPitch*. 2011. Available at: <<http://www.winpitch.com/>>. Accessed on: 23/07/2015.
44. MARTY, Anton. *Gesammelte Schriften*, vol. II, part 1. Abteilung Halle: Max Niemeyer Verlag, 1918.
45. MATHESIUS, Vilém. On linguistic characterology with illustrations from Modern English. *Actes du Premier Congrès International de Linguistes à La Haye*, 1928.
46. MONEGLIA, Massimo; RASO, Tommano. Notes on Language into Act Theory (LAcT). In: RASO, Tommaso; MELLO, Heliana (eds). *Spoken Corpora and Linguistic Studies*. Amsterdam, Philadelphia: John Benjamins, 2014.
47. PAUL, Hermann. *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer, 1880.
48. PRINCE, Ellen F. On the syntactic marking of presupposed open propositions. In: FARLEY, Anne M.; FARLEY, Peter T.; McCULLOUGH, Karl-Erik (eds.). *Papers from the parasession on pragmatics and grammatical theory at the twenty-second regional meeting, CLS, 22, 2*, 1986.
49. RASO, Tommaso; MELLO, Heliana. *C-ORAL-BRASIL I: corpus de referência do português falado informal*. Belo Horizonte: Editora UFMG, 2012.
- REINHART, Tanya. Pragmatics and linguistics. An analysis of sentence topics. *Philosophica*, 27, 1981.
50. RESENES, Mariana Santos de. *Sentenças pseudo-clivadas do português brasileiro*. M.A. Dissertation, Universidade Federal de Santa Catarina, 2009.
51. RESENES, Mariana Santos de. *A sintaxe das construções semiclivadas e pseudoclivadas do português brasileiro*. Ph.D. Dissertation, Universidade de São Paulo, 2014.
52. ROSS, Haj. The Syntax of Emphasis – A Base Camp. Manuscrito. Available at: <<http://www-personal.umich.edu/~jlawler/The-syntax-of-emph-Maurice.pdf>>. Accessed on: 12/07/2015.
53. TITZE, Ingo R. *Principles of Voice Production*. Englewood Cliffs: Prentice Hall, 1994.
- VALLDUVI, Enric. *The information component*. Ph.D. Dissertation, University of Pennsylvania, 1990.
54. WARD, Gregory L. The semantics and pragmatics of preposing. Ph.D. Dissertation, University of Pennsylvania, 1985.
55. WEIL, Henri. *De l'ordre des mots dans les langues anciennes comparées aux langues modernes*. Paris: Joubert Libraire-Éditeur, 1844.

ABSTRACT: The inception of studies on information structure are usually attributed to the notions of *point de départ* and *but du discours* by Weill (1844) and psychological subject and predicate by Paul (1880) and Gabelenz (1891). Later, within the context of the Prague School the notions theme and rheme were introduced (cf. AMMANN, 1928; MATHESIUS, 1929), and this was followed by the adoption of different terminological possibilities to similar categories within diverse theoretical frameworks. Some examples are topic and comment comentário (HOCKETT, 1958; CHAFE, 1970; GUNDEL, 1977); presupposition and focus (CHOMSKY, 1971; JACKENDOFF, 1972), new and given (HALLIDAY, 1976), and topic and focus (LAMBRECHT, 1994). In this paper, following the Language into Act Theory by Cresti (2000) and data from the C-ORAL-BRASIL, a spontaneous speech Brazilian Portuguese corpus, by Raso and Mello (2012), we distinguish focus from comment. Additionally we indicate that semantic focus, a prosodically characterized phenomenon, is exclusive to the textual information units topic and focus and performs a functional role in the recognition of these information units, whereas focalization and emphasis, despite being able to exhibit prosodic prominence, perform different roles in spoken language, mostly related to the semantics of highlighting, and may occur in any information unit through different linguistic devices. **Keywords:** focus; focalization; emphasis; language into act theory; spontaneous speech; information patterning