

Alexandre Delfino Xavier

**Prosody and information status: analysis of production in
Brazilian Portuguese**

Belo Horizonte, 29 de abril de 2014

Alexandre Delfino Xavier

**Prosody and information status: analysis of production in
Brazilian Portuguese**

Tese apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Linguística Teórica e Descritiva.

Área de concentração: Linguística Teórica e Descritiva

Linha de Pesquisa: Processamento da Linguagem

Orientador: Profa. Dra. Maria Luiza Cunha Lima

Co-orientador: Prof. Dr. Pablo Arantes

Belo Horizonte

Universidade Federal de Minas Gerais

2014

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca FALE/UFMG

D349p

Delfino, Alexandre.

Prosody and information status [manuscrito] : analysis of production in Brazilian Portuguese / Alexandre Delfino. – 2014. 141 f., enc. : il., tabs., grafs., color.

Orientadora: Maria Luiza Cunha Lima.

Coorientador: Pablo Arantes.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: Processamento da Linguagem.

Tese (doutorado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f. 112-125.

Anexos: f. 126-141.

1. Língua portuguesa – Entonação – Brasil – Teses. 2. Língua portuguesa – Versificação – Teses. 3. Análise prosódica (Linguística) – Teses. 4. Entonação (Fonética) – Teses. I. Lima, Maria Luiza Cunha. II. Arantes, Pablo. III. Universidade Federal de Minas Gerais. Faculdade de Letras. IV. Título.

CDD : 414



FOLHA DE APROVAÇÃO

Prosódia e Estrutura Informacional no PB: Descrição e análise da influência da informação prosódica no processamento referencial

ALEXANDRE DELFINO XAVIER

Tese submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGÜÍSTICOS, como requisito para obtenção do grau de Doutor em ESTUDOS LINGÜÍSTICOS, área de concentração LINGÜÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Linha D - Organização Sonora da Comunicação Humana.

Aprovada em 29 de abril de 2014, pela banca constituída pelos membros:

Prof(a). Maria Luiza Gonçalves Aragão da Cunha Lima - Orientador
UFMG

Prof(a). Aline Alves Fonseca
UFJF

Prof(a). Elisângela Nogueira Teixeira
UFC

Prof(a). Rui Rothe-Neves
UFMG

Prof(a). Luciana Lucente
UFAL

Belo Horizonte, 29 de abril de 2014.

À vovó Hilda.

Agradecimentos

Aos meus colegas e amigos do Laboratório de Psicolinguística Thaís Maíra, Jefferson Maia, Vanessa Pinha, Luiza Luchesi, Bruna Amaral e vários outros não menos importantes pelo apoio constante e amizade que tornaram o trabalho menos árduo.

Aos amigos e outras pessoas que gentilmente se dispuseram gastar parte de seu tempo para contribuir para os experimentos.

Ao César Reis e a todas as auxiliares de pesquisa do Laboratório de Fonética (Labfon) que tão prontamente atenderam aos meus pedidos de uso do laboratório.

À Beatriz Raposo de Medeiros pelas observações sobre o primeiro experimento durante o VI SETED em 2011.

À Petra B. Schumacher, Manuel Dingl, Hannah Weiland, Stefanie Röhrig, Brita Rietdorf e ao restante da equipe do Laboratório de Neurociências em Mainz, pela recepção calorosa e fundamental apoio durante o experimento lá realizado.

Ao Stefan Baumann, Martine Grice (IfL-Phonetik) e Chiara Gianollo (LAK) em Colônia pelas observações e sugestões, cujas ideias foram certamente incorporadas para o melhoramento deste trabalho.

Ao Miguel Souza, pela extrema generosidade na procura por participantes brasileiros em Darmstadt e pelas agradabilíssimos bate-papos acompanhados de café no Menza.

À Ceriz Bicalho Costa, pela grande amizade desde o mestrado, pelo apoio logístico e constante preocupação enquanto estava afastado do país.

À Cláudia Ligocki, amiga confiante em vários momentos e grande incentivadora da minha carreira.

À Luciana Molina Queiroz, pelo compartilhamento das alegrias e agruras da vida acadêmica e pelas muitas e divertidas horas online que diminuíram o meu isolamento no Velho Mundo.

A Luciana Lucente e Rui Rothe-Neves, pelas valiosas contribuições durante a qualificação.

Ao Ricardo Souza, pelas excepcionais aulas desde a graduação, que me levaram a me dedicar com outros olhos à língua do Bardo.

Aos membros da banca, por terem aceito o convite.

À Maria Luiza Cunha Lima, por me apresentar ao intrigante e maravilhoso mundo da Psicolinguística não só durante as aulas da pós-graduação, mas no convívio ao longo desses anos. Por ensinar na prática como um cientista da linguagem deve ser meticoloso e trabalhar arduamente na sua rotina diária. Pela paciência, pela delicadeza, pela amizade e por uma orientação cuidadosa sem a qual esse trabalho não teria se concretizado.

Ao Pablo Arantes, que antes de se tornar oficialmente co-orientador, sempre contribuiu de forma generosa para várias questões teóricas e técnicas. Por despertar em mim uma visão mais aguçada sobre estatística e raciocínio lógico aplicadas à linguagem. Por também ser um orientador rigoroso e paciente ao mesmo tempo.

A minha mãe, Carmen, e meu irmão, Gustavo, que mesmo sem entender exatamente o que venho fazendo ao longo de quase meia década, sempre apoiaram os meus sonhos e torceram por mim.

À Samantha Alves, paciente e incansável, sempre me mostrando o lado doce da vida mesmo nos momentos em que as dúvidas e o cansaço pesaram mais.

Ao CNPq-DAAD pelo financiamento da minha estadia na Alemanha.

*I believe in intuition and inspiration.
... Imagination is more important
than knowledge. For knowledge is
limited, whereas imagination
embraces the entire world,
stimulating progress, giving birth to
evolution. It is, strictly speaking, a
real factor in scientific research.*

*Albert Einstein, Cosmic Religion :
With Other Opinions and Aphorisms
(1931)*

Resumo

O presente trabalho tem como objetivo principal avaliar quais parâmetros acústicos estão associados à codificação prosódica do status informational no Português Brasileiro (PB). Foram conduzidos três experimentos de produção. No total, 13 pessoas foram instruídas a ler uma série de narrativas cujas palavras-alvo variavam o seu status discursivo em função do contexto (dado, novo ou acessível). Após as sessões de gravação, as palavras-alvo foram armazenadas em arquivos de som. Foram analisadas a duração da palavra-alvo (excluindo-se o artigo), uma medida de tendência central (média de f_0) e duas medidas de dispersão (desvio-padrão e gama de f_0). Os dados para análise foram extraídos por meio de *scripts* do Praat e receberam tratamento estatístico. Os resultados mostram que duração e a média de f_0 foram os dois parâmetros mais relevantes. Em geral, referentes novos possuem maior duração e recebem uma proeminência maior de f_0 nas sílabas pretônicas se comparados aos referentes dados e acessíveis. Esses dois últimos status não mostraram diferenças significativas.

A análise geral dos dados dos estudos indicam que falantes do PB de fato usam informação prosódica para sinalizar o status informacional de entidades no discurso. Apesar da variação encontrada entre os sujeitos, pode-se perceber que as diferenças acústicas entre os status são estáveis intrasubjetivamente. Diferentemente dos resultados de outras línguas (com o alemão), o PB aparentemente não possui uma marcação distintiva para o status acessível, conclusão que é apoiada nos dados de ambos os estudos. Aparentemente, apesar de o status informacional ser um fenômeno inerentemente cognitivo em todas as línguas, a sua manifestação é variável entre as línguas.

Palavras-chave: status informacional, givenness, prosódia, entonação, estrutura informacional

Abstract

This study aims to assess which acoustic parameters are associated to the prosodic encoding of informational status in Brazilian Portuguese (henceforth BP). Three production experiments were conducted. In total, 13 people were instructed to read a series of narratives whose target words varied its status by their discursive context (given, new or accessible). After the recording sessions, the target words were stored in sound files. The duration of the target word (excluding the article), a central tendency measure (mean f_0) and two dispersion measures (standard deviation and range of f_0) of were analyzed. Data for analysis was extracted through Praat customized scripts and received statistical treatment. The results show that the duration and the mean f_0 were the two most relevant parameters. In general, new referents have longer duration and receive a higher f_0 prominence on prestressed syllables compared to given and accessible referents. These last two statuses showed no significant differences between them.

The overall analysis of the data from the studies indicates that the BP speakers actually use prosodic information to signal the informational status of entities in discourse. Despite the variation among subjects, one can notice that the acoustic differences between the statuses are stable within subjects. Unlike the results found in other languages (like German), BP seems to not have a distinctive marking for accessible status, a conclusion which is supported by the data from both studies. Apparently, despite the fact that informational status is an inherently cognitive phenomenon, its manifestation is variable across languages.

Keywords: Informational status, givenness, prosody, intonation, information structure

List of tables

Table 1 - List of status of the Assumed Familiarity Hierarchy (Prince, 1981)	18
Table 2 - Givenness Hierarchy statuses definitions and examples (Gundel et al., 1993)	20
Table 3 - Correlation between linguistic form and highest required status (Gundel et al., 1993: 284).....	21
Table 4 - Labels for the annotation of discourse referents; “r-“ indicates the referential level (Bauman & Riester, 2012: 16)	28
Table 5 - Labels for the annotation of nouns denoting discourse referents; “l-“ indicates the lexical level (Bauman & Riester, 2012: 20)	29
Table 6 - Reaccenting in twelve languages (Cruttenden, 2006: 319).....	49
Table 7 - Summary of statistics (F-values) for mean duration in Experiment 1	68
Table 8 - Summary of statistics (F-values) for mean average f_0 in Experiment 1 ..	70
Table 9 - Summary of statistics (F-values) for mean f_0 SD in Experiment 1	71
Table 10 - Summary of statistics (F-values) for mean f_0 range in Experiment 1	73
Table 11 - Summary of statistics (F-values) for mean duration in Experiment 2 ..	79
Table 12 - Summary of statistics (F-values) for mean average f_0 in Experiment 2 ..	81
Table 13 - Summary of statistics (F-values) for mean f_0 SD in Experiment 2	82
Table 14 - Summary of statistics (F-values) for mean f_0 range in Experiment 2	84
Table 15 - Summary of statistics (F-values) for mean duration in Experiment 3 ..	88
Table 16 - Summary of statistics (F-values) for Mean average f_0 in Experiment 3 ..	90
Table 17 - Summary of statistics (F-values) for mean f_0 SD in Experiment 3	91
Table 18 - Summary of statistics (F-values) for mean f_0 range in Experiment 3	93

List of figures

Figure 1 - Assumed Familiarity Hierarchy (Prince, 1981: 237).....	17
Figure 2 - Givenness Hierarchy scale (Gundel et al., 1993: 275).....	20
Figure 3 - The timing of the activation costs with relation to speaker and listener (Chafe, 1994: 74).....	25
Figure 4 - Cognitive concepts of identifiability and activation in Lambrecht (1994: 109).....	26
Figure 5 - Syntactic domains of the referential (r) and lexical (l) levels for the example phrase on the tall tree (Baumann e Riester, 2012: 15, adapted)	27
Figure 6 - Schematised intonation contours of the sentence ‘The people at the next table called the waiter’ (Baumann, 2005: 9, adapted).....	46
Figure 7. Praat editing screen with oscilogram, spectrogram, syllable tier and duration tier (sound file: ‘tempestade’ with accessible status)	64
Figure 8. Example of application of the time-normalized f_0 contour script. (Arantes et al., 2012)	66
Figure 9 - Mean duration in Experiment 1.....	68
Figure 10 - Mean average f_0 in Experiment 1.....	69
Figure 11 - Mean f_0 SD in Experiment 1	71
Figure 12 - Mean f_0 range in Experiment 1	72
Figure 13 - Mean time-normalized f_0 contours in Experiment 1	74
Figure 14 - Mean duration in Experiment 2.....	79
Figure 15 - Mean average f_0 in Experiment 2.....	80
Figure 16 - Mean f_0 SD in Experiment 2.	82
Figure 17 - Mean f_0 range in Experiment 2	83
Figure 18 - Mean time-normalized f_0 contours in Experiment 2.....	85

Figure 19 - Mean duration in Experiment 3.....	88
Figure 20 - Mean average f_0 in Experiment 3.....	89
Figure 21 - Mean f_0 SD in Experiment 3	91
Figure 22 - Mean f_0 range in Experiment 3	92
Figure 23 - Mean time-normalized f_0 contours in Experiment 3.....	94
Figure 24 - Time-normalized f_0 contour analysis. (Antão et al., 2013, adapted)	98

Table of Contents

Chapter 1	9
Givenness	9
1 Topic, focus, givenness: are they in the same basket?	9
1.1 Topic	10
1.2 Focus	11
1.3 Givenness.....	12
2 Alternative taxonomies for givenness	15
Chapter 2	31
Prosody and Intonation	31
1 Intonation	31
2 Phonetics of Intonation	33
2.1 Pitch	33
2.2 Duration	35
2.3 Stress and accent	36
Chapter 3	40
Prosodic marking and information status	40
1 Introduction	40
2 Encoding information status	45
2.1 De-accenting, re-accenting	46
2.2 Other means of acoustic marking	54
2.3 Latest work in Portuguese.....	56
Chapter 4	61
Production Experiments	61
1 Experiment 1	61
1.1 Materials	61
1.2 Procedures	63
1.3 Acoustic correlates.....	64

1.4 Statistical analyses	67
1.5. Results	67
1.5.1 Duration	67
1.5.2 Average f_0	69
1.5.5 Time-normalized f_0 contours.....	73
1.5.6 Discussion	75
2 Experiment 2.....	77
2.1 Materials	77
2.2 Procedures, acoustic correlates, statistical analyses	78
2.3 Results	78
2.3.1 Duration.....	78
2.3.2 Average f_0	80
2.3.3 SD	81
2.3.4 Range.....	83
2.3.5 Time-normalized f_0 contour	84
2.3.6 Discussion	85
3 Experiment 3.....	86
3.1 Materials	86
3.2 Procedures, acoustic correlates, statistical analyses	87
3.3 Results	87
3.3.1 Duration.....	87
3.3.2 Average f_0	89
3.3.3 SD	90
3.3.4 Range.....	92
3.3.5 Time-normalized f_0 contour analysis	93
4 General discussion.....	95
Conclusion and Outlook.....	104
References	109
APPENDIX A - List of sentences for Production Experiment 1	122
APPENDIX B - List of sentences for Production Experiment 2	130
APPENDIX C - List of sentences for Production Experiment 3	134

Introduction

When human beings engage in conversation, a dramatically dynamic and complex exchange of information comes on stage. Before the speaker could even form a proposition, he must access tiny bits of information and stitch them up in several linguistic, contextual and discourse levels. During this process, he must also consider whether the addressee is aware of the content or at least part of it. When the message is about to come out, the speaker needs to signalize which part of this message is only a hook of previous contents and which part is newsworthy. As an interactive activity, the speaker receives almost instant feedback from the addressee, as the form of the message is adapted to the interlocutors' needs.

Such interplay of how information flows in communication and how it is presented is, *grosso modo*, what we might assume as the Information Structure (IS). As for this present study, we consider IS under the notion of information packaging (Chafe, 1976). In this sense, we can say that the information flow, or the interactive conversation, can be divided into two parts: the message itself, that is, the underlying propositions and their truth values, and the way these propositions are presented to the addressee. According to Chafe, information packaging refers to “how the message is sent and only secondarily with the message itself, just as the packaging of toothpaste can affect sales in partial independence of the quality of the toothpaste inside.” (Chafe, 1976: 28)

As one might see, not only the content of the message i.e. the propositions are important, but the form it is presented plays a role on the cognitive construction

of the discourse model. Chafe also defines the role of speakers' beliefs about the hearers' awareness of the propositional content in this process. Vallduví (1992), in his *The Information Component*, states that

the term information packaging is used [...] to denote this nonsyntactic non-logico-semantic structuring of sentences" (p. 12), that is, "a small set of instructions with which the hearer is instructed by the speaker to retrieve the information carried out by the sentence and enter it into her/his knowledge-store (Vallduví, 1992: 18).

Both definitions are quite similar, though each of them emphasizes a particular aspect of information packaging. Chafe defines it in terms of the independence of the semantic (i.e. truth conditions and proposition) level and presented a cognitive dimension while Vallduví is more explicit about the procedural (set of instructions) that is present in this linguistic component. In both senses, information packaging responds to the immediate communicative needs of interlocutors. I adopt this view within the model of communication as continuous change of the common ground (CG)¹.

On one hand, proposals like the Centering Theory (Grosz et al., 1995) deal with the question of the continuation of reference in the discourse, i.e. how we keep track of which referential expression refers to which discourse entity. In this account, a system of rules and constraints govern the relationship between what the discourse is about and the mapping of referential forms (e.g. full and (in)definite NPs, and pronouns). Another approach that tackles the same issue is found in Givón (1983). While Centering Theory studies discuss about degree of saliency, Givón talks about topicality or topic continuity.

¹ We are not going to enter into details concerning Common Ground (Reinhart, 1982). However, it is important to notice that it is inherent to our notion that part of the communication success lies on information sharing, that is, the information message is shared between the interlocutors.

Imagine a context in which two women, *Jennifer* and *Helen*, were being discussed. The current conversation topic is the hospitalization of Helen when one of the speakers says:

- (1) How did **she** react?²
(Chafe, 1994: 76)

At the time of (1), there was a competition for the pronoun *she*: two referents, both with the same grammatical features. At this moment, however, *she* would have been more appropriate for *Helen*. A very simple explanation based on Centering Theory would be that as *Helen* is the current most salient center of the information, the pronoun *she* would have a back looking center trace that identifies to *Helen*³.

On the other hand, there are proposals that consider the construction of the discourse model based in semantic and pragmatic processing, with a strong interference from memory-based models. In this approach, discourse coherence relies on the degree of activation (or saliency) of referents in the memory (Chafe, 1974, 1976, 1994) and the hearer's level of familiarity with these referents (Prince, 1981, 1992). The choice of referential forms takes into account not only grammatical features in an isolated syntactic module, but also general constraints of the cognition architecture, including working memory limitations. This approach has made room to the investigation of many aspects like the relationship between information saliency with costs of processing in the working memory⁴. At the same time, other studies have exhaustively investigated how the mapping of discourse entities and its linguistic forms occurs in order to establish the mental model in the addressee's mind (Gernsbacher, 1991; Gundel et al., 1993; Baumann & Riester, 2012).

² In the examples used in this thesis, we adopted the following criterion: CAPITAL for acoustic prominence; []_F for focus; *italics* for antecedents and **boldface** for referents.

³ Chafe (1994: 76) used this example to explain accentuation of full NPs.

⁴ For a detailed account, see Accessibility Theory in Ariel (1990) and Informational Load Hypothesis in Almor (1999).

Gernsbacher (1991) investigated a case of pronominal anaphora whose reference number marking is not licensed. The author points out that anaphora resolution is done by heuristics, based on four types of constraints: (1) lexical as number, gender and case marking, (2) syntactic, (3) thematic discourse markers such as topic, focus and (4) pragmatic knowledge provided by the listener and inferable knowledge based on the real world. In principle, the different types of marking guide the hearer to the anaphora resolution. However, there is a specific case in which the antecedent is a word in singular and the pronominal anaphora is in plural:

- (2) A. I want *a new Harley 1200*.
B. **They**'re really huge, but **they**'re gas-efficient.
(Gernsbacher, 1991: 82)

At first glance, this breakdown of the number of lexical marking (a Harley 1200 ~ they) is illegal from a heuristic point of view. The author conducted two experiments to test anaphora resolution with plural pronoun anaphoras. The first experiment evaluated the acceptability level of sentences by participants. The results show that speakers were not surprised to use plural pronouns with singular antecedents. The second experiment evaluated whether in fact the anaphora are resolved lexically or conceptually. If the resolution were lexical, participants would have difficulty in relating the referent and the pronoun, increasing the response time of the test. The results showed no significant difference to indicate the resolution was primarily lexical. The explanation lies on the idea that despite the violation of lexical marking of number, listeners integrate the referents in a discourse context in which the referent singular can mean a class or a set of elements. Pronouns are not thus simple substitutions of the background, but conceptual anaphoras. In addition, the second experiment demonstrated that although there is a lexical violation, it does not hinder the resolution of anaphora. Based on the proposal of Tyler and Marslen-Wilson (1982), the author proposes that the anaphoric resolution is parallel, i.e., the

heuristic resolution of anaphoric operates all types of constraints simultaneously. When the lexical heuristic does not apply (or is violated), the heuristic solves the anaphora pragmatically. Thus, "illegal" plural pronouns do not hinder comprehension. This view of the resolution by pragmatic heuristic leads us to the idea that the listener constructs a mental model that captures situational content of the real world of the utterance.

Discourse understanding involves not only the representation of a textbase in episodic memory, but, at the same time, the activation, updating, and other uses of a so-called situational model in episodic memory: this is the cognitive representation of the events, actions, persons, and in general the situation, a text is about.
(van Dijk & Kintsch, 1983: 11-12)

Mental models⁵ provide a representation of the utterance and allow the understanding of upcoming information. The model is built incrementally, being added and modified as new information is processed.

Earlier studies with Information Structure (Halliday, 1967; Haviland & Clark; 1974; *inter alia*) have traditionally proposed a dichotomy of discourse information into previously mentioned (given) and not previously mentioned (new) referents. Later studies, however, have defended alternative taxonomies in which information can have multiple states according to different aspects such as the hearers' familiarity with the subject, the level of consciousness, and the speakers' pragmatic intentions (Prince, 1981, 1992; Gundel et al., 1993; Allerton, 1978; Baumann & Riester, 2012).

In the last years, the role of phonological information has been proposed among the types of linguistic information available during referential processing. Several studies associate the information structure with prosodic focus e.g. contrastive, broad and narrow focus. Studies like Baumann (2006) show some evidence of prosodic information is associated to the different degrees of

⁵ Mental models here are seen as the dynamic 'mental picture' the addressee forms in this mind as he adds new information to previous discourse during communication (see van Dijk and Kintsch, 1983).

information status beyond the dichotomy new-accented and given-deaccented. One of the findings in Baumann's work is that the degree of accessibility of a referent in the speakers' working memory is correlated to a specific set of pitch accents. The more activated a referent in the memory, the more chance it gets deaccented. The less activated a referent, the more chance it is marked with a high pitch prominence. Studies with visual world paradigm and electrophysiological measures have also provided several evidences for the cognitive nature of givenness and the role of prosodic information in referential processing (Watson et al., 2008; Schumacher & Baumann, 2010; Baumann & Schumacher, 2012). Contrary to a traditional view of dichotomic categories, this current trend sees given and new as both ends of a continuum. Furthermore, crosslinguistic data has shown that the prosodic encoding of this continuum is variable and dependent on many factors.

The present study is thus based on the current discussions about the nature of the informational status and its taxonomy under an approach that takes into account not only the linguistic information but also cognitive constraints, such as memory workload, attention span, and focus of attention. We depart from the vision that there is a relation between linguistic encoding and referential processing (Gundel et al., 1993), which in turn is incremental and based in multiple linguistic sources of information (Altmann & Mirković, 2008).

The main objective here is to investigate which acoustic correlates play a role on the prosodic encoding of the information status in the discourse, specifically in the variant of BP spoken in the metropolitan region of Belo Horizonte.

As secondary objectives, we can list: (i) to study how Chafe's (1976, 1994) activation account of discourse referents can be applied to the prosodic encoding of referents, particularly those referents whose information status lies on the middle of the givenness continuum. In German, each activation level (active, semi-active and inactive) can be related to a specific set of pitch accents

(Baumann, 1996). The idea is to analyze linguistic data to see if the same correlation is present in BP; (ii) to provide data for a discussion on the deaccenting account of given information as a cognitive universal (cf. Halliday, 1967; Chafe, 1976; Cruttenden, 1993, 2006). We intend to relativize this hypothesis through experimental evidences in BP; (iii) to discuss crosslinguistic differences of how prosodic information is integrated in discourse, which allows the improvement of theories and models in an area of investigation traditionally associated to Semantics and Pragmatics; and (iv) to contribute to the general knowledge of the language. Despite being a major Romance language with a large population of native speakers, BP counts with relatively few studies, especially in the specific matter of givenness as a separate phenomenon from Focus. Thus, we intend to provide more information about the linguistic structure of BP.

This study is organized in 5 parts following this Introduction. In **Chapter 1**, we provide a review of the most relevant studies concerning the term givenness. A brief discussion of the conceptual differences between topic, focus and givenness is presented in order to define our research object. Then, we present a review of modern proposals for the taxonomy of givenness. Most of these taxonomies try to account the mental state of discourse entities and referential forms. In **Chapter 2**, a short presentation of prosody, considering mainly the phonetic correlates of intonation and some remarks about intonation structure in BP. **Chapter 3** provides a comprehensive crosslinguistic survey studies that investigated the relationship between information status and prosodic information, considering a discussion of the main mechanisms of prosodic encoding and a review of crosslinguistic data. First, we try to show evidences for the linguistic encoding of information status, and then we discuss the results of several studies on this matter. In **Chapter 4**, we present a series of production experiments that investigated which acoustic cues are involved with the marking of information status. Experimental design, materials, and procedures are described before the discussion of the results. Finally, the **Conclusion & Outlook** provides a summary

of the ideas and results that were explored along this study. We also present an outlook for further research in the area.

Chapter 1

Givenness

In this first chapter, we present the main concepts concerning the central matter of this study: givenness or the information status of referents. It is one major category in the area of IS, though its notion is opaque in many studies. We will try to briefly discuss three major categories - topic, focus and givenness - exploring their scope and basic differences. Finally, we will present a set of state-of-the-art alternatives that explore the notion of information status.

1 Topic, focus, givenness: are they in the same basket?

There are three major terms used in the theory of Information Structure: *topichood*, *focus* and *givenness*. Although different, these terms are often vague and there is a profusion of concepts associated to each term, regarding different perspectives. As they can even overlap in concept, it is important to explore a more clear distinction between them before we can present in more detail the main object of this chapter, that is, givenness.

1.1 Topic

For most authors, the term ‘topic’⁶ is defined by being that entity which the sentence is about. In a similar way, Strawson (1964:104 *apud* Zimmerman and Féry, 2010) defined it as “what is of current interest or concern”. This definition is related to “aboutness”: in any natural language, successful interpretation involves distinguishing between what the speaker intends to communicate something about (topic) and what is actually communicated (comment). It is reasonable to assume, moreover, that this distinction determines how an addressee will assess and store information contained in an utterance. According to Reinhart (1982), communication is not just in the form of unstructured propositions, but is rather associated with entities, just like information in a file card system is associated with individual file cards that bear a particular heading. In this sense, one function of topic is to establish the relevance of an utterance i.e. to ‘hook new information up’ to the ongoing discourse.

- (3) a. [The spy]_{Topic} [sent an encrypted message via email]_{Comment}.
b. [An encrypted message]_{Topic} [was sent via email (by the spy)]_{Comment}.

For example, while (3a,b) express the same proposition, these sentences structures are differently insofar as (3a) should be stored as information about *the spy*, whereas (3b) should be stored as information about *an encrypted message*.

A slightly different approach linked the notion of topic to the mental representation of discourse by the speaker and the addressee. According to the Accessibility Theory (Ariel, 1990), a topic is identified based on its high availability to the addressee. Similarly, Gundel (1976) suggests that shared knowledge and high familiarity of the entity to both the addressee and the

⁶ As our main concern is just to clarify some basic terms in the area, we won’t make much of a distinction between topic and topichood, though they might denote different things. *Grosso modo*, topic is the entity that carries the status (of) topichood. Along this chapter we might use both terms interchangeably.

speaker account for topic status. Thus, according to these approaches, topic is identified as the most salient entity in the mental representations of both speaker and addressee. Hornby's (1972) psycholinguistic experiments suggests that speakers interpret what the sentence is about as corresponding to that aspect of the meaning of a sentence which is in some sense already assumed. In some cases, the topic becomes so redundant that the addressee can answer only with the comment:

- (4) (In a cafeteria)
A: What are you gonna have?
B: [Cappuccino]_{Comment}.

1.2 Focus

According to Krifa (2008: 247), focus “indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions”. The classic pragmatic use of focus is to highlight the part of an answer that corresponds to the wh-part of a constituent question (Paul, 1880).

- (5) A: Who did you talk to?
B: I talked to [PEter]_F.

As we have discussed, the discourse model is not only presented under the form of propositions, but there seems to be a hierarchy of relevance of different entities along the utterance. It is important to distinguish a difference between semantic focus and pragmatic focus. The former relates to factual information, that is, the truth conditions of the proposition and the latter serves to convey the communicative intention of the speaker.

- (6) A: They live in BERlin.
B: They live in [BerLIN]_F!

(Krifka, 2008: 248)

As we might see in (6), focus can be used even below the word level to show some contrast. In this case, the speaker wants to correct his interlocutor's pronunciation.

As focus is generally used as the alternative for a linguistic expression, it is often associated with the presence of new information. It usually occurs when the entity is semantically focused, that is, it has to do with the truth conditions of the proposition. However, there are cases in which even old information can be focused, depending on the speaker's pragmatic intention.

- (7) Grandpa didn't [kick the BUcket]_F, he [passed aWAY]_F.
(Krifka, 2008: 248)

In the example (7), though the two expressions denote the same event, i.e. to die, the speaker uses focus to emphasize the politeness of the last expression.

1.3 Givenness

In general, givenness⁷ refers to the degree to which a certain entity is salient⁸ or present in the immediate linguistic context. As we know, all events and entities of our reality must be denoted using an appropriate linguistic expression or

⁷ The term "givenness" is also referred to as "newness", though looking at opposite points of view, both terms refer to the same phenomenon. For instance, Arnold et al. (2000) defines newness as 'the distinctions among given, inferable, and new information', pretty much based on Prince (1981, 1992). Another common term is referential (or informational) status, but we will reserve this term for the actual state of a referent, and not for the IS component itself. As the majority of authors (and especially the theoretical perspectives we adopted for this study) refer to it as 'givenness', we decided to keep this terminology.

⁸ The concept of saliency has also repeatedly been invoked in the literature, but it lacks a solid definition. Intuitively, saliency is associated to prominence i.e. something stands out compared to its environment. For the purpose of the study, saliency refers to how 'activated' or 'lit up' a certain entity is compared to other entities in its context (Chafe, 1994). This concept is central to our analysis of givenness and will be explored in more details along this chapter.

referent. These referents are stored in a linguistic component i.e. lexicon, and retrieved during conversation. Apart from having a linguistic form, the referents also seem to have a temporary 'state' when used during communication. This 'temporary state' is taken into account under different perspectives. For instance, one might say that the state refers to the degree some information is shared by both addressee and speaker. An alternative is to think that this temporary state is cognitively-constrained, that is, it reflexes the cognitive mechanisms by which the referents are retrieved and stored in the working memory during processing.

So far, we tried to establish a rather clear definition for topic, focus and givenness. However, as one can see, they lack a more comprehensive study in order to have a more solid and straightforward consensus concerning the definition and functions of each category. Givón (1979), for example, presented a notion of topicality "that reflects the status of the referent according to the discourse thus far". This notion, however, has more to do with scales of accessibility than with a traditional view of 'topic'.

Though it can be applied to isolated entities or to entire phrases, it seems that topic is related to the propositional aspect of the discourse. Apart from the propositional content of the message, the speaker is also responsible for defining what the sentence is about. In other words, it is his point of view concerning a certain entity in the discourse. Givenness, on the other hand, is related to the temporary state of an entity as the communication goes on the fly, based on its previous mention or not.

As the topic expresses some shared information between both speaker and addressee, it sounds reasonable to imagine that topic constituents are about old, given information, as they are often easily recoverable from the context. But there are certainly cases of new information in topics.

(8) [A good friend of mine]_{Topic} [married Britney Spears last year]_{Comment}.

(Krifka, 2008: 265)

The sentence above introduces a new referent into discourse and, at the same time, it denotes a topic constituent, as the sentence is about “a good friend of mine”.

In the same way, the notion of focus can be rather confusing with givenness, since focus tends to emphasize newsworthy information. It is important to remind, however, that the primary function of focus is to signalize a point of alternative which serves a conversational purpose in the sentence. Sometimes given information can receive focus:

(9) The press registered the happiness of the ballerinas_i after receiving the award.

In the end, [ONly a few of them_i]_F acknowledged her sponsor.

(10) A: Who stole the cookie, John or Mary?

B: [JOHN]_F stole the cookie.

(Krifka, 2008: 256)

In (9), the focus is projected over the phrase *only a few of them*, referring to a subset of the original referent *the ballerinas*. This type of construction is common with focus-sensitive particles, such as *only*, *even*, *always*, and *also*.

In (10), although *John* had been previously mentioned, it received a contrastive focus in order to emphasize its difference from the alternative ‘Mary’.

As we could see, topic, focus, and givenness are different aspects of information packaging. One common point, however, is that they have traditionally been seen as binary categories: an entity is either topic or not; it can receive focus or not. Until recently, givenness has also been seen as either old information or not (i.e.

new information). In the next section, we will see alternative proposals that show givenness with a more complex spectrum of states.

2 Alternative taxonomies for givenness

Sometimes, a simple dichotomy between given and new information goes against our intuition:

- (11) John had to call the *tow truck* because **the engine** had broken down in the middle of the road.

In (11), the item *engine* can not be classified simply as new information. Despite being first mentioned in the discourse, it sounds “familiar”, as we can infer from the context that *the engine* is part of the car which John called the tow truck for. It is neither given information, since it is firstly mentioned in the speech. In this case, a better classification would have to take into account that the referent is contextually inferable. Therefore, “given” and “new” seems to reflect two ends of a continuum for a whole range of informational statuses of referring expressions.

Several studies in recent years have sought to propose taxonomies of referential expressions that take into account degrees in the middle of this continuum. Prince (1981) presented a seminal work which discusses the nature of the transmission of information and proposes a classification reviewing the classification between given and new throughout several studies in the area. Many of these definitions, though all dualistic in nature, have different perspectives on informational status.

Basically, Prince grouped the different distinctions for information status into three types:

1) Givenness in the sense of predictability and recoverability of information: "The speaker assumes that the hearer CAN PREDICT OR COULD HAVE PREDICTED that a PARTICULAR LINGUISTIC ITEM will or would occur in a particular position WITHIN A SENTENCE." (Prince 1981: 226). Aligned to this, Kuno (1978) defines old-new information in terms of recoverability: An element is old information (predictable) if it is recoverable by the preceding context, if it is not recoverable, it represents new information (not predictable) (Kuno, 1978: 282-283 *apud* Prince, 1981). Even within this perspective, Halliday (1967) defines given and new in terms of intonation. A focus marked or unmarked intonation intentionally identifies new information, given is defined as the complement of the focus marked.

2) Givenness in the sense of saliency: "The speaker assumes that the hearer has or could appropriately have some particular thing/entity/... in his/her CONSCIOUSNESS at the time of hearing the utterance" (Prince 1981: 228). In this view, given information is the knowledge that the speaker assumes to be in the listener's mind at the time of the utterance and new what he assumes is being presented to the listener. Chafe (1994) presents a cognitive approach to the term *activation*. The information given-new can be defined in terms of the cost of activating the speaker has to invest to convey an idea of a stage prior to an active state in the consciousness of the speaker. If the referent is already active at the time of the utterance, it is as if he becomes active in an inactive state, it is new. Chafe also proposes a third state, semi-active, i.e., entities that are not mentioned directly linked to entities in the statement. Chafe calls this state *accessible*.

3) Givenness in the sense of shared information: "The speaker assumes that the hearer 'knows', assumes, or can infer a particular thing (but is not necessarily thinking about it)" (Prince 1981: 230). Given is the information that the speaker believes the listener already knows and accepts as true and again that he believes that the listener does not know yet.

Prince proposes the concept of integration of the various information status according to an Assumed Familiarity Hierarchy (p. 233). The term *Assumed Familiarity* describes the speaker's perspective on certain assumptions of the listener. Prince prefers the term *Familiarity* because the term *Shared Knowledge* seems to suggest a point of view of an omniscient observer who knows what goes on in the mind of the listener. In this taxonomy, the information is divided into three major groups (new, evoked and inferable) which are subdivided, summing up seven types of discourse entities.

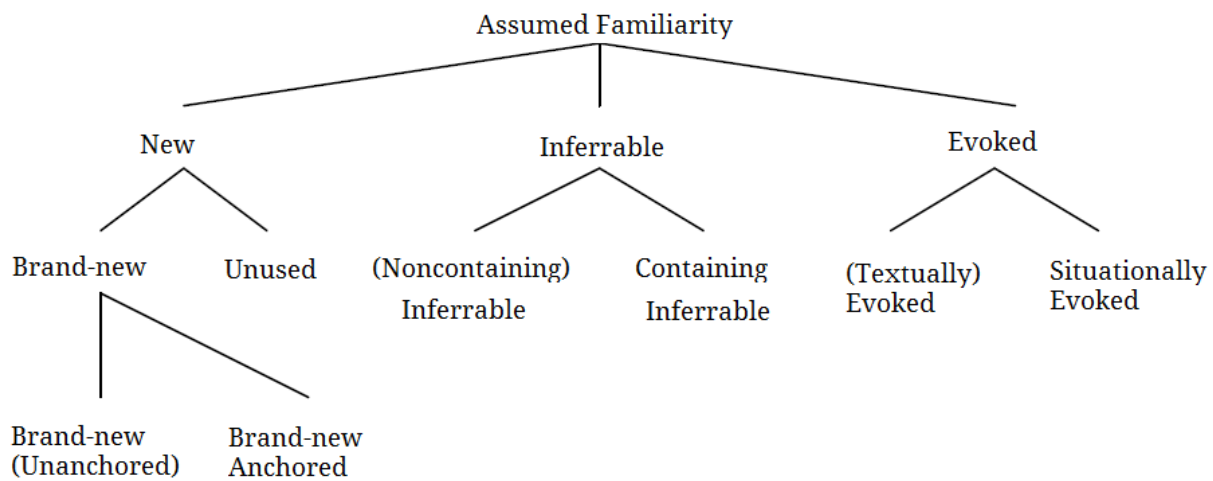


Figure 1 - Assumed Familiarity Hierarchy (Prince, 1981: 237)

More recently, Prince (1992) proposed alterations in her taxonomy combining two levels of givenness. She distinguishes between Hearer-old and Hearer-new entities from Discourse-old and Discourse-new entities. This distinction takes into account the fact that what is new in the discourse is not necessarily new to the hearer (Chafe, 1976; Lambrecht, 1994). Discourse-status refers to what is present (or not) in the discourse while Hearer-status refers to what the speaker believes to be present (or not) in the hearer's knowledge store. A Discourse-old, for instance, entails Hearer-old, since it is something already evoked in the discourse *ergo* present in the hearer's mind. Discourse-new, however, can be either Hearer-

old or Hearer-new, as it is about to be evoked in the discourse but present (new~unused) or not present (new~brand-new) in the hearer's knowledge store⁹.

Table 1 - List of status of the Assumed Familiarity Hierarchy (Prince, 1981)¹⁰

	Brand-new: creates a new entity ("go out and buy a suckling pig")	Unanchored: it is not linked with any Anchor	I got on a bus yesterday and the driver was drunk.
New ("put it on the counter")		Anchored: the NP representing is linked to some other discourse entity	A guy I work with says he knows your sister.
	Unused: corresponding entity in the hearer's own model ("take some staple off the shelf when its presence is suddenly taken for granted in a recipe")		Noam Chomsky went to Penn.
Evoked ("on the counter")	(Textually) evoked: hearer had evoked the entity earlier, on textual grounds, by following instructions from the speaker		A guy I work with says he knows your sister.
	Situationally evoked: hearer evokes by himself, for situational reasons, represent discourse participants and salient features of the extratextual context		Pardon, would you have change of a quarter?
Inferables speaker assumes the hearer can infer it via logical or plausible reasoning from discourse entities already evoked or from other inferables	(Noncontaining) inferrable: what is inferred out of the inferrable NP itself		I got on a bus yesterday and the driver was drunk.
	Containing inferables: what is inferred off is properly contained within the inferrable NP itself		Hey, one of these eggs is broken!

⁹ The 'hearer's knowledge store' notion will be discussed later as a core assumption in Chafe's activation taxonomy.

¹⁰ Prince used the text of a recipe in order to illustrate her taxonomy. That is the reason of the mentioning of cooking terms. This table was created with the description and examples provided by the author along the text.

Prince's Taxonomy, despite its merit for an innovative approach on the informativeness of referents, makes room for some questions. The hierarchy structure allows a closer view of the reality of different referential expressions, but they are mutually exclusive. Unlike this proposal, current ideas suggest there is a gradiency of different types of information. Finally, this classification does not allow for modeling or predictions about what kind of referential expression is related to one or more types of information hierarchy.

Gundel et al. (1993) presents some possible solutions to the problems of taxonomy in Prince (1981). The question that motivates Gundel et al. refers to the relationship between the form and the specific cognitive status information references:

What do speakers/writers know that enables them to choose an appropriate form to refer to a particular object and what hearers/readers know that enables them to identify correctly the intended referent of a particular form? (p. 274)

The authors point out that, despite numerous studies that investigate the linguistic and psychological nature of referents, basic facts about the distribution and understanding of different forms are still unexplained. Gundel et al. present a theory whose main premise is that different pronoun forms and determiners signalize different cognitive states (information location in memory and attentional state), allowing the addressee to restrict the number of possible measures.

The Givenness Hierarchy is an alternative to the taxonomy in Prince (1981). In this hierarchy, a basic assumption is that the form of referring expressions is linked to the cognitive status of the referent. The Givenness Hierarchy proposes six cognitive states corresponding to the form of referring expressions. Each status is necessary and sufficient condition for the proper use of a different form or forms in the same way as seen on Gricean Maxims (Grice, 1975).

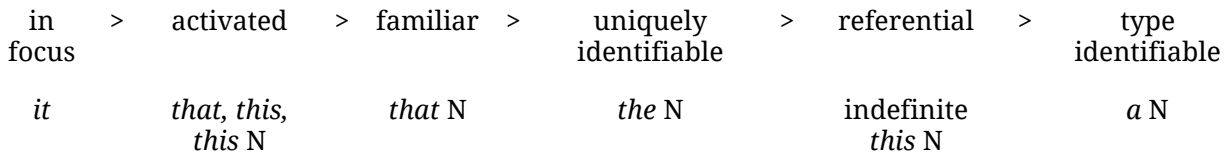


Figure 2 - Givenness Hierarchy scale (Gundel et al., 1993: 275)

By using a particular form, the speaker signals that he assumes that the cognitive status is known. Unlike the proposal of Prince, the Givenness Hierarchy proposes an entailment of lower cognitive status when we use a particular form associated with a particular cognitive state.

Table 2 - Givenness Hierarchy statuses definitions and examples (Gundel et al., 1993)

Type identifiable	Identifies what kind of thing this is.	I couldn't sleep last night. A dog kept me awake.
Referential	Associates a unique representation by the time the sentence is processed.	I couldn't sleep last night. This dog (next door) kept me awake.
Uniquely identifiable	Associates a unique representation by the time the nominal is processed.	I couldn't sleep last night. The dog (next door) kept me awake.
Familiar	Associates a representation already in memory.	I couldn't sleep last night. That dog (next door) kept me awake.
Activated	Associates a representation from working memory.	I couldn't sleep last night. That kept me awake.
In focus	Associates a representation that your attention is currently focused on.	I couldn't sleep last night. It kept me awake.

As seen above, the taxonomy of Prince's Assumed Familiarity Hierarchy considers the categories as mutually exclusive whereas in the Givenness Hierarchy, the categories are entailed i.e., a particular state entails the lower states, but not otherwise. The statuses are thus ordered from most restrictive (in focus) to the least restrictive (type identifiable) with respect to the possible set of related referents that they include.

One clear advantage, beyond the notion of cognitive status and entailment, is the association between specific nominal forms to each cognitive status. This type of classification might allow for empirical research, testing whether there is indeed a relationship between form and mental status.

The Givenness Hierarchy moves forward when it includes terms that allow greater cognitive plausibility. The definition of higher states is based on notions such as "activation" and "presence in working memory." The 'familiar' status, for example, is defined as one that refers to a representation in memory for short or long term. The concept of activation is related to how easy it is to access the referent in memory of the listener.

Table 3 - Correlation between linguistic form and highest required status (Gundel et al., 1993: 284)

	In Focus	Activated	Familiar	Uniquely Identifiable	Referential	Type Identifiable
Chinese	∅ <i>tá</i> 's/he, it'	<i>TA</i> <i>zhè</i> 'this' <i>nèi</i> 'that' <i>zhè</i> N		<i>nèi</i> N		<i>vi</i> N 'a N' ∅ N
English	<i>it</i>	<i>HE, this, that, this</i> N	<i>that</i> N	<i>the</i> N	indefinite <i>this</i> N	<i>a</i> N
Japanese	∅	<i>kare</i> 'he' <i>kore</i> 'this' <i>sore</i> 'that' medial <i>are</i> 'that' distal <i>kono</i> N 'this N' <i>sono</i> N 'that N' medial	<i>ano</i> N 'that N' distal		∅ N	
Russian	∅ <i>on</i> 'he'	<i>ON</i> <i>èto</i> 'this' <i>to</i> 'that'	<i>èto</i> N <i>to</i> N		∅ N	
Spanish	∅ <i>él</i> 'he'	<i>ÉL</i> <i>éste</i> 'this' <i>ése</i> 'that' medial <i>aquél</i> 'that' distal <i>este</i> N	<i>ese</i> N 'that' N medial <i>aquel</i> N 'that' N distal	<i>el</i> N 'the N'		∅ N

Another contribution of the work of Gundel et al. is the universalistic perspective of the scale of familiarity. Prince's taxonomy is based on forms of textual analysis in English, that is, no direct association between degrees of familiarity and their own linguistic forms. The Givenness Hierarchy has a universal application in languages other than English. This allows a deeper connection between the mental representation of the referents and "surface forms".

Prince considers inferables as a separate cognitive status. Gundel et al. (1993: 281) argues, however, that an inferable should be seen as a particular cognitive status achieved by association to an entity previously activated. Clark and Haviland (1977) call inferables *bridging inferences* and Hawkins (1978) calls them *associated anaphors*.

Although this scale represents an advance in Prince's taxonomy, both can be considered as scales of familiarity with the referent. The scale of Gundel et al. is based on how general a reference is for the receiver, going from a status whose representation is more general i.e., a set of possible candidates not identifiable in memory to a single possible entity which works as the discourse focus.

Despite the merit of proposing a hierarchy applicable in more than one language (besides English, the study classified Chinese, Japanese, Russian and Spanish), the results of classification of Gundel et al. showed that the six-status scale does not work isonomically for all languages. For instance, in Japanese, lower levels (types) are set into a single group, given the fact that they have the same status referential expression. This might lead us to conclude that different languages have different strategies or clusters of reference representation. Can we assume that Japanese has less cognitive status than English simply because it has fewer referent expressions?

Other aspects during discourse development might influence referential processing and permit a distinction of different status using the same surface form. As Gundel et al. admit,

while language forms play an important role in determining what will be brought into focus, actual, inclusion in the 'in-focus' set depends ultimately on pragmatic factors, and is not uniquely determined from the syntax. (p. 280)

The Informational Load Hypothesis of Almor (1999) presents an alternative to the issue of the two-way relationship between cognitive status and forms of referential in the Givenness Hierarchy. Almor argues that, despite the evolution of knowledge about the distributional patterns of anaphoric expressions, little is known about the processes underlying these expressions. One possible explanation for the relationship between a given form and salience of the referent in the statement is that it has a communicative function - speakers use anaphoric forms as a clue to help listeners identify the referents. Another possible explanation comes from the Accessibility Theory (Ariel, 1990), which associates different processing costs to different forms e.g., anaphora set of full NPs have a higher processing cost than pronouns.

Based on Ariel (1990), Almor proposes an empirically testable psychological theory that accounts for the relationship between cost of processing and distribution of referential forms. Almor defends the idea that processes underlying the anaphora have the general principle that the additional cost of processing must serve some additional discursive function. This principle is linked to the underlying architecture of the memory system involved in discourse processing. The notion of processing costs relates to the integration of semantic representation of referential expression in speech and the amount of semantic information that is activated by the anaphoric expression. This principle, in the same way as the Givenness Hierarchy, is also based on Gricean Maxims (Grice, 1975).

According to Almor (1999), more generic anaphoric forms represent more active informational status in memory and at the same time, lower syntactic and/or prosodic saliency. Anaphoric forms with more restrictive status are always those with the comparatively least phonetic content¹¹, such as pronouns and clitics.

Knowledge about the world and inferential knowledge can help to integrate a set that is considered new in speech. Situations where information facilitates the integration of extralinguistic referent in the mental model of the listener are not rare.

(12) Luke had a conversation with his friend during the trip. He said the coffee was cold.

In (12), *the coffee* may generate an interpretation error; without additional information, the sentence seems inconsistent or anomalous. However, in other contexts, the interpretation can be made correctly. Imagining a situation where Luke pointed out to the coffee in question, we can say that deictic information can be used to integrate a defined NP in the mental representation.

A cognitive approach of givenness can be seen in the work of Chafe (1976, 1994), whose central ideas are the notion of consciousness and activation. The author claims that “it is ultimately impossible to understand the distinction between given and new information without taking consciousness into account” (1994:72) and givenness is related to the activation cost in order to transfer an idea from a inactive state into an active state in the speaker's or the listener's during communication. If an entity is already active in the listener's memory at the time of the utterance i.e. it has been mentioned recently, it has a given information status; if a referent was inactive and became active, then it has a new

¹¹ Phonetic content here refers to the comparison of lexical forms of full NPs and pronouns. Almor cites *a/the robin* as an example of a referent with low saliency compared to *it*, which is so activated (or shared between the interlocutors) that it requires only a reduced lexical form to be represented.

information status. Additionally, Chafe proposes that accessible referents have an intermediate state "semi-active", which refers to the entities that have not been mentioned but are linked to already active entities in the listener's mind. In order to explain this semi-active state, Chafe suggests the referential processing comprises linguistic and visual processing as well. He cites Baar's (1993: 135) simile noting that as given referents and foveal vision are responsible for processing of focused, salient information in (short-term) memory, peripheral vision and accessible referents correspond to secondary, not directly focused processing of information, though they are as much as important as the former (Chafe, 1994: 54).

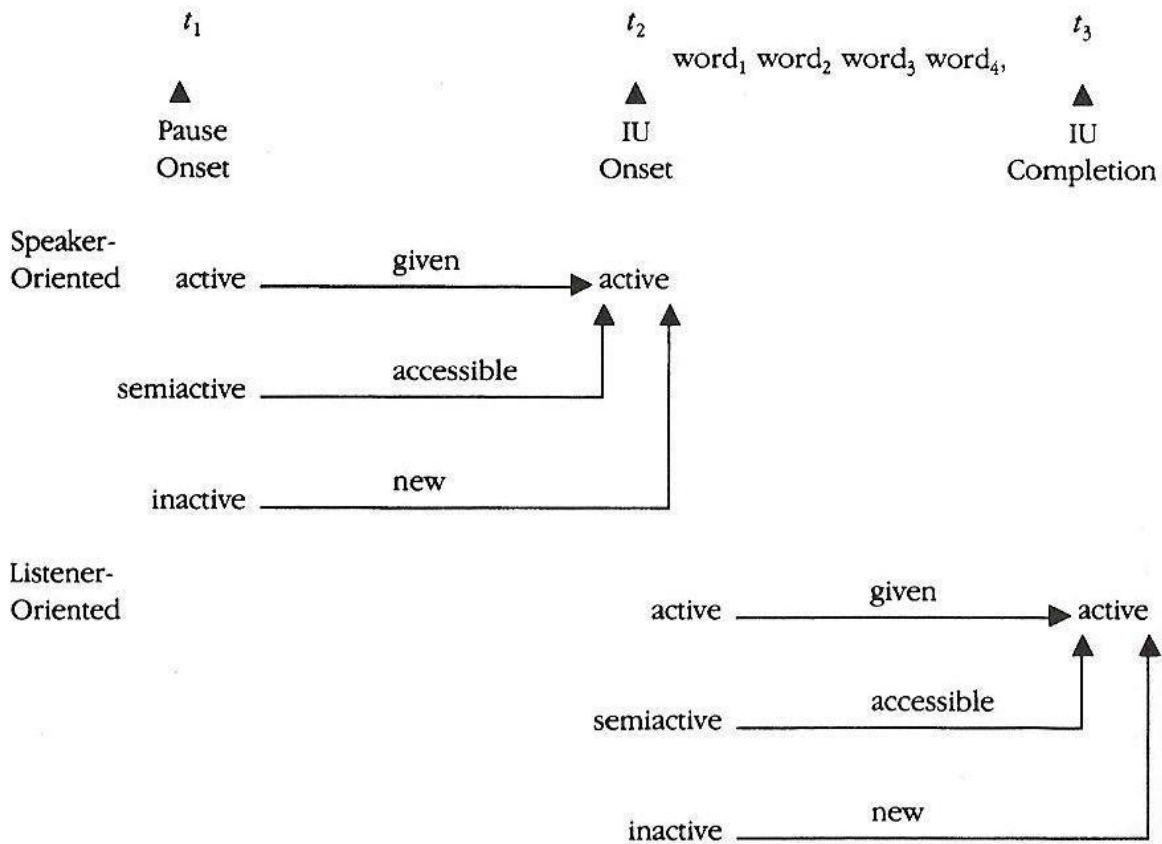


Figure 3 - The timing of the activation costs with relation to speaker and listener (Chafe, 1994: 74)

It is important to notice that Chafe's definition of inactive is restrictive. It refers to referents that are not available at the time of discourse, and in principle count all

entities that are a) known to the listener, b) unknown to the listener or c) non-existent. Whether an item is present in the listener's (long-term) memory is irrelevant for the system of activation (Baumann and Riester, 2012).

Based on this issue, Lambrecht (1994) distinguishes identifiability from activation, "one having to do with knowledge, the other with consciousness" (1994:105). Identifiability does not necessarily entail familiarity, but the ability "to pick [an entity] out from among all those which can be designated with a particular linguistic expression and identify it as the one which the speaker has in mind" (1994:77). A referent may be either unidentifiable or identifiable regarding its presence in the listener's knowledge store i.e. long-term memory.

The speaker must then assume if the referent is present in both (short-term) or (long-term) listener's memory. According to Lambrecht, the identifiability of a referent is a fundamental step for it to become activated in the listener's consciousness. The speaker must indicate through whether the referring expression introduces a new concept of a referent in the textual world - to open a new "'file' in the discourse register to which further elements of information may be added in the course of the conversation and which can be reopened in future discourses" (Lambrecht, 1994: 77). Only after the referent can be distinguished from all potential entity candidates (i.e. identifiable) it is qualified to enter the various activation parameters of consciousness (in this case, meaning activation in the short term ~ working memory). Lambrecht (1994:109) illustrates the relation between identifiability and activation in a diagram:

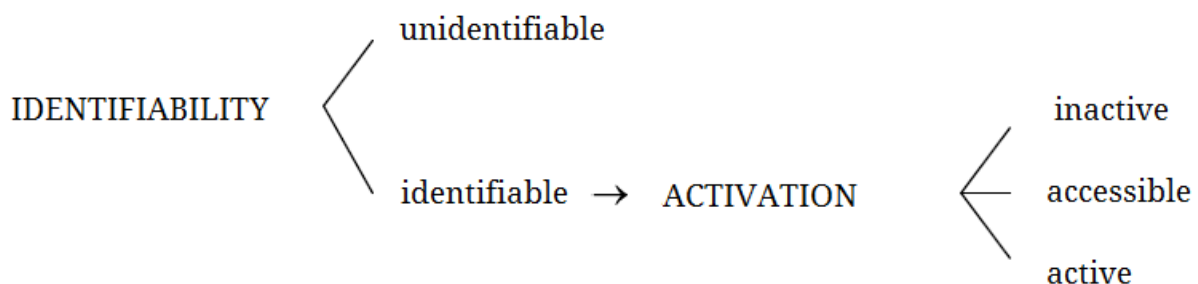


Figure 4 - Cognitive concepts of identifiability and activation in Lambrecht (1994: 109)

Burkhardt (2009) suggests that the process of interpreting a defined NP, for example, consider two components - core D and the complement NP. The core D carries functional traits and encodes definiteness, number, person, gender and case. The information encoded in the nucleus is grammatical in nature, processed by a strict syntax and influences the operation interface between the syntax and the discourse structure. The complement, in turn, carries semantic content and its encoding is lexical.

Baumann & Riester (2012) propose that a proper analysis of the informational status should consider two levels, referential and lexical. Following the line of Chafe (1994), Baumann and Riester claim that only nominal expressions defining ideas of individuals taking part in specific states or events may have informational status.

Similarly to Burkhardt (2009), one phrase has a complex referential status on both levels, although every status applies at different levels. The informational status is associated to the reference level PP or DP, whilst the informational status is associated to the lexical word level, or a modified NP.

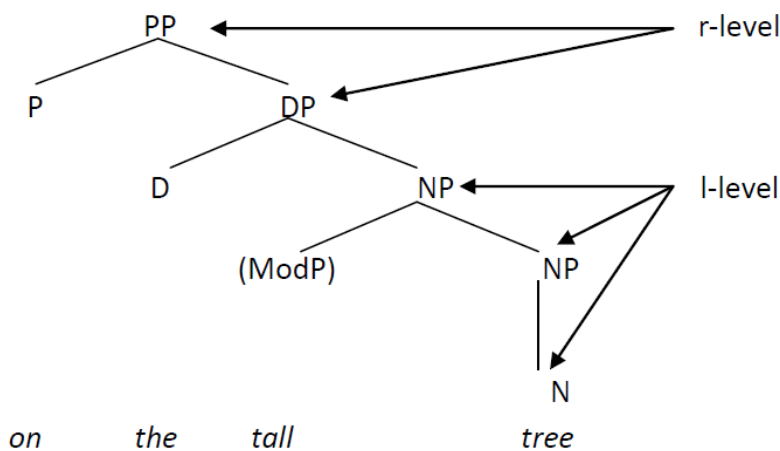


Figure 5 - Syntactic domains of the referential (r) and lexical (l) levels for the example phrase on the tall tree (Baumann e Riester, 2012: 15, adapted)

A breakthrough in the division of informational status on two levels is a finer semantic difference in determining the activation of anaphora. Baumann & Riester present a taxonomy with labels that reflect the status of nominal expressions for both the referential level and the lexical level.

The reference level (r-level) serves to distinguish the level of processing of grammatical features by the addressee:

Table 4 - Labels for the annotation of discourse referents; “r-“ indicates the referential level (Bauman & Riester, 2012: 16)

Definites	
r-given	anaphor corefers with antecedent in previous discourse
r-given-sit	referent is immediately present in text-external context (in particular discourse participants) – symbolic deixis
r-given-displaced	coreferring antecedent does not occur in previous 5 intonation phrases
r-environment	refers to item in text-external context (conversational environment) – gestural deixis / demonstratives
r-bridging	anaphor can be inferred from non-coreferring antecedent or scenario
r-bridging-displaced	non-coreferring antecedent does not occur in previous 5 intonation phrases
r-bridging-contained	anaphor is inferrably anchored to an element within the same expression
r-unused-known	item which is generally known, but not derivable from previous discourse
r-unused-unkown	item which is identifiable from its own linguistic description, but not derivable from previous discourse
Definites or Indefinites	
r-cataphor	item whose referent is established later on in the text

r-generic	abstract or generic item
Indefinites	
r-new	specific indefinite introducing a new referent

The lexical level (I-level) serves to distinguish semantic relations between the referent and anaphora (synonymy, hypernymy, hyponymy, and metonymy):

Table 5 - Labels for the annotation of nouns denoting discourse referents; “l-“ indicates the lexical level (Bauman & Riester, 2012: 20)

l-given	recurrence of same noun
l-given-syn	relation between nouns at the same hierarchical level (synonyms)
l-given-super	noun is lexically superordinate to previous noun (markable is a hypernym or holonym)
l-accessible	two related nouns, whose hierarchical lexical relation is unclear (e.g. in a scenario)
l-accessible-sub	noun is lexically subordinate to previous noun (markable is a hyponym or meronym)
l-new	noun not related to another noun within last 5 intonation phrases

The aim of this chapter is to explore the notion of givenness. First, we presented a tentative discussion on its main differences compared to other IS terms, topic and focus. Both are commonly mistaken as part of givenness, as generally topics tend to be given, shared information and focus tend to highlight newsworthy information. However, the scope of the three categories is rather different. For this study, the central, relevant aspect is referential status; topic and focus are taken as part of information structure, though they are not taken as part of the analysis. Second, we presented an introductory definition for givenness and

explored the main studies in the area that deal with alternative taxonomies to a given-new traditional point of view. As we can see, givenness went from a conversational perspective to a cognitive account.

As the main objective of this study is to investigate the influence of prosodic information on referential processing, we adopted a cognitive perspective, based on Chafe's (1976, 1994) notion of referent activation with Lambrecht's (1994) notion of identifiability. Although Baumann & Riester's (2012) current taxonomy seems to deal with this phenomenon in a more comprehensive way, it seems that a three-way division of the information status is sufficient for our analysis. Besides, Baumann's (2006) and Baumann and Riester's (2012) studies are based also on the same cognitive account present in Chafe's work.

In the next chapter, we will to explore the relationship between referential status and prosodic information.

Chapter 2

Prosody and Intonation

In this chapter, we present fundamental aspects related to prosody and intonation. We will review some phonetic correlates of intonation and a brief discussion of accent.

1 Intonation

In general terms, prosody refers to sound features that extend over individual segments and hence are referred to as suprasegmentals (Cruttenden, 1986). Among the different definitions, some authors include the intimate relation of the acoustic properties of speech to the conveyance of meaning (Crystal, 1969; Lehiste, 1970; Laver, 1994; Cutler et al., 1997). Such acoustic features can extend over since a small group of segments e.g. a syllable or a morpheme to relatively longer chunks, such as phrases or sentences. It is considered as a linguistic component, as it serves not only as the physiological execution of the message, but it has - under many proposals - a proper structure and language universals. Current approaches such as the Prosodic Phonology (Nespor & Vogel, 1986, *inter alia*) propose that the sounds in all languages are organized into non-recurrent,

abstract hierarquical constituents that are implemented by linguistically-principled algorithms. Yet, it is also highly parametric i.e. it has considerable variation on how languages use this structure to convey meaning. For special purposes, we often “slice” this (prosodic) component into minor domains, being intonation a recurrent notion.

Hirst & Di Cristo (1998) state that prosody and intonation are indistinctively used as synonyms and, when there is any distinction, it is not explicit. The term “intonation” has been defined in at least two different ways in the literature. A “narrow” definition relates intonation to ‘speech melody’, restricting it to “the occurrence of recurring pitch patterns, each of which is used with a set of relatively consistent meanings, either on single words or groups of words of varying length” (Cruttenden, 1986: 7; see ‘t Hart et. al, 1990: 2 for a similar definition). A broader account of intonation is equivalent to what is often called ‘prosody’, subsuming different phenomena such as pitch movements and range (speech melody), the division of speech into chunks (phrasing), highlighting at word level (lexical stress) and utterance level (accentuation), the marking of prominence relations (rhythm) and variations in speech rate (tempo) (see Crystal 1969; Moraes, 1993; Ladd, 1996). All these phenomena are realised with a combination of phonetic properties such as perceived pitch, loudness, vowel quality, and relative length of syllables, words and pauses.

One might notice that there is a gap between the physical realization of sounds and its linguistic counterparts. Sound production and perception are not always a one-to-one relation to acoustic properties and are frequently dependent on the speakers’ and hearers’ knowledge of the system.

Such impressions are evidently determined not only by the physical characteristics of the speech signal but also by the speaker’s linguistic knowledge and they somehow straddle the boundary between the physical world and a speaker’s abstract (cognitive) representation of that world.
(Hirst & Di Cristo, 1998: 6)

In many cases, the importance of each property relies on specifics of language systems. For example, in languages in which length serves as a phonemic contrast, loudness tends to play a more important role together with pitch (Cruttenden, 1986).

In spoken language, intonation serves diverse linguistic and paralinguistic functions ranging from simply conveying the sentence modality e.g. statement and question (Hirst & Di Cristo, 1998) to express attitude, such as irony and sarcasm, and emotions, such as surprise and sadness. It can convey a discoursal meaning like inviting a listener to make a contribution to the conversation (Cruttenden, 1986). Intonation also serves to determine the limits of information units, as in many situations intonation and syntactic boundaries overlap. Finally, intonation can signal relevant points of information in discourse, that is, it can signal the focus of information flow. In terms of prosody, the information focus is marked by a change in the pitch contour movement or by a variation of the loudness or length. As we might explore in more detail, we argue that intonation not only signals saliency of referents along the discourse (Bolinger, 1986; Chafe, 1974; Jackendoff, 1972; Pierrehumbert and Hirschberg, 1990; Prince, 1981; Cagliari, 1993), but it does so as a reflection of different cognitive status of the same referents (Chafe, 1994; Schumacher and Baumann, 2010).

2 Phonetics of Intonation

2.1 Pitch

Pitch is perceived by listeners in terms of lower or higher along the utterance and it is considered to be the main component of intonation. The term 'pitch' in fact refers to the perceptual correlate of speech melody. It has also physiological and acoustic correlates. The physiological or articulatory source of pitch variations are changes in the rate at which the vocal folds vibrate. The vibration is a result

of aerodynamic forces acting upon the folds after they have been sufficiently approximated by some of the laryngeal muscles.

Acoustically, pitch is correlated to fundamental frequency (f_0), which is the repetition frequency of a complex, quasi-periodic sound wave, equivalent to the highest common factor of the sound wave's component frequencies or 'harmonics' (Ladefoged 1962). Fundamental frequency is generally measured in Hertz (Hz), or cycles per second.

Fundamental frequency among men varies between 60Hz and 240Hz and among women between 180Hz and 400Hz¹² (Cruttenden, 1986). The rate at which the vocal folds vibrate is determined by their elasticity, length and mass (with vocal folds in females being shorter and lighter than in males, thus producing higher vibration frequencies) but also, more indirectly, by muscular tension and the amount of air pressure below the glottis. Increased tension stretches the vocal folds, which, in turn, thins their effective vibrating portion and leads to higher f_0 values. Enhanced subglottal pressure, on the other hand, primarily increases the amplitude of vocal folds vibration, not its frequency. Nevertheless, higher amplitude induces a greater deformation of the folds, which, in turn, leads to greater mechanical stiffness and, as a consequence, faster vibration of the vocal folds ('t Hart et al., 1990).

Other articulatory gestures e.g. point of articulation in the vocal tract seem to influence in a minor extent the intrinsic pitch of different segments. High vowels tend to have higher intrinsic pitch than low vowels (Whalen & Levitt, 1995), and the f_0 of vowels in general is affected by the voicing of adjacent consonants. Such minor influences, known as 'microprosody' or

¹² The average values of vibration rate reported in the literature seem to be a bit variable among different sources. In 't Hart et al. (1990), for instance, the values for males are 80-200 Hz and for females are 150-400 Hz. Laver (1994) affirms that the average value is 50-250 Hz to males and 120-400 Hz to females. In BP, Behlau and Pontes (1995) points out that the average values are 80-150 Hz to males and 150-250 Hz to females. Apparently, such differences seem to be insignificant and are due to effects of sampling. No study taking into account dialectal or crosslinguistic differences of the average of vibration rate was found.

‘microintonation’, are not considered during the decoding of the intonation contour by the listener.

The higher the frequency of vocal fold vibration - and, in turn, the higher the fundamental frequency of a sound -, the higher is its perceived pitch. However, pitch values can be perceived even if the f_0 is missing, as e.g. in hissed (i.e. voiceless) sounds or in conversations on the telephone, which does not transmit sound waves below 300 Hz. In these cases, the human ear can detect the fundamental frequency by measuring the difference between the harmonics (which are integer multiples of f_0).

Although pitch and fundamental frequency are correlates of the same feature - and often interchangeably used - they do not vary in the same rate. For instance, the perceptual difference between 100 Hz and 200 Hz is the same between 200 Hz and 400 Hz, though variation of the fundamental frequency of the latter was the double of the former. The way listeners perceive the pitch can be better approximated to the semitone scale, whereas the fundamental frequency is usually measured in Hertz. The difference between two physical frequencies can be calculated in the semitone scale by means of the formula below. This formula gives the value of f_{Hz} in semitones relative to a reference level of 100 Hz (st re100).

$$12 \log_2 \left(\frac{f_{\text{Hz}}}{100} \right) \quad (1)$$

2.2 Duration

Acoustic duration refers to the objective temporal extension of a speech segment. It can be derived from visual inspection of its waveform and/or spectrogram, and it is generally measured in seconds (s) or milliseconds (ms). In many languages,

the phonemic distinction between tense and lax vowels can have an effect on duration. As it is one-dimensional, that is, based only on time, it is the simplest prosodic feature. However, it is often necessary to establish some criteria when measuring segments in connected speech, especially in spontaneous register. The boundaries of adjacent segments are often not clear as the articulatory movements are dynamic. Also, voicing of near segments can influence their productions and its relative position in a linguistic unit - segments tend to be longer near unit boundaries or on stressed points (Cruttenden, 1986), though such lengthening does not seem to be relevant from the listener's point of view.

Length is also related to the speech rhythm, which is the result of the interplay between the segments duration and their prominence along the utterance. Lehiste (1970) associates duration to a phonetic interpretation. Length can be stretched or compressed to adjust to the speech rhythm, and it interacts with accent rules, vowel quality, overall intonation and speech tempo.

2.3 Stress and accent

In a general way, stress can be regarded as the prominence of strong and weak syllables (Ladd, 1987:108). Though such definition might be little revealing, it reflects the difficulty to determine which acoustic parameters are associated to it.

On one hand, American phonologists such as Bloomfield (1935), Pike (1945), Trager & Smith (1951) and Chomsky & Halle (1968) thought of accent as associated with a single acoustic parameter. Stress, for instance, was associated with loudness whereas intonation was associated to pitch variations in the utterance. On the other hand, British linguistics such as Jones (1950) and Kingdom (1958) suggested that stress is a multi-dimensional feature, that is, stress and intonation are intrinsically related.

In a series of perception experiments in English, Fry (1955, 1958) used synthesized minimal pairs like 'PERmit' (as a noun) and 'perMIT' (as a verb) in a carrier sentence like "Where is the accent in ___". These words were shown with combinations of different vowel intensities, durations and f_0 patterns. Subjects were instructed to say which syllable was accented. Results suggest that f_0 is the most effective cue to the perception of accent, followed by duration and intensity. Later investigations have shown that, though f_0 is the most relevant correlate to stress, it does not entail the irrelevance of the other phonetic parameters (loudness and length). Nakatani and Aston (1978) showed that the effect of duration and vowel quality was as the same as f_0 in prenuclear position in the sentence. In nuclear position, f_0 was the most effective cue whereas duration was to be the most effective cue in postnuclear position.

Based on Fry's results, Bolinger (1958) proposed that a stressed syllable is the potential location for the *pitch accent*. Pitch accents are expressed by tones like H (high) or L (low) or tone complexes like HL (high-low) or LH (low-high). The tone associated to a stressed syllable is marked with an asterisk¹³ (*). For instance, a H* means that the stressed syllable has a high tone; a H*L means that the stressed syllable has a high tone with a low trailing tone; likewise, a LH* means that the stressed syllable has a high tone with a low leading tone. The location of pitch accents is determined by several factors that are variable across languages from lexical rules - that determine the main stress and thus the location of pitch accent in a word - to information structure - which causes the deletion or dislocation of pitch accents according to focused or given information (Gussenhoven, 2002). As pitch accents are tones associated to a stressed syllable, there are tones associated to the edges of intonational groups; in this case, they are called *boundary tones*. Boundary tones are single tones and are marked with a % symbol (H% or L%) to represent the edge of intonational phrases and with a - symbol (H- or L-) to

¹³ The symbol for the location of the stressed syllable is frequently called 'star' when describing pitch accents.

represent the edge of phonological phrases (Pierrehumbert, 1980 and Pierrehumbert and Beckman, 1988).

Based on Pierrehumbert's (1980) study on the American English intonation system, Silverman et al. (1992) and Pitrelli et al. (1994) proposed a notation system under the name of ToBI (Tones and Break Indices) in order to describe tonal events American English. For the last years, similar language-specific notation systems were developed like G-ToBI (for German, see Grice and Baumann, 2002), Sp-ToBI (for Spanish, see Beckman et al., 2002). In BP, a recent alternative to ToBI notational system is DaTO (Dynamic Tones of Brazilian Portuguese, Lucente & Barbosa, 2008; 2010).

The implementation of pitch accents (and thus the intonation contour) is regulated by universal principles, but its distribution and meaning can be language-specific. That way, the Autosegmental Metrical Theory proposes a phonological analysis to identify contrastive elements of the intonation system whose combination produces pitch contours that are available in the languages. Each intonation unit of information presents a single nuclear stressed syllable. In a neutral utterance, for instance, the nuclear tone (or nuclear accent) is located in the stressed syllable of the last lexical word of the intonational group. It is in the nuclear accent that a meaningful change in the pitch contour occurs (Ladd, 1996).

Among the several studies that have investigated the intonation structure in BP, Moraes (1998, 2007, 2008) presents a comprehensive overview. First, pitch accents are realized by tone complexes with a leading tone followed by a starred one. Leading tones are always realized on the syllable immediately preceding the stressed syllable. Single tones or tone complexes with a tailing tone are not part of the language intonational inventory. Second, different from the poor productivity of the boundary tone, there is a large variety of pitch accents, which means that the intonational contrasts lie specifically on the last stressed syllable and its preceding syllable, specially in the intonational phrase final position.

Finally, there is always a nuclear accent in the final position (last stressed syllable) of an intonational phrase, even in postfocal position. It means that when the element that receives focus is not in the final position, the nuclear accent remains unaltered, entailing the dissociation between focal accent and nuclear accent. The same phenomenon is also present in other languages such as European Portuguese (Frota, 2000) and Italian (Grice et al., 2005)¹⁴.

¹⁴ The phenomenon of keeping the last stressed syllable with an accent is referred to as reaccenting as an opposition to languages in which postfocal elements have their accent dislocated (deaccenting). In the next chapter, we dedicate a whole section to a crosslinguistic review.

Chapter 3

Prosodic marking and information status

In this chapter, we part from the fundamental concepts reviewed at the previous chapters to discuss the marking of information status, especially concerning prosody. Traditionally, the studies in the area have shown a straightforward relation new/prominent and given/non-prominent. We will discuss this idea analyzing crosslinguistic data.

1 Introduction

As we have discussed, information structure (or information packaging) relates to the way the propositional information is organized in discourse. Many studies argue that the surface structure of the sentence (and by extension the discourse as well) reflects general constraints that communicate with linguistic information (Chafe, 1976).

One of the evidences for such account is the linearization of information structure. In general, given information should appear before new information (Halliday, 1967; Clark and Clark, 1977; Clark and Haviland, 1977). Gundel (1988:

229) also proposes the idea that information status is related to word order in sentence under the "given-before-new principle". Furthermore, Arudd et al. (2000) shows that structural complexity and discourse status have independent influence on constituent ordering in sentences. As a consequence, complex phrases and newsworthy material tend to be placed in the end of the sentences.

A second evidence is that prosodically prominent words play a role in signalling the discourse status in discourse (Jackendoff 1972; Chafe 1974; Prince 1981; Bolinger 1986; Pierrehumbert and Hirshberg, 1990). Terken and Nootboom (1987) observed that listeners had more difficulty to process given information when it was accented. The authors remarked that given information can be accented, but usually when they are related to focus/contrastive information¹⁵. The authors presented a processing account for deaccenting with an algorithm for accent patterns in reference resolution based on perception experiments with (de-)accenting of given and new referents. According to them, deaccented expressions are interpreted within the set of already activated entities in the discourse. Accented entities, on the other hand, are interpreted chiefly on the basis of the information contained in the speech signal. In this way, new information needs to be more acoustically salient or "well pronounced". When an addressee listens to a deaccented new referent, he will unsuccessfully try to map the entity onto the set of activated referents. Reversely, accented given referents would lead the listener to construct an interpretation from the bottom-up input. In the last two scenarios, the listener will have to reanalyze the information to adjust the entities status to the context of the discourse, which increases the cost in referential processing.

Recent studies using modern techniques such as eyetracking (Van Hooijdonk et al., 2007; Dahan et al., 2002; Watson et al., 2008; Ito et al., 2011) and electrophysiological measures (Schumacher and Baumann, 2010; Baumann &

¹⁵ Brown (1983) points out that when a given entity is re-introduced into discourse after some time, it is marked by a pitch accent. It seems that, in this case, recency is a major aspect concerning activation of referents and, thus, their prosodic marking.

Schumacher, 2012) have comparable results.

Using visual world paradigm in an eyetracking study, Watson et al. (2008) examined the role of acoustic prominence during referential processing. A secondary aim of the study is to understand whether acoustic prominence is speaker- or listener-centered. According to the authors, speakers might aim to produce intelligible language when there is less information from the context to help the listener, as in the unpredictable words (Lieberman 1963). On the other hand, predictable words might be less prominent because they require less effort by the speaker. Speakers face the challenge of preparing and uttering their conversational contributions in real time, often while completing other, nonlinguistic tasks. When these demands require speakers to plan complex utterances or prepare upcoming words, they tend to produce hesitations, repetitions, intonational phrase boundaries. These demands also result in longer word duration. All in all, it suggests that when speech is effortful, word durations are longer, contributing to the impression of acoustic prominence. Such results are related to cognitive approaches that associate more complex information to increase in processing (Ariel, 1990; Almor, 1999)

Schumacher and Baumann (2010) investigated how the informational status is processed in combination with prosodic cues using electrophysiological measures (EEG). In this experiment, the participants listened to sentences with accessible referents (whole-to-part) like in (13):

- (13) Sabine repariert einen alten Schuh. Dabei zerschneidet sie **die Sohle**.
Sabine repairs an old shoe. In doing so, cuts she **the sole**.
(Schumacher e Baumann, 2010: 618)

The target referents were manipulated into three conditions: referents with manipulated given pitch accent (deaccented, with a L+H* pitch accent associated to the verb anteceding the target referent), referents with manipulated new pitch

accent (H*), and referents with no manipulation (H+L*). Participants' reactions were monitored using an EEG. The authors analyzed two Event-Related Potential components: N400 and late positivity. The N400 component is commonly associated with the effort to access the lexical semantic referents, i.e. how the referents are activated in memory¹⁶. The late positivity¹⁷ is associated with the integration effort for the new discursive model. The working hypothesis is that a divergence of prosodic information expected by participants hinders access of lexical referents and, consequently, their integration to the discursive model. The experiment results show that when there is a change of the typical accent pattern of the target item (addressee listens to an accessible referent with either a given or new accent pattern), participants produce an N400 and a late positivity components with higher amplitude compared to the same referent with the expected accent pattern. These results are an evidence that prosodic information is taken into account during referential processing: "The data show that the prosodic information guides accessibility of reference computing and may result in integration costs when accentual patterns are found less suitable" (Schumacher and Baumann, 2010: 621).

Gussenhoven (2002, 2004) defends the existence of biological codes that relates to the meaning of intonational contours. According to him, these biological codes explain the universality of the interpretation of pitch prominence, for instance. The Effort Code relates the amount of energy put into speech production to the amount of emphasis of the message. If the speaker puts more effort when producing speech, it means the message importance increases proportionally.

¹⁶ More recent studies (Kutas & Federmeier, 2011) have shown that, in fact, the N400 component relates to the addressee's level of expectancy of upcoming information. In an incremental perspective of referential processing, the addressee updates its discourse model as he receives more linguistic material and his comprehension system tries to "predict" new referents. The higher the match between the expected information and actual received one, the lower the amplitude of the component.

¹⁷ Late positivity was early known as the P600 component. As the peak of this component can be registered in a longer time window and it is traditionally associated with syntactic well-formedness, some authors have decided to adopt the term **late positivity** to avoid misinterpretation. It is lately also associated with the integration of referents in the sentence or in the discourse. The easier to integrate a referent in the discourse, the lower the amplitude of the component.

The acoustic result is greater articulatory effort (e.g. articulatory precision) and an increase of frequency of the vocal cords vibration, that is, higher pitch values.

It seems that contextual information also has a deal with the way information is signalized. Gregory (2002 *apud* Watson et al., 2008) observed that the more predictable a word, the more likely it is to be produced without a pitch accent. Bolinger (1963, 1981) suggests that when talkers utter words that are unusual in their contexts, they lengthen them, i.e. talkers lengthen words that have little contextual support, or more generally, that have little or no other information than their acoustic signal. Likewise, Lieberman (1963) found that words with favoring context are less intelligible than words that are produced in an uninformative context, that is, speakers tend to make less articulatory effort when they know that the context will save the hearer.

In a series of experiments to examine the duration of repeated words, Fowler (1988) observed that the simple repetition of a lexical item does not guarantee shortening. According to the results, words repeated in lists have no shortening while reductions were observed for the same words in a meaningful context. Homophones also did not suffer shortening. Finally, repeated words produced in a communicative context underwent more reduction than the same discourse transcribed and read. Fowler suggests that the shortenings reflect the addressee's awareness of the redundancy of the entity in the discourse, and not simply a mapping to the lexical form. The differences found in spontaneous speech compared to the read one also suggest that the speaker behaves more cooperatively by the presence of a listener, which may reflect in how much the acoustic correlates are manipulated. These results also corroborate Gussenhoven's proposal. Similar results were found in other studies (Fowler and Housum, 1987; Gregory et al., 1999, Bell et al., 2003)¹⁸.

¹⁸ We will see more details on the results found in Fowler and Housum when we discuss the effect of information status on duration.

2 Encoding information status

As Vallduví (1992) points out, there is considerable crosslinguistic variation in how languages encode information packaging. In Japanese, given information is marked morphosyntactically by the particle *wa* and new information is marked by the particle *ga* as in (14):

- (14) Mukashi mukashi, **ojii-san ga** sunde imashita.
Once upon a time, **a man** old there lived.
Ojii-san wa totemo shinsetsu deshita.
He very kind was.

Another form of morphosyntactic marking is through the use of determiners. In general, definite articles are used to mark already mentioned NPs whereas indefinite articles are used for newly presented NPs¹⁹. In some cases, however, syntax has to work alone: as Turkish does not have articles, accusative case is said to mark definite nouns thus given information (Görgülü, 2008).

A number of studies (Contreras, 1976; Vallduví, 1991, 1995, Zubizarreta, 1998, *inter alia*) have observed a two-way distinction on how languages prosodically encode information structure with the PLASTICITY parameter of intonation (Vallduví, 1991).

In the so-called *plastic languages*, intonational proeminences may be shifted to different positions in the utterance to mark the point of information relevance while the syntactic structure tends to remain unaltered, that is, the information status of discourse referents is encoded by the distribution of pitch accents along the utterance. As we are about to see in the next section, typical plastic languages

¹⁹ It is important, however, to point out that definiteness is not a synonym for givenness. Both are distinct categories, though they may have a high correlation. Inferable (accessible) information, for instance, can be marked by definite articles even not being previously mentioned in discourse:

¹⁹A artista vendeu um dos quadros sem **a moldura**. (The artist sold one of the paintings without **the frame**.)

rely on phenomena such as deaccenting in order to signal given information.

On the other hand, *non-plastic languages* have the intonational prominences fixed in the utterance, that is, the distribution of pitch accents is not affected by the referent's information saliency. The management of information structure is subject to word order and relies more on syntactic operations. One direct consequence to this characteristic is that languages in this group tend to block deaccenting of given referents. Instead, reaccenting of referents (the maintenance of pitch accent) appears where it is expected deaccenting in plastic languages. Because they have fixed pitch accent distribution, non-plastic languages use more f_0 magnitude variation as a cue to signalize information status.

2.1 De-accenting, re-accenting

By definition, deaccenting is the lack of an accent in a position where it would be expected (Ladd, 1996). Generally, it means that the accent is dislocated to a near constituent and the referent gets a flattened contour.

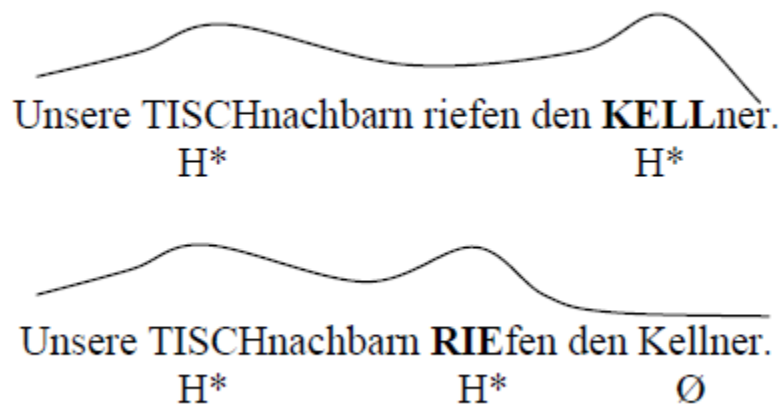


Figure 6 - Schematised intonation contours of the sentence 'The people at the next table called the waiter' (Baumann, 2005: 9, adapted)

In figure 6, the prominence on *den Kellner* (the waiter) is dislocated to the verb *riefen* (called) and the NP gets deaccented.

Many researchers have proposed that there is a close correspondence between information status and (de-)accentuation (Halliday, 1967; Cruttenden, 1993, 2006; Ladd, 1980, 1996). As observed by Cruttenden (2006: 314), this phenomenon has been described for a long time in English: “[t]hose things with which we suppose our hearers to be pre-acquainted we express by such a subordination of stress as is suitable to the small importance of things already understood” (Walker, 1781: 14) and “[i]n phrases and sentences all words that express ideas new to the context are distinguished by accent; and all words that have been previously stated or implied are unaccented” (Bell and Bell, 1879: 11).

Chafe (1976: 31) states that “the principal linguistic effects of the given-new distinction in English, and perhaps all languages, reduce to the fact that given information is conveyed in a weaker and more attenuated manner than new”.

Cruttenden (1993: 16) also notes that de-accenting might easily be understood as some sort of cognitive universal: “we do not wish to re-accent repeated information because [...] it is already in the consciousness of the speaker”.

In West Germanic languages, the pitch accent is used to mark referential status: inactive referential expressions (new) are often marked by an accent (H*), and active items (given) are typically deaccented (Ladd, 1996; Pierrehumbert and Hirshberg, 1990). Brown (1983) used the Assumed Familiarity Taxonomy (Prince, 1981) to describe variations in intonational prominence in English task-oriented speech. She found that speakers tend to place pitch accents on new information, while marking given information by deaccenting. Other studies in English (Terken & Hirshberg, 1994), in German (Féry & Ishihara 2005; Baumann et al., 2006; Féry & Kügler, 2008, *inter alia*), Danish (Thorsen 1985; Grønnum 1992) and Dutch (Kruyt, 1985) have found comparable results.

The prosodic marking of accessible information remained, however, with no consensus. Baumann (2006) proposes that there are different accentual patterns related to different degrees of informational status, beyond the dichotomy new-accent/ given-deaccented. The idea is that the type of accessibility of a referent correlates with the type of accent used to mark it. The more active a referent, the greater the chance it is deaccented. The less active a referent, the greater the chance that it receives an accent type H*²⁰. When the referent is between the ends of the activation spectrum, the greater the chance it receives an intermediate accent type (H+L*). Baumann analyzed the intonation pattern by reading a corpus of German MULI (Multilingual Information Structure, Baumann et al., 2004), in addition to performing two experiments in perception. The author's conclusion is that there is a high correlation between the activation of the referent and its prosodic marking in the statement. As also observed in Allerton (1978), the preference for marking intonation of referring expressions varies according to the relation between the antecedent and its correferent (metonymy, synonym, hypernym, hyponymy).

Nevertheless, as other studies have pointed out, deaccenting might not occur in the same way for all languages (Ladd, 1996; Cruttenden, 2006). It seems that deaccenting works well for most English varieties, though dialectal variation might affect it; it seems that deaccenting does not occur in Hawaiian English:

- (15) Forty t'ree per cent is gavment owned, and fifty seven per cent is privately OWNed.
(Vanderslice & Pierson, 1967)

In an expansion of the study in 1993, Cruttenden (2006) created a cross-linguistic study in order to observe if the deaccenting rule proves to be universal. He

²⁰ The notation for pitch accents (H*, H+L*) are from GToBI (German Tones and Break Indices, Grice and Baumann, 2002), an adaptation of the original ToBI system for American English to annotate intonation events along the sentence (Pierrehumbert, 1980; Silverman et al., 1992).

applied ten different setting-response dialogue types where the response involved a lexical item repeated from the setting.

- (16) A: If you don't hurry up, you'll be late.
 B: I don't care if we are late. / (I don't care if we're late)
 (Cruttenden, 2006: 318)

Table 6 - Reaccenting in twelve languages. The numbers on the first row indicate the number of each setting-response (Cruttenden, 2006: 319)

	1	2	3	4	5	6	7	8	9	10a	10b	10c
Albanian (4)	0	1	0	0	3	2	0	0	3	3	2	-
Arabic (4)	1	0	0	4	3	1	3	4 ^a	3	3	4	3
English (7)	7	7	7	7	7	7	7	7	7	7	7	7
French (14)	3	6	13	3	11	3	2	14 ^a	14	14	10	11
German (4)	0	1	0	0	0	0	0	0	-	-	0	0
Greek (4)	0	1	1	1	3	4	0	0	-	-	-	0
Italian (11)	0	0	0	5 ^b	-	3	1	-	10	10	8	7
Lithuanian (4)	0	0	0	0	0	0	1	0	-	-	4	2
Macedonian (2)	1	1	1	2	0	2	0	0	-	0	0	1
Russian (4)	1	0	0	0	1	0	0	-	1 ^c	-	3	3
Spanish (4)	3	1	4	4 ^b	4	4	3	2	4	3	-	-
Swedish (4)	3	4	3	0	3	2	2	3	-	-	2	1

(Total number of informants shown in brackets)

^a Involves the change of gender

^b Involves a change to a causative verb

^c Only one out of two (other two responses irrelevant)

The dialogues were translated from English into 11 different languages and read by native speakers: French (principally Provence), Italian (principally Sicily), Lithuanian (Vilnius), Arabic (Tunisia), Albanian (Tirana), Russian (Moscow,

Dushambe), Spanish (Chile, Castilian, Catalanian Spanish), Greek (Athens), Macedonian (Skopje), Swedish (Skåne, Östersund), and German (Westphalia, Munich, Tübingen/Hamburg). After the recording, the intonational nucleus was annotated by the author and another phonology expert.

The results in Table 6 show that the deaccenting behaviour is in fact not too clear-cut. It seems that it is rather a spectrum in which languages vary the degree to which a repeated lexical item de- or re-accent. Most languages behaved similarly to English, but the amount of deaccented items varied. In French, eleven speakers put the last accent on *fiche* and three re-accented *retard*:

- (17) Je m'en fiche d'être en retard.
(Cruttenden, 2006: 319)

In Spanish and Swedish, three out of four speakers reaccented the word *tarde*:

- (18) No me importa si llegamos tarde.
(Cruttenden, 2006: 319)

It is important to point out that the repetition of a lexical item does not necessarily mean that both items refer to the same entity in the world, that is, the use of the same lexical form does not ensure coreferentiality.

- (19) A: You need a pair of black SHOES for the wedding.
B: I've already GOT a pair of black shoes.
(Cruttenden, 2006: 320)

Note that in (19A), the phrase *a pair of black shoes* denotes a generic or non-specific referent, whereas in (19B) it denotes a specific referent and can be considered as new (which is supported by its morphosyntactic marking by an indefinite article, cf. Baumann & Schumacher, 2012). Given, in this case, is the

repetition of the lexical form, not the referent itself. In natural speech, given information tends to be pronominalized, though sometimes other lexical choices can be made.

(20) A: Did you see Dr. Cremer to get your root canal?

B: Do not remind me. I'd like to [STRANgle]_F the butcher.

Do not remind me. I'd like to [strangle the BUTcher]_F.

(Baumann & Riester, 2012: 12, *apud* Buring, 2007: 4)

In (20), *the butcher*, despite being new in the speech, does not have accent as it is interpreted as coreferent to *Dr. Cremer*. Allerton (1978) has observed that the semantic relationship of nouns has an influence on accenting.

(21) Bach wrote many pieces for viola. He must have LOVED string instruments.

(Baumann and Riester, 2012: 13)

In (21), despite the lack of lexical identity, there is a hypernymy relationship, which allows the addressee to map *string instruments* to *viola*. The other way round, however, does not seem to work well:

(22) Bach wrote many pieces for string instruments. He must have loved the viOla.

Bach wrote many pieces for string instruments. He must have LOVED the viola.

(Baumann and Riester, 2012: 14)

Apparently, when the anaphor is a hyponym, it brings new information to discourse, which in this case must be marked (Allerton, 1978; Nolan & Jónsdóttir, 2004; Baumann & Riester, 2012). Such problem has been justified the division of informational status into two levels (cf. RefLex system from Baumann & Riester,

2012 as seen in the first chapter). As we can see, the relationship between accentual marking of an item depends not only on linguistic material presented (r-level) as the semantic relationship between the referent and its coreferent (l-level).

Romance languages have been a challenge for the universality account of deaccenting of given information. Swerts et al. (2002) conducted a crosslinguistic study with Dutch and Italian in order to assess deaccenting in both languages. The analysis of map-task dialogues reveals that Dutch has a strong tendency to deaccent repeated nouns whereas the Italian counterparts were all reaccented. The authors concluded that the lack of de-accenting in Italian is due to a rule that prevents pitch accents from “moving [...] inside syntactic constituents in order to convey contextual givenness” (Swerts et al., 2002: 632). They point out that an alternative strategy by which Italian could mark information status prosodically would be to use a different nuclear accent shape or contour, notwithstanding in the same position.

In a study for Roman Italian using task-oriented dialogues, Avesani and Vayra (2005) concluded that, contrary to Germanic languages, the vast majority of the repeated mentions are accented irrespective of their information status (93% of repeated nouns). As the authors point out, deaccenting is possible in Italian, but only in special occasions such as contrastive focus:

- (23) con tre segnetti?
with three little signs?
s/ con DUE segnetti.
yes with TWO little signs.
(Avesani & Vayra, 2005, adapted)

Swerts (2007) used a similar protocol of dialogue task as in Swerts et al. (2002) in a crosslinguistic study with Romanian and Dutch. The results show that, likewise

Italian, Romanian has a very low rate of accent shift for given and contrastive information.

On discussing nuclear tone in BP, Crystal (1975: 44) concluded that “this tendency [of acceting] applies even when one has repeated items in coordinate constructions”.

(24) Este livro custa cinco dólares e este aqui três DÓlares. (adapted)²¹

Languages with tones have also shown also examples where reaccenting is much more regular than deaccenting (Ladd, 1996). Hindi and Mandarin Chinese do not have a complete deletion of the accent target, but a register compression that allows tones to be realized (Patil et al., 2008; Xu, 1999). In Akan, a level tone language spoken in Ghana, the observed pitch contour lowering is triggered by a post-focal effect (Kügler & Genzel, 2011). As tonal languages have lexical and grammatical distinctions encoded by tones, the pitch compression still permits the expression of them where it would be expected to have deletion in other languages.

In a comparative study with English and Icelandic, Nolan & Jónsdóttir (2001) analyzed read examples of four speakers of each language where given information was repeated (given) in a setting that would provoke deaccenting. Contrary to their English counterparts, Icelandic speakers kept the accent where the English speakers deaccentend, except for a small minority that would do so in cases of hypernym substitution. Valgeirsson (2011) found the same reaccenting pattern in an analysis of Icelandic newsreaders.

Hellmuth (2005) used Swerts et al. (2002) map task protocol to evaluate

²¹ A remark is important for the example above. Though it shows reaccenting of repeated mention, the nouns are in contrast of information (cinco vs três) and therefore such example might not be the best example to show accent distribution in BP.

deaccenting of given information in Cairo Arabic. Results demonstrate that given items positioned after an in-situ focus are not deaccented.

As one can see, experimental data reveals that it is far from being a cognitive universal. As Ladd (1980: 98) remarks, “the deaccenting 'rule' in the more usual sense of the word rule is simply the focus rule - accent goes on the most accented syllable of the focus constituent - together with our knowledge of the relative accentability of different items in different contexts.” Deaccenting seems rather one variable than a universal across languages. Furthermore, it has been shown that it is more gradual than categorical, as some languages use de- and re-accenting in a more or less extension. In English, for instance, deaccenting is almost obligatory whereas Italian strongly favors reaccenting.

2.2 Other means of acoustic marking

As we discussed in the beginning of this chapter, non-plastic languages tend to use acoustic variation of correlates such as duration, f_0 and vowel quality than accent distribution in order to mark focus and information status.

Wang & Xu (2011) observed that information status in sentence initial position seems to affect duration in Mandarin Chinese. On average, new referents are 5% longer than given referents). Pan et al. (2005) analyzed the production of NPs with given, new and contrastive information status in spontaneous Taiwan Mandarin speech. Production experiments showed that when the NPs carried new or contrastive information, the duration was lengthened. Results on perception of this correlate showed, however, that listeners were not able to use duration as a distinctive cue for given and new information.

Even typical plastic languages (like English) are described as using acoustic variation. Fowler & Housum (1987) observed that the second mention of nouns in English tend to be “attenuated”, that is, they are less intelligible and shortened. As

the authors suggest, listeners make use of priming of already presented words in discourse, that is, as they have already figured out the referent, they just need to use the acoustic signal to retrieve the referent from memory (a similar explanation was also presented in Terken & Nootboom, 1987). Other studies also observed duration effects on information status (Eady & Cooper 1986; Wolters, 1999; Beaver et al., 2007). German has also presented a relation of lengthening of newly presented nouns (Baumann et al. 2006; Kügler, 2008) and second occurrence focus (Féry & Ishihara 2005).

Besides lengthening, f_0 variation is also a major acoustic cue for information status, especially for languages in which accent is lexically determined. Speakers of Tokyo Japanese tend to mark new information with higher pitch range than given information. However, it seems that listeners are not sensitive to this pitch variation when assigning information status (Swerts et al., 2000).²² In Peking Mandarin, expansion of f_0 range seems to be a more salient marker than duration. In Taiwan Mandarin, however, duration is reported as being a more salient marker for information status (Pan et al., 2005).²³

Kang (1996) has found that speakers of Seoul Korean use acoustic correlates such as duration and f_0 variation to mark information status, but also qualitative changes in accent phrasing. Similar findings are also reported in other studies for Korean (Ladd, 1996; Lee, 2012) and Japanese (Venditti, 2000; Venditti et al, 2008).

Norlin (1989) conducted a pilot study of Cairo Arabic intonation and found evidence of gradient rather than categorical marking of focus condition. He found that pre- and post-focal pitch accents underwent a compression instead of pitch deaccenting. Similar effects have also been reported by Chahal (2001, 2003) for Lebanese Arabic: the contrast between focussed and out-of-focus elements in

²² As we have discussed before, it might be explained to the fact that Japanese relies on explicit morphosyntactic marking (with the particles [wa] and [ga]).

²³ Such difference has been observed due to influence of another language (Taiwan Min) on Taiwan Mandarin. In Min, the salience of duration in broad and narrow focus is dramatically different than f_0 range.

a sentence was enhanced both increasing acoustic cues in focus position (enlarging f_0 excursion) and decreasing cues in out-of-focus positions.

Miglio et al. (2014) conducted an experiment in order to explore pitch variation in Mexican Spanish and Chicano Spanish²⁴ in natural speech. They used two corpora of semi-directed interviews, analyzing 20 speakers per group. The results show that new information is marked by broader pitch excursion and lengthening of stressed vowels. For given information, however, there seems to be an interaction between vowel length and pitch excursion. Vowels shorter than 110 ms had no pitch excursion whereas vowels longer than that were also marked by pitch excursion. The authors also observed differences in accenting: Mexican Spanish showed flatter accents (L-L) for given and rising (L-H) for new information; Chicano Spanish showed an H-L accent for both given and new.²⁵

2.3 Latest work in Portuguese

Most studies in BP and EP (European Portuguese) which explore the prosody manifestation in information structure tend to explain its relation with topic and focus. Information status, however, is often overlooked and taken under other perspectives rather than the activation one.

Moraes (2006) briefly investigates prosodic manifestations of focus in Rio de Janeiro dialect. Though he remarks on the terminology profusion around the subject, he ends up assuming a direct relation of theme and rheme as “respectively, its portion which constitutes the previously given information, or inferable, thus irrelevant (or less relevant) to communication, and the part which corresponds to its central, new information.”²⁶

²⁴ According to the authors, it is a variant spoken by English-Spanish bilinguals from California.

²⁵ As the complete study was not available at the time of the preparation of this thesis, we were unable to check which notation system was adopted to describe accents in both variants of Spanish.

²⁶ “[...] respectivamente, sua porção que constitui a informação previamente dada, ou inferível, portanto não (ou menos) relevante para a comunicação (o tema) e a parte que corresponde a sua

In a spontaneous speech analysis, Stein (2005) discussed the relation between introduced entities into discourse and acoustic prominence. The author's theoretical discussion included Chafe's (1994) activation notion and Lambrecht's (1994) identifiability notion as well. Based on Lambrecht's proposal, Stein concluded that identifiability is a more appropriate term than definiteness, as the latter is categorial (it is either definite or indefinite) and the former can be gradual. It is not clear, however, whether the activation is also gradual nor how it is related to identifiability. It seems that givenness is associated with identifiability though. Also, his hypothesis includes the term focus to express the acoustic prominence assigned to newly introduced referents. The author used a relatively small corpus of spontaneous speech: there were primarily 12 extracts, only 8 being used in the final analysis. The experiment consisted in the comparison of the occurrence of prominences along the pitch contour and a perceptual impression of 25 listeners to "where they judge the sentence focus were". The author states that about 40% of listeners agreed in the focus position assignment, and no statistical test was carried out.

(25) ...chegando de Israel e aí... ele tinha interesse ele já tinha trabalhado lá... um pouco com essa área de processamento digital da voz né:... que é uma área que teve bastante::: né:::... *produtiva* éh: por aí na UFRJ na PUC também. (Stein, 2005: 6)

... coming from Israel and then... he had an interest he had already worked there... a little with this voice digital processing area isn't it:... which is an area that has been really::: isn't it::: *productive* uh: around UFRJ and PUC too.

The extract above is an example of the sentences analyzed in the experiment. As one may see, the analysis of spontaneous speech is rather complicated due to the low frequency of comparable conditions. The impressionist analysis showed to be ineffective due to an unsatisfactory level of agreement among the listeners.

informação central, nova (o rema, ou foco)."

Besides, spontaneous speech is full of repairs and hesitation pauses, which have a great deal on how the speaker signals relevant information to the addressee. The author points out a pitch prominence in 'productive' as an evidence of the introduction of a new entity. We may disagree with this conclusion, as we would expect a prominence in the beginning of the NP (especially in BP in which the modifier usually comes after the noun). One possible explanation for the prosodic marking in 'productive' is the presence of a long hesitation before it. It is known that the speaker can make use of intonation to compensate a longer pause in speech (Delfino e Magalhães, 2010).

In his study on secondary stress in BP, Arantes (2010) investigated which acoustic parameters are related to givenness. Different from previous studies in the area, he isolated referential status from topichood and focus. As he describes, along with the repetition of the same referent, givenness in BP can be marked with determiners. Definite determiners (o, a, os, as) are traditionally considered as given status markers whereas indefinite determiners (um, uma, uns, umas) are related to new status marking. He also acknowledges the existence of a continuum of referential statuses, though in his dissertation he adopted the dichotomy given-new. As one can see in the sentence below, the first instance of 'deputy' introduces a new referent; its second occurrence has, thus, a given status.

(26) *Um deputado* foi depor no plenário da Câmara Federal.

A deputy went to the National Congress to make a deposition.

O deputado foi evasivo em suas declarações.

The deputy was evasive in his declaration.

(Arantes, 2010: 89)

Arantes analysed the duration and the fundamental frequency of the target words. The results show that the most expressive effects are the global variance of f_0 and, in a lesser account, the quantity and alignment of f_0 extremes along the target words. Besides, duration also showed to be a relevant parameter to mark

givenness. Another important remark concerns the size of the target words. As the author observes, the acoustic parameters were more expressive in longer words, as the features could 'spread over' more sound material. Such difference is especially explicit in the pitch contour on prestressed syllables.

Lucente et al. (2011) investigated the distribution of information status and its relation to intonation using the VoCE spontaneous speech corpus²⁷. The authors classified the nouns as given or new following Prince's (1981) Assumed Familiarity Taxonomy and then carried out the data analysis. Preliminary results showed that there is a high correlation between information status and grammatical function, that is, given information tends to occur in subject position and new information in object position. The acoustic prominence vary according to the position of given and new referents in the utterance or the presence of an adverb or another referring expression before the target word. The pitch accent analysis showed that given NPs are usually marked by a LH or >LH contours while new NPs are marked by a LHL or HL contours²⁸.

Arantes et al. (2012) conducted a production experiment in which they tried to verify which acoustic cues play a role in encoding of referential status in BP. One of the merits of this study is the detailed control of the acoustic parameters. According to them, as most studies regard f_0 as the primary correlate of information status, they decided to investigate other phonetic parameters: duration, spectral emphasis and voice quality. For the experimental design, target words were controlled for the number of prestressed syllables (two to four) and stress position (last or second to last syllable lexical stress). In total, 270

²⁷ The VoCE corpus is a corpus essentially composed of spontaneous speech data from radio broadcasts and podcasts from the internet. This corpus has been being collected since 2006 and contains approximately 10 hours of conversation among announcers and guests of interviews in mp3 format from Rádio Você AM (Americana, Brazil) and Rádio CBN (São Paulo, Brazil).

²⁸ DaTo annotation system intends to represent the BP intonation features using a function-oriented dynamic analysis. LH (rising) indicates a high peak reached on the stressed vowel; >LH (late rising) indicates a high peak tailing the stressed syllable; LHL (rising-falling) indicates a high tone with a second tone always lower than the preceding one; HL (falling) indicates a low tone on the stressed vowel.

instances were produced, considering that the combination of the levels for each condition had five repetitions. The subject was a male graduate student from the city of Piracicaba, in the countryside of Sao Paulo state. The sentences were read aloud and recorded for analysis. The results show that some of the parameters had little or no significant role, like spectral emphasis and long-term average spectrum. However, parameters such as duration and f_0 revealed to be relevant parameters. As the authors conclude, the scope of the results was due to some limitations. First, the use of read sentences instead of spontaneous material could have affected the overall results. Second, the analysis of one single informant can increase the difficulty of the generalization of results to the population. It is known from other studies that this phenomenon is susceptible to subject (Wolters, 1999) and dialectal (Pan et al., 2005) variability.

Given the lack of detailed descriptions of prosodic parameters and their relationship with the referential status in BP, this work presents a description of accentual patterns in BP and its relation to the structure of referential expressions. Important to say, the aim of this study is not a phonological analysis of givenness in BP. We understand that a more detailed description of the phenomenon from the production point of view is necessary, taking into account the analysis of acoustic correlates that then would serve as a basis for a phonological proposal.

In the next chapter, we present the production experiments that explore the acoustic correlates of referential status.

Chapter 4

Production Experiments

In this chapter, we present the production experiments relating acoustic correlates to referential status. As we have stated in the previous chapters, we adopted the activation taxonomy proposed by Chafe (1994) with three degrees. New referents are associated with inactive status; given referents are associated with active status; and accessible referents are associated with semi-active status.

1 Experiment 1

1.1 Materials

For this experiment, we designed a list with 30 target words, each of them in three conditions (given, new and accessible statuses), summing up a total of 90 narratives in short paragraphs. Each target word was embedded in a control phrase. Sentences preceding the control phrase provide textual context that determines the information status²⁹ of the target word³⁰.

²⁹ Though we adopted Chafe's proposal of activation status, it is comparable to some in Prince's (1981) givenness taxonomy. The information status of the test words in all experiments are brand-

The following paragraphs in (28) are examples of the three conditions investigated in the experiment. Target words are in **bold** and antecedents (for given and accessible statuses) are in *italics*.

(28a) New: **Um cavaleiro** garantiu sua vitória na corrida de hoje mesmo estando há um ano longe dos treinos. Um trágico acidente forçou o afastamento do cavaleiro, que chegou a anunciar sua aposentadoria prematura.

A **rider** guaranteed his victory in the race today despite being a year away from the training. A tragic accident forced the rider's leave, who even announced his early retirement.

(28b) Given: A cavalgada pelo sertão do estado terminou ontem com a chegada de um *cavaleiro* mato-grossense. O **cavaleiro** garantiu sua vitória fazendo o percurso em menos tempo.

The cavalcade through the state hinterland finished yesterday with the arrival of *a* mato-grossense *rider*. **The rider** guaranteed his victory by running the route in less time.

(28c) Accessible: Os brasileiros festejaram o fim do campeonato no *hipódromo* de Jacarepaguá. O **cavaleiro** garantiu sua vitória ao completar a prova com a pontuação máxima, e foi premiado com a medalha de ouro.

new for new referents, textually given for given and inferentially accessible (whole-to-part) for accessible ones.

³⁰ It is a current trend to advocate for more experiments that produce data from real-world situations e.g. communicative- or goal-centered tasks and semi-(spontaneous) speech. However, controlled lab speech has still its importance, especially when the phenomenon to be investigated is affected by many levels. In our case, it seems that for a detailed phonetic description of information status, we had - at least for now - to control the maximum number of factors as possible. Xu & Xu (2005) argued that accentuation itself is a concept that confuses between focus and lexical stress, so that accented units are usually stressed syllables that are focused. Thus acoustic differences solely due to given-new contrast cannot be known unless focus is explicitly controlled. Baumann (2006) also remarks about the possible influence of focus on the activation level of givenness, regardless the cognitive status of the referent. One example is for contrastive focus, in which even given referents can be prominent. Besides, the distribution of pitch accents (thus prominence) is also influenced by rhythmic/metric structure of the sentence e.g. long NPs force the creation of two intonational phrases, each with a phrase accent (Baumann, 2006: 97-9). For a detailed defense of lab speech, see Xu (2010).

The Brazilians celebrated the end of the championship at the Jacarepaguá *racecourse*. **The rider** guaranteed his victory after finishing the track with perfect score, and was awarded with the gold medal.

All target words have four syllables, and the lexical stress is on the second to last syllable. Arantes (2010) observed that prosodic effects related to referential status tend to be more evident in longer words, as the suprasegmental features spread over more syllables. The stress on the the second to last syllable is the most frequent word stress pattern in BP. Such configuration of the target words allows a detailed observation of intonation marking, especially on prestressed syllables.

The distance between the referent and its antecedent (for given and accessible items) was controlled with a mean of 5-6 syllables. Brown (1983) noticed in her study that when a given entity is re-introduced into discourse after some digressions, it is marked by a pitch accent. A possible explanation relies on the idea that the working memory has a cache/buffer with a fixed number of information units. As the listener receives more input, older information units need to be discarded and already activated items "fade", that is, they are no longer salient and need to be reactivated³¹.

1.2 Procedures

Four subjects (one male) from the city of Belo Horizonte were presented the stimuli, displayed one by one on a computer screen in a pseudorandomized fashion. The experimenter instructed the subjects to read each narrative silently before reading it aloud to ensure they would be aware of the content of the paragraphs and minimize hesitation. The subjects were recorded in separate sessions in a sound treated room at the Phonetics Laboratory (Labfon) in UFMG.

³¹ Chafe's cognitive account is totally compatible with this *recency* feature as well as other studies (Ariel 1990; Arnold 1998; Birner & Ward, 1998; Gundel et al 1993; Almor, 1999).

After each recording session, sound files were edited and labeled. For each sound file, the target NP (determiner plus noun) was manually segmented into syllables and boundaries and was stored in Praat metadata files. Two interval tiers were created in Praat, one separating the syllables and the determiner, and the second separating only the determiner from the noun. All acoustical analyses were performed with the help of Praat scripts (Arantes, 2010). The figure below shows the segmentation of the target words into intervals.

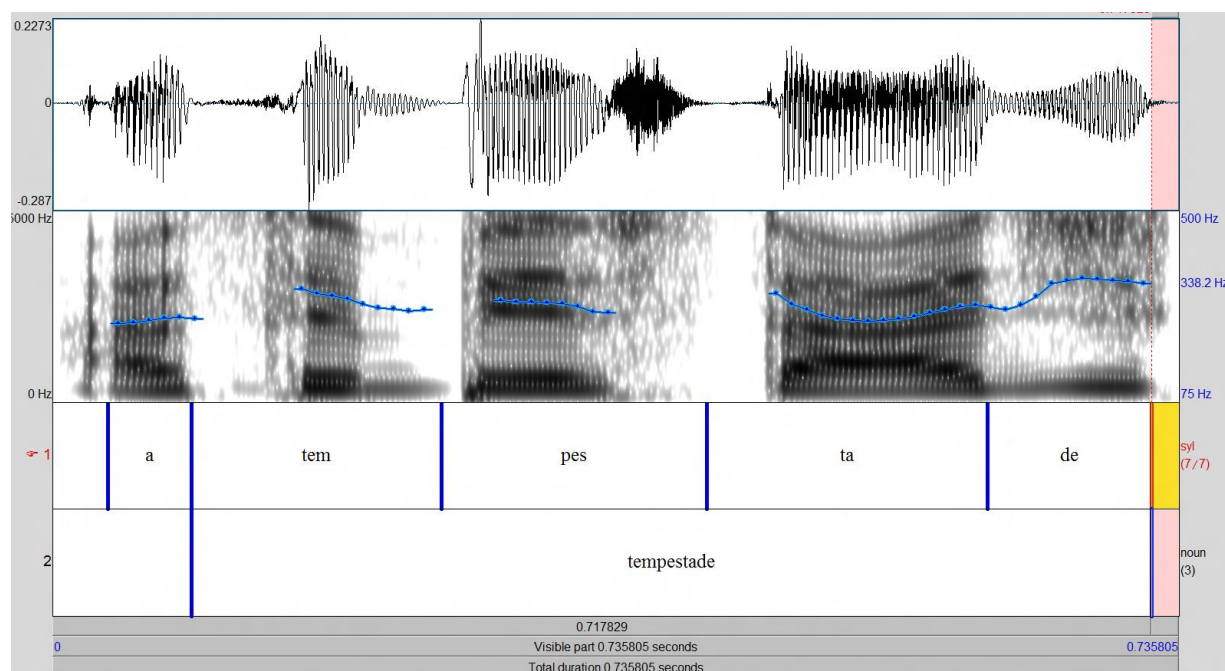


Figure 7 - Praat editing screen with oscilogram, spectrogram, syllable tier and duration tier (sound file: “tempestade” with accessible status)

1.3 Acoustic correlates

Because definite and indefinite determiners have different number of syllables in BP (*o(s)*, *a(s)*, *um(ns)*, *uma(s)*), we decided to measure the duration of the noun instead of the whole NP. The duration values were extracted from the noun interval on the second tier.

The fundamental frequency was analyzed in different ways. It was divided into one central tendency measure (a) f_0 mean, and two dispersion measures (b) standard deviation (SD) and (c) range. For (a), we extracted the average value of f_0 of all target NPs (determiner plus noun) and calculated the mean of these average values. This measure can be interpreted as the average pitch level of the f_0 contours. For (b), we extracted the values of f_0 standard deviation for each target NP and calculated their means. This measure can be interpreted as an estimator of how variable the f_0 contour is. Finally, for (c), we calculated the difference between the maximum and minimum pitch values in each NP. Then, we calculated the mean range of all target words. This final measure can be interpreted as the extension of the pitch excursion along the target word. For the range calculation, we converted the final result from Hertz into semitones. In the formula below, maxHz stands for the maximum value of f_0 and minHz stands for the minimum value of f_0 :

$$12 \log 2 \left(\frac{\text{max Hz}}{\text{min Hz}} \right) \quad (2)$$

The time-normalized f_0 contours of the target words were also analyzed. We applied a Praat script from Arantes (2010, which in turn which was adapted from Xu, 1993). The f_0 curves were smoothed so any bumps or sharp edges in the f_0 contour greater than 4 Hz were eliminated and voiceless portions were interpolated. Finally, the script extracted five samples of f_0 values for each syllable interval. These procedures allows the comparison of sound segments with different duration values, as the pitch contours get aligned with an equal number of samples for each interval.

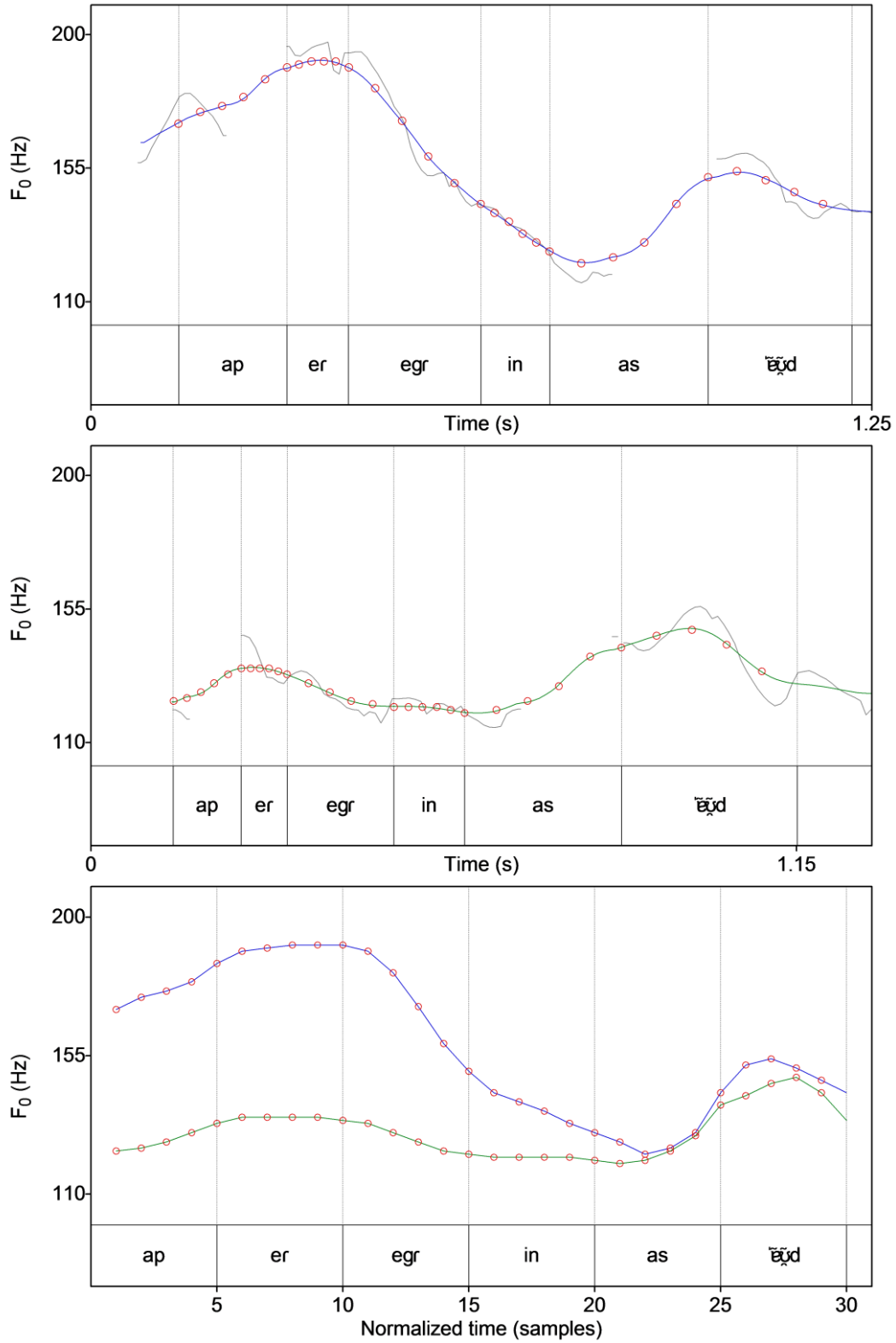


Figure 8. Example of application of the time-normalized f_0 contour script. The first two contours show the raw pitch contours (grey lines) and the smoothed, interpolated contours (blue and green lines). The red dots are f_0 value samples for each syllable interval. The last picture shows the final contours overlaid after script application (Arantes et al., 2012)

1.4 Statistical analyses

The data generated by the four subjects were analyzed separately. Referential status with three levels (given, new, accessible) was the independent variable for all analyses. Analysis of variance (ANOVA) was used to determine if differences in mean values of the acoustic parameters were statistically different among the levels of the independent variable. Bonferroni *post-hoc* tests were carried out later to evaluate the interaction of the levels between themselves.

1.5. Results³²

1.5.1 Duration

The Figure 9 shows that, in average, new referents are the longest and given the shortest ones. Accessible referents showed an intermediary duration value. As we expected, new referents are longer because they are the first mention of the entity in the discourse. Therefore, they tend to be more prominent. Given referents, especially in the case where the referent is repeated, are expected to be shortened.

The statistical analysis showed that there is a highly significant difference between new and given conditions. However, there seems to have no difference neither between new and accessible nor given and accessible conditions.

From the production point of view, it seems that duration is a relevant acoustic parameter to prosodically encode referential status.

³² Samples of acoustic data and correspondent Praat metafiles (pitch and textgrid files) of the three experiments are available at <http://goo.gl/Yz2luC>.

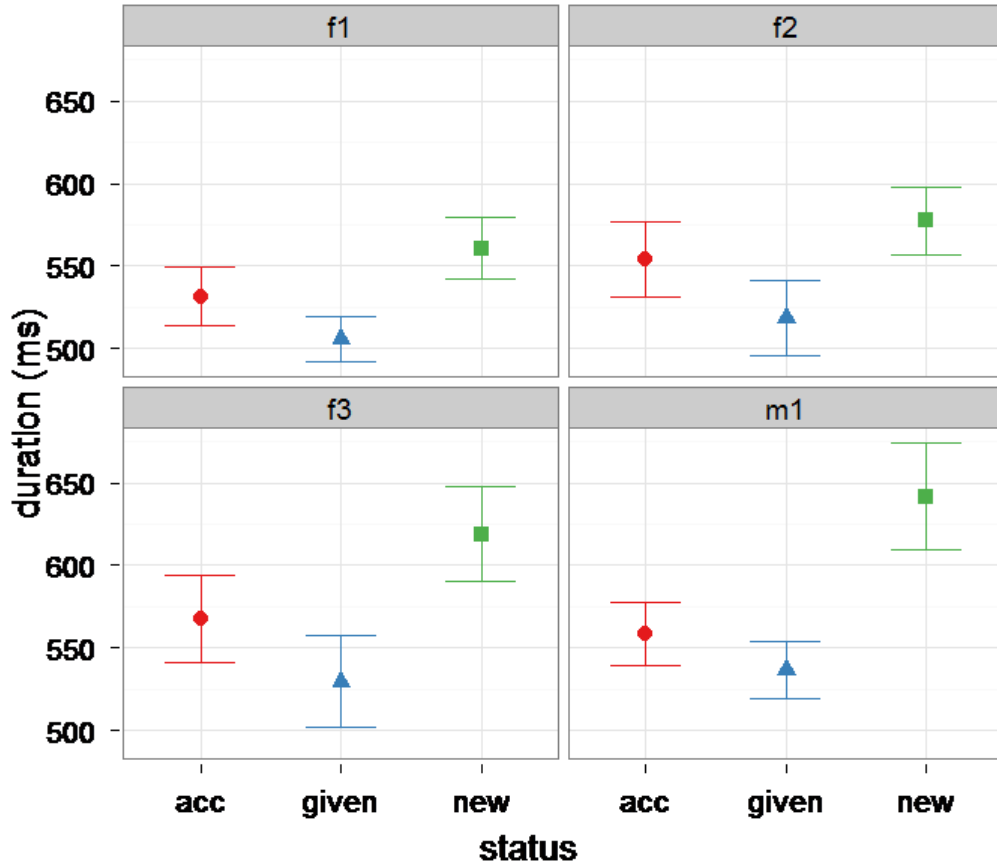


Figure 9 - Mean duration in Experiment 1. Whiskers indicate 95% confidence intervals around the mean.

Table 7 - Summary of statistics (F-values) for mean duration in Experiment 1

subjects	F(2, 84)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	9,96	< 0,001	< 0,001	n.s.	n.s.
f2	6,99	< 0,001	< 0,001	n.s.	n.s.
f3	20,52	< 0,001	< 0,001	n.s.	n.s.
m1	10,4	< 0,001	< 0,001	n.s.	< 0.1

1.5.2 Average f_0

In three of the subjects, the new condition showed higher average values than the given and accessible conditions. Such results indicate that the speakers produce referents with an average higher pitch level when they are presented as new information. Given and accessible referents are produced with a significantly lower pitch level, what is expected from previous studies. Only one subject (f3) showed a different behaviour, producing accessible referents with higher pitch levels. Such differences, when average values are compared, are irrelevant.

The statistical analysis showed that in three of the subjects, the average f_0 mean is significantly different for new compared to given and accessible conditions. There is no difference between given and accessible conditions.

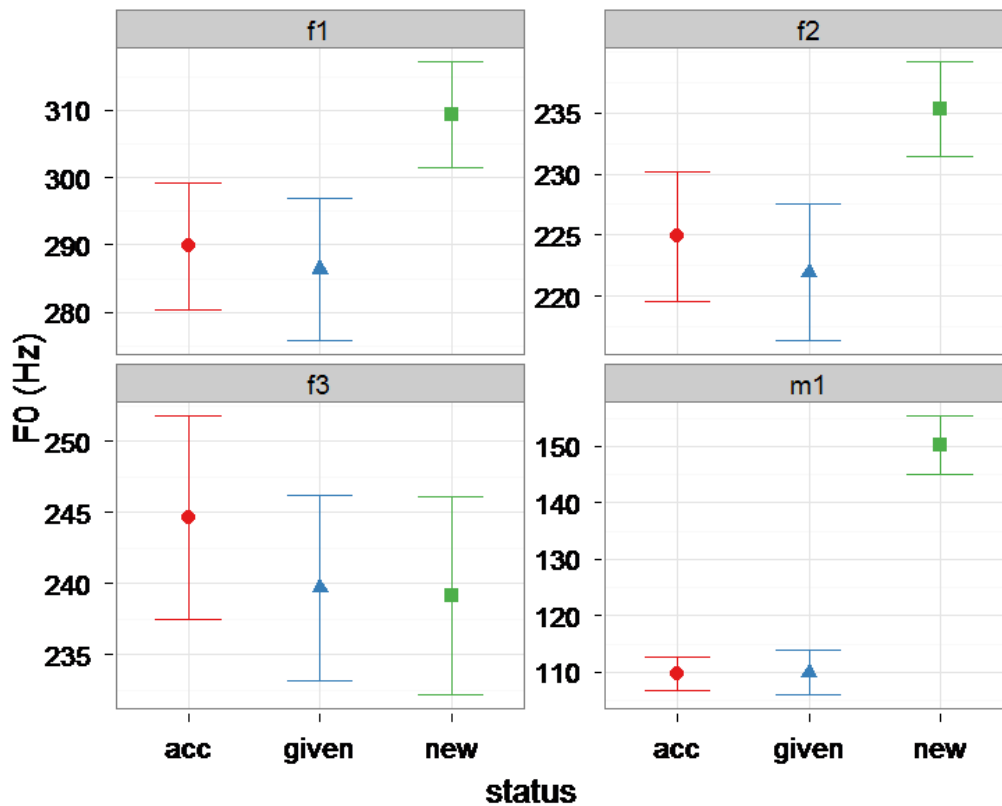


Figure 10 - Mean average f_0 in Experiment 1

Table 8 - Summary of statistics (F-values) for mean average f_0 in Experiment 1

subjects	F(2, 84)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	6,64	< 0,01	< 0,001	n.s.	n.s.
f2	7,61	< 0,01	< 0,01	< 0,05	n.s.
f3	0,73	n.s.	-	-	-
m1	123,73	< 0,001	< 0,001	< 0,001	n.s.

1.5.3 SD

There is a tendency for new referents to show higher mean standard deviation values along NPs than given and accessible referents, that is, new referents tend to be produced with more peaks and valleys in its contour, and given and accessible tend to be produced with a smoother contour. However, in only one subject (m1) such differences are significant (see table with the ANOVA results).

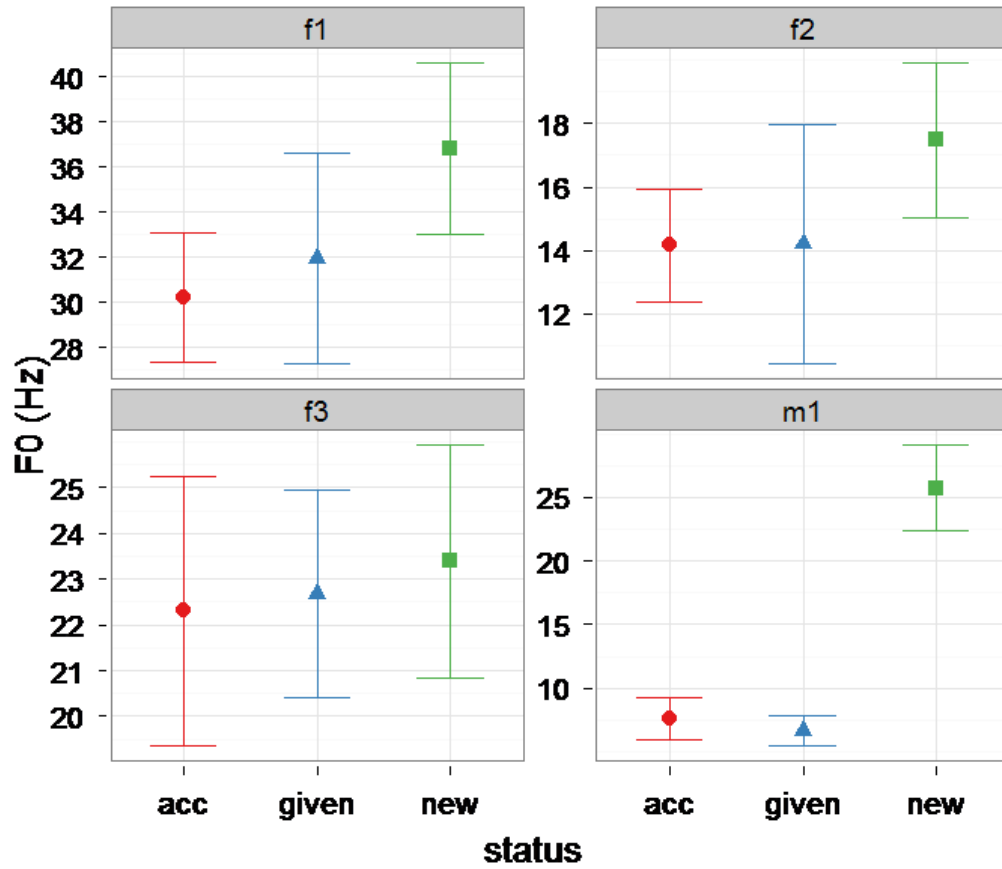


Figure 11 - Mean f_0 SD in Experiment 1

Table 9 - Summary of statistics (F-values) for mean f_0 SD in Experiment 1

subjects	F(2, 84)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	3,00	n.s.	-	-	-
f2	1,79	n.s.	-	-	-
f3	0,17	n.s.	-	-	-
m1	85,99	< 0,001	0,001	0,001	n.s.

1.5.4 Range

Results of the range analysis showed that there is a slight tendency for new referents to have a broader pitch excursion than given and referents. In one subject, new referents showed even lower range values than given and accessible conditions. Thus, speakers go from the lowest to the highest pitch points along the NP regardless its referential status. In only one subject (m1), the new condition is significantly different.

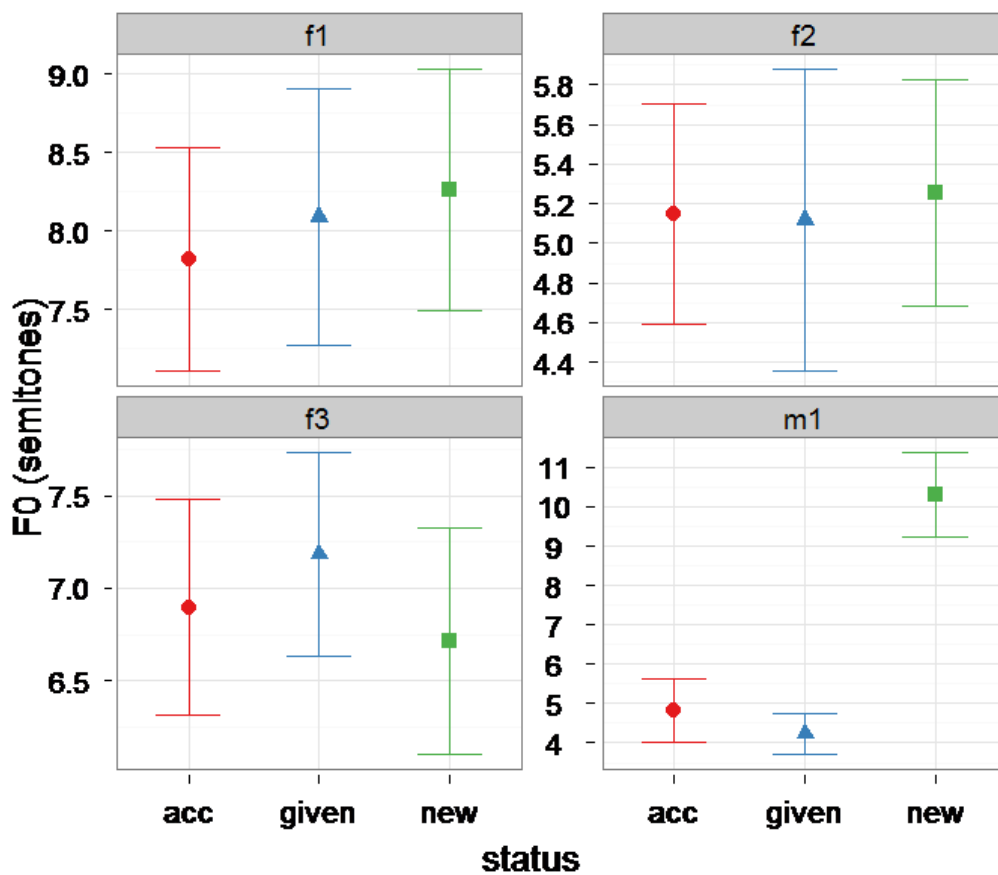


Figure 12 - Mean f_0 range in Experiment 1

Table 10 - Summary of statistics (F-values) for mean f_0 range in Experiment 1

subjects	F(2, 84)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	0,32	n.s.	-	-	-
f2	0,03	n.s.	-	-	-
f3	0.57	n.s.	-	-	-
m1	0.57	< 0,001	< 0,001	< 0,001	n.s.

1.5.5 Time-normalized f_0 contours

For subjects f1 and f2, contours of the three status are similar except for the fact that new referents have an ascending f_0 movement that spans over the determiner and the prestressed syllables. Given and accessible statuses showed a much smaller excursion aligned to prestressed syllables. Given and accessible referents have a broader range, though new referents presents a more variable contour along the NP. In subject f3, it seems that referential status does not have influence on the shape of f_0 contours. In average, referents in the three conditions showed a similar pitch excursion and prominences aligned to the stressed syllable. The pitch variation along the NP is also very little. Subject m1 presents the most extreme values for new referents, as it holds the highest pitch values, range and pitch excursion. Given and accessible referents presented a pretty similar flat contours.

On the whole, the results revealed that f_0 contours of new referents are different from the given and accessible ones. Despite the individual variability, new referents are characterized by the presence of two major pitch peaks, one extending over the chain of prestressed syllables and other aligned to the stressed syllable. Given and accessible referent contours are very similar to each other.

Both tend to present just one peak, either aligned to the pre-stressed syllables or to the stressed syllable.

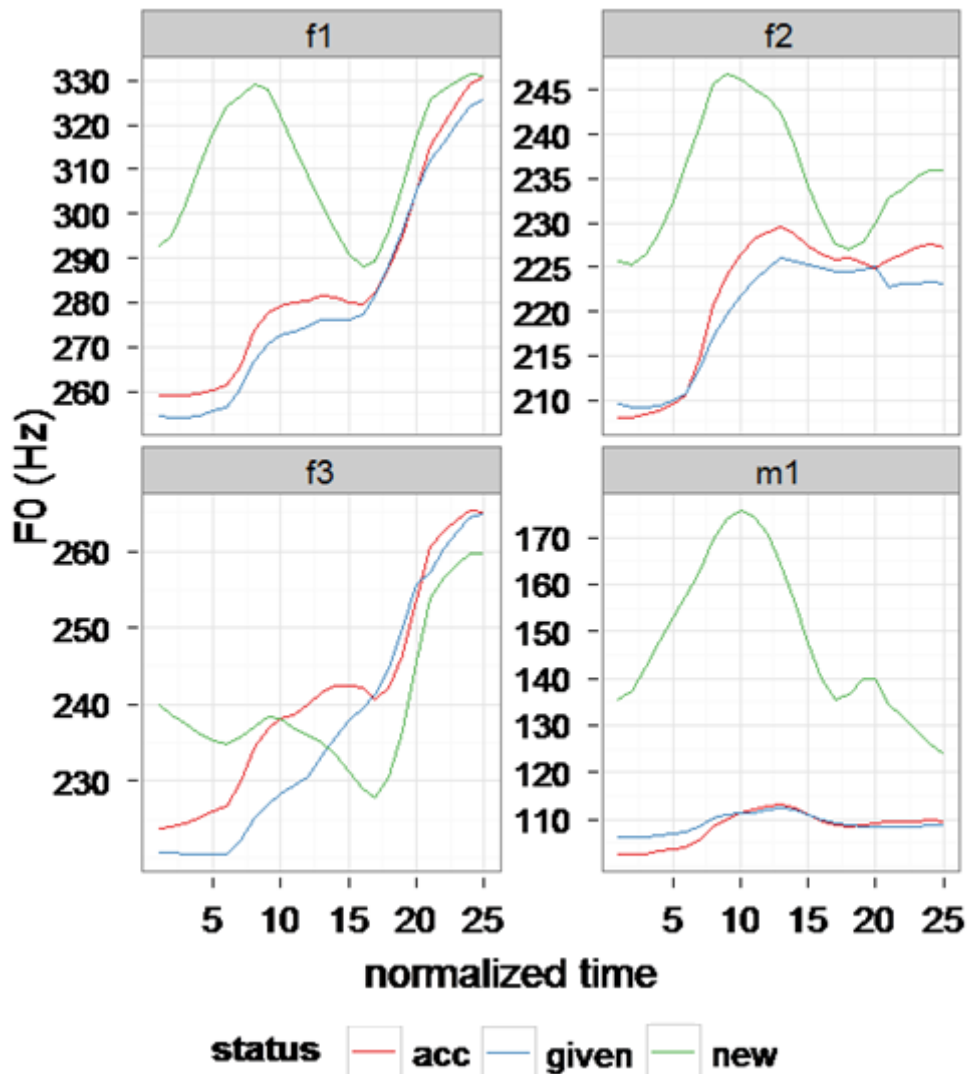


Figure 13 - Mean time-normalized f_0 contours in Experiment 1. Line type indicates different referential status. f_0 samples are displayed in the horizontal axis - each five samples correspond to a syllable in the test NP. The article extends over the first five samples; the prestressed syllables extend from sample five to sample 15 and the stressed syllable from sample 15 to 20

1.5.6 Discussion

The results in Experiment 1 showed that, in general, there is positive evidence that acoustic parameters are affected by the referential status contrast. Duration seems to be the most robust correlate of the referential status distinction, for it is the only parameter that was affected by the status variable in all subjects. In addition, average pitch value is also affected by the referential contrast, with new referents being assigned higher average values. Except for subject m1, standard deviation and range do not seem to be sensitive to the referential status contrast.

The time-normalized f_0 analysis shows that new referents not only present higher pitch level, but also peaks aligned to prestressed syllables. Accessible and given referents are relatively flat in this region. In BP, the chain of prestressed syllables (including the NP determiner) seems to be the main locus of pitch differences among the status levels. Apparently, there is a general set of rules related to the saliency of referents and prosody – as in other languages (German and English), new referents tend to be acoustically prominent and given referents tend to be flattened.

Another point is to understand whether the prosodic encoding of referential status works as an independent account. The lack of a clear distinct acoustic “signature” for the accessible status can be evidence that the prosody works in a general way, associated to other types of information (e.g. syntactic position, semantic relationship, register, focus). Baumann (2006) observed that accessible status does not have a clear prosodic marking, and semantic information i.e. semantic relation between the referent and its antecedent might have a great deal. The clearest pitch accent marking accessible referents were when there was a whole-to-part relation e.g. *car* ~ *engine*. As we did not control the semantic relation on accessible referents, it can explain part of our results.

Another hypothesis is that such acoustic marking of referents is crosslinguistically constrained or encoded, that is, different languages might use diverse information to signal different degrees of givenness.

Besides semantic information, it can be suspected that sentence position is also relevant. In this experiment, the target words under the new condition were in absolute paragraph-initial position, and target words under given and accessible conditions were in sentence-initial, paragraph-mid position. Wang & Xu (2011) observed some paragraph-initial position effects on intonation in Mandarin. There is a chance that the difference found between new referents and given and accessible ones might be biased by sentence position. Also, as usually the locus of information focus is in the end of the sentence - and all target words were in the beginning -, the prosodic marking might have been affected by these factors.

Variation between subjects also seems to play a role. For three of the recorded participants, the results of most parameters were pretty similar. However, the last subject (the male one) presented a great range of values for almost all parameters. As all subjects were about the same age, with similar economic and educational backgrounds, other factors (personal reading style, focus, gender) can be considered as possible explanations for such differences. It is remarkable that even with the variation among subjects, the acoustic differences among the referential status were consistent, that is, all subjects used the parameters in slightly different ways to create the same distinctions among the status levels.

What is clear, though, is that speakers need to signal to their addressee that new referents are being included into the discourse, and as given referents are already playing their parts, they need little or no saliency.

In Experiment 2, we tried to explore some of these issues.

2 Experiment 2

2.1 Materials

For this experiment, we created a list with 72 narratives, dividing the target words in a similar way to Experiment 1. We decided to create shorter paragraphs, in a chance to avoid the decrease of subjects' level of attention during the task. The sentences below are examples of paragraphs used in this experiment.

- (29a) New: O homem foi à loja para comprar uma **bateria**.
The man went to the store to buy a **battery**.
- (29b) Given: O rapaz comprou uma *bateria* nova para o seu notebook.
Quando ele sentiu um cheiro de queimado, ele foi à loja devolver a **bateria**.
The lad bought a new *battery* for his notebook. When he smelled something burning, he went to the store to return **the battery**.
- (29c) Accessible: Ao sair de casa, o rapaz não conseguiu ligar o *carro*.
O frio havia congelado a **bateria**.
When leaving home, the lad couldn't start *the car*. The cold had frozen **the battery**.

As we saw in the item discussion of Experiment 1, the prosodic marking of accessible referents seems to be very sensitive to its semantic relation to the antecedent (Baumann, 2006). Because of this, we decided to control the semantic relation between the target word and its antecedent. Baumann observed that different semantic relations for accessible referents can be marked with different pitch accents, being the relation whole-to-part the most stable. Thus, all target words in this experiment have a whole-to-part relation when in the accessible condition. Philipp et al. (2008) observed that animacy can play a role on information order in sentences. Because of that, we also controlled the animacy of all target words.

Another difference concerns the syntactic position in the sentence: in Experiment 1, all target words were in subject position, and new referents were in absolute paragraph-initial position. Because the final position is usually point of focus of attention (it usually receives focus and is the point of nuclear accent), we decided also to control all target words in object position, avoiding paragraph-initial bias. As we did in Experiment 1, all target words have four syllables with the stress on the second do last syllable.

2.2 Procedures, acoustic correlates, statistical analyses

The procedures during the experiment are the same of Experiment 1. Five subjects read a list of randomized experimental sentences along with more 24 distractors on a computer screen. The acoustic analysis took also the same acoustic parameters (duration, average f_0 , SD, range and time-normalized contour). Acoustic analyses were also carried out using the same steps and scripts on Praat. The statistical analyses (three-way ANOVAs) were also carried out the same way (including Bonferroni post-hoc tests).

2.3 Results

2.3.1 Duration

The results showed that duration is not a relevant parameter. The statistical analyses showed no difference between the conditions.

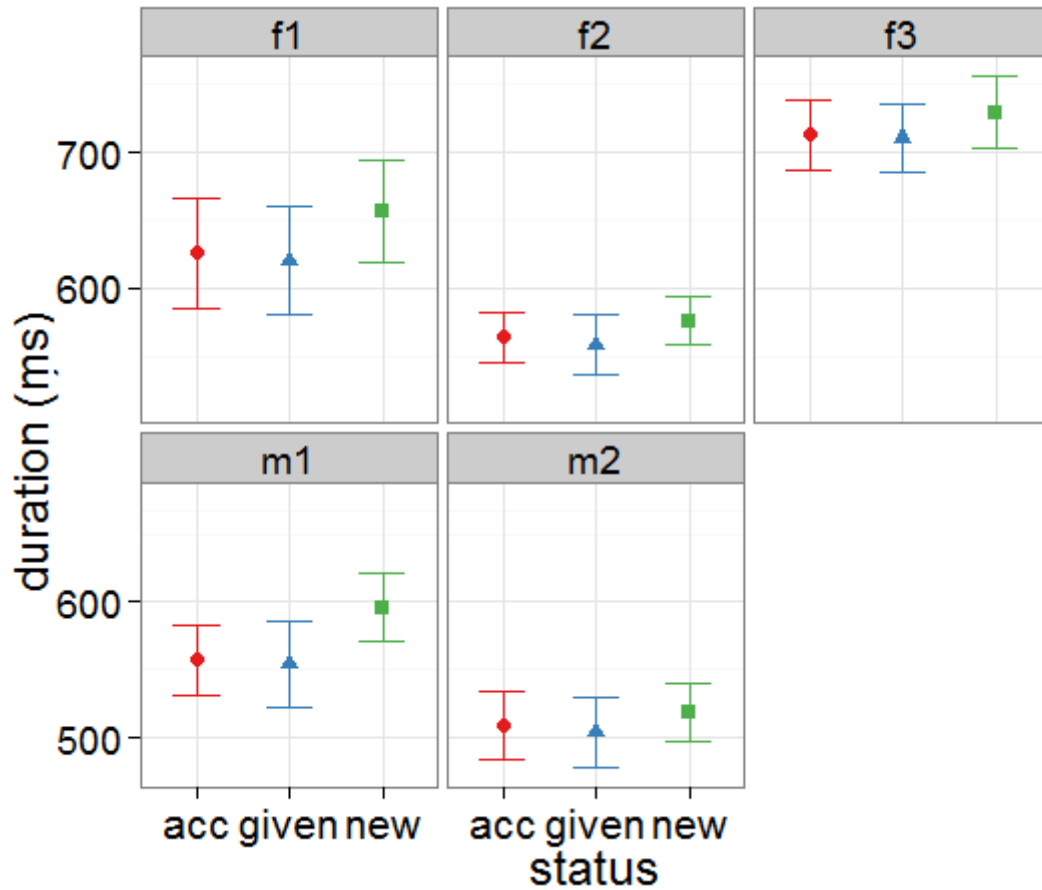


Figure 14 - Mean duration in Experiment 2

Table 11 - Summary of statistics (F-values) for mean duration in Experiment 2

subjects	F (2, 69)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	0.5018	0.6076	-	-	-
f2	0.5005	0.6084	-	-	-
f3	0.606	0.5484	-	-	-
m1	0.3627	0.6971	-	-	-
m2	0.3035	0.7392	-	-	-

2.3.2 Average f_0

The analysis of mean average f_0 also showed little difference between conditions. Only in one subject (f1), the ANOVA showed a marginal difference. Apparently, the pitch level of all referents are produced around the same average level, regardless its referential status.

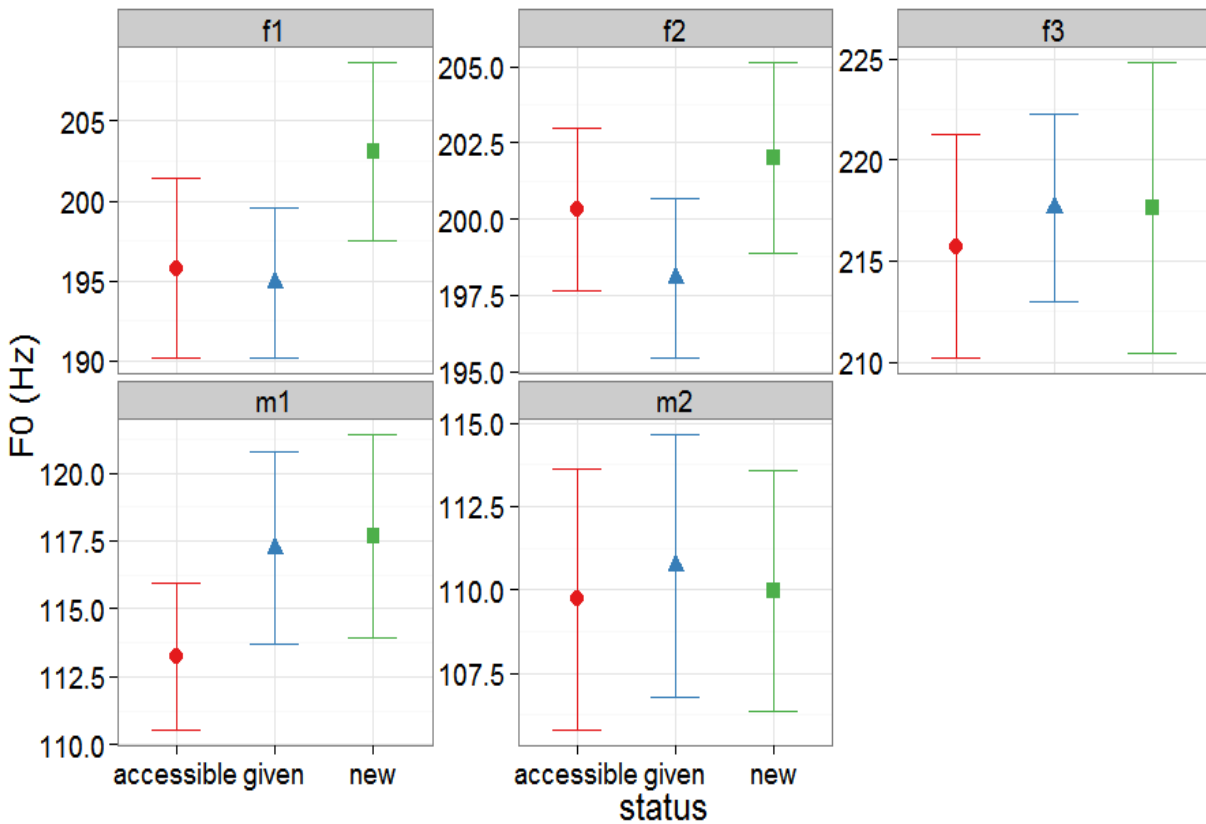


Figure 15 - Mean average f_0 in Experiment 2

Table 12 - Summary of statistics (F-values) for mean average f_0 in Experiment 2

subjects	F (2, 69)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	2.7515	0.07085	0.11	0.18	n.s.
f2	1.8926	0.1584	-	-	-
f3	0.1368	0.8724	-	-	-
m1	0.0743	0.9285	-	-	-
m2	2.0514	0.1363	-	-	-

2.3.3 SD

Likewise the results of average f_0 , the results of SD showed no significant difference between the conditions. In only one subject (f3), the statistical analysis showed to be significantly different. The post-hoc test showed that the SD contrast is between new and given conditions.

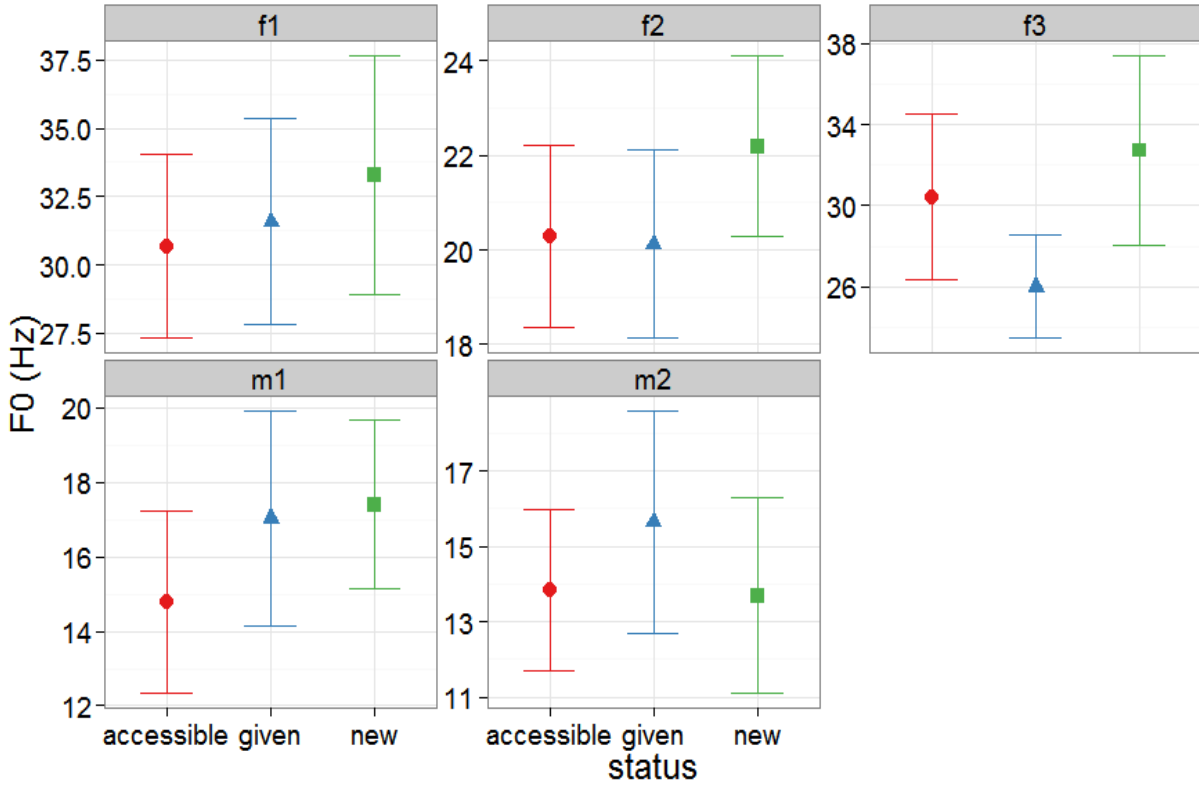


Figure 16 - Mean f_0 SD in Experiment 2.

Table 13 - Summary of statistics (F-values) for mean f_0 SD in Experiment 2

subjects	F (2, 69)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	0.4571	0.635	-	-	-
f2	1.3449	0.2673	-	-	-
f3	2.9997	0.05634	0.056	n.s.	n.s.
m1	1.1738	0.3153			
m2	0.6846	0.5077	-	-	-

2.3.4 Range

The results showed also that range is not strongly influenced by referential status. In only two subjects (f2 and m1), the statistical analyses showed marginal results. Such results are not sufficient to make any conclusion of a positive effect on range though.

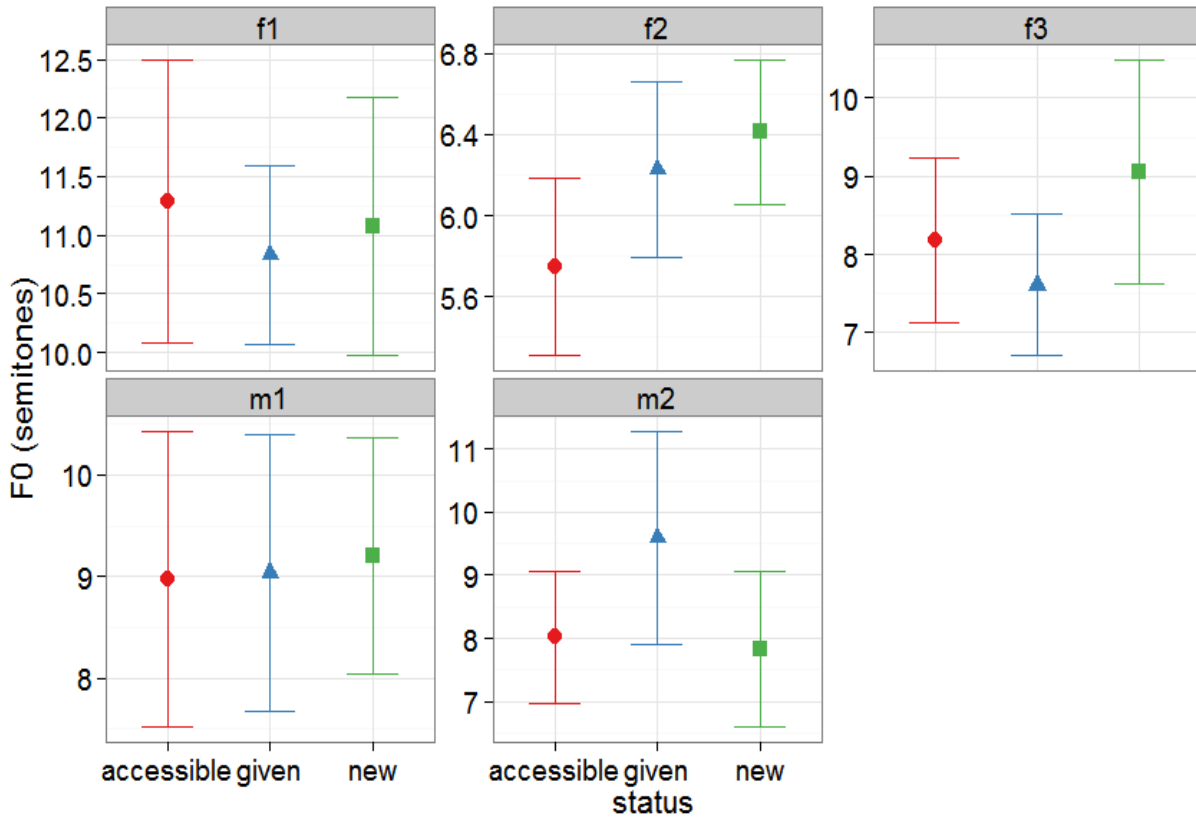


Figure 17 - Mean f_0 range in Experiment 2

Table 14 - Summary of statistics (F-values) for mean f_0 range in Experiment 2

subjects	F (2, 69)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	0.3318	0.7188	-	-	-
f2	2.7133	0.0734	n.s.	0.082	n.s.
f3	1.4215	0.2483	-	-	-
m1	2.6855	0.07531	0.13	n.s.	0.17
m2	0.0625	0.9395	-	-	-

2.3.5 Time-normalized f_0 contour

In most subjects the pitch contour are very similar in the three conditions. In subjects f1 and f2, the new condition presents a slightly higher pitch value aligned to the prestressed syllables, though this difference is not significant. In one subject (m1), the accessible condition presented a flat contour compared to the new and given conditions. Notice, however, that the contour aligned to prestressed syllables is the same in the three conditions. Subject m2 also presented very similar results for the three conditions; apart from the female subjects, he presented a lower prominence along the prestressed syllables and a raise of contour on the poststressed syllables.

One interesting point is to observe the variability of pitch behaviour among subjects. Female subjects tend to project a prominence aligned to prestressed syllables whereas male subjects tend to raise the pitch contour at poststressed syllables. Again, such differences of the prosodic marking of NPs apparently undergoes personal choices.

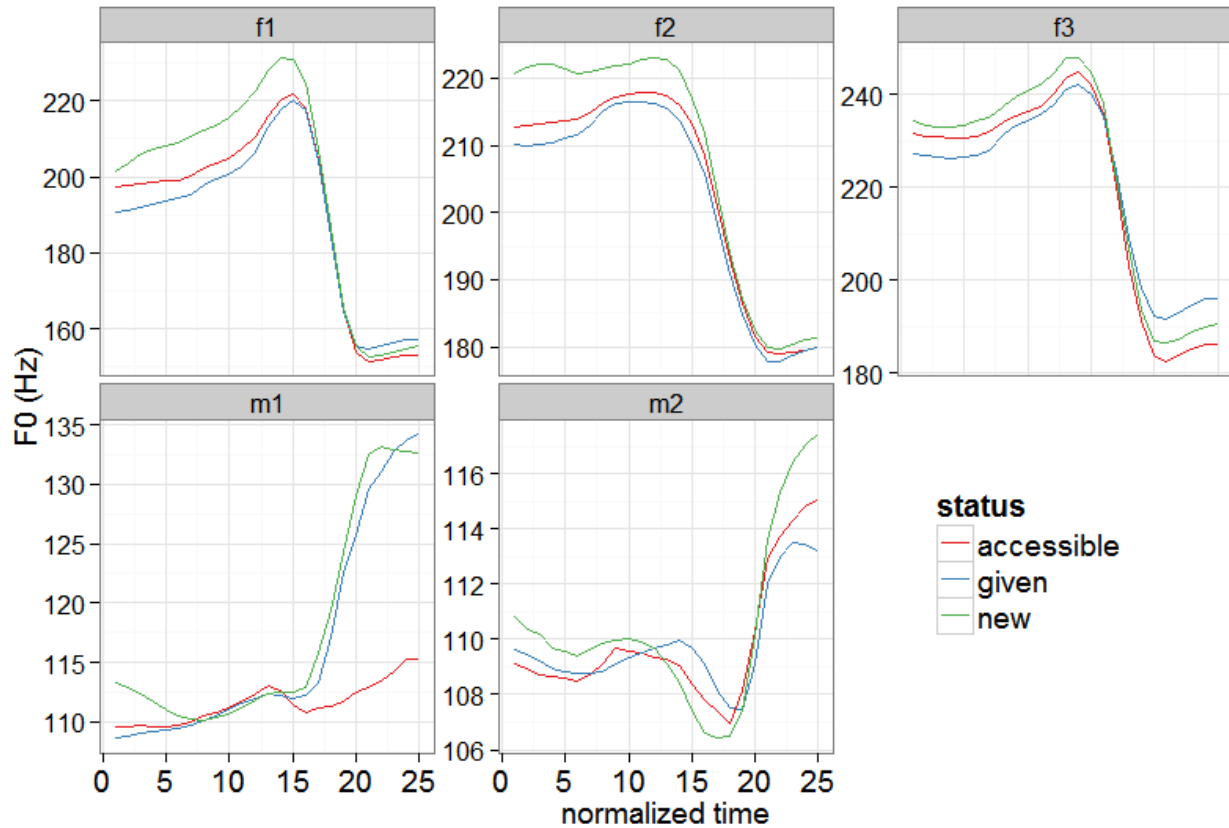


Figure 18 - Mean time-normalized f_0 contours in Experiment 2

2.3.6 Discussion

Compared to Experiment 1, the main differences of Experiment 2 were the position of the target word in the end of the sentence and a more strict semantic control on accessible referents.

However, even the strong differences found in Experiment 1 i.e. duration and pitch level on new referents compared to given and accessible were cancelled.

One might conclude that the position of referents in the sentence has a great account on the prosodic marking of referential status. The results on duration, which we expected to behave in an independent fashion from pitch contour, also showed no relevance in final position.

If the final position in the sentence is the place for information focus, one might expect that the speaker would make use of a stronger contrast between new and given information in this place. If the information is new, the speaker would produce a more prominent pitch contour; if the information is given, the speaker would then deaccent the referent by either dislocating pitch prominence to a previous word (e.g. verb) or shortening its duration.

Such results would make us refer back to one alternative explanation of the results in Experiment 1. It is possible that the prosodic encoding of referential status is based on general rules of information structure and implemented in different ways across languages. For instance, it is known that in German speakers have a preference to start sentences with topic (and indirectly given information), and explicit case marking also plays a role on sentence order (Hung and Schumacher, 2012). In BP, it seems that the contrast between referential statuses is lessened in sentence final position, maybe due to the expected descending movement of the f_0 in the end of utterances.

Our last experiment tries to explore some of these questions.

3 Experiment 3

3.1 Materials

For this experiment, a list with 72 mini-stories was created, similar to the list of Experiment 2. The main difference between this list and the list of Experiment 2 is that the place of the target words in subject position. In order to control any paragraph-initial position effect, a phrase was placed in the beginning of each paragraph. The NP syllable structure and semantic controls for the target words were kept.

- (30a) New: Em média, **uma ferradura** custa perto de cinquenta reais.
In average, **a horseshoe** costs nearly fifty reais.
- (30b) Given: Um apostador ganhou uma *ferradura* de presente. Ironicamente, **a ferradura** trouxe má sorte.
A gambler got *a horseshoe* as a gift.
Ironically, **the horseshoe** brought bad luck.
- (30c) Accessible: O dono do sítio viu o *cavalo* mancando.
A ferradura teve que ser trocada.
The farm owner saw *the horse* limping.
The horseshoe had to be changed.

3.2 Procedures, acoustic correlates, statistical analyses

The procedures during the experiment are the same of Experiment 1 and 2. Four subjects read a list of randomized experimental sentences along with more 24 distractors on a computer screen. The acoustic analysis took also the same acoustic parameters (duration, average f_0 , SD, range and time-normalized contour). Acoustic analyses were also carried out using the same steps and scripts on Praat. The statistical analyses (one-way ANOVAs) were also carried out the same way (including Bonferroni post-hoc tests).

3.3 Results

3.3.1 Duration

In three subjects, the referential status showed influence on duration. In one subject (m1), the differences were highly significant. Post-hoc tests showed that there is a difference between new and given conditions in two subjects (f1 and m1) and between new and accessible in one subject (m1).

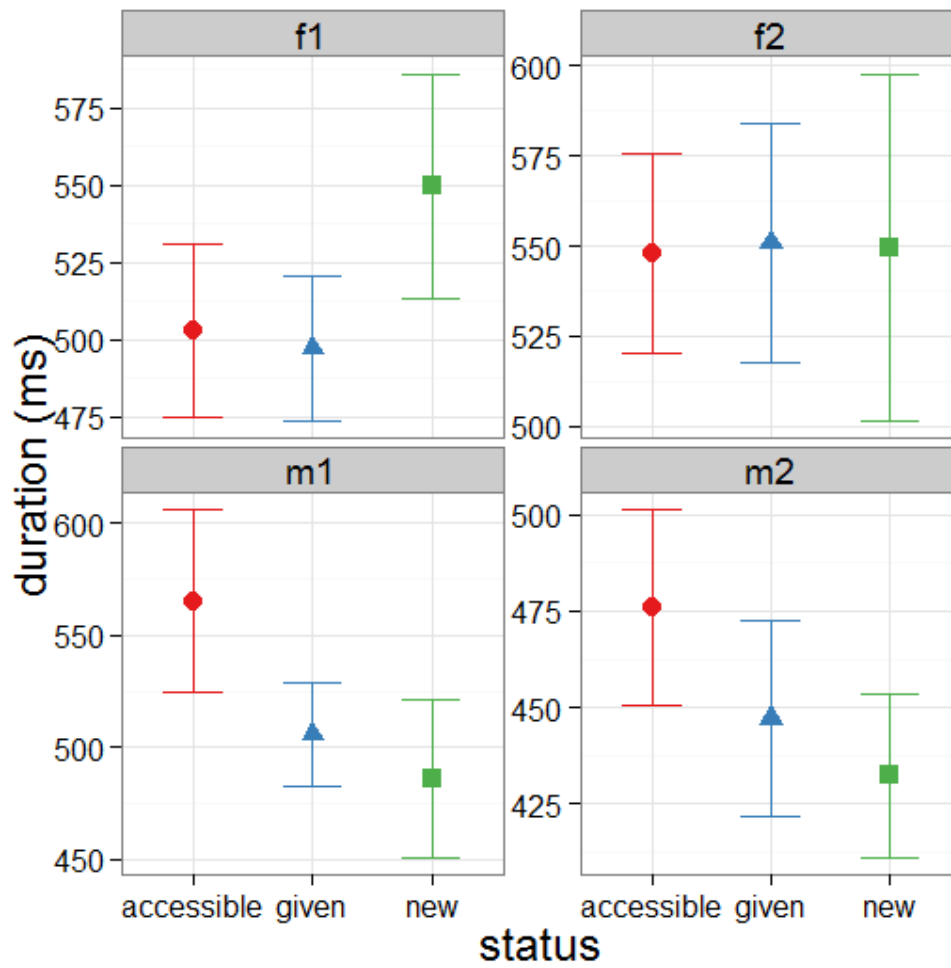


Figure 19 - Mean duration in Experiment 3

Table 15 - Summary of statistics (F-values) for mean duration in Experiment 3

subjects	F (2, 69)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	3.6303	0.03169	0.05	n.s.	n.s.
f2	0.4963	0.6109	n.s.	n.s.	n.s.
m1	5.691	< 0,001	< 0,01	0.05	n.s.
m2	3.2178	0.04612	n.s.	n.s.	n.s.

3.3.2 Average f_0

The results on average f_0 showed a highly significant difference in three subjects. Post-hoc tests showed also that there is a significant difference between new and given conditions, and between new and accessible conditions. According to the statistical data, we can conclude that subjects produced a relevant difference with pitch levels.

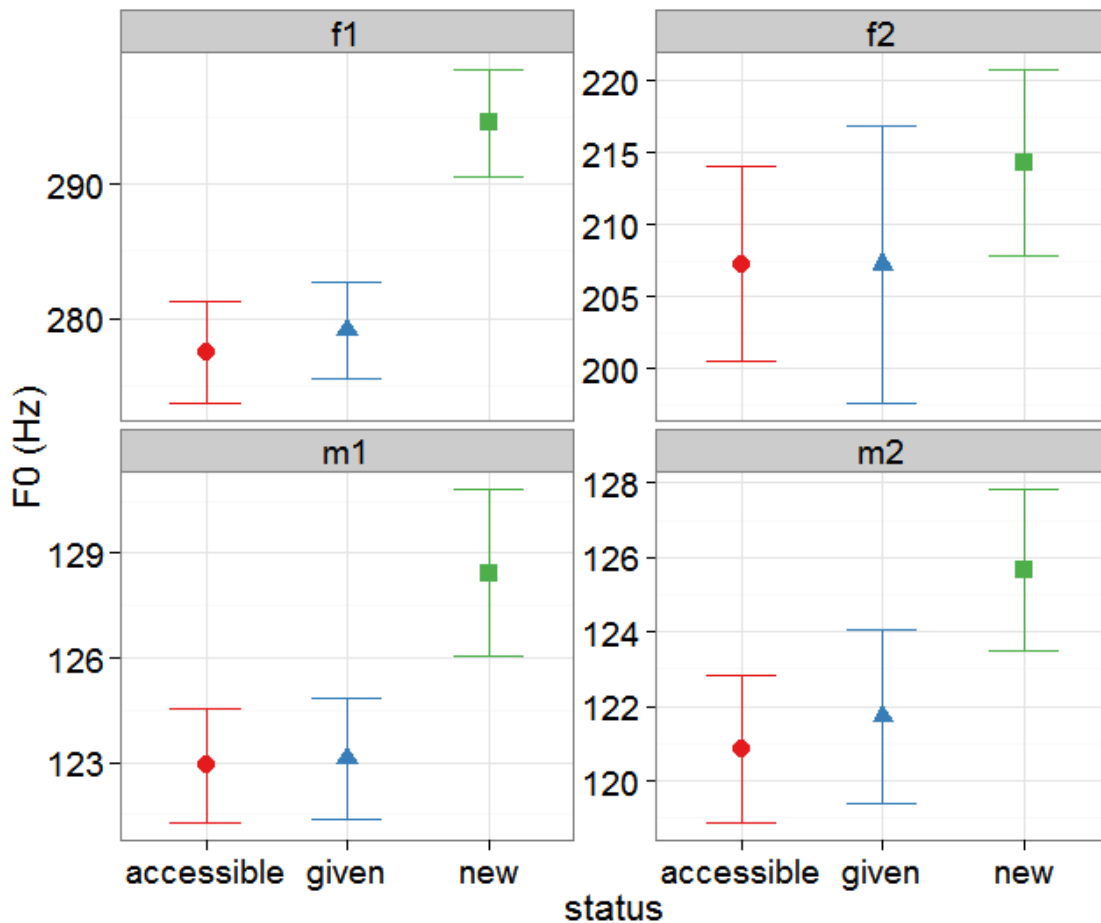


Figure 20 - Mean average f_0 in Experiment 3

Table 16 - Summary of statistics (F-values) for Mean average f_0 in Experiment 3

subjects	F (2, 69)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	23.608	< 0,001	< 0,001	< 0,001	n.s.
f2	1.057	0.353	n.s.	n.s.	n.s.
m1	9.8039	< 0,001	< 0,001	< 0,001	n.s.
m2	5.327	< 0,001	< 0,05	< 0,01	n.s.

3.3.3 SD

Likewise the results in Experiment 1, new referents tend to present a higher variability of f_0 values along the target words. However, the statistical analysis results showed that in only one subject (f1) there was a significant difference between conditions. The new condition was different from given and accessible.

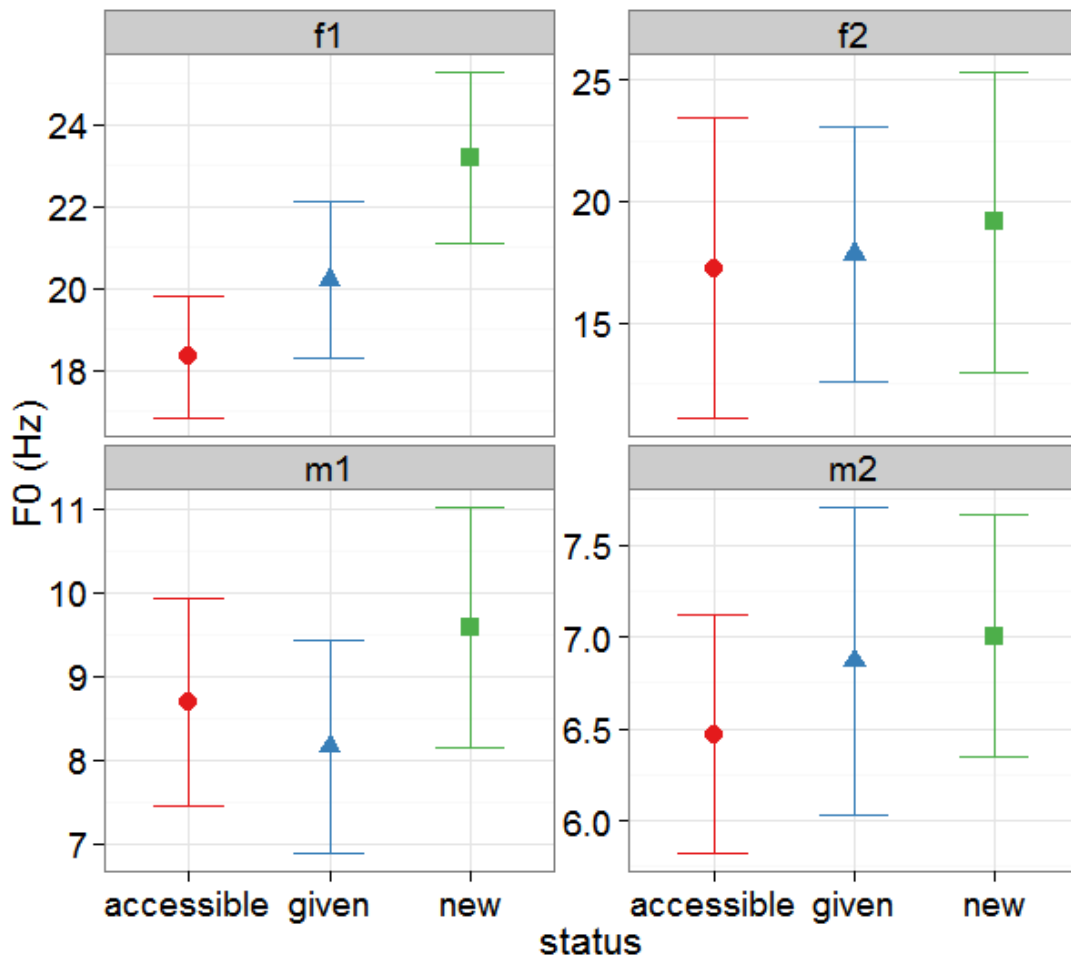


Figure 21 - Mean f_0 SD in Experiment 3

Table 17 - Summary of statistics (F-values) for mean f_0 SD in Experiment 3

subjects	F (2, 69)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	6.7643	< 0,001	< 0,001	< 0,001	n.s.
f2	0.051	0.9003	n.s.	n.s.	n.s.
m1	1.1349	0.3274	< 0,001	< 0,001	n.s.
m2	0.564	0.5715	< 0,05	< 0,01	n.s.

3.3.4 Range

In two subjects (f1 and m1), new referents had slightly higher pitch excursions. The results of range analysis showed no difference between conditions.

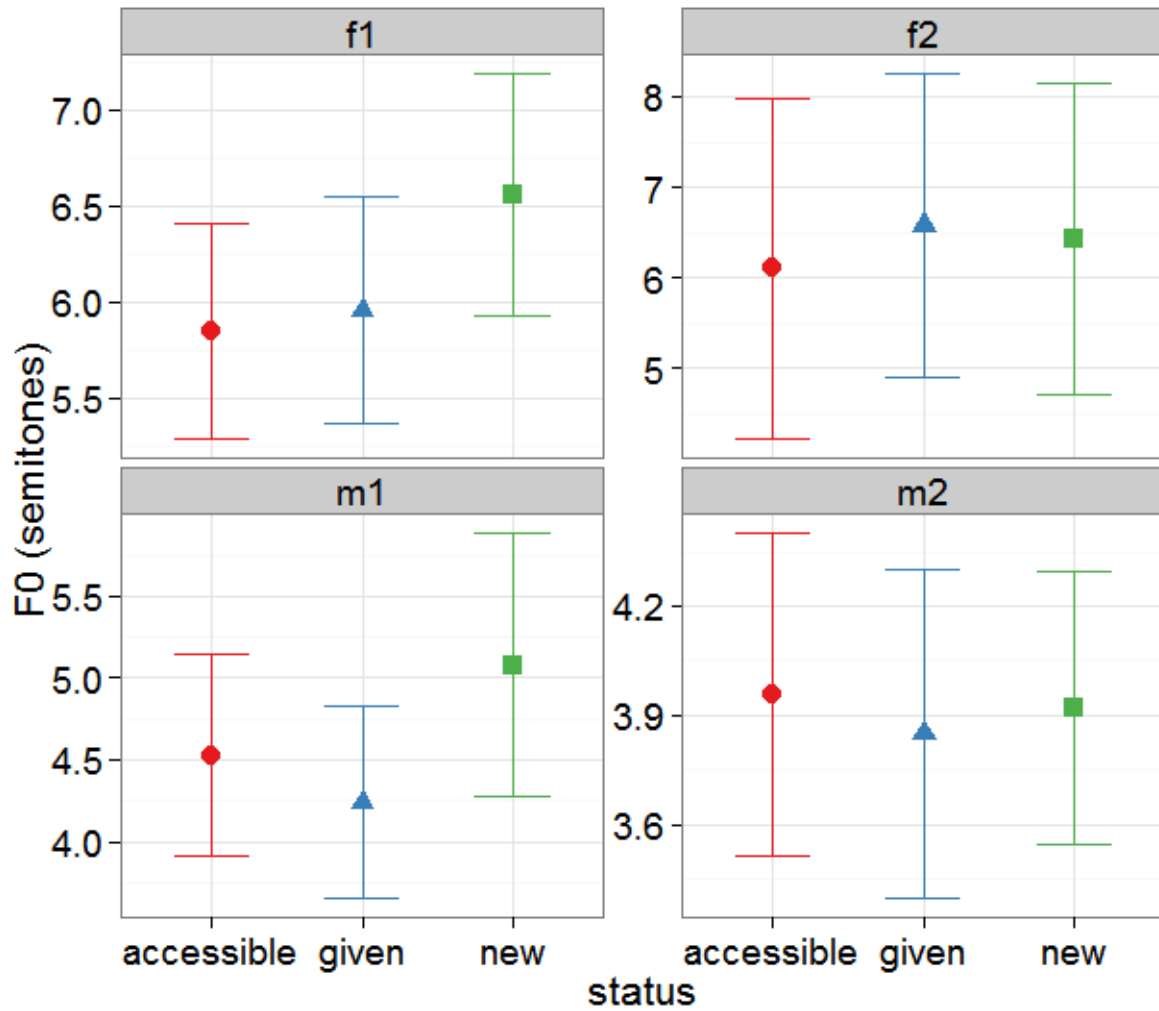


Figure 22 - Mean f_0 range in Experiment 3

Table 18 - Summary of statistics (F-values) for mean f_0 range in Experiment 3

subjects	F (2, 69)	p-value	post-hoc		
			new-giv	new-acc	acc-giv
f1	1.7019	0.1899	< 0,001	< 0,001	n.s.
f2	0.0213	0.9789	n.s.	n.s.	n.s.
m1	1.6818	0.1936	< 0,001	< 0,001	n.s.
m2	0.07	0.9325	< 0,05	< 0,01	n.s.

3.3.5 Time-normalized f_0 contour analysis

In general, we can observe a stronger contrast of the new condition compared to the given and the accessible ones. Given and accessible referents behave in a very similar way. In all subjects, new referents presented in average higher pitch levels. They also presented higher contour variability and excursion than given and accessible referents. In only one subject (m1), given and accessible referents presented higher pitch excursion. New referents tend to present a major prominence aligned to prestressed syllables, though it is possible to observe minor peaks aligned to prestressed syllables (subjects f2 and m2).

In all subjects and conditions, the contours aligned to the stressed syllables are rather similar. On the other hand, the contour associated to prestressed syllables may be a reliable cue to signal new referents.

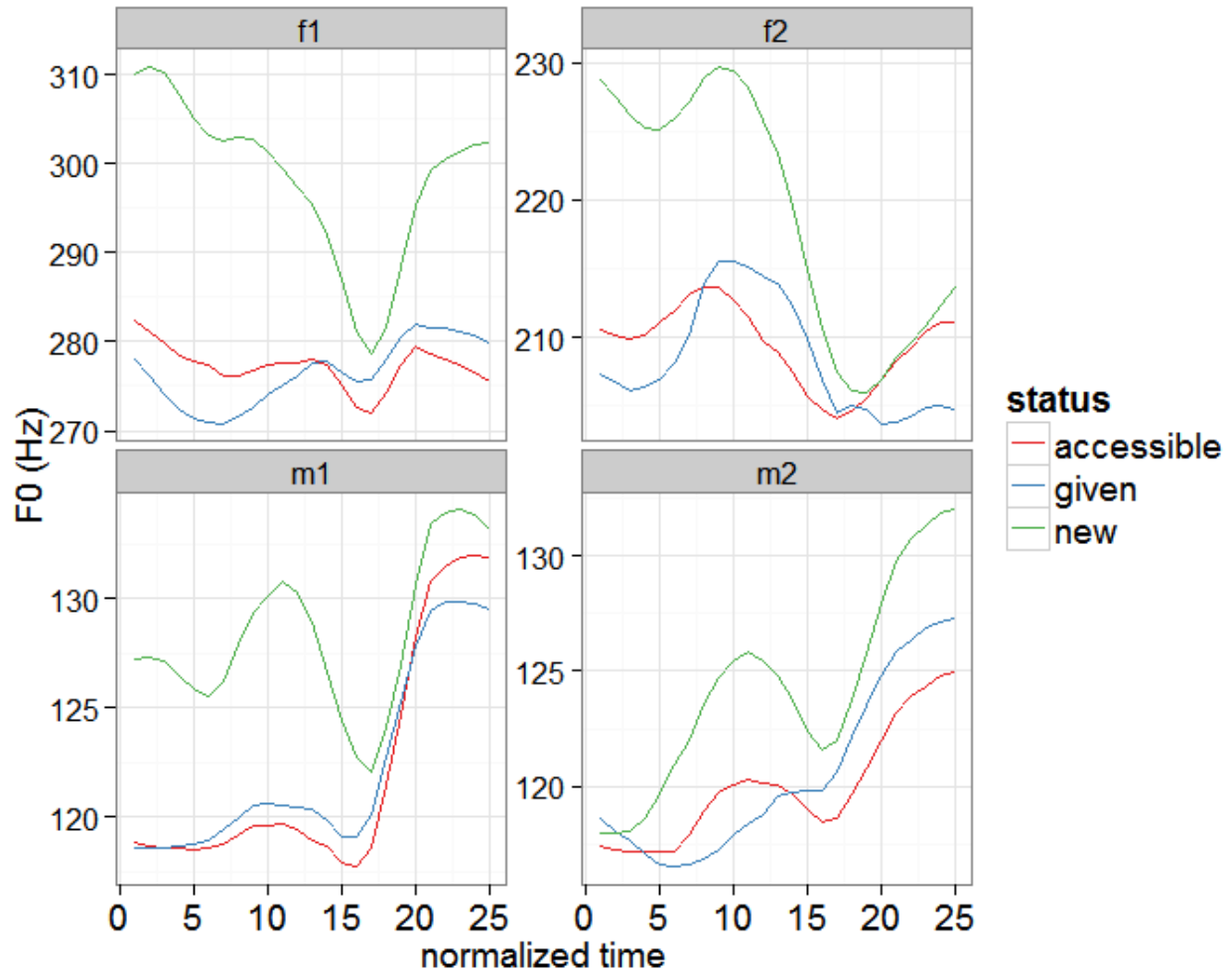


Figure 23 - Mean time-normalized f_0 contours in Experiment 3

3.3.6 Discussion

The results on Experiment 3 confirmed the results found in Experiment 1. Speakers use acoustic cues to contrast the introduction of new referents in the discourse. The contrast of new and given information is highly significant. Given and accessible, however, showed very similar results in all parameters. In Experiment 2 and 3, we controlled the semantic relation between the target word and its antecedent. This control does not show any effect on accessible status marking.

Accessible status seems to behave as an intermediate category, not only in terms of referential status but also in terms of prosodic marking. From the point of view of activation (Chafe, 1994), accessible referents behave as newsworthy entities. However, its prosodic marking makes it behave similarly to given referents.

Sentence position had also a major effect, though we did not control it as an experimental variable. Compared to the results found in Experiment 2, we can conclude that at least in BP speakers tend to contrast givenness in sentence initial position. An explanation for such effect is yet to be explored. So far, we may observe that all target words, regardless its referential status, seem to have their pitch contour aligned in sentence final position, and the 'givenness effect' is lost.

Duration and pitch level (average f_0) are the most relevant parameters to express referential status of referents in BP. Despite intersubjective variability, the differences among conditions remained the same.

4 General discussion

The results found in the three production studies corroborate our assumptions about the prosodic marking of information status in BP. In general, speakers use acoustic cues to signalize to his addressee the status of new referents in the discourse in contrast with already activated referents in the addressee's working memory. However, it is worth discussing a number of points as such phenomenon seems to be subject to some factors.

Apparently, the position of the referent in the utterance seems to play an important role on the way speakers use prosody to mark status information. In Experiments 1 and 3, differences in duration and average f_0 were found to be

very significant to mark information status in initial position whereas in Experiment 2, none of the acoustic correlates differed relating to information status in final position.

When we see the results of similar experiments from other languages such as German or English, it would be expected a lowering of the pitch configuration for given referents (i.e. deaccenting). Our results indicate, however, that BP behaves similarly to Italian or Spanish, in which intonational prominence always falls on clause-final position and deaccenting does not occur (Vallduví, 1991; Zubizarreta 1998).

In this case, the boundary tone might have an effect on the overall contour of the referent, forcing a lowering for new referent or even the nuclear pitch accent might overrule the expected pitch accent for the referent in final position. Syntactic approaches such as Zubizarreta (1998) take into account Nuclear Stress Rule (NSR) and Focus Stress Rule (FSR) to explain such behavior. In that proposal, focused elements (i.e. discourse salient referents) tend to occur to the rightmost position in the clause, attracted by the nuclear accent of the sentence.³³

This approach can justify the lack of difference for acoustic prominence between new and given referents in the data produced in Experiment 2. However, it does not explain other cases. Sometimes, referents occur in the middle of intonational phrases, that is, they are not placed at the rightmost position of the utterance. Examples of this might be sentences with post-object material (e.g. adverbial phrases) such as:

³³ Notice that the proposal for the present study is a discussion of a phonetic account of the phenomenon. However, it is sometimes necessary to cite phonological approaches for the sake argument structuring.

(31) A prefeitura vai emitir *um comunicado* no noticiário que mudará a política em uma cidade do interior.

The city hall will release *a statement* on the news that will change the politics in a town in a countryside.

A população aguarda ansiosa para ver **o comunicado** na televisão.

The population waits anxiously to see **the statement** on television.

(Antão et al., 2013)

In the second sentence in (31), the target word *comunicado* occurs in the middle of a prosodic constituent thus the nuclear accent falls on the last noun phrase *a televisão*. If indeed given information tends to be prosodically less salient, then the target word here would have shorter duration and less f_0 prominence than *um comunicado*. Empirical data showed, however, that referents with grammatical function of object have very similar prominence levels, regardless of referential status. Antão et al. (2013) carried out a production study using a similar experimental design to the one used in this thesis. The authors considered three factors: information status (given and new), referent size (4-, 5-, and 6-syllable words), and sentence position (subject and object)³⁴. All target words in object position were non-final. The results show that only in subject position new and given referents present significant differences in terms of mean f_0 , but not for standard deviation and range.

Figure 24 shows that their results of time-normalized f_0 contour analysis were similar to ours: in subject position, new referents are marked by an increase of mean f_0 value and a peak along the prestressed syllables, and word size seems to be a factor that boosts such behavior. In object position, however, the contours for new and given referents are very similar, even overlapping in many points regardless word size.

³⁴ Though subject and object are not in fact sentence positions, this study has considered the canonical SVO order in BP as a direct relation between position and grammatical function.

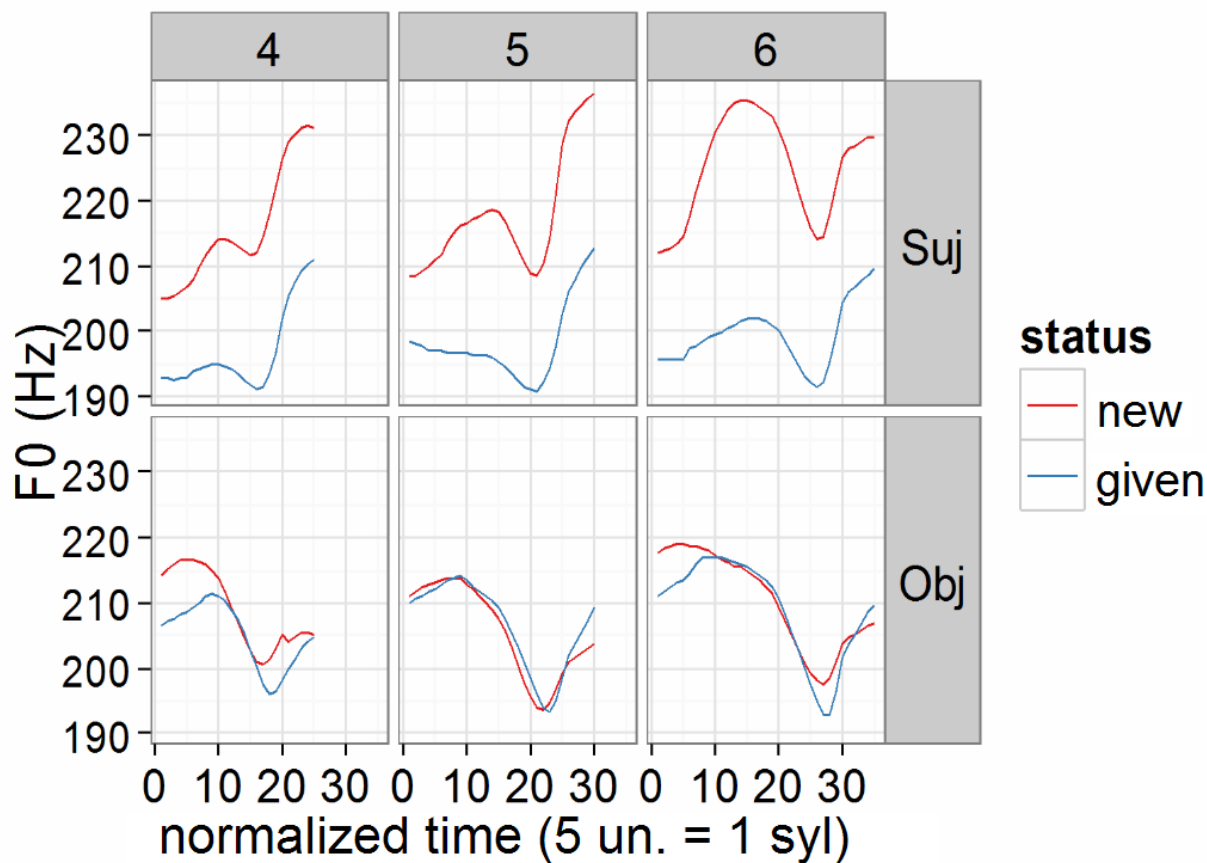


Figure 24 - Time-normalized f_0 contour analysis. Results are divided by word size (number of syllables) and sentence position. The analysis took five samples of f_0 per syllable of all target words (Antão et al., 2013, adapted)

Despite the fact that in Experiment 2 the target words were in sentence-final position, all referents behaved similarly to the experiment described above: there it seems that new referents lack an acoustic prominence when they have the grammatical function of object.

We propose that the general organization of information structure might explain it. As we have seen, the given-before-new principle (Gundel, 1988) predicts that old information (i.e. given referents) tends to occur in the beginning of the utterance and new information is set to occur in later positions³⁵. The Syntax-

³⁵ Gundel (1988) is frequently cited as the main reference related to this principle of linearization of information. However, such idea has already been presented in Halliday (1967) under the

Discourse Model (Burkhardt, 2005, see also Avrutin, 1999 for a debate on syntax-discourse interface) also proposes that information in the beginning of sentences tends to be given, as it serves as a link to previous discourse units. Because the object position is the place *par excellence* for new information, the need for an acoustic prominence marking of new referents in that position seems less necessary. It is as if the listener already expects newsworthy information in this position, thus no prosodic prominence is necessary to signalize discourse relevance. In this case, listeners would rely more on other linguistic information (e.g. morphosyntactic information) to assign the right information status.

Reversely, for subject position, listeners expect given information as a hook for previous discourse. In this case, Gussenhoven's Effort Code (2002, 2004) might explain the need to signalize new, discourse-relevant information with acoustic prominence. According to this idea, speakers and listeners share a grammaticalized biological code in which relevant information is signaled with more articulatory effort. The speaker has to indicate that the relevant information is placed where nonprominent information is expected to occur, that is, the placement of new information in initial position is marked and must receive acoustic prominence in order to signalize that it is in fact novel information. As our results show, it is done mainly by an increase in duration, higher average f_0 level, and a different overall f_0 contour in BP.

The analysis of accessible referents indicates that they do not have any distinctive acoustic correlate from given referents. Apparently, accessible referents behave as their given counterparts (at least in a prosodic account), leading to a number of questionings.

First, it is possible that readers of the three experiments were guided also by morphosyntactic cues (e.g. definiteness of determiners and identity of lexical

name *Given-New Principle* and studies from the Prague School under the name *Communicative Dynamism* (Firbas, 1971, inter alia).

forms) when encoding a prosodic marking. As we have seen, new information in BP can be marked by indefinite articles (*um, uma, uns, umas*) while given information can be marked by definite articles (*o, a, os, as*). In the experiment stimuli, the definite article in front of accessible target words could have driven the prosodic encoding towards a given status-like pattern. The similar behavior of the correlates of duration and f_0 for both given and accessible information is an evidence for this hypothesis. Nevertheless, such argument clashes with the given-before-new principle. If speakers are indeed guided more by local, grammatical features than to global, discourse-driven constraints, then we would expect new referents (i.e. with indefinite articles) to receive acoustic prominence even in sentence-final position. As we have seen before, major differences between the information statuses are related to a general organization of the discourse structure in the utterance.

Second, because evidence from other languages shows that accessible status can be variable in terms of its acoustic marking, there is no consensus on a possible prosodic marking. Baumann (2006) points out that accessibility is very sensitive to semantic relationships between the antecedent and the referent thus its prosodic marking is variable. In this study, he observed that there is not a typical pitch accent that would characterize most types of semantic relationships, that is, referents can be marked by several different pitch accents (e.g. H*, H+L*). According to him, the only case in which a semantic relationship is marked by a typical pitch accent is when the referent and its antecedent have a whole-to-part relation (see example 13 in Chapter 3). Findings in Baumann & Grice (2004) also indicate that accessible information cannot be treated as a uniform category - at least in terms of a consistent prosodic marker. We argue that, in this case, listeners might map the information status onto his mental model by a set of complementary cues e.g. contextual and/or semantic information. Such claim is aligned with some other proposals, e.g. Lambrecht (1994), who denies a direct phonological correlate of accessible information. Chafe (1994:75), when discussing about the prosodic prominence associated to each activation status,

suggests that in most cases accessible information is marked with an accent as new information. Such claim might be related to the idea that accessible referents are in a semi-active state, thus being regarded as newsworthy information. Contrastively, our empirical results indicate that accessible status is treated as much as given information.

Third, when we compare the results found in BP to the German studies, it is possible that accessible status does not have a typical acoustic signature because the prosodic encoding is subject to crosslinguistic variation. As we have seen in Chapter 3 for example, languages can encode prosodic information in different ways according to the plasticity parameter. In this scenario, typological differences between the two languages might explain why in BP accessible referents tend to behave as given ones and in German they tend to be marked with an acoustic prominence.

Finally, we must also consider that the results found by Baumann might be biased by his experimental design: his study was based mainly on an impressionistic approach, which possibly has affected his results. In this case, the aspect of reproducibility of similar results in other circumstances should be regarded as an important matter. In our series of experiments, we controlled a wide range of phonetic/phonological (word size, stress pattern), semantic (animacy, antecedent-referent semantic relationship) and syntactic (recency, grammatical function, sentence position) variants. Most studies considered only part of these factors in their analyses. Also, statistical treatment has proved to be a powerful tool to observe the significance of several potential correlates for the linguistic encoding of information status. Additionally, our results show that it is possible to elicit prosodic differences between different referential statuses even in a lab speech setting. The results reported here are comparable to those found in studies which used question-answer elicitation, map task, and annotated corpora from (semi-) spontaneous speech. To sum up, we are confident of the reproducibility of our results because of the number of subjects and the design of

the items used in the experiments.

To solve this matter of accessible referents, more exploratory studies are necessary, in which accessible referents were considered as a condition. So far, our conclusions are only comparable to the German studies (Baumann, 2006). Such phenomenon deserves more detailed phonetic descriptions in order to provide comparable results for crosslinguistic investigations. We propose then that further investigations should adopt an approach of experimental design similar to the present study, considering an analysis of quantitative controlled data and proper statistical treatment.

Another issue which is worth noting in the experiments is individual variability. Consistent prosodic contrast between new and given referents was always found within subjects, despite differences between them. In the three experiments, subjects used differences in F0 contour and duration; in many cases, range was also a relevant parameter. We assume that other factors can also be subject to individual variability, such as register, personal style, emphatic focus, and dialect. Wolters (1999), in a similar study with radio broadcast corpus of Boston English, also found out that

[...] if there are clear prosodic correlates of givenness, we should not expect them to be the same across speakers (Donzel & Beinum, 1997), which raises the question if some speakers use prosody less optimally than others. Our results show that speakers can vary greatly the way they use prosodic cues to mark status³⁶.

As we have seen in our review of crosslinguistic studies in Chapter 3, the PLASTICITY parameter of intonation (Vallduví, 1991) divides languages on how they prosodically encode information structure.

The results found in the present study might suggest that BP has features of non-plastic languages such as Italian, Spanish, Catalan and Romanian. In Experiments

³⁶ Subject variability is also discussed in Wolters & Byron (2000).

1 and 3, we observed that duration and changes in f_0 contour were the main acoustic correlates for the encoding of information status. New referents tend to have significantly longer duration and higher average pitch level than given and accessible referents. Contrary to the results found in Japanese and Mandarin, information status seems to have no significant effect on pitch range in BP, but an overall increase in mean pitch values due to a difference of f_0 movement along the article and the prestressed syllables (such results have been already observed in Arantes, 2010).

Nevertheless, further analysis of other aspects such as (re/de)accenting is necessary for a better understanding of the effect of this parameter in BP. Studies such as Cruttenden (1993, 2006) permit a direct comparison of the effect of information status in deaccenting. The work of Swerts (2000, 2002, 2007) in different languages, adopting the same map task protocol, also permits fruitful insights in crosslinguistic data. In this case, we propose also experimental analysis that could indicate if BP behaves as other nonplastic languages.

Conclusion and Outlook

The central idea investigated in the present study is how different degrees of information status are encoded by prosodic correlates in the discourse under the premisses of taxonomies that take into account linguistic and cognitive constraints (Chafe, 1976, 1994; Ariel, 1990; Gundel et al., 1993; Lambrecht, 1994; Almor, 1999; Baumann & Riester, 2012).

The results of the production experiments show that, indeed, speakers of BP use prosody to mark different levels of information status. Depending on the point of view, one can say that speakers tend to lengthen NPs that introduce newly presented information into the discourse or that they shorten already mentioned and contextually accessible entities. Moreover, speakers also produce a pitch prominence along the prestressed syllables to signalize which information is new from given and accessible information. Apparently, speakers mark the information status (considering the level of activation) by varying the strength of prominence in secondary stress (see Bolinger, 1986 and Baumann, 2006 for similar discussion). These findings corroborate the results found in Arantes (2010) on the marking of given and new information. The lack of a significant differences in standard deviation and range of the fundamental frequency contrasts to the results found in other languages e.g. Japanese (Swerts et al., 2000) and Mandarin Chinese (Pan et al., 2005), which reveal the role of pitch excursion

as an acoustic correlate for information status. Our analysis might suggest that BP is figured in the so-called *nonplastic languages*, which make use of variation of acoustic correlates (mainly f_0) and reaccent given information. More evidences are necessary for to confirm such hypothesis. Yet, it would be another evidence relativizing deaccenting as a feature of group of languages, not as a universal.

One interesting point is the analysis of accessible information. In our crosslinguistic review presented in Chapter 3, most of the studies considered only three types of information: new, given, and contrastive. Overall, the results indicate a tendency for contrastive information to receive the highest pitch prominence followed by new information. None of these studies analyzed accessible information as a distinct status. In a trailblazing study, Baumann (2006) has found that, at least in German, accessible information can in fact be encoded with a typical pitch accent. Chafe (1994:75), when discussing the prosodic prominence associated to each activation status, suggests that accessible information is in most cases marked with an accent as new information. Contrastively, our results indicate that accessible status does not seem to have a one-to-one or single acoustic signature and this prosodic encoding is subject to crosslinguistic variation. At least in our three production experiments, accessible information is treated as much as given information.

If we consider Gussenhoven's (2002) Effort Code - which suggests that languages have gramaticalized a biological code to signalize relevant information during communication -, it is possible to say that languages use the same set of linguistic sources to manifest this phenomenon on the surface of the utterance but in particular ways. Such differences can be parametric (like deaccenting) or due to structural pressure (given-to-new principle).

Some considerations are important to be made about the hold-downs of this study. First, it seems that speakers of BP vary in a considerable degree how they use the different acoustic cues to signalize the information status to his listener,

which prevents a direct comparison of the parameters in a between-subjects analysis. Still, we could find relevant results in the separate analyses of the subjects. One of the advantages of using a larger number of subjects in production experiments is the possibility of stronger generalizations for the population if compared to results found in studies with fewer subjects, even with repeated output.

Another point is the nature of the task and the way subjects were instructed. Sociolinguistic studies have claimed that subjects would act naturally if not aware of his monitoring and only real life data would allow to have a transparent vision of the behaviour of the languages. However, lab speech can also provide useful data for the understanding the language (Xu, 2010). For instance, since the objective of our study was the analysis of acoustic cues in an experimental approach, it was fundamental to filter out as many potential biases as possible in order to obtain more confidence from the statistical results. Still, our findings has shown that it is possible to find a balance in order to achieve reliable results.

Finally, researchers often confound categories such as focus and givenness, which creates overlapping effects in the results. For instance, as in most cases the new referent is the point of focus in the utterance, it is very ease to take them in a biunique relation. We suggest future studies in the area to follow more strict theoretical and methodological criteria to provide comparable data.

Departing from our current results, we are leaned to make the following questions:

- 1) Once addressers use acoustic cues to signalize information status, do these cues have a correlation or they act independently? Are any of these correlates more relevant to the addressee's point of view?
- 2) Do different information statuses have specific inventory of pitch accents in BP?

3) In this study, we considered the marking of simple noun phrases. The map task protocol seen in some studies (Swert et al., 2000; Swerts et al., 2002; Swerts, 2007) considered the analysis of more complex noun structures e.g. adjective+noun. How would the information status be encoded in these cases? Is the prominence projected to the whole phrase or only on the head?

4) In two experiments, we saw that whole-to-part accessible information is regarded as given information. Would there be any difference with other semantic relations e.g. hyponym-to-hypernym and hypernym-to-hyponym?

5) Do new, given and accessible information have neural correlates in BP as those found in German (Schumacher and Baumann, 2012)?

6) We found that in BP the prosodic encoding of givenness is more explicit in subject position. How then grammatical function or position affect prosodic encoding? Are there any other factors to be considered? Is the given-before-new principle verifiable in BP?

Behavioral experiments such as self-paced listening will be able to evaluate the relevance of the acoustic parameters in an independent account. Also, they might show evidence to understand the position effect on the prosodic encoding of referents.

A phonological analysis can provide interesting findings concerning the relation between information status and intonation structure. The employment of notation systems such as DaTo (Lucente, 2008), which has a dynamic approach, will permit the observation of distribution of pitch accents as well as phrasing associated to givenness.

Further production experiments can consider other aspects of prosodic encoding

such as spontaneous speech, pronominalization of given information, and other semantic relations to accessible referents. Goal-oriented and/or collaborative experiments would allow the participants to engage in the task, being more expressive with a relative control of production at the same time.

Finally, the use of non-intrusive online techniques such as eyetracking and EEG would permit the observation of the prosodic encoding during referential processing with a great time resolution. Behavioral experiments are usually cheaper and faster to be executed, but they lack the possibility of analysing linguistic phenomena in a deeper level or without the necessity of attentional focus on the task.

References

- [1] 't Hart, J., Collier, R., Cohen, A. (1990). *A Perceptual Study of Intonation: An Experimental Phonetic Approach to Speech Melody*: Cambridge: Cambridge University Press.
- [2] Allerton, D. J. (1978). The notion of givenness and its relations to presupposition and to theme. *Lingua Amsterdam*, 44(2-3), 133-168.
- [3] Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, 106(4), 748.
- [4] Altmann, G., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33(4), 583-609.
- [5] Arantes, P. (2010). *Integrando produção e percepção de proeminências numa abordagem dinâmica do ritmo da fala*. PhD Thesis]. Campinas: UNICAMP.
- [6] Arantes, P., Lima, M. L. C., & Barbosa, P. A. (2012). Some prosodic correlates of referential status in Brazilian Portuguese. *Revista Diadorim*, 12.
- [7] Ariel, M. (1990). *Accessing noun-phrase antecedents*: Routledge London.

- [8] Arnold, J. E., Losongco, A., Wasow, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 28-55.
- [9] Avesani, C., & Vayra, M. (2005). *Accenting, deaccenting and information structure in Italian dialogue*. Paper presented at the 6th SIGdial Workshop on Discourse and Dialogue.
- [10] Baumann, S., Brinckmann, C., Hansen-Schirra, S., Kruijff, G.-J., Kruijff-Korbayova, I., Neumann, S., & Teich, E. (2004). *Multi-dimensional annotation of linguistic corpora for investigating information structure*. Paper presented at the Proc. NAACL/HLT" Frontiers in Corpus Annotation.
- [11] Baumann, S. (2006). The intonation of givenness. *Tübingen: Niemeyer*.
- [12] Baumann, S., Grice, M., & Steindamm, S. (2006). *Prosodic marking of focus domains-categorical or gradient*. Paper presented at the Proceedings of speech prosody.
- [13] Baumann, S., & Riester, A. (2012). Referential and lexical givenness: semantic, prosodic and cognitive aspects. *Prosody and Meaning 25, De Gruyter Mouton, Berlin*.
- [14] Baumann, S., & Schumacher, P. B. (2012). (De-) Accentuation and the Processing of Information Status: Evidence from Event-Related Brain Potentials. *Language and speech*, 55(3), 361-381.
- [15] Beckman, M., Díaz-Campos, M., McGory, J. T., & Morgan, T. A. (2002). Intonation across Spanish, in the Tones and Break Indices framework. *Probus*, 14(1), 9-36.
- [16] Behlau, M., & Pontes, P. Avaliação e tratamento das disfonias. São Paulo: Ed. Lovise; 1995. *Links*.
- [17] Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America*, 113, 1001.

- [18] Birner, B. J., & Ward, G. L. (1998). *Information status and noncanonical word order in English* (Vol. 40): John Benjamins.
- [19] Bolinger, D. L. (1963). Length, vowel, juncture. *Linguistics*, 1(1), 5-29.
- [20] Bolinger, D. L. M. (1958). *A theory of pitch accent in English*.
- [21] Bolinger, D. L. M. (1986). *Intonation and its parts: Melody in spoken English*: Stanford University Press.
- [22] Brown, G. (1983). Prosodic structure and the given/new distinction. *Prosody: Models and measurements*, 67-77.
- [23] Burkhardt, P. (2008). Dependency precedes independence: Online evidence from discourse processing. *Pragmatics And Beyond New Series*, 172, 141.
- [24] Burkhardt, P., & Domahs, U. (2009). Experimental evidence for minimal structure. *Lingua*, 119(10), 1373-1379.
- [25] Cagliari, L. C. Prosódia: algumas funções dos supra-segmentos. *Cadernos de Estudos Lingüísticos*, 137-151.
- [26] Chafe, W. L. (1970). *Meaning and the structure of language*. Chicago, Ill.: University of Chicago Press.
- [27] Chafe, W. L. (1974). Language and consciousness. *Language*, 111-133.
- [28] Chafe, W. L., & Li, C. N. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View in Subject and Topic.
- [29] Chafe, W. L. (1987). Cognitive constraints on information flow. *Coherence and grounding in discourse*, 11, 21-51.
- [30] Chafe, W. L. (1994). *Discourse, consciousness and time: the flow and displacement of conscious experience in speaking and writing*. Chicago, Ill.: University of Chicago Press.

- [31] Chahal, D. (2001). Modeling the intonation of Lebanese Arabic using the autosegmental-metrical framework. *Unpublished dissertation, University of Melbourne.*
- [32] Chahal, D. (2003). *Phonetic cues to prominence in Arabic.* Paper presented at the Proceedings of the 15th International Congress of Phonetic Sciences.
- [33] Chomsky, N., & Halle, M. (1968). The sound pattern of. *Eng/isa, Cambridge, Mass.*
- [34] Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. *Discourse Production and Comprehension. Discourse Processes: Advances in Research and Theory, 1,* 1-40.
- [35] Cruttenden, A. (1979). *Language in infancy and childhood: a linguistic introduction to language acquisition.* New York: St. Martin's Press.
- [36] Cruttenden, A. (1986). *Intonation.* Cambridge Cambridgeshire; New York: Cambridge University Press.
- [37] Cruttenden, A. (1993). *The de-accenting and re-accenting of repeated lexical items.* Paper presented at the ESCA Workshop on Prosody.
- [38] Cruttenden, A. (1997). Intonation. *Cambridge textbooks in linguistics.* Cambridge U.K.; New York, NY, USA: Cambridge University Press.
- [39] Cruttenden, A. (2006). The de-accenting of given information: A cognitive universal?. *Empirical Approaches To Language Typology, 20(8),* 311.
- [40] Crystal, D. (1975). Prosodic features and linguistic theory. *The English tone of voice: Essays in intonation, prosody and paralanguage,* 1-46.
- [41] Crystal, D. (1976). *Prosodic systems and intonation in English (Vol. 1):* CUP Archive.
- [42] Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech, 40(2),* 141-201.

- [43] Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47(2), 292-314.
- [44] Delfino, A., & de Magalhães, J. O. (2010). Estudo prosódico das disfluências de reparo. *ReVEL*, 8 (15).
- [45] Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America*, 80, 402.
- [46] Féry, C., & Ishihara, S. (2005). Interpreting second occurrence focus. *Manuscript. Potsdam. Paper submitted for publication.*
- [47] Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, 36(4), 680-703.
- [48] Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26(5), 489-504.
- [49] Fowler, C. A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *Language and speech*, 31(4), 307-319.
- [50] Frota, S. (2000). *Prosody and focus in European Portuguese: Phonological phrasing and intonation*: Psychology Press.
- [51] Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America*, 27, 765.
- [52] Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech*, 1(2), 126-152.
- [53] Gernsbacher, M. A. (1991). Comprehending conceptual anaphors. *Language and cognitive processes*, 6(2), 81-105.
- [54] Givón, T. (1979). *On understanding grammar*: Academic Press New York.
- [55] Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study* (Vol. 3): John Benjamins Publishing.

- [56] Görgülü, E. (2009). On definiteness and specificity in Turkish.". *MIT Working Papers in Linguistics*, 58.
- [57] Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). *The effects of collocational strength and contextual predictability in lexical production*. Paper presented at the Chicago Linguistic Society.
- [58] Grice, H. P. (1975). Logic and conversation. 1975, 41-58.
- [59] Grice, M., & Baumann, S. (2002). Deutsche intonation und G-ToBI. *Linguistische Berichte*, 191, 267-298.
- [60] Grice, M., D'Imperio, M., Savino, M., & Avesani, C. (2005). Strategies for intonation labelling across varieties of Italian. *Jun, S.-A. (Ed.). Prosodic typology: the phonology of intonation and phrasing*, Oxford University Press, Oxford, 55-83.
- [61] Grønnum, N. (1992). *Perceptual invariance in Danish stress group patterns*. Paper presented at the Nordic Prosody VI, Papers from a symposium, Stockholm.
- [62] Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational linguistics*, 21(2), 203-225.
- [63] Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274-307.
- [64] Gussenhoven, C. (2002). *Intonation and interpretation: phonetics and phonology*. Paper presented at the Speech Prosody 2002, International Conference.
- [65] Halliday, M. A. (1967). Notes on transitivity and theme in English: Part 2. *Journal of linguistics*, 3(02), 199-244.
- [66] Halliday, M. A. K. (1967). Some Aspects of the Thematic Organization of the English Clause.

- [67] Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of verbal learning and verbal behavior*, 13(5), 512-521.
- [68] Hawkins, J. A. (1978). *Definiteness and indefiniteness: a study in reference and grammaticality prediction*: Croom Helm London.
- [69] Hellmuth, S. (2005). No De-accenting in (or of) Phrases: Evidence from Arabic for cross-linguistic and cross-dialectal prosodic variation. *Prosodies*, 99-112.
- [70] Hirst, D., & Di Cristo, A. (1986). Unités tonales et unités rythmiques dans la représentation de l'intonation. *Actes des 15èmes Journées d'Etude sur la Parole, Aix en Provence*, 93-95.
- [71] Hirst, D., & Di Cristo, A. (1998). *Intonation systems: a survey of twenty languages*: Cambridge University Press.
- [72] Hornby, P. A. (1972). The psychological subject and predicate. *Cognitive Psychology*, 3(4), 632-642.
- [73] Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language*, 66(1), 265-284.
- [74] Jackendoff, R. (1972). *Semantic interpretation in generative grammar*: MIT press Cambridge.
- [75] Jackendoff, R., & Chafe, W. (1996). *On language and consciousness*. Amsterdam: John Benjamins Publishing Company.
- [76] Jones, D. (1950). *The pronunciation of English* (Vol. 369): Cambridge University Press.
- [77] Kingdon, R. (1958). *The groundwork of English intonation*: Longmans.
- [78] Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3), 243-276.

- [79] Kruyt, J. G. (1985). *Accents from speakers to listeners: an experimental study of the production and perception of accent patterns in Dutch*: JG Kruijt.
- [80] Kügler, F. (2008). *The role of duration as a phonetic correlate of focus*. Paper presented at the Proceedings of the Speech Prosody 2008 Conference.
- [81] Kulk, F., Odé, C., & Woidich, M. (2003). *The intonation of colloquial Damascene Arabic: a pilot study*. Paper presented at the Institute of Phonetic Sciences, University of Amsterdam.
- [82] Kutas, M., & Federmeier, K. D. (2009). N400. *Scholarpedia*, 4(10), 7790.
- [83] Ladd, D. R. (1980). *The structure of intonational meaning: Evidence from English*: Indiana University Press Bloomington.
- [84] Ladd, D. R. (1987). *A model of intonational phonology for use in speech synthesis by rule*. Paper presented at the European Conference on Speech Technology.
- [85] Ladd, D. R. (1996). *Intonational phonology* (Vol. 79): Cambridge University Press.
- [86] Ladefoged, P. (1996). *Elements of acoustic phonetics*: University of Chicago Press.
- [87] Lambrecht, K. (1996). *Information structure and sentence form: topic, focus, and the mental representations of discourse referents* (Vol. 71). Cambridge university press.
- [88] Laver, J. (1994). *Principles of phonetics*: Cambridge University Press.
- [89] Lee, Y.-c. (2012). Prosodic Correlation Between the Focusing Adverb Ozik ‘Only’ and Focus/Givenness in Korean. *Journal of Speech Sciences*, 2(2), 85-111.
- [90] Lehiste, I. (1970). *Suprasegmentals*. Oxford: MIT Press.
- [91] Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and speech*, 6(3), 172-187.

- [92] Lucente, L. (2008). DaTo: um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos: ênfase no foco e na fala espontânea.
- [93] Lucente, L., & Barbosa, P. A. (2008). *Narrow focus in Brazilian Portuguese: spatial and temporal constraints*. Paper presented at the Proceedings of Fourth Conference on Speech Prosody.
- [94] Lucente, L., & Barbosa, P. A. (2010). *The role of alignment and height in the perception of LH contours*. Paper presented at the Speech Prosody 2010-Fifth International Conference.
- [95] Lucente, L., Hirschberg, J., & Barbosa, P. (2011). Intonation, discourse structure and information status in spontaneous speech. *Experimental and Theoretical Advances in Prosody, 2*.
- [96] Marslen-Wilson, W., Levy, E., & Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. *Speech, place, and action, 339-378*.
- [97] Miglio, V. G., Gries, S. T., & Harris, M. J. (2014, January 2-5). *New/given information in Mexican & Chicano Spanish intonation*. Paper presented at the 88th Annual Meeting of the Linguistic Society of America, Minneapolis, MN.
- [98] Moraes, J. A. (1998). Intonation in Brazilian Portuguese. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, 170-194
- [99] Moraes, J. A. (2006). *Variações em torno de tema e rema*. Paper presented at the Cadernos do IX Congresso Nacional de Lingüística e Filologia.
- [100] Moraes, J. A. (2007). Intonational Phonology of Brazilian Portuguese.
- [101] Moraes, J. A. (2008). *The Pitch Accents in Brazilian Portuguese: analysis by synthesis*. Paper presented at the Proceedings of Speech Prosody.
- [102] Nakatani, L. H., & Aston, C. H. (1978). *Acoustic and linguistic factors in stress perception*.

- [103] Nava, E., & Zubizarreta, M. (2010). Deconstructing the Nuclear Stress Algorithm: evidence from second language Speech. *The Sound patterns of Syntax*.
- [104] Nespor, M., & Vogel, I. (1986). *Prosodic phonology: with a new foreword* (Vol. 28): Walter de Gruyter.
- [105] Nolan, F., & Jónsdóttir, H. (2001). *Accentual Patterns in Icelandic*. Paper presented at the van Dommelen, W. A. & T. Fretheim Nordic Prosody. Proceedings of the VIII Conference, Trondheim.
- [106] Nooteboom, S. G., & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *The Journal of the Acoustical Society of America*, 82, 1512-1524.
- [107] Norlin, K. (1989). A preliminary description of Cairo Arabic intonation of statements and questions. *Speech Transmission Quarterly Progress and Status Report*, 1, 47-49.
- [108] Pan, H.-h., Huang, W.-T., & Huang, Y.-H. (2005). *Perception of new, given and contrastive information in Taiwanese*. Paper presented at the ISCA Workshop on Plasticity in Speech Perception.
- [109] Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., & Vasisht, S. (2008). Focus, word order and intonation in Hindi. *Journal of South Asian Linguistics*, 1(1), 55-72.
- [110] Paul, H. (1880). *Prinzipien der sprachgeschichte* (Vol. 6): Walter de Gruyter.
- [111] Philipp, M., Bornkessel-Schlesewsky, I., Bisang, W., & Schlesewsky, M. (2008). The role of animacy in the real time comprehension of Mandarin Chinese: Evidence from auditory event-related brain potentials. *Brain and Language*, 105(2), 112-133.
- [112] Pierrehumbert, J. (1990). The Meaning of Intonational Contours in the Interpretation of Discourse Janet Pierrehumbert and Julia Hirschberg. *Intentions in communications*, 271.
- [113] Pierrehumbert, J., & Beckman, M. (1988). Japanese tone structure. *Linguistic Inquiry Monographs* (15), 1-282.

- [114] Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Massachusetts Institute of Technology.
- [115] Pike, K. L. (1945). *The Intonation of American English*.
- [116] Pitrelli, J. F., Beckman, M. E., & Hirschberg, J. (1994). *Evaluation of prosodic transcription labeling reliability in the tobi framework*. Paper presented at the ICSLP.
- [117] Prince, E. F. (1981). Toward a taxonomy of given-new information. *Radical pragmatics*.
- [118] Prince, E. F. (1992). The ZPG letter: Subjects, definiteness, and information-status. *Discourse description: diverse analyses of a fund raising text*, 295-325.
- [119] Reinhart, T. (1981). Pragmatics and Linguistics: An Analysis of Sentence Topics in Pragmatics and Philosophy I. *Philosophica anc Studia Philosophica Gandensia Gent*, 27(1), 53-94.
- [120] Reinhart, T. (1982). Pragmatics and Linguistics: An Analysis of Sentence Topic. *Philosophica*, 27.
- [121] Schumacher, P. B., & Baumann, S. (2010). Pitch accent type affects the N400 during referential processing. *NeuroReport*, 21(9), 618-622.
- [122] Schumacher, P. B., & Hung, Y.-C. (2012). Positional influences on information packaging: Insights from topological fields in German. *Journal of Memory and Language*, 67(2), 295-310.
- [123] Selkirk, E. (2002). Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in English. Paper presented at the Speech Prosody 2002, International Conference.
- [124] Stein, C. C. (2011). Aspectos da manifestação prosódica em relação com os processos de ativação cognitiva no português. *Ciências e Cognição/Science and Cognition*, 5, 2-18.
- [125] Swerts, M. (2007). Contrast and accent in Dutch and Romanian. *Journal of Phonetics*, 35(3), 380-397.

- [126] Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics*, 30(4), 629-654.
- [127] Swerts, M., Taniguchi, M., & Katagiri, Y. (2000). *Prosodic marking of information status in Tokyo Japanese*. Paper presented at the INTERSPEECH.
- [128] Terken, J., & Hirschberg, J. (1994). Deaccentuation of words representing 'given' information: effects of persistence of grammatical function and surface position. *Language and speech*, 37(2), 125-145.
- [129] Terken, J., & Nootboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and cognitive processes*, 2(3-4), 145-163.
- [130] Thorsen, N. G. (1985). Intonation and text in Standard Danish. *The Journal of the Acoustical Society of America*, 77, 1205.
- [131] Trager, G. L., & Smith, H. L. (1951). *An outline of English structure* (Vol. 3): Battenburg Press.
- [132] Valgeirsson, G. (2011). The Re-accenting of Given Information in Icelandic. A Comparative Study of English and Icelandic Intonation patterns.
- [133] Vallduvi, E. (1992). *The informational component*. New York: Garland.
- [134] Van Dijk, T. A., Kintsch, W., & Van Dijk, T. A. (1983). *Strategies of discourse comprehension*: Academic Press New York.
- [135] Van Hooijdonk, C., Commandeur, E., Cozijn, R., Krahmer, E., & Marsi, E. (2007). *The online evaluation of speech synthesis using eye movements*. Paper presented at the Proc. of the 6th ISCA Workshop on Speech Synthesis (SSW-6).
- [136] Vanderslice, R., & Pierson, L. S. (1967). Prosodic features of Hawaiian English. *Quarterly Journal of Speech*, 53(2), 156-166.
- [137] Venditti, J. J. (2000). *Discourse structure and attentional salience effects on Japanese intonation*. PhD dissertation. The Ohio State University.

- [138] Venditti, J. J., Maekawa, K., & Beckman, M. E. (2008). Prominence marking in the Japanese intonation system. *Handbook of Japanese linguistics*, 456-512.
- [139] Wang, B., & Xu, Y. (2011). Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. *Journal of Phonetics*, 39(4), 595-611.
- [140] Watson, D. G., Arnold, J. E., & Tanenhaus, M. K. (2008). Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 106(3), 1548-1557.
- [141] Wolters, M. (1999). *Prosodic correlates of referent status*. Paper presented at the Proc. ICPHS.
- [142] Xu, Y. (2010). In defense of lab speech. *Journal of Phonetics*, 38(3), 329-336.
- [143] Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, 33(2), 159-197.
- [144] Zimmermann, M., & Féry, C. (2010). *Information structure: theoretical, typological, and experimental perspectives*: Oxford University Press.
- [145] Zubizarreta, M. L. (1998). *Prosody, focus, and word order* (Vol. 213): MIT Press Cambridge, MA.

APPENDIX A - List of sentences for Production Experiment 1

ACC	Começam a funcionar as instalações elétricas depois que o <i>granizo</i> atingiu uma subestação da Cemig. A tempestade castigou uma região habitada , e os moradores ficaram sem fornecimento de energia elétrica desde então.
NEW	Uma tempestade castigou uma região habitada , próxima ao deserto do Atacama na noite de domingo. Equipes de busca tentam achar várias pessoas desaparecidas.
GIV	O governo chileno pediu ajuda à ONU depois que uma tempestade devastou uma região inteira. A tempestade castigou uma região habitada, que possui mais da metade da população do país.
ACC	Depois de alguns problemas na correção das provas, a Prefeitura de Contagem suspendeu um <i>concurso</i> público realizado no mês passado. A candidata pediu uma revisão , e pediu que todos que se sentiram prejudicados façam o mesmo.
NEW	Uma candidata pediu uma revisão dos resultados de um concurso depois que surgiram denúncias de favorecimento por membros da banca. Os membros foram indiciados e as provas serão reanalisadas.
GIV	Examinadores de uma faculdade particular desclassificaram uma <i>candidata</i> que usava um celular durante as provas. A candidata pediu uma revisão , alegando que havia pegado o celular para atender uma chamada de emergência.
ACC	Ontem foi a última apresentação do grupo artístico este ano. Todos aplaudiram o <i>espetáculo</i> com um entusiasmo contagiante. A dançarina comemorou emocionada o encerramento da temporada, e anunciou que haverá novidades para o próximo ano.
NEW	Uma dançarina comemorou emocionada depois da indicação de seu nome para o prêmio de "Artista do Ano". O anúncio foi feito durante uma coletiva concedida pela associação que promove a premiação.
GIV	O diretor de eventos do grupo Corpo recebeu um prêmio pelo espetáculo apresentado nesta temporada. O diretor anunciou que o prêmio vai para a <i>dançarina</i> mais nova do grupo. A dançarina comemorou emocionada , e fez uma homenagem ao diretor ao final da premiação.
ACC	Foi aprovada verba para a reforma de um casarão antigo na região central da cidade. O casarão foi tombado há dez anos, mas um problema na fundação do <i>imóvel</i> impedia a sua reforma. O projetista propôs uma solução para o problema, e a obra iniciará no próximo mês.
NEW	Um projetista propôs uma solução depois de várias reuniões sobre a divisão da área comum do prédio. O síndico convocou os moradores para que o problema fosse resolvido, mas nada foi feito até que alguém tivesse a ideia de contratar um especialista para analisar o problema.
GIV	Terminou essa semana a conferência que definiu o projeto para a construção do novo estádio. Os dirigentes dos clubes de futebol consultaram o <i>projetista</i> para um projeto inovador. O projetista propôs uma solução , e o comitê liberou a verba para a finalização das obras.

- ACC O governador decidiu fechar a faculdade de medicina após crime ocorrido. **A professora procurou a reitoria**, e solicitou apoio dos colegas para evitar que 600 alunos fiquem sem aulas. O secretário de Educação disse que analisará o pedido.
- NEW **Uma professora procurou a reitoria** depois que vândalos fecharam o prédio da sede administrativa. Ela exigiu que alguma medida fosse tomada para que as atividades ali realizadas não fossem prejudicadas.
- GIV Depois de uma reclamação por assédio moral, o conselho gestor da faculdade expulsou a professora denunciada sem aviso prévio. **A professora procurou a reitoria**, alegando que não houve apresentação formal de queixa.
- ACC A Corregedoria da Polícia Civil investigará uma denúncia anônima. De acordo com o inquérito, um garoto de treze anos ficou detido por horas em uma cela. **O delegado pediu uma boa explicação** após perceber o que ocorria na delegacia.
- NEW **Um delegado pediu uma boa explicação** sobre o caso de correspondência adulterada ocorrido na última semana. Uma empresa especializada em transportes de valor entregou um envelope que estava parcialmente aberto, o que gerou suspeitas sobre a autenticidade do documento.
- GIV O Tribunal de Justiça está investigando um esquema de subornos denunciado no mês passado. A defesa intimou o delegado a comparecer para esclarecimentos. **O delegado pediu uma boa explicação**, já que o seu nome foi indicado como suposto chefe do esquema. O delegado nega as acusações e se diz vítima de uma armação.
- ACC Há algum tempo, o Procon tem recebido várias queixas do péssimo atendimento de um restaurante na zona sul da capital. **A garçonete comentou o problema**, e explicou que o baixo número de funcionários prejudica a qualidade do atendimento. O dono do restaurante prometeu resolver o problema para reabrir o local.
- NEW **Uma garçonete comentou o problema** gerado por dois clientes na noite de quinta-feira. Um dos clientes arremessou uma cadeira, acertando outro cliente do bar que estava sentado próximo ao local da briga. A polícia foi chamada para registrar a ocorrência.
- GIV O fórum determinará hoje a pena para o advogado acusado de assédio. De acordo com o processo, o advogado assediou uma garçonete durante uma festa promovida pela sua empresa. **A garçonete comentou o problema**, e alegou que o advogado insistiu em chamá-la para sair depois da festa, mesmo depois de inúmeras recusas.
- ACC A Guarda Municipal está intensificando a vigilância nos arredores de uma escola municipal em Pradópolis. Alunos do curso noturno invadiram a escola para roubar comida. **A pedagoga questionou a atitude** dos assaltantes, já que a maioria dos alunos conta com a merenda da escola todos os dias.
- NEW **Uma pedagoga questionou a atitude** de uma lanchonete no centro da cidade. A lanchonete dobrou o preço de todos os itens do menu, fazendo com que muitos estudantes de uma escola situada perto da lanchonete não pudessem comer no intervalo das aulas.
- GIV A Secretaria de Educação está analisando um caso de denúncia de violência física em uma creche da cidade municipal. Alguns pais perceberam marcas roxas nas crianças e registraram queixa contra a pedagoga da creche, que foi afastada de seu cargo. **A pedagoga questionou a atitude**, alegando que as marcas são resultado das brincadeiras das crianças.

ACC	Confrontos violentos foram registrados nesta quinta-feira em Santiago, depois de um protesto de várias minorias étnicas para exigir um processo de imigração mais justo. Houve uma tentativa de invasão da embaixada do Brasil que fica no centro da cidade. O diplomata convocou as autoridades , que responderam com bombas de efeito moral.
NEW	Um diplomata convocou as autoridades responsáveis pela proteção das fronteiras do país depois da descoberta de mais um cartel de drogas. A Polícia Federal espera que a sua ação, juntamente com o apoio da embaixada, possa conter o avanço do tráfico de drogas nas fronteiras.
GIV	O governo americano anunciou que manterá as tropas no Oriente Médio até o fim do ano. Depois da ameaça de grupos terroristas diante da longa ocupação, um diplomata solicitou um programa gradativo de retirada. O diplomata convocou as autoridades , que se reunirão com Chefe de Estado árabe para uma transição de governo pacífica.
ACC	A polícia iniciou a investigação do roubo do carregamento de bauxita que era transportado em um comboio para o porto de Santos. O maquinista contou sua versão do ocorrido e fez uma descrição dos assaltantes. A polícia anunciou que já tem uma pista de onde está a carga.
NEW	Um maquinista contou sua versão sobre o atropelamento de três turistas no bonde de Santa Tereza, no Rio de Janeiro. O funcionário disse em depoimento oficial que o freio do bonde falhou, e não houve como evitar o acidente.
GIV	Após um novo problema registrado em um trem nesta tarde, os passageiros que seguiam no ramal Japeri precisaram andar sobre os trilhos depois que o maquinista liberou as portas de emergência. O maquinista contou sua versão , e disse que a rede elétrica que alimenta o sistema foi interrompida durante a viagem.
ACC	A Secretaria de Segurança pública anunciou hoje que irá tomar providências mais severas para conter a rebelião na cadeia estadual em Contagem. O criminoso pediu auxílio , e um advogado de uma firma famosa na capital ofereceu seus serviços para cuidar dos seus interesses.
NEW	Um criminoso pediu auxílio depois de ter sido ameaçado de morte dentro da cadeia de Ribeirão das Neves. O advogado do criminoso conta que ele foi colocado em uma ala junto a membros de uma facção rival. O diretor da cadeia prometeu resolver a situação.
GIV	Uma perseguição policial terminou com um morto na noite desta quinta-feira. Segundo a polícia, um caixa eletrônico foi arrrombado por um criminoso quando a polícia foi acionada. O criminoso pediu auxílio para um comparsa que estava no carro, mas foi preso logo em seguida. O comparsa foi preso poucas horas depois.
ACC	A reitoria da universidade readmitiu o aluno expulso depois que o Colegiado de Graduação rejeitou o seu pedido de continuidade de estudos. A faculdade reconheceu o erro e fez um pedido formal de desculpas ao aluno, que sofre de dislexia e, portanto, possui dificuldade em aprender.
NEW	Uma faculdade reconheceu o erro na correção das provas e decidiu suspender o resultado. A Comissão Permanente de Vestibular disse que haverá uma reunião para decidir se fará convocação para novas provas. Essa decisão determinará o futuro de 20 mil candidatos.
GIV	Professores e alunos fizeram uma manifestação pacífica depois que uma estudante foi baleada dentro do campus da faculdade por um assaltante. A faculdade reconheceu o

erro, e disse que fará uma licitação para contratar uma empresa de segurança especializada para evitar novos crimes.

- ACC Depois de duas semanas de chuvas ininterruptas, a comunidade de São Conrado pode receber socorro. **A moradora mobilizou-se para ajudar**, chamando toda a vizinhança para ajudar no recolhimento dos entulhos das casas destruídas.
- NEW **Uma moradora mobilizou-se para ajudar** um sem-teto que transita no bairro Santo Agostinho há alguns anos. Vários moradores da região, sensibilizados com a campanha de uma senhora que vive no bairro há mais de 30 anos, resolveram colaborar e conseguiram um lar para o sem-teto.
- GIV A Associação de Moradores do Centro fará uma campanha para que a prefeitura adapte as calçadas da cidade para pessoas com deficiência motora. A Associação pediu uma moradora idosa para ser a garota-propaganda da campanha. **A moradora mobilizou-se para ajudar**, e outras instituições manifestaram apoio ao projeto.
- ACC A equipe Williams de F1 anunciou a renovação do contrato de seus pilotos para o campeonato de 2012. **A temporada com lucros começou** com a vitória nos quatro primeiros grandes prêmios, e a diretoria da equipe quer garantir a dupla de sucesso no próximo ano.
- NEW **Uma temporada com lucros começou** para os jogadores da principal liga de basquete nos Estados Unidos. A associação dos jogadores conseguiu um acordo para limitar em 10% a comissão dos agentes dos jogadores. Alguns agentes se mostraram indignados.
- GIV O porta-voz da rede de televisão NBC anunciou que haverá grandes mudanças na programação da temporada do próximo ano. **A temporada com lucros começou** com a venda dos direitos de alguns programas para outros países, e o presidente da rede espera que o dinheiro possa ser investido na produção de novos programas.
- ACC Um novo escândalo reabriu o debate sobre a ética de publicação de livros. Uma editora proibiu a publicação de uma biografia de um político, considerada sensacionalista. **A redatora demonstrou insatisfação**, e anunciou que o livro será publicado em uma outra editora.
- NEW **Uma redatora demonstrou insatisfação** depois da demora na entrega de sua compra por uma famosa loja de eletrodomésticos. A loja fez uma promoção relâmpago, mas várias pessoas tiveram que aguardar a entrega posterior da encomenda. O gerente da loja disse que o problema será resolvido em breve.
- GIV Será decidida hoje a polêmica sobre a censura do jornal de maior circulação no país, depois que uma coluna denunciou um esquema de superfaturamento do governo. O editor-chefe advertiu a redatora pela conduta, suspendendo a série de reportagens. **A redatora demonstrou insatisfação**, e disse que é um absurdo a censura das reportagens que mostram a verdade por trás da licitação de obras.
- ACC Quatro homens armados tentaram assaltar, no final da manhã desta segunda-feira, uma agência bancária no centro de Farroupilha. **O segurança reagiu à ameaça** e foi baleado na perna. Segundo a polícia, os criminosos chegaram ao local em um automóvel roubado.
- NEW **Um segurança reagiu à ameaça** feita por manifestantes na sede do governo e foi agredido no início desta manhã. Desde o início da greve, manifestações violentas têm causado temor no Secretário de Saúde, que solicitou aumento na segurança dos prédios públicos.

- GIV Uma briga numa boate no centro da capital terminou com um ferido a bala. Dois jovens tentaram sair sem pagar a conta, e um dos jovens apontou uma arma para o segurança responsável por resolver o problema. **O segurança reagiu à ameaça** e foi baleado na perna. A Polícia Militar foi acionada logo depois do incidente.
- ACC A rodovia MG10 ficará interditada até o fim da tarde. A causa é o tombamento de um caminhão carregado com areia. **A condutora causou um tumulto** depois que vários moradores dos arredores tentaram saquear a carga. Um trator foi chamado para liberar parte da pista.
- NEW **Uma condutora causou um tumulto** depois de ter atropelado dois ciclistas na manhã de domingo. Testemunhas contam que a condutora invadiu a ciclovia, forçando várias pessoas a correrem do local. A condutora alegou que o freio do carro falhou na hora.
- GIV A Polícia Rodoviária Federal flagrou uma condutora portando uma carteira de habilitação falsificada. **A condutora causou um tumulto** ao receber a voz de prisão. Ela alega que a carteira é legítima porque comprou de um funcionário do Detran.
- ACC A SPFW bateu o recorde de desfiles este ano. **A passarela permitiu o acesso** de várias modelos a contratos milionários desde a primeira edição do evento. Outros profissionais também se beneficiam do evento, como produtores, estilistas e fotógrafos.
- NEW **Uma passarela permitiu o acesso** aos destroços de uma casa depois do temporal de ontem. Os bombeiros conseguiram resgatar os moradores através de uma passarela de concreto, que passava do lado da casa. Nenhum morador sofreu ferimentos graves.
- GIV Depois de anos de abandono, moradores da periferia comemoraram a reinauguração da passarela sobre a rodovia. Desde a sua construção, **a passarela permitiu o acesso** dos moradores ao outro lado do bairro, onde existe um hospital e uma creche.
- ACC As críticas ao novo espetáculo foram péssimas para os produtores e artistas do teatro mais tradicional da cidade. **O dramaturgo provocou uma confusão** depois que a multidão começou a esvaziar o teatro, e houve comentários de que a temporada da peça se encerraria mais cedo do que o previsto.
- NEW **Um dramaturgo provocou uma confusão** ao ser confundido com um preso foragido durante uma blitz da polícia militar. O homem só foi liberado depois que a polícia constatou que os dois tinham o mesmo nome.
- GIV Uma coletiva de imprensa terminou em constrangimento depois do tumulto causado por um respeitável dramaturgo que conceceu entrevista ontem. **O dramaturgo provocou uma confusão** quando o jornalista insinuou que a sua peça era um plágio.
- ACC No próximo sábado, haverá uma passeata de ex-funcionários do exército em função dos baixos salários. **O veterano queixou-se da situação humilhante** pela qual vários pensionistas passam, chegando a depender da ajuda de instituições de caridade.
- NEW **Um veterano queixou-se da situação humilhante** causada por um funcionário de um centro de compras. Algumas pessoas que presenciaram o ocorrido contaram que o funcionário fez comentários sobre a aparência do veterano, que não teve como se defender.
- GIV Uma manifestação gigantesca fechou o acesso da faculdade na manhã de segunda. Estudantes da faculdade exigem melhores condições de funcionamento depois que um veterano ficou preso no elevador quando a energia foi cortada. **O veterano queixou-se da**

situação humilhante, mas nada foi resolvido até agora.

- ACC Atletas de várias modalidades participaram da competição beneficente no último final de semana. **A corredora concorreu ao prêmio** e anunciou o nome da instituição que receberá o dinheiro.
- NEW **Uma corredora concorreu ao prêmio** de 'atleta da década' em uma premiação patrocinada pelo COB. Contudo, o prêmio foi dado para um nadador, que sofreu um grave acidente há cinco anos e conseguiu bater vários recordes na última Olimpíada.
- GIV Foi realizada no domingo a famosa corrida de São Silvestre. Um jornal local disse que uma corredora do Quênia estaria preparada para vencer facilmente. **A corredora concorreu ao prêmio** e dedicou sua vitória ao seu preparador físico.
- ACC Estudiosos da Sismologia têm procurado analisar os dados de tremores para prever novas ocorrências. **O terremoto causou a destruição** do norte do Japão sem que nenhum centro de climatologia pudesse perceber qualquer alteração no subsolo.
- NEW **Um terremoto causou a destruição** de boa parte da costa leste dos Estados Unidos. Várias cidades não tinham um programa de evacuação, o que deu trabalho para as equipes de resgate.
- GIV O governo decidiu pelo fechamento da usina nuclear após o terremoto ocorrido no mês passado. **O terremoto causou a destruição** de vários componentes do reator, aumentando o risco de contaminação da região.
- ACC Está marcada para o próximo mês uma reunião entre os empregados e o clube de futebol para definir os termos do acordo. **O dirigente solicitou alterações** no acordo original, e os empregados rejeitaram a nova proposta.
- NEW **Um dirigente solicitou alterações** na lei que define a punição para políticos corruptos. Vários partidos apoiaram a solicitação, que será votada na próxima semana.
- GIV A CBF analisará a arbitragem de doze jogos do último turno. Depois de constatadas várias falhas, um dirigente acusou os juizes de fraude. **O dirigente solicitou alterações** na tabela de pontuação, e alguns times que seriam rebaixados poderão continuar na primeira divisão.
- ACC O aumento da renda dos brasileiros têm gerado mudanças na rotina das famílias. Por exemplo, as lojas têm registrado um aumento na venda de eletrodomésticos considerados mais caros. **A lavadora promete uma mudança** na rotina das donas de casa, que têm mais tempo para fazer outras coisas.
- NEW **Uma lavadora promete uma mudança** no mercado de eletrodomésticos. A Brastemp anunciou esta semana que a nova linha de produtos terá como foco a economia de água e energia.
- GIV O governo está fazendo uma campanha para conscientizar as comunidades ribeirinhas sobre o despejo de lixo nos rios. A líder do grupo que promove palestras é uma lavadora de uma dessas comunidades. **A lavadora promete uma mudança** nos hábitos das mulheres da região.
- ACC Uma discussão acalorada se transformou em constrangimento durante a reunião do governo na tarde de ontem. **O presidente perdeu a razão** ao tentar justificar os atos cometidos durante a ditadura. Segundo ele, algumas prisões foram necessárias para manter a ordem.

NEW	Um presidente perdeu a razão na reunião dos chefes de Estado na sede da ONU. Quando questionado sobre a onda de violência causada pela guerra civil, o presidente desligou o microfone e se retirou do local.
GIV	Depois de uma declaração controversa durante um debate, dois grandes partidos discutem a renúncia do presidente da câmara na reunião de hoje. O presidente perdeu a razão quando permitiu que os seus assessores fizessem acordos sigilosos com partidos menores para fortalecer a sua base de apoio.
ACC	Várias organizações internacionais anunciaram o apoio ao país em dificuldades financeiras. O governante divulgou a ideia de que o dinheiro deveria ser utilizado para estabilizar a economia e favorecer as classes mais pobres.
NEW	Um governante divulgou a ideia de angariar fundos para a construção de um hospital na região mais carente do Estado. Várias ambulâncias daquela região vêm todos os dias para a capital em busca de socorro para pacientes.
GIV	Os refugiados da região da Somália receberão ajuda graças a um governante dos Emirados Árabes. O governante divulgou a ideia em um jantar beneficente promovido para diminuir os efeitos devastadores da guerra que assola a África há anos.
ACC	O Tribunal Superior autorizou a transmissão ao vivo do julgamento mais polêmico do ano. A testemunha mudou a declaração a favor do réu, e disse que outras testemunhas também estão comprometidas.
NEW	Uma testemunha mudou a declaração feita durante a investigação sobre a morte do empresário. A polícia suspeita que a testemunha foi comprada pelos réus.
GIV	Uma reviravolta no caso de corrupção do Ministério dos Transportes. A CPI foi adiada depois que uma testemunha foi chamada pelo relator. A testemunha mudou a declaração que havia feito antes, e alegou que estava sob forte pressão para confirmar as acusações.
ACC	Várias pessoas ficaram sem transporte público depois da greve iniciada nesta segunda-feira. O sindicato divulgou a notícia sobre a paralisação depois que alguns empregados foram advertidos por panfletarem na porta da empresa.
NEW	Um sindicato divulgou a notícia da recontração dos funcionários demitidos pela fábrica de automóveis. A demissão em massa foi uma manobra da empresa para minimizar os efeitos da recessão econômica.
GIV	Morreu hoje o mais antigo membro do sindicato dos metalúrgicos de São Paulo. O sindicato divulgou a notícia depois de receber um telefonema do assessor do sindicalista.
ACC	O governo enviou uma nota de apoio ao exército depois da retirada das tropas. O fuzileiro celebrou a conquista e declarou que pretende se casar quando voltar à sua cidade natal.
NEW	Um fuzileiro celebrou a conquista da tomada da capital da Líbia durante a madrugada. Alguns jornalistas registraram o momento em que o soldado festejava no topo da sede do governo.
GIV	O general determinou a reintegração do fuzileiro expulso por desobediência. O fuzileiro celebrou a conquista , e disse que fará de tudo para merecer a nova chance.
ACC	Os brasileiros festejaram o fim do campeonato no hipódromo de Jacarepaguá. O cavaleiro garantiu sua vitória ao completar a prova com a pontuação máxima, e foi premiado com

a medalha de ouro.

NEW **Um cavaleiro garantiu sua vitória** na corrida de hoje mesmo estando há um ano longe dos treinos. Um trágico acidente forçou o afastamento do cavaleiro, que chegou a anunciar sua aposentadoria prematura.

GIV A cavalgada pelo sertão do estado terminou ontem com a chegada de um cavaleiro matogrossense. **O cavaleiro garantiu sua vitória** fazendo o percurso em menos tempo.

APPENDIX B - List of sentences for Production Experiment 2

ACC	Ao sair de casa, o rapaz não conseguiu ligar o <i>carro</i> . O frio havia congelado a bateria .
NEW	O homem foi à loja para comprar uma bateria .
GIV	O rapaz comprou uma <i>bateria</i> nova para o seu notebook. Quando ele sentiu um cheiro de queimado, ele foi à loja devolver a bateria .
ACC	O turista foi à feira e comprou uma <i>lagosta</i> para o jantar. O comerciante explicou como quebrar a caparaça .
NEW	Todos os insetos possuem uma carapaça .
GIV	O acúmulo de sujeira do oceano formou uma <i>carapaça</i> dura no casco do navio. Os marinheiros tiveram que retirar a carapaça .
ACC	O soldador gastou várias horas tentando consertar uma <i>solda</i> . Mais tarde, ele percebeu que bastava trocar o eletrodo .
NEW	Ao abrir a máquina, o técnico descobriu que a ferrugem havia corroído um eletrodo .
GIV	O médico viu que não havia sinal de um <i>eletrodo</i> durante o exame. Ele pediu à enfermeira que trocasse o eletrodo .
ACC	O mecânico verificou o estado do <i>motor</i> e passou o orçamento para a moça. Ele disse que precisaria retificar o cabeçote .
NEW	Como não havia peças disponíveis no mercado, o técnico fez um cabeçote .
GIV	O balconista vendeu um <i>cabeçote</i> genérico porque a peça original era muito cara. Depois de duas semanas, o cliente teve que trocar o cabeçote .
ACC	A alta no preço do condomínio afetou o bolso de alguns moradores do <i>prédio</i> . Um dos moradores mais antigos se viu obrigado a vender a cobertura .
NEW	O jogador de futebol disse na entrevista que quando recebeu seu primeiro salário, ele comprou uma cobertura .
GIV	Com a crise no mercado, a imobiliária estava com uma <i>cobertura</i> há meses em proposta. A solução foi abaixar o preço para vender a cobertura .
ACC	O casal comprou um <i>aquecedor</i> usado mas que não funcionava. Eles tiveram que chamar um técnico para consertar o termostato .
NEW	O técnico aprendeu no curso como calibrar um termostato .
GIV	O aquecimento do apartamento novo contava com um <i>termostato</i> digital. O senhor inexperiente não sabia como mudar o termostato .
ACC	A polícia conseguiu prender os bandidos que arrombaram a <i>porta</i> . Durante o interrogatório, um dos suspeitos descreveu como eles abriram a fechadura .
NEW	O presidiário ensinou ao novato como abrir uma fechadura .

GIV	O homem instalou uma <i>fechadura</i> dourada na entrada da casa. Contudo, a mulher achou que a cor não combinava com a fachada e pediu para retirar a fechadura .
ACC	O dono do sítio viu o <i>cavalo</i> mancando e chamou um veterinário. Ele recomendou ao dono que trocasse a ferradura .
NEW	O ferreiro ensinou ao seu aprendiz como fazer uma ferradura .
GIV	Um apostador ganhou uma <i>ferradura</i> de presente para melhorar a sorte. Depois de perder tudo o que tinha, ele devolveu a ferradura .
ACC	Na manhã de natal, a criança recebeu um <i>presente</i> muito especial. Ela estava muito eufórica e acabou rasgando a embalagem .
NEW	A senhora foi à papelaria e comprou uma embalagem .
GIV	A artesã criou uma <i>embalagem</i> para presentear uma amiga no seu aniversário. A amiga ficou surpresa e elogiou a embalagem .
ACC	O aluno chegou atrasado na escola e contou que havia um problema com a <i>bicicleta</i> . Ele havia caído em um barranco e empenou a pedaleira .
NEW	Um famoso ciclista foi à loja recém-inaugurada para comprar uma pedaleira .
GIV	O ciclista ganhou uma <i>pedaleira</i> nova do seu patrocinador. Durante a corrida, o atleta elogiou a pedaleira .
ACC	Um mulher ameaçou se jogar de uma janela do décimo andar do prédio . Para realizar o resgate, o bombeiro teve que se subir o parapeito .
NEW	O arquiteto redefiniu a fachada e incluiu um parapeito .
GIV	O diretor chamou os pais dos alunos picharam um <i>parapeito</i> da escola. Os alunos terão que limpar o parapeito .
ACC	Durante a reforma do apartamento, os pedreiros tiveram que fechar a <i>área de serviço</i> . A empregada reclamou que não pôde usar a lavadora .
NEW	A mulher foi à loja de eletrodomésticos e comprou uma lavadora .
GIV	A família conseguiu comprar uma <i>lavadora</i> depois de fazer uma vaquinha. A mãe ensinou às filhas menores como usar a lavadora .
ACC	O aposentado comprou uma <i>vara de pesca</i> nova no feriado. Quando chegou no lago, ele percebeu que havia esquecido a carretilha .
NEW	O pescador ensinou ao filho como limpar uma carretilha .
GIV	O representante levou uma <i>carretilha</i> recém-lançada para o encontro de pescadores. Ele queria mostrar como manusear a carretilha .
ACC	O médico tirou um <i>medidor de pressão</i> da bolsa para examinar o paciente que estava em convulsão. O médico teve dificuldade de ajustar a braçadeira .
NEW	O capitão do time de futebol deve usar uma braçadeira .

GIV	O maratonista ganhou uma <i>braçadeira</i> nova que monitora o rendimento durante a corrida. O atleta estava ansioso para estrear a braçadeira .
ACC	O caminhoneiro teve que ir à mecânica porque uma <i>roda</i> estava frouxa. O funcionário disse que bastava apertar o parafuso .
NEW	O mecânico pediu ao funcionário que comprasse um parafuso .
GIV	A velhinha pediu ao neto que procurasse por um <i>parafuso</i> perdido no carpete. Depois de vários minutos, o garoto encontrou o parafuso .
ACC	Na época de reprodução, o salmão tem que nadar o <i>rio</i> até a nascente. Muitos peixes morrem no caminho por não conseguirem subir a correnteza .
NEW	O especialista em sobrevivência mostrou aos novatos como encarar uma correnteza .
GIV	Durante a tempestade, o morador caiu em uma <i>correnteza</i> que se formou na rua. Os bombeiros resgataram o homem que enfrentava a correnteza .
ACC	O motorista teve que levar o carro em uma mecânica para consertar a <i>porta</i> depois da batida. O impacto empenou a maçaneta .
NEW	O professor de física disse que, durante uma tempestade, deve-se evitar encostar em uma maçaneta .
GIV	A senhora comprou uma <i>maçaneta</i> para o portão novo. Ela pediu ao porteiro do prédio que instalasse a maçaneta .
ACC	Durante a celebração, várias pessoas suspeitas transitavam pela <i>igreja</i> . O padre pediu ao coroinha que trancasse a sacristia .
NEW	Os fiéis da paróquia fizeram uma quermesse que levantou fundos para construir uma sacristia .
GIV	A igreja possuía uma <i>sacristia</i> muito ornamentada com vários objetos valiosos. Um meliante se disfaçou de padre e roubou a sacristia .
ACC	Esta semana será reaberta a feira de exposição com um <i>presépio</i> vivo. Vários turistas vêm à região para fotografar a manjedoura .
NEW	Durante a ceia, o pai explicou ao filho por que Jesus nasceu em uma manjedoura .
GIV	Com a chegada do natal, a família fez uma <i>manjedoura</i> para enfeitar a sala de estar. Uma das crianças ajudou a enfeitar a manjedoura .
ACC	O homem percebeu que haviam roubado os documentos do armário. O ladrão tinha arrombado o cadeado .
NEW	Logo depois de comprar uma bicicleta, o rapaz resolveu comprar um cadeado .
GIV	O empregado havia perdido as chaves de um <i>cadeado</i> que trancava o portão. Ele foi obrigado a chamar um chaveiro para arrombar o cadeado .
ACC	A criança se assustou com o escuro do quarto e acabou fazendo xixi no <i>berço</i> . A mãe teve que usar uma toalha para secar o colchonete .

- NEW Antes de acampar, o casal decidiu comprar um **colchonete**.
- GIV O vizinho atencioso emprestou um *colchonete* para um desabrigado. Antes de devolver, o homem enrolou o **colchonete**.
- ACC A chuva de granizo danificou o *carro* estacionado na rua. O motorista pagou caro para consertar o **para-brisa**.
- NEW O dono do lava-jato ensinou ao novato como limpar um **para-brisa**.
- GIV Um homem atirou uma pedra e acertou um *para-brisa* durante uma briga no trânsito. Depois de detido pela polícia, o culpado falou que iria trocar o **para-brisa**.
- ACC O rapaz queria consertar a porta empenada. Ele teve que substituir **a dobradiça**.
- NEW O casal queria repintar a janela mas não queria sujar *uma dobradiça* de tinta. Ele demorou um bom tempo para desparafusar **a dobradiça**.
- GIV O menino prestava atenção enquanto o pai ensinava o modo correto de instalar uma **dobradiça**.
- ACC Depois do incidente aéreo ocorrido no último feriado, a Aeronáutica liberou uma vistoria do *jatinho*. Os peritos suspeitam de sabotagem e estão analisando a **fuselagem**.
- NEW Mesmo depois de estudar aeromodelismo, o menino não sabia como montar uma **fuselagem**.
- GIV Os peritos fizeram uma inspeção no hangar e relataram falhas em uma *fuselagem* do novo projeto. Os técnicos tiveram que descartar a **fuselagem**.

APPENDIX C - List of sentences for Production Experiment 3

ACC	Ao sair de casa, o rapaz não conseguiu ligar o <i>carro</i> . Durante a madrugada, a bateria havia congelado.
NEW	De acordo com o manual, uma bateria é o suficiente para ligar o brinquedo novo.
GIV	O rapaz comprou uma <i>bateria</i> para o aparelho. Depois de alguns minutos, a bateria apresentou um cheiro de queimado.
ACC	Um turista foi à feira e comprou uma <i>lagosta</i> . Por mais uns trocados, a carapaça foi limpa pelo vendedor.
NEW	De acordo a enciclopédia, uma carapaça envolve todos os insetos.
GIV	O acúmulo de sujeira no estaleiro formou uma <i>carapaça</i> dura no casco do navio. Com muito custo, a carapaça foi totalmente retirada.
ACC	O técnico gastou várias horas tentando consertar a <i>solda</i> . Aparentemente, o eletrodo estava corroído por dentro.
NEW	Sem causa aparente, um eletrodo causou pane no servidor da empresa.
GIV	O médico viu que não havia sinal de um <i>eletrodo</i> durante o exame. O eletrodo foi substituído pela enfermeira.
ACC	O mecânico abriu o <i>motor</i> para achar o problema. O cabeçote foi enviado para a retífica.
NEW	Sem recursos, um cabeçote foi a solução encontrada pelo técnico para consertar a máquina.
GIV	O balconista vendeu um <i>cabeçote</i> genérico para o cliente. Duas semanas depois, o cabeçote foi devolvido.
ACC	A alta dos impostos afetou a vida dos moradores do <i>prédio</i> . Hoje de manhã, a cobertura foi posta à venda.
NEW	Atualmente, uma cobertura pode ser uma opção para quem deseja privacidade.
GIV	A imobiliária estava com uma <i>cobertura</i> há meses no mercado. A cobertura foi vendida pela metade do preço.
ACC	O casal comprou um <i>aquecedor</i> usado mas que não funcionava. O termostato estava desregulado.
NEW	Em muitos casos, um termostato pode ajustar a temperatura do ambiente automaticamente.
GIV	O técnico incluiu um <i>termostato</i> digital na reforma da casa. O termostato contava com vários recursos.
ACC	A polícia conseguiu prender os bandidos que arrombaram a <i>porta</i> . Durante o roubo, a fechadura foi aberta com um maçarico.

NEW	Sem manutenção, uma fechadura pode se emperrar em pouco tempo.
GIV	O homem instalou uma <i>fechadura</i> nova no apartamento. A fechadura não combinava com a cor da porta.
ACC	O dono do sítio viu o <i>cavalo</i> mancando. A ferradura teve que ser trocada.
NEW	Em média, uma ferradura custa perto de cinquenta reais.
GIV	Um apostador ganhou uma <i>ferradura</i> de presente. Ironicamente, a ferradura trouxe má sorte.
ACC	Na manhã de natal, a criança recebeu o <i>brinquedo</i> mais esperado do ano. A embalagem foi destruída em segundos.
NEW	No marketing, uma embalagem deve ser pensada para atrair mais consumidores.
GIV	A artesã criou uma <i>embalagem</i> para presentear uma amiga. A embalagem tinha várias fitas coloridas.
ACC	O garoto mostrou o defeito da <i>bicicleta</i> . Durante a queda, a pedaleira havia empenado.
NEW	Aparentemente, uma <i>pedaleira</i> é uma boa opção para quem quer fazer manobras radicais.
GIV	O ciclista ganhou uma <i>pedaleira</i> nova do seu patrocinador. Durante a corrida, a pedaleira foi muito elogiada.
ACC	Uma mulher ameaçou se jogar da <i>janela</i> do último andar. O parapeito foi usado pelo bombeiro no resgate.
NEW	Na manhã de domingo, um <i>parapeito</i> da construção histórica apareceu pichado.
GIV	O diretor notificou os alunos que depredaram um <i>parapeito</i> do terraço da escola. O parapeito foi consertado pelos culpados.
ACC	Os pedreiros tiveram que isolar a <i>lavanderia</i> durante a reforma do apartamento. A lavadora ficou indisponível por vários dias.
NEW	Em época de chuvas, uma lavadora é um eletrodoméstico bastante útil.
GIV	A família conseguiu comprar uma <i>lavadora</i> depois de muito esforço. A lavadora era de um modelo simples.
ACC	O aposentado comprou uma <i>vara de pesca</i> na loja do shopping. A carretilha veio de brinde.
NEW	Sempre que possível, uma carretilha deve ser limpa depois de cada pescaria.
GIV	O vendedor apresentou uma <i>carretilha</i> nova no encontro de pescadores. A carretilha possui uma nova trava de segurança.
ACC	O médico tirou um <i>medidor de pressão</i> para examinar o homem em convulsão. A braçadeira não parava no braço do paciente.
NEW	Na antiguidade, uma <i>braçadeira</i> era símbolo de status social.

GIV	O maratonista ganhou uma <i>braçadeira</i> nova que monitora o rendimento durante a corrida. A braçadeira contava com um GPS.
ACC	O motorista não conseguiu trocar a <i>roda</i> sem precisar de ajuda. O parafuso estava muito apertado.
NEW	Durante o teste final, um parafuso soltou de uma das rodas do protótipo.
GIV	A velhinha pediu ao neto que procurasse por um <i>parafuso</i> no carpete. O parafuso estava em um canto da sala.
ACC	Na época de reprodução, o salmão tem que subir o <i>rio</i> por vários dias. A correnteza é o maior obstáculo para muitos peixes.
NEW	Durante a tempestade, uma <i>correnteza</i> atingiu as áreas mais pobres da região.
GIV	Os bombeiros tentavam resgatar um morador que caiu em uma <i>correnteza</i> no córrego da cidade. A correnteza arrastou o homem por vários metros.
ACC	A síndica pediu ao porteiro para consertar o <i>portão</i> da entrada principal. A maçaneta estava enferrujada.
NEW	No Japão, uma maçaneta deve sempre ser instalada do lado direito da porta.
GIV	O motorista foi a uma mecânica para consertar uma <i>maçaneta</i> empenada na batida. A maçaneta não abria por fora.
ACC	Várias pessoas suspeitas transitavam pela <i>igreja</i> durante a cerimônia. A sacristia foi trancada a pedido do padre.
NEW	Durante uma escavação, uma <i>sacristia</i> foi desenterrada por um grupo de arqueólogos belgas.
GIV	A paróquia possuía uma <i>sacristia</i> cheia de objetos valiosos. A sacristia foi roubada por um meliante disfarçado de padre.
ACC	Os organizadores montaram um <i>presépio</i> vivo para a feira anual. A manjedoura é o item mais caro da exposição.
NEW	Todos os anos, uma manjedoura é posta no exato local onde Jesus nasceu.
GIV	Com a chegada do natal, a família fez uma <i>manjedoura</i> para enfeitar a sala. A manjedoura foi decorada pelas crianças.
ACC	O homem percebeu algo de estranho no <i>armário</i> do depósito. O cadeado havia sido arrombado por alguém.
NEW	Tradicionalmente, um cadeado em uma ponte é uma simpatia para manter o amor do casal.
GIV	O empregado havia perdido as chaves de um <i>cadeado</i> que trancava o portão. O cadeado teve que ser aberto por um chaveiro.
ACC	A criança se assustou e fez xixi no <i>berço</i> durante a noite. O colchonete ficou com uma mancha enorme.

- NEW Entre outros itens, **um colchonete** é equipamento indispensável pra quem deseja acampar por muitos dias.
- GIV O vizinho emprestou um *colchonete* para um desabrigado. **O colchonete** foi devolvido em perfeito estado.
- ACC A chuva de granizo danificou o *carro* no estacionamento. **O para-brisa** foi totalmente destruído.
- NEW Em Copacabana, **um para-brisa** protegeu o motorista de um tiro durante um assalto.
- GIV Um homem atirou uma pedra e acertou um *para-brisa* durante uma briga no trânsito. **O para-brisa** será pago pelo seguro.
- ACC O rapaz queria consertar a *porta* empenada. **A dobradiça** teve que ser substituída.
- NEW Por muito tempo, **uma dobradiça** enferrujada impedia a abertura da janela da choupana.
- GIV O marceneiro tinha que trocar a *dobradiça* da janela com urgência. **A dobradiça** custou a ser desparafusada.
- ACC Depois do acidente, a Aeronáutica liberou a vistoria do *jatinho*. Pelo relatório, **a fuselagem** tinha sido sabotada.
- NEW Hoje de manhã, **uma fuselagem** foi encontrada perto de uma fazenda abandonada no Mato Grosso.
- GIV Os peritos relataram falhas em uma *fuselagem* do novo projeto militar. **A fuselagem** rachou em uma das juntas da cauda.