



UNIVERSIDADE FEDERAL DE GOIÁS (UFG)
FACULDADE DE LETRAS (FL)
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS E LINGUÍSTICA (PPGLL)

CLEITON RIBEIRO E OLIVEIRA

**A constructional analysis of the expressiveness of the
pronominal subject in Brazilian Portuguese and in European
Portuguese**

Goiânia
2023



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE LETRAS

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese Outro*: _____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

Cleiton Ribeiro e Oliveira

3. Título do trabalho

A constructional analysis of the expressiveness of the pronominal subject in Brazilian Portuguese and in European Portuguese

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;

- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Cleiton Ribeiro E Oliveira, Usuário Externo**, em 14/03/2023, às 15:27, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Vania Cristina Casseb Galvao, Professor do Magistério Superior**, em 28/03/2023, às 10:10, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3594385** e o código CRC **FD3F3EE2**.

Referência: Processo nº 23070.067711/2022-50

SEI nº 3594385

UNIVERSIDADE FEDERAL DE GOIÁS (UFG)
FACULDADE DE LETRAS (FL)
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS E LINGUÍSTICA (PPGLL)

CLEITON RIBEIRO E OLIVEIRA

**A constructional analysis of the expressiveness of the
pronominal subject in Brazilian Portuguese and in European
Portuguese**

Thesis presented to *Programa de Pós-graduação em Letras e Linguística* from *Faculdade de Letras* of the Federal University of Goiás as a requirement to obtaining the title of *Doutor em Letras e Linguística*.

Concentration Area: Estudos Linguísticos

L.P.4

Thesis Supervisor: Dra. Vânia Cristina Casseb-Galvão

Goiânia
2023

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Oliveira, Cleiton Ribeiro e
A constructional analysis of the expressiveness of the pronominal subject in Brazilian Portuguese and in European Portuguese [manuscrito] / Cleiton Ribeiro e Oliveira. - 2023.
CXXXIX, 139 f.: il.

Orientador: Profa. Dra. Vânia Cristina Casseb Galvão.
Tese (Doutorado) - Universidade Federal de Goiás, Faculdade de Letras (FL), Programa de Pós-Graduação em Letras e Linguística, Goiânia, 2023.

Bibliografia.

Inclui siglas, abreviaturas, gráfico, tabelas, lista de figuras, lista de tabelas.

1. Constructional Perspective. 2. Brazilian Portuguese and European Portuguese. 3. Syntactical-prosodical interface. 4. Subject Pronoun. I. Galvão, Vânia Cristina Casseb, orient. II. Título.

CDU 81



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE LETRAS
ATA DE DEFESA DE TESE

Ata Nº 2 da sessão de Defesa de Tese Cleiton Ribeiro e Oliveira que confere o título de Doutor em Letras e Linguística, na área de concentração em Estudos Linguísticos.

Aos dezoito dias do mês de janeiro do ano de dois mil e vinte e três, a partir das nove horas, de forma híbrida, realizou-se a sessão pública de Defesa de Tese intitulada "A constructional analysis of the expressiveness of the pronominal subject in Brazilian Portuguese and in European Portuguese". Os trabalhos foram instalados pela Orientadora, Professora Doutora Vânia Cristina Casseb Galvão (PPGLL-FL-UFG) com a participação dos demais membros da Banca Examinadora: Professor Doutor Leosmar Aparecido Silva (PPGLL-FL-UFG), membro titular interno; Professora Doutora Gláucia Vieira Cândido (PPGLL-FL-UFG), membro titular interno, Professora Doutora Déborah Magalhães de Barros (POSLLI /UEG) membro titular externo; Professora Doutora Marília Silva Vieira (POSLLI /UEG) membro titular externo. Durante a arguição os membros da banca não fizeram sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese tendo sido o candidato aprovado pelos seus membros. Proclamados os resultados pela Professora Doutora Vânia Cristina Casseb Galvão, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos dezoito dias do mês de janeiro do ano de dois mil e vinte e três.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Vania Cristina Casseb Galvao, Professor do Magistério Superior**, em 21/03/2023, às 13:12, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leosmar Aparecido Da Silva, Professor do Magistério Superior**, em 21/03/2023, às 15:30, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Glauca Vieira Cândido, Professor do Magistério Superior**, em 21/03/2023, às 15:34, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **MARILIA SILVA VIEIRA, Usuário Externo**, em 22/03/2023, às 16:37, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Déborah Magalhães de Barros, Usuário Externo**, em 24/03/2023, às 17:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3611409** e o código CRC **BD7D9046**.

Referência: Processo nº 23070.067711/2022-50

SEI nº 3611409

CLEITON RIBEIRO E OLIVEIRA

**A CONSTRUCTIONAL ANALYSIS OF THE EXPRESSIVENESS
OF THE PRONOMINAL SUBJECT IN BRAZILIAN PORTUGUESE
AND IN EUROPEAN PORTUGUESE**

EXAMINERS

Prof. Dr. Vânia Cristina Casseb Galvão
UFG - Universidade Federal de Goiás
President

Prof. Dr. Glaucia Vieira Candido
UFG - Universidade Federal de Goiás

Prof. Dr. Leosmar Aparecido da Silva
UFG - Universidade Federal de Goiás

Prof. Dr. Marília Silva Vieira
UEG – Universidade Estadual de Goiás

Prof. Dr. Déborah Magalhães de Barros
UEG – Universidade Estadual de Goiás

ALTERNATES

Prof. Dr. Sebastião Carlos Leite Gonçalves
UNESP – Campus São José do Rio Preto

Prof. Mônica Veloso Borges
UFG – Universidade Federal de Goiás

To Aparecida, Luiz and Henrique.

ACKNOWLEDGMENT

Firstly, I would like to thank my parents, Luiz and Aparecida, and my brother and sister-in-law, Henrique and Daniela, for always being there and backing me up no matter how crazy the plan was. Without your support and your faith in me, none of this would happen.

I also thank Vânia Cristina Casseb Galvão for believing in me and seeing a potential in me that I must yet acknowledge, without your guidance none of this would be possible. All my gratitude to Vito Fiume as well. Vânia and Vito, thanking you for welcoming me into your house and for allowing me to build another home.

Thank you, Déborah, for all the years of friendship, for every piece of advice, for all the dreams we dreamt together.

Professors Leosmar Aparecido Silva and Sebastião Carlos Leite Gonçalves, you have my sincere gratitude, for being understanding and for having patience and for taking care of my ideas in a nourishing and tender way.

Professor Marília Vieira, thank you for guiding me during my masters, you will always be part of every scientific paper I write. Professor Glaucia Vieira Candido and Professor Mônica Veloso Borges, thank you in advance for your time and your valuable contributions.

Thank you, Professor Barbara Gili Fivela, Università Del Salento, my doctoral sandwich supervisor, for having believed me across the ocean and guided me overseas.

I also need to express my affection for those who have somehow helped me walk this path and encouraged me every bit of the way. Giovanna, you have a part of my heart and all my gratitude; Gustavo, thank you for being one of the reasons I still believe in the humanity; Marco, you've been part of my family way before proposing to Gi; Gabriel, thank you for being a haven when it was needed. Lincoln and Soninha, thank you for everything.

Marcos, Giovana, Heloíse and Kawã, you were a safe shore during darkest moments and helped me see a light at the end of the road, a reason to keep going. Thank you for always believing in me.

To the Inácio family, thank you for never being judgmental, for having a blind faith and me and for making me feel special. I am not afraid of going anywhere in the world, because I know I have a family waiting for me to come back.

To Fisk and Coopecigo family, it is because you understand what I needed to do and allow me to do it that this moment is possible. I specially thank Marília for enduring me for more than a decade, and Rosi for taking care of for a lifetime.

Thank you, Rosania, for all the support, all the pep talks and all the time that you understood what I needed without my even having to speak up. I am so not sorry Michely, for tricking you into following the same path I did so we could mutually help each other.

To Laylla and Bruno, soul friends I was blessed with during one of the hardest parts of my life. You made the pandemic time lighter and bearable and made me believe it was possible to get to the other side.

To Jungle Yoon, for your kind and warm words.

My forever gratitude to all of you.

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001*.

one is always responsible for one's position as subject

Jacques Lacan

ABSTRACT

This thesis is linked to the project *Rede/Itália - Português Brasileiro em contexto Italiano, aspectos sociais, políticos e linguísticos* which has the goal to analyze linguistic phenomena and contribute in the Portuguese teaching area. The thesis describes the expressiveness of the pronoun subject in Brazilian Portuguese (BP) and European Portuguese (EP) (TARALLO, 1993; KATO, 1999; KATO; DUARTE, 2014; NEVES; GOULART, 2017; DE ROSA, 2019; OLBERTZ, 2020) to investigate how these two varieties of Portuguese dealing with the necessity of expressing the subject via a pronoun and the relation of this phenomenon with the phonological level of the constructional. In order to have a more detailed idea of the phenomenon, we use Construction Grammar to decompose it and analyze it considering the six levels that form a construction – phonological, morphological, syntactic, semantic, pragmatic and discursive – and how one level can interfere in the others, to make this possible, we use the theories of Croft and Cruse (2004) and Croft (2001), Goldberg (2006), Traugott and Trousdale (2013) and Traugott (2015) to propose a scheme and a network for the expressiveness of the subject pronoun. We also analyzed the cognitive processes that are related to the speakers need to express the subject via a pronoun – iconicity, perspective, informativity, analogy, rich memory, markedness – with the postulates of Bybee (2010; 2015), Givón (1991; 2011), Lakoff (1987) and Langacker (1987; 2000; 2008; 2013). To analyze the phonological level, we used a phonological approach to investigate the intonational pattern of the occurrences in which the prosody falls on the subject pronoun, in order to do it, we used a computer program, PRAAT, and the analysis methodology of Gili Fivela (2002; 2008; 2018), Scarpa and Fernandes-Svartman, (2012) and Pietro and Roseano (2018). The corpora of analysis come from the project *Português Falado - Variedades Geográficas e Sociais*, a project by the Centro de Linguística da Universidade de Lisboa (CLUL) which has 47 interviews from BP and EP from which were selected 20 (10 from Brazil – 5 from the 80s and 5 from the 90s; and 10 from Portugal– 5 from the 80s and 5 from the 90s). We analyzed 820 occurrences, 212, 24 prosodically marked, in EP and 608, 53 prosodically marked, in BP, classifying them following Olbertz (2020) proposal into: referential pronoun; topic pronoun; reactivation of topic; pronouns with no apparent motivation. Our hypothesis is that although there are similarities regarding the expressiveness of the subject via a pronoun in both varieties of Portuguese, that are some aspects that set them apart. The main differences found were: 1) the intonation pattern used by the speakers and its function, in BP, we found the following 8 patterns: H*; L + H; L+ H*; L* + H; H*+L; H + L*; H + H*; H* + H; and, in PE, we find 6: H*; H*+ L ; L + H*; H* + H%; L*+H + H%; L + H* + L; 2) BP uses subject pronouns to refer to generic entities, while EP does not; 3) BP speakers emphasize subject pronouns that are motivated by syntax; while EP speakers emphasize the subject pronouns that profile the role of the new topic. Thus, it became clear that, although it is possible to analyze the expressiveness of the subject in the six levels of a construction, the main differences between the analyzed varieties are more salient in the phonological level.

Keyword: Constructional Perspective; Brazilian Portuguese and European Portuguese; Syntactical-prosodical interface; Subject Pronoun.

RESUMO

Esta tese está ligada ao projeto *Rede/Itália - Português Brasileiro em contexto Italiano, aspectos sociais, políticos e linguísticos* que tem como objetivo analisar os fenômenos linguísticos e contribuir para o ensino do Português. A tese descreve a expressividade do sujeito pronominal no português brasileiro (PB) e no português europeu (PE) (TARALLO, 1993; KATO, 1999; KATO; DUARTE, 2014; NEVES; GOULART, 2017; DE ROSA, 2019; OLBERTZ, 2020) para investigar como essas duas variedades de português estão lidando com a necessidade de expressar o sujeito e a relação desse fenômeno com a nível fonológico da construção. Para termos uma ideia mais detalhada desse fenômeno, utilizamos a Gramática da Construção para decompô-lo e analisá-lo considerando os seis níveis que formam uma construção – fonológico, morfológico, sintático, semântico, pragmático e discursivo – e como um nível pode interferir nos outros, para tornar isso possível, utilizamos as teorias de Croft e Cruse (2004) e Croft (2001), Goldberg (2006), Traugott e Trousdale (2013) e Traugott (2015) para propor um esquema e uma rede para a expressividade do sujeito pronominal. Nós analisamos também os processos cognitivos que estão relacionados à necessidade do falante de expressar o sujeito por meio de um pronome – iconicidade, perspectiva, informatividade, analogia, memória rica, marcação – com os postulados de Bybee (2010; 2015), Givón (1991; 2011), Lakoff (1987) e Langacker (1987; 2000; 2008; 2013). Para analisar o nível fonológico, utilizamos uma abordagem fonológica para investigar o padrão entonacional das ocorrências em que a prosódia recai sobre o pronome sujeito, para isso, utilizamos um programa de computador, PRAAT, e a metodologia de análise de Gili Fivela (2002; 2008; 2018), Scarpa e Fernandes-Svartman, (2012) e Pietro e Roseano (2018). Os corpora de análise são provenientes do projeto *Português Falado - Variedades Geográficas e Sociais*, um projeto do Centro de Linguística da Universidade de Lisboa (CLUL) que conta com 47 entrevistas de PB e PE das quais foram selecionadas 20 (10 do Brasil – 5 dos anos 80 e 5 dos anos 90; e 10 de Portugal – 5 dos anos 80 e 5 dos anos 90). Analisamos 820 ocorrências, 212, 24 marcadas prosodicamente, em PE e 608, 53 marcadas prosodicamente, em PB, classificando-as segundo a proposta de Olbertz (2020) em: pronome referencial; pronome tópico; reativação de tópico; pronomes sem motivação aparente. Nossa hipótese é que, embora existam semelhanças quanto à expressividade do sujeito via pronome nas duas variedades do português, existem alguns aspectos que as diferenciam. As principais diferenças encontradas foram: 1) o padrão entoacional utilizado pelos falantes e sua função, no PB, encontramos os seguintes 8 padrões: H*; L + H; L+ H*; L* + H; H*+L; H + L*; H + H*; H* + H; e, em PE, encontramos 6: H*; H*+ L ; L + H*; H* + H%; L*+H + H%; L + H* + L; 2) BP usa pronomes sujeitos para se referir a entidades genéricas, enquanto EP não; 3) falantes do PB enfatizam os pronomes sujeitos que são motivados pela sintaxe; enquanto os falantes do EP enfatizam os pronomes sujeitos que caracterizam o papel do novo tópico. Ficou claro que, embora seja possível analisar a expressividade do sujeito nos seis níveis de uma construção, as principais diferenças entre as variedades analisadas são mais salientes no nível fonológico.

Palavras-chave: Perspectiva Construcional; Português Brasileiro e Português Europeu; Interface sintaxe-prosódia; Sujeito Pronominal.

LIST OF FIGURES AND TABLES

FIGURES

Figure 1 – Construction Anatomy	54
Figure 2 – Ditransitive Construction	55
Figure 3 – Ditransitive Patterns: To and For	56
Figure 4 – Gradient of hierarchical relationships between constructions	58
Figure 5 – Perspective Based on Spatial Points	71
Figure 6 – Perspective Based on Time Points	72
Figure 7 – Example of Graph Generated by Praat	89
Figure 8 – BP H* Pattern	108
Figure 9 – BP L*+H Pattern	109
Figure 10 – BP H*+L Pattern	111
Figure 11 – BP L+ H* Pattern	111
Figure 12 – BP L+H Pattern	112
Figure 13 – BP H + H* Pattern	113
Figure 14 – BP H + L* Pattern	115
Figure 15 – EP H* Pattern	116
Figure 16 – EP H*+H% Pattern	117
Figure 17 – EP L + H* Pattern	118
Figure 18 – EP L + H* Pattern 2	119
Figure 19 – EP L*+H + H% Pattern	120
Figure 20 – EP H*+L Pattern	121
Figure 21 – EP L + H* + L Pattern	121
Figure 22 – Expressiveness of the Subject Network	126

TABLES

Table 1 – Portuguese Subject Pronouns Paradigm	37
Table 2 – Evolution in Portuguese Inflectional Paradigm	40
Table 3 – BP Hypothetical Paradigm	44
Table 4 – Present Indicative Inflection in Portuguese	46
Table 5 – Construction Dimension	61
Table 6 – Topicalizing Through Prosody	66
Table 7 – Pitch Accents Patterns.	85
Table 8 – General Panorama of the Subject Pronoun Analyzed	92
Table 9 – Subject Pronouns in EP – 80s	93
Table 10 – Subject Pronouns in EP – 90s	93
Table 11 – Subject Pronouns in BP – 80s	94
Table 12 – Subject Pronouns in BP – 90s	94
Table 13 – Pronouns with Generic Reference in the 80s	100
Table 14 – Pronouns with Generic Reference in the 90s	100
Table 15 – Categories of Subject Pronouns in BP	104
Table 16 – Categories of Subject Pronouns in EP	104
Table 17 – Brazilian Prosody Pattern for The Subject Pronoun	115
Table 18 – Ep Prosody Pattern for The Subject Pronoun	122

LIST OF ABBREVIATIONS AND ACRONYMS

- BP – Brazilian Portuguese
- CR – Cross-reference
- EP – European Portuguese
- EPP – Extended Projection Principle
- EM – Emphatic Modifier
- lm – landmark
- NP – Noun Phrase
- O – Object
- PL – plural
- PP – Principles and Parameters theory
- R – Referential
- S – Subject
- SG – singular
- SP – subject pronoun
- tr – trajectory
- UG – Universal Grammar
- V – verb
- VP - vantage point

SUMMARY

ACKNOWLEDGMENT _____	10
RESUMO _____	14
LIST OF FIGURES AND TABLES _____	15
LIST OF ABBREVIATIONS AND ACRONYMS _____	16
SUMMARY _____	17
INTRODUCTION _____	19
A Brief Introduction To The Theoretical Foundation _____	22
Hypothesis _____	26
Objectives _____	27
Methodology approach _____	28
Justificative _____	29
Thesis Organization _____	32
PART I – CONSTRUCTING THE CONCEPT OF SUBJECT _____	34
1.The subject _____	34
2.Subject pronominally expressed _____	36
3.Expressiveness of the subject _____	38
3.1 Expressiveness of the subject in Brazilian Portuguese _____	40
3.2 Expressiveness of the subject in European Portuguese _____	46
PART II – THEORETICAL FOUNDATION _____	50
4. A cognitive and functional view of language _____	50
4.1 Language Conception _____	50
5. Construction Grammar _____	52
5.1 Fundamentals of Construction Grammar _____	53
5.2 Linguistic change under a construction view _____	60
5.3 The integration of a construction [Form + Meaning] _____	64
6. Usage-Based Linguistics _____	66
6.1 Iconicity _____	69
6.2 Perspective _____	71
6.3 Informativity _____	75
6.4 Analogy _____	76
6.5 Rich memory _____	80
6.6 Markedness _____	81
7. Phonological Analysis Paramaters _____	83
PART III – METHODOLOGY _____	88
8. Methodology _____	88
8.1 The <i>corpora</i> _____	88
8.2 PRAAT _____	90
8.3 Analyses criteria _____	91

PART IV – CONSTRUCTIONAL ANALYSES	93
9. Expressiveness of the subject under constructional perspective	93
9.1 Expressiveness of the subject – Morphological Level	93
9.2 Expressiveness of the subject – Syntactical Level	97
9.3 Expressiveness of the subject – Semantic-Pragmatic-Discursive Level	99
Semantic Trait	99
Referentiality	101
Olbertz’s proposal	103
9.4 Expressiveness of the subject – Phonological Level	108
10. Expressiveness of the subject: a constructional representation in Brazilian Portuguese and European Portuguese	125
CONSIDERATIONS AND FORTHCOMING WORKS	128

INTRODUCTION

This thesis is linked to the Project *Rede Itália – Português Brasileiro em contexto italiano, aspectos sociais, políticos e linguísticos*, which is in its 3rd phase: *Epistemes e tradições linguísticas e literárias para o ensino de Português Brasileiro em contexto italiano* financed by the public notice Universal/CNPq/2018 and by CAPES/SECADI. Rede/Itália links both Italian (D'Annunzio University of Chieti - Pescara, University of Salento, University of Pisa, and Roma Ter University) and Brazilian universities (Federal University of Goiás, State University of Goiás, University of Brasília and Federal University of Mato Grosso – Barra do Garças).

This Project focus on researching linguistic and literary phenomena in the broad domain of Portuguese which involves comparative studies and the application of the results of the research in Portuguese teaching in different context: as mother language, second language, foreign language, and heritage language.

Rede/Itália made it possible to develop research in partnership with the University of Salento – IT being aided by the professor Barbara Gili Fivela. Professor Gili Fivela has studied phonetic and phonological phenomena and has coordinated programs focusing on prosody studies both in Italian and cross-linguistic. Gili Fivela has also worked with Luso-Brazilian Association Speech Science regarding *Prosody, variation and contact*.

Our own passion about grammar, the language functionality and adaptability and the one Brazilian Portuguese were our biggest motivations to carry out our research. Looking for a grammatical phenomenon to dive into, initially, we thought about investigating the *pro-drop* phenomena, phenomena already exploited by Gerativism, through a new perspective, the Construcional Grammar, but throughout our research, we soon realized that just considering the *pro-drop* would fall out of construction works, then, we opted to investigate the *expressiveness¹ of the subject via a pronoun* and its implication in the six linguistic levels considered by constructionalists.

Therefore, as part of Rede/Itália and believing that the language can only be studied when we observe its use, it was possible to carry out a study a linguistic phenomenon already exploited by diverse linguistic approaches with a different one, the

¹ In order to describe our object, we opted to use the term *expressiveness* regarding the subject which is phonologically fully expressed.

constructional approach. And Gili Fivela lab and guidance in Unisalento enabled us to analyze our phenomena since its phonological realization.

Based on Tomasello's studies (2008; 2009), it is safe to say that human language has developed at the same pace as the society itself. Every time a new necessity arises or a new path for communication is created, we adapt the language to attend to our interaction needs. This trend of thought it is also present in the linguistics studies that consider the language fluidity, its adaptability, the different contexts of meaning production and the human cognitive apparatus involved in the language usage. All these elements are key factors to linguistic analysis with the perspective of Usage-Based Linguistics and the Construction Grammar approaches.

Regarding these approaches, it is only possible to access linguistic changes if we work with the language in use, because it is only due to interaction and communicational necessity that we adapt the language to our needs and give the language its meaning. Because of these reasons, we adopted the Usage-based linguistics theory and the Construction Grammar proposal to the development of this study.

These approaches to language also state that we must consider every linguistic level that is part of the language production when we are carrying out a linguistic analysis, as well as we can focus on one level at time if we keep in mind that nothing in the language works unassociated. It means that we must consider the phonetic, the morphologic, the syntactic, the semantic, the pragmatic and the discursive levels of the language and a change in any of these levels can result a further change.

Beyond that, we have also to understand the importance of the cognitive system to the linguistic system. How we perceive the world and how we conceptualize it are deeply linked to how we structure the language, that is why the cognitive abilities such as perspective, iconicity and analogy can influence the linguistic system to an extend that even grammar structure can be altered to better express the user's perspective of the world.

In the Grammar Construction view, the grammar is as susceptible to change as the lexicon. It happens because this view consider that the language can be conceived as a set of constructions – pairs of form and meaning – from its basic units to its most complex patterns, from the basic abstract schema of sentence organization to single words like *car*. We can say that there is no distinctive separation between grammar and lexicon in the sense that the linguistic change can occur changing a lexical item to a grammar item via

a grammaticalization process or changing a grammar item to a lexical item via a degrammaticalization process.

These changes, independently of the process, are dependable on different factors. Therefore, all these factors – linguistic, extralinguistic and cognitive – may be behind every change in the grammar structure or in the lexicon.

As for language changes, a phenomenon that have been studied and which is present in both Brazilian Portuguese (BP) and European Portuguese (EP) is the way the subject is expressed, mainly pronominally.

Gerativism perspective have studied this phenomenon under the label of *pro-drop* language. A *pro-drop* language is a language in which the subject is liable to be recovered by the inflectional paradigm and, in this way, the subject does not need to be made explicit by the speaker unless it is the focus of the information. (DIK, 1989).

To better understand a *pro-drop language*, it is necessary to review the Principles and Parameters theory (PP) which established the “rules” used to classify a language as such. According to Kato (2002), the initial Gerativism studies assumed that the learners had an internal grammar, biologically programmed, innate – this grammar is the so-called Universal Grammar (UG). Thus, “UG is defined as the set of invariant Principles that govern natural languages plus the Parameters of interlinguistic variation, conceived as present [+ or 1] or absent [- or Ø] options” (KATO, 2002, p.311).

Kato (2002) also affirms that the principles are believed to be the unequivocal rules that dictate how the language is organized, they cannot be learned; the Parameters, also part of set of rules, they are related to the elements that can be present or absent of the language, hence, they are defined by their valor [+1] or [-/Ø].

From the Principles, there is one that is directly linked to the subject – Extended Projection Principle (EPP). Kato (2002) says that this principle determines that every clause must have a subject, by which languages would have expletive subject. However, there are languages that do not have an explicit subject in some contexts, Romance languages, for example. These languages, such as Portuguese, violated the Extended Projection Principle and the alternative, according to Kato (2002), was to propose a *null subject* (*pro*), that could be identified by the verbal agreement – this phenomenon is known as *zero anaphora* or *null anaphora*. A *null anaphora* happens when a pronoun can be omitted because its information can be somehow recovered – this is what, basically, characterizes a *pro-drop language*.

Thus, Portuguese has been classified as a *pro-drop* language; however, a lot of research have pointed out that the Brazilian variety has undergone a change in its pronominal system that has affected the inflectional system or the change in the inflectional system may reorganize the language making the subject to be expressed. This change happens by reducing the inflectional affixes which has enabled a new classification to Brazilian Portuguese, *partial pro-drop language* – it means that, in some contexts, there has been a tendency of pronominally expressing the subject.

On the other hand, it seems that these changes that has happened to the Brazilian variety have not occurred to the European variety at the same pace and, as such, their speakers are dealing with these changes in different ways. This thought was what originally motivated us to investigate the pronominal expressiveness of the subject in both varieties, what most linguistic studies have called *null subject*.

Since the Construction Grammar states that every construction is multilayered and that its layers are connected to one another, we propose to investigate the expressiveness of the subject pronoun as a construction, so it will enable us to analyze each level separately (the three levels of form: phonological; morphologic and syntactic; and the tree of meaning: semantic; pragmatic and discursive), as well as to analyze whether one of the levels has any effect on other linguistic levels and propose a network that link all the kinds of expressed subjects. A brief theoretical foundation of this thesis is introduced on the next section.

A Brief Introduction To The Theoretical Foundation

Regarding this research goals, we will align this investigation with the theoretical proposal of Usage-Based Linguistics. This theory refers to the conversion between Functionalism and Cognitive Linguistics. This perspective supports the idea that the language is structured based on its use, it is acknowledged that there's a cycle in which grammar enables the use while the use modifies the language structures; it is also acknowledged that the cognitive abilities to perceive the world and act upon it and store data gathered through experiences also have an important impact on the way the language is structured. In this sense, language fluidity and the dynamicity occur within a range that goes beyond the linguistics levels, i.e., speakers' communicational necessity has an important role on the structure.

Thinking about the processes that are related to the language formation and that are important for carrying out this research, we can mention: iconicity², perspective, analogy, rich memory, markedness, topicalization and the linguistic economy principle³.

By *iconicity*, we understand that meanings have influence over the forms that are used to represent them. There is, indeed, a close relation between the way the world is conceived in our minds and the way we conceptualize it and transform these world experiences into linguistic experiences. As we are dealing with the expressiveness of the subject pronoun and the prosodic phenomena, the iconicity studies can help us understand the implications that they may have on meaning production.

Perspective is related to our perception of the world, it is our ability to read the same event in different ways, from different points of view. The relation between what is seen and who sees it is what we call the perspective of the event, and it can impact how we communicate. Hence, in languages that have two possibilities of representation of the same event – omitting or not the subject pronoun –, opting by one of the possibilities over the other may be an indication of how we perceive the event. After perceiving one event, the speaker needs to choose the information that is going to be passed on to the hearer. During this process, one will evaluate different kind of information: who is being talked to; previous knowledge of the event; the linguistic level of the participants involved in the communicational event. Everything that is perceived plays an important role in expressing the subject because, depending on what is intended to be informed, the subject pronoun can be used, or it can be omitted.

Analogy is the domain-general process that, applied to the language domain, allows the speaker to create a new utterance based on patterns acquired through previous experiences. This process is part of the linguistic creativity and is also linked to other processes such as categorization (BYBEE, 2010). Being linked to categorization makes this process also related to frequency, in the sense that while an item is frequent enough to become a central member of a category, it also acquires power to become a model for other items to follow.

² The relation between the iconicity and the economy principle was analyzed by Guerra (2021) and how these processes are linked to the transparency and opacity level of the expressiveness of the first and second pronoun subjects – Her findings will be discussed on the PART I of this research.

³ These processes will be fully explored on Part I – Chapter 2 where they will be associated with the *expressiveness of the subject* phenomenon.

So, understanding the domain-general process of analogy can be a key factor to understand the change in the morphological level, mainly in BP, once this variety is becoming a language with a more reduced inflectional system compared to EP and this reduction may be impacting the way the subject is expressed.

Rich memory is the human capacity to retain detailed information of world experiences. Due to the amount of input that is constantly received, the data storage is only possible through generalizations and schema formations, it means that only part of the information is truly stored. The *rich memory* is also responsible to link different cognitive processes that are accessed during the communication so that the speaker can decode the messages meanings. As phonetical pattern information is also stored and it can be related to other phenomena through *rich memory*, it is possible that the *prosody* system linked to the syntactic system can have impact on the meaning production, it means that when the subject is fully expressed by a pronoun, there is a possibility that the phonologically focused part of the speech to be the subject.

Another phenomenon that can be associated with meaning production is *topicalization*. Topicalization is related to the process of selecting which information is more relevant to the interlocutor and bringing it to the first part of the sentence as a topic. This process is described as *foreground* and *background* by Garcia (1996), being the foreground the most relevant information, normally associated with nouns and verbs, while the background is related to accessory information, generally adverbs, adjectives, articles, etc.

Once this thesis proposes the relation of one linguistic level on the other to analyze the expressiveness of subject pronouns, it is thought that which element of the clause is localized in the apex of the melodic curve can be also considered a case of *topicalization*.

The selecting of the information that will occupy the foreground may codify a marked structure. *Markedness*, according to Givón (1991) and Lakoff (1987), is the process in which some part of the speech is somehow highlighted by the speaker in a functional and informative perspective. This phenomenon is highly linked to frequency distribution and structural and cognitive complexity; relating to grammar structures, usually, marked structure shows a less frequent structure with high structural and cognitive complexity, whereas a non-marked structure is more frequently activated by the speakers and has less structural and cognitive complexity.

So, not expressing the subject pronoun may have been a neutral structure in Portuguese in the past, however, given the frequency that the pronoun is expressed in both BP and EP, it is possible to consider that expressing the subject is becoming a non-marked structure.

The *linguistic economy principle* is the mechanism the speaker uses to transmit more information with less cognitive effort. The language itself operates on one side with efficiency and on the other with expressiveness, the speaker must know how to balance both sides to communicate with success. This principle is a main aspect of expressing the pronoun subject in both EP and BP, while in EP the non-obligation of expressing the subject pronoun can be considered an economy, in BP the reduction in the pronominal system and in the inflectional suffix system can also be considered an economy.

All these phenomena, according to Construction Grammar, are related to one another somehow, and it is the main reason the postulates of this model of grammar is necessary to carry out this study, mainly to describe the relation between the different phenomena that may be associated to the expressiveness of the subject pronoun and also because a construction involves, as stated by Cruse (2004, p.258), six levels that can be analyzed, three levels of form – phonological, morphological and syntactic – and three levels of meaning – semantic, pragmatic and discursive-functional. An alteration at any of these levels can result in new construction.

Goldberg (2006) argues that constructions can be seen at any level of the language, from phonological and morphological aspects to more complex structuring schemes, such as *if-clauses*. Construction Grammar conceives that language structures emerge from the individual's need to interact with the world, thus, more basic needs result in more basic structures that, in a way, are similar, while different experiences foster the need to elaborate different structures.

The structures present in language are organized in a speaker's mind through an intertwined network. The network is formed by schemes which, in turn, have its bases in generalizations processes. The schemes are a representation of a language phenomenon and, generally, represent constructions and network of constructions.

From the perspective of Construction Grammar, a construction, according to Traugott and Trousdale (2013), it is the result of the pairing of form and meaning: the link between these poles is arbitrary to an extent – by all means, it can be implied that, although the link has arbitrary properties, the factors that connect the poles can be motivated. So, it is important to emphasize that such arbitrariness is accompanied by the

influence of the context, whether social or cultural, in which the individual is inserted, and which can circumstance the language.

For an analysis of the parameters of the Construction Grammar, it is necessary to consider all levels of the language, focusing on one at a time, and understand how these levels are, in some way, interconnected. Therefore, studying the prosody of the language, especially regarding the expressiveness of the pronominal subject and its focus, can be a crucial factor to understand these two languages, BP and EP, and how they relate to the subject pronoun.

Construction Grammar allows both a study of inter-levels and a deepening in a specific level that helps to clarify a certain phenomenon of language use. Therefore, the phonological level will be widely investigated through a technological tool: PRAAT; in addition, an analysis of the phonological level is pertinent since the oral modality was chosen as a constituent of the corpora.

The oral modality was chosen for two basic reasons, the first is associated with the fact that speech is a representation of the language fluidity and dynamicity and of its capacity of renewal and innovation; the second is related to a particular objective of analyzing how the prosody happens in BP and in EP when the pronominal subject is expressed.

Hypothesis

Primary, it is a fact that Brazilian Portuguese and European Portuguese are two different varieties of Portuguese that share common antecessors, histories, and linguistic phenomena; however, it is also a fact that there are some contrastive points between them. One of the phenomena that can be perceived as contrastive is the *expressiveness of the subject pronoun*, while in BP there is a higher frequency of the subject pronoun EP has been showing a great number of occurrences as well.

Considering that BP and EP are in different degrees regarding the expressiveness of the pronominal subject, this difference may reflect on the phonological level of the constructional organization because this level integrates a part of the pairing in which the other levels of form and meaning are interrelated.

Frota et al. (2015) analyzed both BP and EP to label the intonational pattern in these linguistic varieties and discovered that, indeed, there a few differences between

them. Taking these assumptions into account, one of our first hypothesis is that the distinctions in the intonational pattern when the subject is expressed are apparent when contrasting BP and EP and may be one of the differences between them.

Also, being BP in a more advanced route in terms of expressing the pronominal subject, we expect that there is a higher frequency of occurrences of it in BP than in EP, which may indicate that the expressiveness of the subject pronoun is already the non-marked structure in BP and may be the future of EP.

Thus, the non-marked form is the expressiveness of the subject pronoun in BP, then, differently from the null subject sentences in which the prosody always falls on the predicate, when the pronominal subject is expressed there is a possibility that the prosody to be marked in elements out of the predicate, including the subject. Hence, we expect that there is a higher number of times in which the prosody happens on the subject in BP than in EP, this change may be explained through semantic and pragmatic motivations.

These considerations are possible when we consider the language as a significant whole, as does the Construction Grammar and therefore when modifying one of the linguistic levels, other levels can also be impacted. In addition, it is possible, as the scholars of Construction Grammar postulate, to describe language phenomena from the simplest patterns to the most complex patterns through constructions, so we have the possibility of representing the phenomenon under study through constructions.

Objectives

The general goal of this research is to analyze the expressiveness of the pronominal subject in BP and in EP and its relation to the phonological level of the constructional organization to verify a possible distinction between these two varieties and how this distinction may be also perceived in other levels of the language organization.

For doing so, smaller goals must be achieved. So, it is necessary to verify the expressiveness of the subject pronoun in both varieties. It is also necessary to investigate which linguistic levels may be linked to this phenomenon; hence our constructional analyzes focusing on the form of the construction – the morphological, phonological, and syntactic levels – and its relation to the meaning production – the semantic, pragmatic, and discursive levels.

It is also part of our goals to analyze in the occurrences in which the subject pronoun is expressed which discourse person is related to the expressiveness, the nature of the subject which is being expressed and how they can be described in the intonational pattern.

Because we consider the language as a holistic process, one that involves cognitive abilities, the investigation of the cognitive process that are involved in linguistic phenomena may come in hand to explain the expressiveness of the subject pronoun.

Methodology approach

To carry out this study, occurrences from two corpora will be analyzed, one from Portugal - representing the EP - and one from Brazil - representing the BP. Both corpora are part of the *Português Falado - Variedades Geográficas e Sociais*, a Project by the *Centro de Linguística da Universidade de Lisboa* (CLUL) data base. This project has been dedicated to studying the different Portuguese around the world with a focus on morphological, lexical, syntactic and discursive processes. Among the corpora available for analysis, those referring to Portuguese from Portugal and Portuguese from Brazil will be used, this contrast was chosen due to the existing history between these two varieties especially in terms of historical-political-social relationship.

As to promote Lisbon project and because part of the thesis was carried in European context, we opted to work only with corpora from this database. Both corpora represent data from the spoken language and date from the 70s to the 90s, thus, it enables us to investigate our object in a *real-time methodology* with a short-term span following Labov's (1994) sociolinguistics patterns.

Most of the data from project, and which will be used, were obtained through interviews in the model of Dialogue between Informant and Documenter (DID). These corpora contain male and female informants between the ages of 20 and 85. The level of education is also variable, there are informants who attended school for four years and there are informants who attended higher education. The corpora are available in audio files and in their transcription, both will be used on our analysis.

The phenomena related to prosody will be analyzed from the data obtained by the computer program PRAAT⁴ – *doing phonetics by computer* in its version 6.1.54. We opted to use this computer program as an auxiliary tool to our investigation because, so far, we couldn't find in Construction Grammar an analysis base that emphasize the phonetical-phonological level. This software was developed by Paul Boersma and David Weenink, from the Institute of Phonetic Sciences, University of Amsterdam, and aims to analyze, among other aspects, the melodic curve, pitch, and the intensity of the speech data.

After carrying out the proposed analyzes, the two varieties of the Portuguese language – BP and EP – will be contrasted to ascertain the hypothesis that there is a distinction between them that result in a different way to express the subject pronoun. Possibly, this distinction reverberates in the prosodic system of the language and reflects the production of different meanings.

Justificative

One of the main the Traditional Grammar arguments in favor of not considering EP and BP as distinct languages involves the non-perception of syntactic difference between these languages, given that the syntactic level is the structure level, a hard level, not very sensitive to radical changes. However, considering that it is already common ground that speakers of these languages already distinguish them when they use it, and likewise when they teach and study it, it is worth promoting an analysis that offers the possibility of making this distinction scientifically, considering the general formatting of the linguistic system and the relationship between the levels of organization in that system.

Therefore, this study intends to investigate whether there is the syntactic difference between Brazilian Portuguese and European Portuguese, considering a fundamental syntactic aspect, namely, the *expressiveness of the subject pronoun*. This phenomenon may indicate a distinction between these two varieties and can corroborate in stating the Brazilian Portuguese has proper lexicon, structures, systems and

⁴ PRAAT – stands for the imperative form of to speak in Dutch. – The program can be downloaded from the following link: <https://www.fon.hum.uva.nl/praat/>

phonological variations and, as such, is not simply a variation of European Portuguese, but it is its very own language.

This phenomenon has been exploited by Generativism under the label *parameter pro-drop* (TARALLO, 1993; DUARTE, 1995; KATO, 1999; DUARTE, KATO, BARBOSA, 2001; KATO, DUARTE, 2014), by the functional perspective (PAREDES DA SILVA, 1988; NEVES; GOULART, 2017; DE ROSA, 2019) by the discursive-functional perspective (GUERRA, 2017; 2021; OLIVEIRA, 2018; OLBERTZ, 2020) and with a constructional approach (GONÇALVES, FARIA, 2021).

These studies have postulated that the pro-drop status of BP has undergone changes and, today, BP can be considered a partial pro-drop language, a language in which there are cases where the referential subject pronoun cannot be omitted.

Generative studies have considered factors internal to the language – the subject's semantic trait, the class of verbs involved in the process of making the subject explicit, among others - while the studies of factors external to the linguistic system – topicalization, perspective, informativeness, the linguistic economy principle, frame semantics, prosody – have only started recently.

De Rosa (2019), for example, argues that the expressiveness of the pronoun subject in BP occurs due to the change that the pronominal paradigm of that language had already undergone. Cyrino, Duarte and Kato (2000) state that 1st and 2nd singular persons in most sentences already appear with expressed referential subject and, as De Rosa (2019) alludes, so does the 3rd person.

Olbertz (2020) also studies the 3rd person pronoun, in her approach, she considered the possibility of BP becoming opaquer, like French, or more transparent. She also investigated the functionality of the subject pronoun when it is expressed, she found out three possibilities – topic, referent, reestablish the referent – plus one – when there is no apparent reason for not using a null subject.

Guerra (2017; 2021) investigates, with a diachronic approach, the linguistic transparency level in PB regarding the pronominal expression of the subject, mainly *você* and *a gente* slowly co-occurring along *tu* and *nós*, respectively, and its reflection into the verb agreement.

Oliveira (2018) also uses the Functional Discourse Grammar to investigate the degree of transparency and opacity to better understand the subject expression in BP. The author proposes that BP has becoming a more transparent system when compared to other Romance languages because there is a necessity, in some contexts, to express the subject

because in BP there has been a tendency of simplification of the verbal agreement system having only one morpheme to express the subject in a one-to-one relation.

Neves and Goulart (2017) explore the relationship between the exclusive inflectional suffix of the pronoun *eu* and the need the speaker feel to make it explicit. The authors' analysis reveals that the speaker feels the need to express the pronoun *eu* even when it is possible to retrieve the pronominal information by the verb inflectional suffix, and this need is even more present when the verbal agreement is not exclusive.

Paredes da Silva (1988) had already established that there are few contexts in which the speakers feel like expressing the subject. One of them, to avoid misunderstanding and assure that the hearers can follow the speakers' trail of thought is when the subject is expressed to avoid ambiguity, it happens when the verbal agreement is not specific to only one person of the discourse and can, therefore, be associated to more than one.

Gonçalves and Faria (2021) analyzes both BP and EP focusing a post-constructionalization change in the pronominal system – after the implementation of *você* and *a gente*, the verbal agreement were also impacted and became a little simpler, with more than one person having the same agreement desinential affixes. Thus, the impossibility of determine by the subject by the verbal agreement, it is required that the subject to be fully expressed by the speaker.

By obtaining new data from the analysis of semantic traits, De Rosa (2019), Neves and Goulart (2017), and Olbertz (2020) make it evident that all levels of the language must be considered when analyzing a linguistic phenomenon. As this is the basis of the studies supported by the Construction Grammar, this will be the theoretical approach used to support this research. In addition to the linguistic levels – phonologic, morphologic, syntactic, semantic, pragmatic, and discursive – it is also possible, in this theoretical perspective, for the researcher to analyze the impact of cognitive phenomena on the language structure.

In this way, since the proposed research considers that all levels of linguistic organization are interconnected, it is possible to analyze, from the constructional perspective, how the different linguistic levels of organization play a specific role that enables the expressiveness of the subject to be an ongoing change, and it innovates when it proposes an analysis of the interface syntactical-prosodical in BP and EP.

For our specific purpose, since the Constructional Grammar allows it, even though we recognize the importance of analyzing all the six levels of a construction, we will

focus on the form of the construction: *expressiveness of the subject pronoun* because it is on the phonological, on the morphological, and on the syntactic level that this phenomenon is more salient and, just latter one, we will consider the meaning pole.

That being said, this research not only contributes to the studies about the *subject via pronouns*, but it also enforces the idea that, although BP and EP share the same ancestors in their formation, they are independent languages, and this affirmation can, in further research, have a teaching potential regarding Brazilian Grammar.

Thesis Organization

This thesis is divided into four parts. On the first part – *The subject in a constructional perspective* – we analyze the concept of subject (chapter 1) and the expressiveness of the subject in a general way (chapter 2); then, we analyze this phenomenon both in Brazilian Portuguese and in European Portuguese, contrasting the points in which one variety differs from the other (chapter 3). We will discuss studies that have analyzed this phenomenon, mostly of them under the Generativist theory, and ponder how the Construction Grammar can contribute to the explanation of this phenomenon.

On the second part – *Theoretical Foundation* – we discuss the theoretical base of this study: our view on the language (chapter 4); the theory of Construction Grammar and its viewing the language as construction with six interrelated different levels – phonologic, morphologic, syntactic, semantic, pragmatic and discursive (chapter 5); the Usage-based linguistic and the importance of considering the use as a key factor to investigate linguistic phenomena as well as some cognitive abilities linked to the linguistic production – iconicity, perspective, informativity, analogy and rich memory, markedness – and linguistic change (chapter 6); the phonological analysis parameters used to describe the prosody, one of the analysis criteria of the thesis (chapter 7).

The third part – *Methodology* – it is where we describe the corpora that we adopted to carry out this research: *Português Falado - Variedades Geográficas e Sociais*, a project by the *Centro de Linguística da Universidade de Lisboa* (CLUL); the tool we used to investigate speech samples: PRAAT – *doing phonetics by computer* and which criteria were considered to reach our goals (chapter 8).

The fourth and last part of the thesis – *Constructional analyses* – is where we discuss the constructional view of our phenomenon (chapter 9), focusing on the form of the expressiveness of the pronoun subject. Then, we propose a constructional representation to this phenomenon (chapter 10). Finally, we propose some considerations about BP and EP in relation to its expressiveness of the subject and how this study can be amplified.

PART I – CONSTRUCTING THE CONCEPT OF SUBJECT

1. The subject

In this initial chapter, before we discuss about how the subject is pronominally expressed, we need to define what we understand as *subject*. To do so, we use the following studies: Casseb-Galvão (2020), Duarte (2018), Lobo (2018), Neves (2018), Raposo (2018) and Raposo e Miguel (2018).

Casseb-Galvão (2020) as well as Raposo (2018) defend that the concept of *subject* is highly dependent on which methodological approach is chosen to work with, even among the linguistics authors, the definition of this category is not unanimous, mainly because the definition of what a *subject* is revolves around the notion of not only syntactical features, but also semantic and pragmatic ones.

Neves (2018), regarding the position of the *subject* in a sentence, says that the most common position it can assume is a pre-predicate position and that it has a syntactical function so closely linked to the verb that it determines the verbal agreement. Therefore, to characterize the *subject*, a few things must be considered, such as the *predication*. Raposo (2018) affirms that *predication* is related to how the world information is linguistically organized to convey a message, this organization is done by the speaker and involves two terms: an entity about which the speaker says something (the subject) and a comment, or judgment, that it is expressed about that entity (the predicate).

However, Casseb-Galvão (2020) argues that this definition alone is related only to the informational level and can be confused as *theme* or as what has been considered a psychological subject.

Also filling this theoretical gap, Raposo (2018) argues that, in Portuguese, there are at least two kinds of subjects: one that makes the agreement relation with the verb (being in number and person) – the *grammatical subject*; and one that introduces the entity about who the persons say something – the *semantical subject*. In most cases, according to the author, both roles refer to the same entity, but it is not always the case.

This distinction, as affirmed by Raposo (2018), is what propitiated in linguistic studies the definition of *topic* and *comment*, being the *topic* generally related to the *semantical subject* and the *comment* related to the *predicate*.

In Raposo's words, *topic* is, thereby, the initial syntagm of a sentence that represents the entity about which anything can be said and that can represent, or not, the *grammatical subject*. And, according to the argument structure, as proposed by Raposo (2018), it corresponds to the external argument, since in the most prototypical clauses, the active ones, the subject being expressed out of the verbal phrase.

About the notion of *topic*, Duarte (2018) utters that even though pre-verbal subjects are generally considered *topics* (non-marked topics when refer to both semantical and grammatical subject; and marked topic when only refers to the semantical one), other properties must be considered to distinguish the *subject* from the *topic* when they happen in a marked context and these distinctions may help us better understand the notion of the subject itself.

Duarte (2018) proposes that, from the **morphosyntactic** point of view, the subject obligatorily triggers the verbal agreement process, while the marked topic does not; regarding **word order**, and considering only simple clauses, while the subject can occupy several internal positions in the sentence, the marked topic can only occupy peripheral positions to the left or to the right of the verb, normally separated by comma; regarding the **argument structure**, except for the cases in which the subject is an expletive pronoun, the subject is always semantically selected by the verb, that is, it is always an argument of the verb, but the marked topic may or may not be associated with arguments of the verb.

Lobo (2018) adds that *subjects* can be classified according to its semantical properties that depends on the semantic value assigned by the predicator – the verb – and by having or not a specific reference, hence, the possibility of classifying the *subjects* into three kinds: *argument subject*, *non-argument subject* and “almost” *argument subject*.

The *argument subjects*, as proposed by Lobo (2018), correspond to the grammatical subjects that have a semantic value, such as *agent*, *theme*, *experiencer etc.*; *non-argument subjects* correspond to grammatical subjects that has no semantic value since their function are purely grammatical, such as impersonal verbs, expletive subjects or with non-referential value, context in which the traditional grammar has regarded the inexistence of a subject; the almost argument subjects are the ones whose referential content is minimal – *subjects* of meteorological verbs, verbs that designate atmospheric phenomena.

Lobo (2018) argues that in Portuguese all these subjects can happen in a null context and just the *argument subject* can, generally, alternate with a pronominal form phonetically expressed.

Lobo (2018) also postulates that the argument subject can be divided into two categories: subjects that have a defined reference, designate a specific entity, and subjects that are arbitrary, which means that they designate a non-specified, indeterminate or generic entity. Traditionally, the 3rd plural person, when not expressed, represent an arbitrary subject, however that is not always the case.

A *noun phrase* (NP) has referential value, according to Raposo and Miguel (2018), when it is used to describe a particular entity or a particular group of entities in the discourse universe, identified perceptually or cognitively by the speaker; although, when the NP has a generic value, or the entity cannot be identified in an immediate context, it can be said to be non-referential.

All things considered, it is possible to say that define what a *subject* is, indeed, more complex than the tradition, in general, has done. Using Casseb-Galvão's words (2020), a *subject* has properties that belong to different linguistic levels, *argument* and *agent* belong to the level of the predicate; the concept of *subject*, to the predication; *theme* belongs to the informational level; and *topic* belongs to pragmatics. All these functions can be profiled by a unique item, and it can be represented by same referent, or not.

Thus, for this work, we will consider only the argument subjects once it can be alternated with a fully expressed subject pronoun and consider the subject having the function⁵ of *referential pronoun*; *topic pronoun*; *reactivation of topic*; and the cases in which the *pronouns have no apparent motivation*.

2. Subject pronominally expressed

This chapter is dedicated to approach the context in which the subject is pronominally expressed, the contexts are presented by Neves (2018; 2011) and Raposo and Miguel (2018).

⁵ These four categories were proposed by Olbertz (2020) while contrasting Portuguese, Spanish and French and their cases of expressing the subject via a pronoun. Olbertz's proposal is fully described in the section 3.1.

Raposo and Miguel (2018) and Neves (2018) defend that the personal pronouns have, by default, as the NP that they can replace, referential capacity, it means that its basic function is to refer to a discourse person.

According to Neves (2018), discourse persons are entities referred in an interactional situation, the interaction can happen directly – who speaks and who is spoken to – or indirectly – who / what is being talked about. The term *person* cannot be identified as a human being, it is a grammatical indication: discourse persons are therefore grammatical persons.

Neves (2018) states that there are three discourse persons as it follows:

The 1st person – which is the person who refers to itself – refers to speaker itself, the speaker uses the verb in the 1st person – when dealing only with itself, the speaker uses the 1st person singular pronoun (*eu*), and when dealing with other people along it, the speaker uses the 1st person plural pronoun (*nós*), with its formal variations, according to its function.

The 2nd person – which is the person who is being spoken to – refer to its interlocutor, the speaker uses the verb in the 2nd person – when dealing with a single interlocutor, its use the 2nd person singular pronoun (*tu/você*), and when dealing with other people besides the interlocutor, it uses the 2nd person plural pronoun (*vós*, very rarely, *vocês*), with its formal variations, confirm its function.

The pronouns *você* and its plural form *vocês* are derived from *a treatment pronoun* (*vossa mercê*), have the function of 2nd discourse person but their verbal agreement, as occurs to *treatment pronouns*, is related to the 3rd person.

The 3rd person – which is the person that is being talked about – refers to which or whom is being talked about, it uses the verb in the 3rd person: when it is a single element, it uses the 3rd person singular pronoun (*ele/ela*), and, when it is more than one element (and the speaker decided to use a pronoun), it uses the 3rd person plural pronoun (*eles/elas*).

Summarizing, Neves (2018) proposes the following table:

	Singular	Plural
1 st person	Eu	Nós
2 nd person	Tu, Você	Vós, Vocês
3 rd person	Ele, Ela	Eles, Elas

Table 1: Portuguese Subject Pronouns Paradigm (NEVES, 2018 p.471)

Even with its referential capacity, there are times when the personal pronouns can be used with a generic reference. Neves (2011) discussed different contexts in which different discourse persons can be used with a generic referent. The pronoun *você*, for example, even though it is a person involved in the discourse, can have a generic reference – it can be used to refer to any person and, when it is the situation, it acquires a high level of indeterminacy. At the same level, the pronoun *eu*, which is by default full of referential content, can also be used with a generic referent.

All in all, these are the pronouns that are traditionally used as subjects alternating with NP. However, it is necessary to make some considerations regarding the necessity of expressing them or not.

3. Expressiveness of the subject

In this third chapter, we analyze the possibility of expressing or omitting the subject and how both BP's and EP's speakers are dealing with this double-possibility. We also consider the degree of linguistic transparency definition and how it affects the necessity of expressing the subject. The possibility of expressing/omitting the subject is based on the postulates of Tarallo (1993), Chomsky (1981), Roberts (1993), Rizzi (1998) and Neves (2018, 2011). BP and EP varieties were analyzed with the aid of Barbosa, Duarte e Kato (2005), Cyrino, Duarte and Kato (2000), Duarte (1993, 1995, 2000, 2004, 2008), Galves (1993), Kaiser (2009) and Holmberg, Nayadu e Sheehan (2009). Hengeveld and Leufkens (2018) and Leufkens (2013) were used to better understand the degree of transparency and Olbertz (2020) provided our main analyses categories.

According to Neves (2018), there are few contexts in which there is no necessity to fully express the subject of a clause because the subject information can be recovered and identified by the verbal inflectional suffixes. This is what we call *subject ellipse* or *zero subject* and what Grammar Tradition has been called *null subject*.

The expressiveness or not of the subject by a NP or by a pronoun is a choice made by the speakers. Neves (2018) defends that this choice is not arbitrary, the speakers make it accordingly to the kind of information that is intended to be communicated since the omission of any of the clause constituent can only happen if the information that is being omitted can be somehow recovered.

This is, according to the author, pragmatical choices of the syntactical construction and are dependent on how the speakers organize the information. As proposed by Neves (2018), for example, in transitive constructions with verbal complements, it is commoner that clause topics are kept through more than one clause sequence and, therefore, the expressiveness of the subject is not necessarily required while the object, a new information, is.

Other case that can enable the omission of the subject is when the verb has a specific verbal agreement suffix and, a specific case, one of the reasons for not expressing the 3rd person plural pronoun, *eles*, is its generic function and the same can be said about the 1st person plural pronoun *nós* (NEVES, 2011).

This is the starting point where there seems to be some contrasting between the Brazilian and the European varieties of Portuguese. Tarallo (1993) attested that there is an asymmetry between syntactic structures of both varieties regarding the pronominal paradigm and its omission and explicitation of the subject and the object.

The author verified that while BP has been favoriting the explicitation of the pronoun in the slot of the subject and the omission of the pronoun in the slot of the object, mainly the clitics, the EP does exactly the opposite, it has a tendency of omitting the subject pronoun and has kept the clitics. To exemplify these phenomena, Tarallo (1993) assumes that to answer the question *Paulo viu Maria ontem?*, we would have two different replies – *Sim, ele viu*, for BP – and – *Sim, a viu*, for EP.

Languages that are inflectionally rich can normally be considered *pro-drop languages*, that is, they are languages that can recover the subject information through the verbal inflectional suffix, through the agreement between the subject and the verb. Chomsky (1981) precisely used the Principles and Parameters theory (PP) to start separating the languages, those that had the ability to represent a world event through a null subject were called pro-drops languages, while languages that could not represent an event without the subject's slot filled were considered non-pro-drop languages.

In this same direction, Roberts (1993) argues that languages that have a rich inflectional paradigm, such as Portuguese, can establish the principle “*avoid pronoun*”,

while languages that have a “simplified” inflectional paradigm needs to make the referential pronoun explicit.

This principle, according to Chomsky (1981), can be seen as a subcase of the conversational principle – it means that the speaker won’t say more than necessary to its communicational purpose; or it can be seen as an extinction of the exclusion principle, or even relate to the recoverability principle, all in all, it is a principle of grammar.

Rizzi (1998) advocates that this principle means that the speaker will make explicit only the necessary linguistic material to convey its idea, to make evident the meaning it is intended to convey, respecting the restrictions of the universal grammar. Thus, given the opportunity to use a null subject, the speaker will make it explicit only when it is the foreground, the topic, or when it is in contrastive position.

This principle is exactly how we can define the null subject parameter, hence, the languages with pro-drop system. However, what would happen if a language started to go through a process of simplification in its pronominal paradigm? If, by the principle of linguistic economy, forms were simplified, and agreement could no longer recover all pronominal paradigm? If through a process of analogy, the verb forms started to be used in a syncretic way?

This is exactly what begins to happen in BP, the principle *avoid the pronoun* of Chomsky begins to stop being a reality and the full pro-drop parameter of Brazilian Portuguese is compromised. We attend these questions in the next topic.

3.1 Expressiveness of the subject in Brazilian Portuguese

According to Duarte (1993), there is the reduction of forms in the inflectional paradigm of BP caused by the loss (in almost all regions of the country) of the pronominal forms *tu* and *vós*, replaced by *você(s)* and *o(s) senhor(es)*.

The use of 3rd person forms (singular and plural) as the only possibilities of reference for the 2nd person makes BP lose its regularity in the option for the null subject. The simplification of BP’s inflectional system is aggravated when the expression *a gente* co-exist with the pronoun *nós*, which leads us to a paradigm with only three distinctive forms which, according to Galves (1993), also helped simplify the verbal inflectional system in BP.

Given these changes, Duarte (1993, 1995) draws attention to the simplification of the inflectional paradigm of BP as it can be seen on the table below:

Person	Number	Paradigm 1	Paradigm 2	Paradigm 3
1 st	singular	cant-o	cant-o	cant-o
2 nd direct	singular	canta-s	-----	-----
2 nd indirect	singular	canta-0	canta-0	canta-0
3 rd	singular	canta-0	canta-0	canta-0
1 st	plural	canta-mos	canta-mos	canta-0
2 nd direct	plural	canta-is	-----	-----
2 nd indirect	plural	canta-m	canta-m	canta-m
3 rd	plural	canta-m	canta-m	canta-m

Table 2: Evolution in Portuguese inflexional paradigm (Source: Duarte, 1993, p.109)

It is possible to verify in Duarte (1993) that first (paradigm 1) there were six discourse persons and six inflectional suffixes with two syncretical forms (2nd direct and indirect persons, with the indirect using the inflectional suffix of the 3rd persons); then, (paradigm 2), there were only four discourse persons, losing the 2nd direct persons, both plural and singular, and their inflectional suffixes; lastly, (paradigm3), with the insertion of the pronoun *a gente* as 1st personal plural pronoun (with the suffix used by 2nd and 3rd singular pronouns) co-existing with the form *nós*, the inflectional suffixes were reduced to only 3. This last two paradigms co-exist and are not excluding.

In her research, Duarte (1993) concluded that the last texts she analyzed, produced in 1992, in comparing to the ones that were first analyzed, dating back to 1845, showed a great decline in the use of null subject in BP.

Even in languages that are considered full pro-drop, there is a special rule regarding the 3rd person, when the subject is known, expect or can be automatically recovered, the use of the null subject is preferable, but if the subject is not known, when it is used to introduce a new topic, the rule is that the subject pronoun must be used. This restriction, however, is not feasible in BP. Duarte (1993) verified that even when the topic is known, there's a preference of using the subject pronoun.

It starts to show that BP is changing its null subjects, hence, it is ceasing to be a full pro-drop language, however, as noted by Duarte (1993) there is still a great number

of occurrences of null subjects, mainly for the 1st person, perhaps it is due to the fact there is no syncretism in the flecional system for *eu* for most of the verbal tenses. Neves and Goulart (2017, p.339), however, noted that for the imperfect tense of the indicative, there are occurrences times where only one verbal form is used for all singular pronouns *eu*, *tu*, *voce* and *ele/ela* – like in:

‘eu’, ‘tu’, ‘ele’ / ‘você cantava’

Due to this change in the pronominal paradigm, a lot of research arose to verify the pro-drop system in PB and the cases in there were motivation to make the subject pronoun explicit.

De Rosa (2019) focused on the 3rd person subject in the movie speech in BP, observing the semantic traits [\pm human], [\pm animated], [\pm specified] and has attested that there are cases in which the subject, even when it is not referential, is explicitated. His research followed the pattern proposed by Barbosa, Duarte e Kato (2005):

1. Sentential Pattern A: the antecedent is in the same period and is subject to the preceding clause
2. Sentential Pattern B: the antecedent is in the preceding period and is a structural/discursive topic or has the syntactic function of a subject.
3. Sentential Pattern C: the antecedent is in the preceding period with another syntactic function.
4. Sentential Pattern D: the antecedent is also subject, but there is at least one intervening clause between the clause containing the pronoun under analysis and the clause in which its antecedent is located.
5. Sentential Pattern E: the subject's antecedent is found in a non-adjacent clause, as in Pattern D, with another syntactic function.

Barbosa, Duarte e Kato (2005) concluded that the sequential patterns A and B, the subject is half the time explicit, while in the patterns C, D and E, the percentage for the explicitation is a little higher, being E almost 70%.

Based on these assertions, De Rosa (2019) states that there a new non-marked form arose from the changes in the pro-drop system, the basic order of the sentence is

now SV(O) – (Subject-Verb-(Object)) with the object not always available, which validates what Tarallo (1993) had previously confirmed.

Kato and Duarte (2014) established that BP has share a few characteristics with French, meaning that when the subject is explicit it tends to be a weak pronoun⁶; but different from French, the null subject can still occur when it is possible to recover the information via an anaphora, a characteristic seem in languages like Chinese.

Kato and Duarte (2014) confirmed what had been proposed by Cyrino, Duarte and Kato (2000), although, PB still lacks a uniform typology when the null subject can or cannot happen, it already has standardization, the more referential a subject is, the more the possibility of it being explicit, being less referential the subjects that lack the feature [+human]. Thus, as stated by Holmberg, Nayadu e Sheehan (2009), PB can be classified as a partial null subject language like Finish and Marathi, even without a typological uniformity.

It is also important to address that Kato and Duarte (2014) observed that due the changes in the necessity of making the subject pronoun explicit, sentences that previously followed the null subject pattern have appeared with more frequency like with verbs that express natural phenomena. The authors use the following examples to indicate it:

a. \emptyset_{expl} **chove** muito nessas florestas.

b. *Essas florestas chovem muito.*

Usually, Portuguese language speakers uses an expletive structure of null subject with impersonal verbs, i.e., to describe natural phenomena (a); however, a variation starts occurring with the speaker expressing an element as the subject in the same context of this kind of impersonal verb (b).

Neves and Goulart (2017) studied the 1st person of singular, *eu*, and its exclusive inflectional suffix in PB. Even having an exclusive inflectional suffix, the data analyzed by the authors showed that 77,7% (1321 occurrences) happened with an explicit pronoun subject and only 22,3% (380 occurrences) happened with a null subject, being 75,6% of the occurrences with exclusive inflectional suffix. It means that even when there is no necessity of explicating the subject pronoun, in BP, it still happens, factor that reinforces

⁶ A *weak pronoun* is, as proposed by Cardinaletti and Starke (1999), a pronoun that can, prototypically refer to non-human entities.

that the BP is losing is null subject pattern and the explicitation of the subject is, basically, a “rule” and not any longer a “mistake”.

Another factor linked to subject expression that has altered BP is the degree of linguistic transparency – a phenomenon that was studied by Olbertz (2020) whose research analyzed Brazilian Portuguese, Spanish, and French for comparative purposes, focusing the analysis on the expressiveness of the pronoun *ele/ela* in BP and changes in the degree of transparency of this variety.

According to Functional Discourse Grammar, linguistic transparency can be defined, normally, as a one-to-one relationship between the parts of form and those of meaning, however, it is necessary that this definition occurs, as studies functionalists have proposed, considering different degrees that make up any linguistic phenomenon. Thus, it is not possible to define that a language is either transparent or opaque. In addition, it still must be considered that to define the degree of transparency that applies to language, it is also necessary to analyze all the levels that make up linguistic acts - interpersonal, representational, morphosyntactic and phonological (GUERRA, 2017).

From these postulates, it is possible to credit that a language is totally transparent when the four levels of the Grammatical Component have a unique representation in the linguistic universe – thus, a one-to-one relationship – and, as alluded by Guerra (2017), it can be understood as one-to-one-to-one-to-one.

According to Hengeveld and Leufkens (2018), one of the phenomena that violates the principle of transparency, especially between the interpersonal and representational levels, is *apposition*. Authors claim that a fully transparent mapping between these two levels occurs when a single act at the interpersonal level corresponds to a single representational category at the representational level. When the same entity has two or more references in the representational plane, the language starts to become opaque.

Hengeveld and Leufkens (2018) exemplify this phenomenon through the example *O irmão de João, Peter, se mudou para Noruega*⁷, in this fabricated sentence, it is possible to see that *O irmão de João* and *Peter* refer to the same entity. It means, following Functional Discourse Grammar guidelines, that two referential (R) subacts at the interpersonal level correspond to a single individual at the representational (X) level.

Very similar to apposition, the cross-reference (CR) process also reveals a type of transparency violation, however, when the CR happens, one of the references is

⁷ *John's brother, Peter, has moved to Norway*

performed by a lexical form, normally, through an optional noun phrase, and the other reference is through an affix of the verb.

With this perspective of transparency and opacity, Olbertz (2020) realized that there are three main reasons why the speaker feels the need to fully express the subject through a pronoun: 1) when there is a change in the referring subject (being of semantic use because it is related to the descriptive content and not to the referential act); 2) when the referent is the topic of the sentence (being a pragmatic use since the referring subject is related to topicalization and also a discursive and functional use, since the emphatic process is at the interpersonal level); 3) when there is reactivation of the referring subject, making it topic again.

These findings reinforce the idea of the constant loss of the null subject in BP, although, it is the uses that apparently have no functional purpose - the subject that is explained simply by the lack of a complex pronoun paradigm and verbal inflectional suffixes, and the 3rd person of the pronoun *ele/ela* used with referents that do not have the [+human] [+animated] feature – which shows a further change of the paradigm.

Olbertz (2020) goes further in her hypothesis and consider the possibility of PB becoming a fully transparent language, with one nominal referent.

Reference	Pronoun	Agreement	Verbal Form
1 st person of singular	<i>Eu</i>	1 _{SG}	<i>falo</i>
-	-	2 _{SG}	-
2 nd person of singular	Você o senhor/a senhora	3 _{SG}	<i>fala</i>
3 rd person of singular	Ele/ela	3 _{SG}	<i>fala</i>
1 st person of plural	A gente	1 _{PL}	<u>fala</u>
-	-	1 _{PL}	-
2 nd person of plural	Vocês	3 _{PL}	<u>fala</u>
3 rd person of plural	eles/elas	3 _{PL}	<u>fala</u>

Table 3: PB Hypothetical Paradigm (Source: OLBERTZ; 2020, p.41)

As it can be perceived, the underlined forms represent the generalization of forms that are accidentally used, as it happens in her examples:

- a. *Eles tava...* ali perto da/sabe?... ali no plantão do América lá
 b. *Eles não pode* invadir a área...

The 3rd person of plural is already accidentally used with the inflectional suffix of the 3rd person of singular; *a gente* as the 1st person of plural may have a possibility of becoming the standard pronoun, replacing *nós*; the 2nd person of singular, *você*, deriving from a *treatment pronoun*, is already used with the suffix of the 3rd person of singular; *vocês*, as a plural form of *você*, following the same pattern, could also be potentially used with the verbal agreement of the 3rd person of singular.

Hence, the hypotheses of the author that if it weren't for the specific inflectional form for the 1st person of singular, *eu*, which is very frequent in BP, as it can be confirmed by Neves and Goulart (2017), and has a very pragmatical and informational function, BP could be in the process of becoming a language with a simpler pronominal paradigm.

Thus, it reinforces the idea that PB is no longer a full pro-drop language, because, as defined by Olbertz (2020, p.43): the subject pronoun is frequently used; the verbal paradigm, already reduced, is becoming even more reduced; the subject pronoun of 3rd person of singular is losing its semantic impact, being used with less referentiality; and it is becoming pragmatically immediate in its use as apposition as a topic mark. But BP is not an entirely non-drop language as well, because it still retains the cases where there is a possibility of the subject pronoun being omitted.

3.2 Expressiveness of the subject in European Portuguese

As it has been perceived by some authors, Barbosa, Duarte and Kato (2005), Holmberg, Nayadu and Sheehan (2009), Duarte (1993, 1995, 2000, 2004, 2008), the main difference between the pronominal paradigm of BP and EP is the further degree of simplification of PB paradigm while the EP resist going through the same process and maintaining a full and functional pronominal paradigm and verbal agreement suffix.

Kaiser (2009) defends that EP has the following pronominal paradigm:

Person	Old Portuguese				Modern European Portuguese				Modern Brazilian Portuguese			
		<i>cant</i>				<i>cant</i>				<i>cant</i>		
1sg	<i>eu</i>			<i>o</i>	<i>Eu</i>			<i>o</i>	<i>Eu</i>			<i>o</i>
2sg	<i>tu</i>		<i>a</i>	<i>s</i>	<i>Tu</i>		<i>a</i>	<i>s</i>	<i>você</i>		<i>a</i>	-
3sg	<i>el(l)e, el(l)a</i>		<i>a</i>	-	<i>ele, ela</i>		<i>a</i>	-	<i>ele, ela</i>		<i>a</i>	-
1pl	<i>nos, nós</i>	<i>cant</i>	<i>a</i>	<i>mos</i>	<i>nós</i>	<i>cant</i>	<i>a</i>	<i>mos</i>	<i>nós</i>		<i>a</i>	<i>mos</i>
									<i>a gente</i>		<i>cant</i>	<i>a</i>
2pl	<i>vos, vós</i>		<i>a</i>	<i>des</i>	<i>vocês</i>		<i>a</i>	<i>m</i>	<i>vocês</i>		<i>a</i>	<i>m</i>
3pl	<i>el(l)es,</i>	<i>cant</i>	<i>a</i>	<i>m</i>	<i>eles,</i>	<i>cant</i>	<i>a</i>	<i>m</i>	<i>eles,</i>	<i>cant</i>	<i>a</i>	<i>m</i>
	<i>el(l)as</i>				<i>elas</i>				<i>elas</i>			

Table 4: Present indicative inflection in Portuguese (Source: KAISER, 2009, p. 139)

This paradigm already shows some changes in EP, one syncretical form for 2pl and 3pl, that could lead, down the road, to the loss of the null system and, this way, the status of full pro-drop language.

However, the Chomsky principle of *avoid the pronoun* is still functional for EP, according to Holmberg, Nayadu e Sheehan (2009), the pronoun subject is made explicit only if it has a topic or contrastive function.

Rubio (2012) analyzed both BP and EP using a sociolinguistic approach and verified that there is an alternance between the uses of *a gente* and *nós* as a 1st person of plural, even though BP has a higher frequency of *a gente*, it is already possible to see some uses of *a gente* in EP. The verbal agreement with *a gente* and *nós* varies as well, being fewer the cases in which the pronoun *a gente* is used with verbal suffixes other than the ones of 3rd person of singular in BP, and, in EP, almost a fourth of the occurrences analyzed by the author presented *a gente* with verbal agreement of 1st person of plural.

Rubio's results showed that although EP is still somewhat more resistant to the simplification of the pronominal paradigm than the Brazilian variety, the process of simplification has already started.

The author also verified the process of making the pronouns *a gente* / *nós* explicit in some contexts: subjects that are explicit in their own actions; subjects that are not explicit, have their information retrieved from the verbal agreement suffix or in the immediate context; subjects that are neutralized. Rubio (2012) found out that there is a higher tendency of explicitating the subject pronoun *a gente* than *nós*, mainly due to the

specific verbal inflectional suffix to the 1st person of plural *nós*, while *a gente* uses the verbal inflectional suffix of 3rd singular person.

Soares da Silva (2006) analyzed the changes EP regarding the subject and conclude that there is a percentage of use of explicit subject pronouns that present the feature [-animated], the percentage correspond to 6% of the data analyzed; the subject pronoun is also present in the specific context of relative sentences in which the explicit referential subject pronoun is favorited for the 1st and the 3rd person both of singular and plural forms.

Carvalho (2009) contrasted both BP and EP to verify the pro-drop parameter in both languages and concluded that EP can no longer be considered a prototypical null subject language, that it has frequent *pro-drop* features, but in some circumstances the subject pronoun must be explicit to avoid confusion in the informational level.

Both varieties, according to Carvalho (2009), share some similarities when, for example, the subject is [+animated] and when there is a correferenciality in the subject; the most contrastive point, though, comes from the verbal agreement, EP has a strong relation with the agreement, and it may have influenced the capacity of retrieving the subject information directly from the inflectional system and helped EP to keep the null subject parameter.

Regarding the transparency degree, Leufkens (2013), when studying European Portuguese, determined that this variety of the Portuguese language is an example of a cross-referencing language. To exemplify this assertion, the author uses the sentence *eu chegue-i* (eu cheguei), in this sentence, there is a reference to the same entity twice, through the pronoun *eu* and the verbal suffix *-i*. Thus, it is possible, in a sentence like this, that only the use of the verb with its suffix that refers to the 1st singular composes a complete predication, and the use of the pronoun, in this sentence, would make it redundant.

Summarizing, due to the changings that are already happening in in EP, perhaps it would not be prudent to call it “a consistent pro-drop language” as Holmberg, Nayadu e Sheehan (2009, p.6) suggests. It is not the case right now, but EP can also, hypothetically, goes through the same process of simplification in its pronominal paradigm as BP.

Therefore, being two varieties that show high frequency of subject pronouns being expressed, it leaves in charge of verifying the different kinds of subjects that are being fully expressed, as Olbertz (2020) did; the possible motivations for this expressiveness;

and, finally, check how the prosody pattern can help us to differentiate both varieties. Thus, the next section covers the theoretical base used to accomplish the forementioned goals.

PART II – THEORETICAL FOUNDATION

4. A cognitive and functional view of language

The first chapter of this section aims to introduce the language concept we chose as foundation of our research and analysis. In a cognitive and functional view to language, the grammar that enables the language has a cognitive base in a way that when one uses the language, one mobilizes a whole set of cognitive processes at the same time. In this view, the language is holistic, and the grammar is usage-based.

As it views the language as a whole set, it is possible to analyze all the linguistic levels, from the form to the meaning, and how one of the levels affects the others. Hence, it may help us to understand how the change in a linguistic aspect of the language, such as the prosody, can reveal further linguistic changes, such as its syntactic organization. To define the concept of language and grammar, we used Bybee (2010), Filippi (2020), Goldberg (1995, 2006), Neves (1997) and Tomasello (2008, 2009).

4.1 Language Conception

The communicational need to get involved and to involve the other in a communicative event is part of the human being's constitution and, as a general cognitive feature, it is present in interactional relationships. Tomasello (2008) argues that human beings, even in preverbal times, used gestures to express their intentions and intervene in the actions of their peers, whereas Filippi (2020) enforces that the intonation system is an important factor of communication since preverbal times.

In both cases, to learn how to convey an idea to other human beings, it is necessary to learn how to be linguistically social, for the intonation part, in special, as Filippi (2020, p.408) proposes, it is important to understand the “(1) producing and identifying phonemes; (2) processing and learning compositional rules in vocal utterances; (3) associating unfamiliar spoken words with their meaning”, to this research in specific, the second process is crucial because it is directly linked to the prosody phenomena.

In fact, the more complex the event in which the individual is inserted, the greater will be the complexity of code used. Thus, due to the need to engage the other in events that are even more complex than showing a predilection for an object and the biological

predisposition of human beings to develop the phonatory system, language becomes associated with gestures, since only gestures they could not, for example, argue to convince. Thus, it is possible to infer, from Tomasello (2008; 2009), that it is in the interaction with others and in the need for cooperation that the basis for communication and for the development of linguistic systems is found.

To describe the language and its system, Bybee (2010) uses the metaphor of sand dunes. For the author, just like the dunes, the languages show regularities in form and structure, but they also present variations, gradients, and the ability to change.

Bybee (2010) determines that among the similarities they show toward each other is that they all use domain general abilities and domain specific abilities. The general abilities are the ones we use outside linguistic contexts, such as the categorization ability, while the specific abilities are specific to language production, such as the ability to turn sounds into phonemes and, consequently, into speech.

These abilities are used to act in the world and interact with it via language, which has the capability to be transformed and organized to better serve its purpose: communication. It can be seen as such, as a complex and adaptative system.

According to Neves (1997) the grammar a language that is in constant change is to also be ready to accommodate the change, being emergent itself and being in constant variations due to the discursive necessities. The syntax of the language, in this concept, is accessed by semantic via pragmatics.

Grammar is, by default, a system of organization of the languages and, even though every language has basic concepts that are universals, such as the concept of *subject* and *verbs*, the way the world information is registered and uttered is subjective and dependent on linguistic and extralinguistic experiences. Neves (2002) postulates that is the use that enables the grammar and is organized by the grammar. Following the same lines, Oliveira and Votre (2009) defend that there is an intimate relation between how the world is perceived and conceptualized and how it is linguistically expressed, and the languages are shaped.

As Goldberg (1995, 2006) had stated, Bybee (2010) says that, to understand the process of language formation, it is necessary to go beyond surface forms and observe the circumstances involved in the formation of the analyzed patterns.

Bybee (2010) still recognizes that, although languages differ, they share some basic principles in their formation. The author utters that, even if the statements are different, there are recognizable remnants of a similar structure. These principles make

us think that there are factors that shape the language and go beyond the linguistic structure, to which Bybee (2010) calls general domain processes.

Recognizing these processes as a forming part of language is not denying the importance of linguistic structure in the process of language change and formation, but rather admitting that there are other factors and forces, shared by speakers of all languages, that also operate in the language system.

One of these forces operating in the formation of language resides in the individual's need not only to communicate a world event, but also to involve another individual in one of these events; even a single gesture, a basic interaction social cognitive skill, such as pointing, or the intonation in which a sentence is uttered, is complex enough to convey more than one message and which, to be understood, depends on understanding its production context.

This notion that a single gesture can have different meanings, depending on the context in which it is used, supports the notion that a word can have different meanings, to be understood in relation to the context in which it is used, a prerogative adopted by Usage-Based Linguistics.

5. Construction Grammar

This chapter is dedicated to the theory known as Construction Grammar, a current approach that portrays the cognitive processing capacity that allows the individual to retain local information and, based on them, formulate generalizations about the functioning of language. The authors that aided this chapter are Barros (2016), Croft and Cruse (2004), Furtado da Cunha et al (2013), Goldberg (1995, 2006, 2019), Langacker (1987), Martelotta (2011), Traugott (2015) and Traugott and Trousdale (2013).

As defined by Goldberg (2006), constructions can be used to analyze the language at all levels, from its phonological and morphological aspects to its complex structuring schemes.

In addition, the constructionist perspective also considers that the grammar of a language emerges from the individual's communicative need and the use he gives to the language, and, when faced with new instances, through various cognitive processes, such as analogy, categorization/generalization, frequency, rich memory, and

conventionalization, we have new possibilities to represent these instances based on already existing forms.

To define the assumptions of the Construction Grammar, we used the theory defended by Goldberg (2006; 2019) and by Traugott and Trousdale (2013) for the representation of a constitution.

5.1 Fundamentals of Construction Grammar

The Construction Grammar is allied to the pertinent studies of Usage-based Linguistics. Therefore, it also relies on both the dogmas of Functional Linguistics and those of Cognitive Linguistics. Studies that consider language grammar as emerging from usage have adopted this perspective in their analyses.

Goldberg (2006) recognizes a grammatical universalism, but does not add it to biological factors, but to the individual cognitive processing that is developed ontologically and philologically, from the observation of the inputs with which it has contact. need to express the same type of message, but how this happens can vary. Therefore, a constructionist approach defends and investigates this viability of forms.

Furthermore, Goldberg (2006) assures that language has more general constructions, but it also has semi-idiosyncratic patterns and, therefore, not all of them can be considered universal. Those that are regular tend to occur more often and are therefore easier to learn.

Language, for the constructionist theory, as advocated by Langacker (1987), is constituted by a network of nodes associated by their similar characteristics. In this sense, the Construction Grammar proposal conceives that all levels of language (phonologic, morphologic, syntactic, semantic, pragmatic, and discursive) are somehow interconnected. That is why, to analyze a language phenomenon, it is necessary to observe the system at its every component levels.

This is one of the most relevant reasons to broaden the scope over the *expressiveness of the subject pronoun* phenomenon in BP to a constructional extent. Being it possible to analyze it from the form to the meaning levels, it will be possible to understand not only the motivations to use null or explicit subject, but also how these motivations can impact the language in different ways.

Furthermore, as argued by Martelotta (2011), the constructionist proposal argues that there is no rigid distinction between lexicon and grammar. The notion adopted in this theory is that of gradient, of *continuum* and that schematic patterns organize usage, just as usage rearranges schematic patterns, an assumption that will be discussed in the next section.

Another principle that governs the Grammar of Constructions is to refute the autonomy of syntax. In this sense, Furtado da Cunha et al (2013) assert that syntax is in favor of use and is used to structure the language. Usage, in turn, is influenced by factors external to the language, which can alter the syntax. Thus, what exists is a cycling process: while syntax organizes use, the speaker's use of language organizes syntax.

According to Goldberg (2006), it is possible to conceive any linguistic pattern as a construction. For this to occur, some aspect of its form or function cannot be completely predictable based on the analysis of isolated parts or even the constructions that already permeate the language. Furthermore, even the structuring patterns of a language sentence are likely to be analyzed in a constructional perspective, since it is the semantics and/or syntactic information specified by the verb that determines the form and interpretation of the basic patterns.

A construction can be defined, from the perspective of Traugott and Trousdale (2013), as the pairing of form and meaning: the link between these two instances is arbitrary and resumes the discussion proposed by Saussure about the arbitrariness of the sign. It is important to emphasize, however, that such arbitrariness is accompanied by the influence of the social and cultural space that circumscribe the language.

As reported by Barros (2016), the construction represents a schematic pattern that guides the use and is fed by it. The constructions are cognitive processes stored in the mind and, according to the speaker's communicative need, they are accessed for the formulation of utterances. Thus, it is possible to infer, as the author does, that meanings are constructed at the time of use from the fusion between form and meaning.

Croft and Cruse (2004, p. 258) present a symbolic structure for the anatomy of a construction:

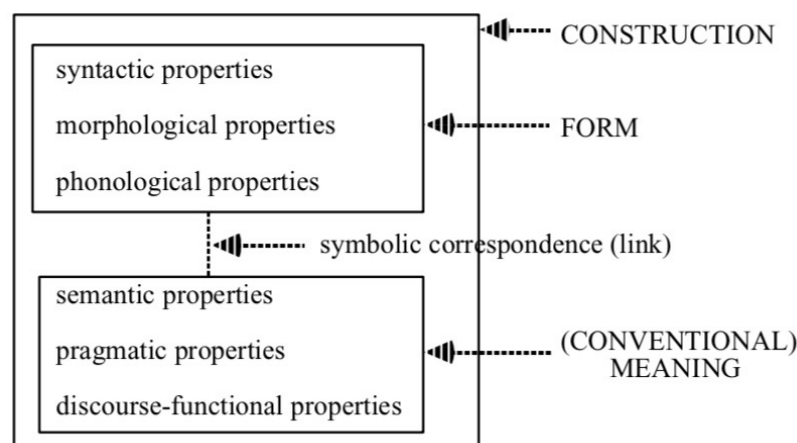


Figure 1 – Construction Anatomy (Source: CROFT; CRUSE 2004, p. 258)

The authors define that the pair of form and meaning is at least partially arbitrary. Furthermore, they emphasize that, through the symbolic link of correspondence, the properties of form and conventional meaning are closely linked and have a direct relationship in the materialization of language, that is, in use. Croft and Cruse (2004) understand conventional meaning as the representation of all aspects of the construction function, as well as the properties of the discourse in which the utterance is found.

Goldberg (2006) states that constructions are learned based on input, world experiences, and general cognition mechanisms, such as analogy. In addition, the author emphasizes the importance of recognizing knowledge of specific items that exist alongside generalizations.

By absorbing knowledge of specific constructions, it is possible to abstract a schematic pattern, a generalization, which occurs, for example, when we record information about how a particular verb is used in argument structures. Goldberg (2006) assumes that the formulation of generalizations of a pattern for the verb occurs due to three factors: a) issues related to the partial productivity of constructions; b) evidence that children are conservative in their use of argument structures; c) the frequency with which specific verbs appear in a specific argument structure influences the speaker's understanding. These generalization formulation processes are commonly related to contact with surface shapes.

Goldberg (2006) also states that the surface shape does not need to specify a particular word order, not even grammatical categories, although there are constructions that specify these aspects, to exemplify this notion, Goldberg (2006) the ditransitives.

Such structures involve a predicate with three arguments, usually agent, patient, and theme, but this does not mean that these roles are static. They are determined by the meanings of the constructions. In the case presented by the author, the predication informally indicates an act of *giving*. In this specific situation, the established roles can be different depending on the verb used in the construction.

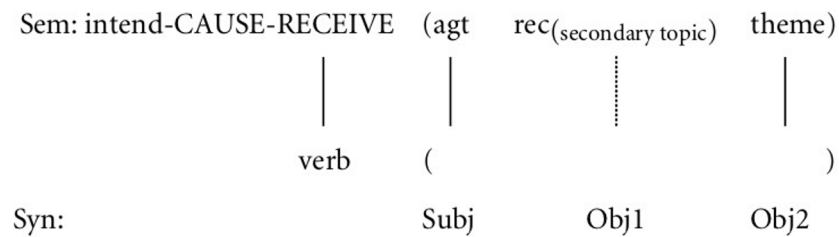


Figure 2 – Ditransitive construction (Source: GOLDBERG, 2006, p. 20)

Goldberg (2006) mentions is that roles are closely linked to construction semantics too. The profiled role and the argument are seen as unified, as shown by the lines in Figure 2, although there are cases in which this role is not pre-established by the verb itself, but by the construction.

The author highlights the importance of realizing that the reference to shape in the definition abstracts from specific surface shapes that can be attributed to other constructions, that is, an expression usually involves a combination of different constructions that can be freely combined, provided that are not in conflict.

This possibility of freely amalgamating constructions, if they are not in conflict, is related to the infinite creative potential of language, a trait shared with generative theories. But the constructionist approach postulates that what generates the sentences is not the grammar, but the speaker, a process that can be seen in the analysis of ditransitive constructions.

Still on ditransitive constructions, analyzing possible paraphrases, Goldberg (2006) highlights that it is possible to note that inputs could be grouped together, in the same way that outputs could also be grouped, although inputs and outputs cannot be grouped. Inputs share many properties but are systematically different from their paraphrases.

Mina bought a book for Mel. – Mina bought Mel a book.
 Mina sent a book to Mel. – Mina sent Mel a book.
 (GOLDBERG, 2006, p. 26)

Despite having the same schematic construction pattern, both are ditransitive, thus forming a group in which there is a separation between the sentences instantiated by *to* and by *for*. A sentence structured with *to* has a dative, Mel, in this construction, is the recipient of the action; on the other hand, the sentence structured by *for* indicates that Mel is the beneficiary of the action, although she is also its receiver. The perspective proposed by Goldberg (1995) regarding the profiling of roles of the participants in a sentence can help us to better understand this distinction.

As the author proposes, the verb has the power to open slots in the argument structure and to determine the profile of the arguments that can fill these slots. Based on this perspective, it is possible to determine, for example, that both the verb *buy* and the verb *send* can accommodate three participants, in the case of buying: the buyer, the purchased item, the receiver; in case of send: the sender, the item sent, the receiver. For sentences instantiated by *to* and *for*, however, it is possible to see that not only prepositions, but topicalization also have the power to change the profile of the participants, and only in *Mina bought a book for Mel*, Mel can be expressly interpreted as being receiver and beneficiary, which is not expressly determined in the paraphrase without the proposition *for* *Mina bought Mel a book*, sentence in which Mel assumes the role of receiver, the same role outlined in *Mina sent a book to Mel* and in *Mina sit Honey a book*.

The profiling of the participants reinforces the idea that, despite the schematic pattern being the same, the surface shapes cannot be grouped into the same group, as can be seen in Figure 3:

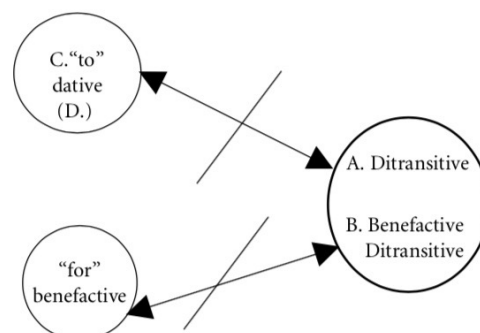


Figure 3 – Ditransitive patterns: *to* and *for* (Source: GOLDBERG, 2006, p. 26)

Furthermore, as the author reiterates, the very semantics of the verb can prevent ditransitive constructions from being seen as universal. The verbs: give, refuse, or take, for example, can be used in ditransitive constructions and share similarities among themselves and, because they have specific characteristics, are categorized into distinct groups.

Verb semantics also play an important role when analyzing the construction. Thus, it is important to note that the same verb can present different schematic patterns, depending on the context in which it is used, and some verbs, in different contexts, can assume different meanings.

From this perspective, as Goldberg (2006) argues, it is important to recognize the surface generalizations around the argument structure, because, thus, it is possible to recognize the existence of generalizations in the language. In line with this idea, it is equally important to admit that the meaning of a sentence is more than the meaning of the construction of the argument structure used to express it. When analyzing the sentence, it is necessary to observe the individual verbs, the arguments, and the specific contexts. When generalizing between different expressions and pointing out the differences between similar constructions, the analyst needs to carefully observe the verb, since it can be interpreted differently depending on the perspective adopted.

However, Goldberg (2006) warns that the paraphrase itself should not be taken as superior to other expressions, as it is sometimes not possible to determine cases of derivation and/or independent constructions. Thus, each construction must be analyzed in isolation, even in the formation of generalizations, because by describing a vast amount of isolated surface forms, it is possible to make broader generalizations, in the form of argument structure constructions.

It is important to emphasize that, for the Grammar of Constructions, all processes inherent to the language are interconnected, since, just as there is a union between form and meaning/function, there is a connection between cognitive processes. Thus, there is no concrete separation between the levels of the language, what exists is a continuum that involves from the smallest units of the language to complex patterns. This assertion, as advocated by Traugott and Trousdale (2013), strengthens the idea that language is formed by interconnected networks of constructions.

Linguistically, as Traugott and Trousdale (2013) show, the schemas are organized in levels, being the most abstract and highest-level schemas. Within a schema, we can

find different subschemas; subschema is a lower level in relation to the schema. It is a more specific group that can accommodate different groups of microconstructions. A microconstruction brings together different *tokens* that represent the same function, or similar meanings. Normally, the tokens gathered in the group of a microconstruction can be used in the same discursive context. *Token* is the lowest level of a scheme and is the representation of use. The schematic network of quantifiers illustrates this explanation:

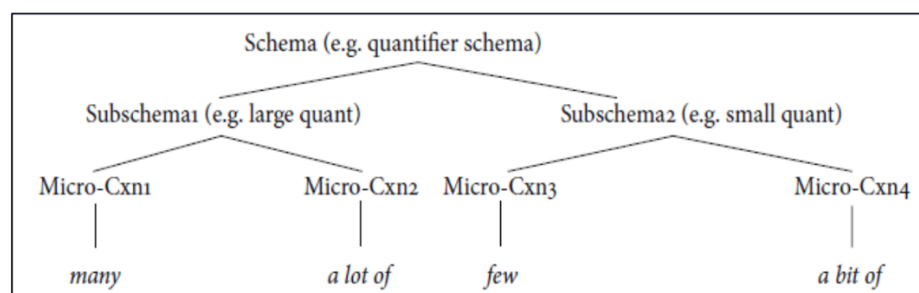


Figure 4 – Gradient of hierarchical relationships between constructions (Source: (TRAUGOTT; TROUSDALE, 2013, p. 17))

It is possible to notice that there is a larger and more abstract scheme, which groups together all types of quantifiers. Because it is more abstract and more general, it is found at a higher level in the hierarchy. Below, we find two subschemas, still abstract, because they only represent the general characteristics used as criteria for the creation of the subscheme, but with sufficient specificities to form two different groups. A subscheme groups quantifiers that express large quantities and another subscheme groups quantifiers that express small quantities.

At the base of the network, we find different microconstructions, which we categorize the construct uses (tokens), grouped by specific uses and functions, being that the constructs of a microconstruction, normally, cannot be used in the same context of use of another construct as the same worth.

Being a network interconnected, when any aspect of the network undergoes some change, it is likely to impact other points in the network. These changes do not occur immediately, they happen in a scalar way, on a continuum, and this is one of the reasons that confirm the idea that the levels of the language are fluid and gradient and, therefore, changes are possible.

5.2 Linguistic change under a construction view

The theory of Construction Grammar understands that language is constituted within a categorical continuum, in which there is no rigid distinction between the lexical and the grammatical. Therefore, it is possible for a word to be used in contexts in which it was not common before, and it may play a new role, being used with a different meaning.

Some research from a functional perspective (MARTELOTTA, VOTRE and CEZÁRIO, 1996; GONÇALVES, et al, 2007; FURTADO DA CUNHA et al, 2013; ROSÁRIO, 2015) conceive these changes that some words go through as a grammaticalization process – when a lexical word becomes a grammar word.

Goldberg (2006) defends the opening and widening of the scope of grammaticalization studies as the change that takes place within a continuum of constructions, in which there is no absolute distinction between lexical and grammatical items. Therefore, it is possible not only for a word to change from a lexical to a grammatical category, but it can also happen that a less grammatical word becomes more grammatical.

These two changes, from lexical to grammatical and from less grammatical to more grammatical, are present in what Traugott (2015) advocates as constructionalization. For the author, constructionalization is the process of forming a new construction, in which change occurs through a neoanalysis that forms a new pair of form and meaning, that is, constructionalization happens when the two sides of the construction are changed, both the form and the meaning, and the new construction starts to integrate a new node in the constructional network. This process is accompanied by gradual changes in schematicity, productivity and compositionality.

In addition, the author presents grammatical constructionalization, a process that also includes cases of formation of schematic constructions and, in this perspective, the entire scheme is grammaticalized, not just the item.

On the other hand, constructional change happens when only one of the construction poles - form or meaning - undergoes variation. Thus, there is no formation of a new node in the constructional network, the "new" construction integrates an existing node. However, it is valid to highlight that a variation is only a possible constructional change, not all variations become an eventful change, some are just an on-point variation.

The fact is that, regardless of the type of change, Traugott (2015) emphasizes that it must occur through five microsteps: 1) innovation: the listener interprets a construct and analyzes it differently from the speaker's analysis; 2) reuse: the listener who reanalyzed the construct becomes a speaker and reuses this construct with a new meaning; 3) conventionalization: occurs when there is a semantic reanalysis and other speakers agree with the conventional relationship between the original form and the new meaning; 4) constructionalization: when semantic and morphosyntactic reanalysis occurs, a new construction is created and enters the constructional network; 5) post-constructionalization: it can occur after the construction is formed and can go through a process of expansion, reduction of shape and even disappearance.

As one of the factors that allows the language to be subject to change is the gradient and fluidity of the categories, it is important to talk about the gradient existing in the very dimension of constructions: size, phonological specificity, and type of concept.

Regarding size, Traugott and Trousdale (2013) show that a construction can be atomic, monomorphic, such as suffixes, or words that do not have affix or inflectional markings; complex, when there is a strong connection between the items of an expression, which makes it impossible to analyze them isolatedly; and intermediate, when only part of the expression can be analyzed in isolation.

Phonological specificity analyzes whether a construction is substantive, schematic, or intermediate. Traugott and Trousdale (2013) explain that a substantive construction is completely phonologically specified as, for example, lexical items or crystallized expressions; constructions that are formed from abstractions and present a degree of schematicity are schematic, such as the inversion of the subject and auxiliary verb - SAI; constructions that have a noun part and a schematic part are considered intermediate, for example, the word formation scheme verb + ed (*played*).

The type of concept, according to Traugott and Trousdale (2013), involves the classification of a construction as being of content (lexical) or procedural (grammatical). Content constructions are those that can be used referentially, such as nouns, verbs, and adjectives; procedural constructions are those that signal linguistic relationships, such as verbal endings, demonstratives, and pronouns. The authors emphasize that there is no absolute division between these two groups. In addition to being a gradient division, it is possible that there are changes from one category to another, as in the grammaticalization

process, when a lexical item is used with procedural value. An example of this change is the use of the verb *go*, which today is also used to mark the future (procedural).

The constructions that have both content and procedural properties are intermediate, such as the way-constructions, which, formed by some *verb+way*, prototypically establish the way a path was carried out: *force his way through*; *elbow his way through*, and *giggle his way through*. As you can see, they have enough content for us to be able to distinguish them, but they have aspect marks, which is related to their procedural property.

To summarize, the authors present the table below and point out that a construction can be analyzed considering three dimensions: size, phonological specification, and type of concept:

Size	Atomic <i>red, -s</i>	Complex pull strings, on top of	Intermediate bonfire
Specificity	Substantive <i>dropout</i> <i>-dom</i>	Schematic N, SAI	Intermediate <i>V-ment</i>
Concept	Content <i>red, N</i>	Procedural <i>-s, SAI</i>	Intermediate <i>way-</i> construction

Table 5: Construction Dimension (Source: TRAUGOTT; TROUSDALE, 2013, p. 13)

In the languages which there is a possibility to express or omit the subject, we can propose a very schematic representation for the basic organization construction: [(SUBJECT) V_{subject} (X)]⁸. (SUBJECT) represents any argument in the role of the subject, for our study purpose, we are considering only its pronominal representation; V_{subject} represents the verb with its inflectional suffix; (X) representing one possibility of an argument element to be used in the role of the object and, for some utterances. For the omission of the subject, in a less schematic construction, we can omit (SUBJECT) in the scheme:

[V_{subject} (X)]

⁸ This scheme will be revisited in Chapter 10.

Langacker (2009) defends the idea that constructional networks are formed by categorizing constructions. The cognitive ability to generalize information allows the individual to form groups that are more general or more specific, but that are somehow related.

Linguistically, as defended by Traugott and Trousdale (2013), schemas are organized into levels, with schema being the most abstract and highest-level group. Within a schema, we can find different subschemas; subschema is a lower level in relation to the schema. It is a more specific group that can house different groups of micro-constructions, at the base of the network, we find different micro-constructions, which categorize the constructs (tokens), grouped by specific uses and functions, and the constructs of a micro-construction usually cannot be used in the same context of use of another construct with the same value.

It is important to note that new subschemas can be created over time. These are the new nodes likely to appear in the constructional network. Just as new micro-constructions can become part of an already existing subschema, so subschemas can no longer exist.

For the formation of a schematic network, the speaker must have contact with a significant number of inputs. Thus, the greater the individual's contact with samples of language use, the greater the chance that an abstraction will occur and form a schema. But, as Bybee (2015) and Goldberg (2019) highlight, there is no way we can measure how many inputs are necessary to cause abstraction, since it is a cognitive process and it is dependent on different factors, the frequency of inputs encounters is not operable.

For the formation of the network, Goldberg (2006) defends the idea that the speakers can abstract schemas from the use of several similar micro-constructions. In the scheme of ditransitives, for example, from contact with constructions of the type: *I_{subject} gave_v a cake_{object1} to João_{object2}*; *I_{subject} sent_v a letter_{object1} to my friend_{object2}*; and *I_{subject} passed_v the salt_{object1} to the guest_{object2}*, the speaker can abstract the scheme X gives Y to Z, of the type: X CAUSE Y TO RECEIVE Z.

Goldberg (2006, p. 7) is also aware that the same verb can be categorized into different subschemas, for example:

He sliced the bread. (transitive)
 Pat sliced the carrots into the salad. (caused motion)
 Pat sliced Chris a piece of pie. (ditransitive)
 Emeril sliced and diced his way to stardom. (construction-way)
 Pat sliced the box open. (resultative)

In addition to the verb that can trigger different schemes, Goldberg (2006, p. 21) also says that the same communicative event also has this property. When enunciating *a dozen roses, Nina sent her mother!*, the speaker activates the construction schemes: a) ditransitives, b) topicalization, c) noun phrases, d) verbal phrases, e) indefinite determinant, f) plural, g) twelve, rose, Nina, send and mother.

Another point to be highlighted is the productivity of a construction or a scheme. Traugott and Trousdale (2013) and Bybee (2015) define *productivity* as the probability that an item is used within a specific context, and it is associated with the frequency with which the item is used. The more frequent an item is, the more likely it will be used in that context in the future. Thus, when an item is used in a new context, it is its productivity and frequency of use that can indicate whether the new form will only configure a variation or a change.

It is important to notice, though, that the changes may happen, as previously noted, in different parts of the construction and the change in one of the levels of the construction may motivate further changes, this changes in how the speaker express the subjects will be analyzed in the chapter 9.

5.3 The integration of a construction [Form + Meaning]

Linguistic changes in a constructional perspective are motivated by the use and the speaker's necessity to communicate. Every change happens in microsteps, and the change can happen at any level of the construction. Bearing this in mind, we need to consider how the change in one of the levels of the construction can cause the other levels of the construction to also change. This goal is a key factor to a constructionist approach; since a construction is by default a pairing of form and meaning and every linguistic phenomenon can be conceived and analyzed as a construction, a study that self-proclaims itself to be constructionalist must analyze both form and meaning.

When asked if syntax and semantics, for example, could be seen as two sides of the same coin, Pulvermüller, Cappelle and Shtyrov (2013, p.14) answered that a

Cognitive and Construction Grammar see “an integration machinery for form and meaning” and explained that a change in one aspect of leads to others.

Even if we conceive the construction as the metaphor of a two-side coin, one being the side of form and the other the side of the meaning, as a coin *per se*, we must look at it as a whole item, being the two sides of equal importance and one couldn't even exist without the other.

Semantics and pragmatics, according to Fried (2013), are two factors that cannot be seen as two disassociated factors either, in fact, the author states that there's an integration between semantics and pragmatics and one cannot be perceived without the other. Fried (2013) reinforces that this unity has its manifestations in new linguistic structures, by all it means, we can dictate that the innovations in the language happen in the use, the syntax is activated by semantics through pragmatics.

Steels (2013) defends the idea the syntactic structures should always be semantically analyzed and that is what makes the integration between two poles – form and meaning – increasingly tighter compared to other theoretical approaches to linguistic analyzes.

Pulvermüller, Cappelle and Shtyrov (2013), Fried (2013) and Steels (2013), they all must agree then that there is, in constructional approaches, a great integration between all the interfaces of a construction: phonological features are linked to morphological ones that are linked to syntactic ones that are linked to semantical ones that are linked to pragmatical ones and they are finally linked to the discursive features – one feature influencing the other and being influenced by it.

After assuming that all the levels of a construction are connected to one another, it is necessary to discuss what precisely link them, especially what link the poles of form and meaning. Croft and Cruse (2004) address that there is a symbolic link that put these two poles together.

This symbolic link that joints the elements of a construction and how it is perceived is what set the constructional approach apart from the general syntactic theories, Croft (2001) states this by saying that different than the general theories that see the symbolic link external to the form and the conventional meaning, the constructional approach sees it as internal to a construction.

Thus, a symbolic link joins an element of the syntactic structure of a construction to a component of the semantic structure of that construction. There is also a symbolic link joining the whole syntactic structure to the whole semantic structure. [...] Each element plus corresponding component is a part of the whole construction (form + meaning) as well. That is, the construction as a symbolic whole is made up of symbolic units as parts. (CROFT, 2001, p.21)

This symbolic link that holds a construction and its components so tightly together is exactly what let us affirm that the change in any part of the construction affects the construction as whole. Considering the *expressiveness of the subject* as a construction, this also let us question which other aspects in BP and in EP have been going through changes due to its need to make the subject explicit. These two questions will be addressed in chapter 10.

6. Usage-Based Linguistics

In our sixth chapter, we discuss the Usage-Based linguistics (BYBEE, 2015; FURTADO DA CUNHA ET AL., 2013; LANGACKER, 2013; MARTELOTTA, 2011; TOMASELLO, 2008) and the following general cognitive processes that are linked to the language: iconicity (GIVÓN, 1984; 2001), perspective (LANGACKER, 2008), informativity, (GARCÍA, 1996; LANGACKER, 2013), analogy (BYBEE, 2015), rich memory (BYBEE, 2010; NADER et al., 2000), we also discuss about the markedness phenomenon (GIVÓN, 1995; LAKOFF, 1987).

Usage-Based Linguistics has as its basic assumption the notion that language is structured according to the needs of those who use it. Thus, the speaker's intentions, during the structuring of the communicational act, can change the way he structures the language. This theoretical field, which arises from the union of Functional Linguistics and Cognitive Linguistics, considers the relationship between form and meaning and the notion that language is constituted and can be understood if its real use is considered.

It is possible to see that the structuring of clauses is highly linked to the information that the speaker wants to communicate, for example, when the speaker prosodically emphasis a specific element of the clause, or when the speaker topicalizes it, placing it in the foreground, the speaker does so for believing that this element is more important than the others, as it is the central information of the predication.

The *topic* always appears at the beginning of the sentence, and it is it that receives

the informative force. However, when talking about the prosody, it is possible to change the focus of the sentence to other elements other than the subject. Therefore, depending on what the speaker wants to communicate, there are different constructions for the same world event, as exemplified in our table in the sentences below, which reveal how prosody can function as a tool for topicalizing the components of the sentence:

	Sentence	Possible emphatic meaning
1a	Ø não bati nele	(Neutral)
1b	EU não bati nele	The speaker indicates that who practiced the action was some else.
1c	Eu NÃO bati nele	The speaker negates practicing the action.
1d	Eu não BATI nele	The speaker indicates that he practiced other action different than <i>bater</i> .
1e	Eu não bati NELE	The speaker indicates that who was inflicted by the action was some else.

Table 6 – Topicalizing through prosody

Thus, it is necessary to analyze not only the context of the form, but it is also necessary to observe the content context, that is, it is necessary to go beyond the phonologic, morphologic, and syntactic relations and consider the semantic, pragmatic, and discursive-functional relations. This fact becomes evident when we emphasize that the same constituents, arranged in different ways from 1a to 1d, have different informational loads and their distributions in the sentence reveal the central information that the speaker wants to communicate.

Syntactic structuring is, thus, organized by the need to communicate an event. Therefore, not only structuring, but all grammatical relations emerge from use and are likely to be influenced by factors external to the structure of the language. Some examples of these factors are the communicative intention; the interaction between the participants in the event; the degree of intimacy between the participants and the knowledge shared by them – all these factors can be confirmed by Barros (2016, p. 24):

[a]treлар a concepção de gramática ao uso e ao discurso implica assumir que também fazem parte do processo de linguagem duas dimensões básicas da formação humana e, exatamente, por isso, elas são constitutivas da gramática, a saber: i) a interação social; e ii) a cognição (BARROS, 2016, p. 24).

Assuming this premise for the formation of grammar is to assume that the communicative need and the goals we intend to achieve are reflected in the way we structure the language. The politeness strategy, for example, even depends on the social position occupied by the event's participants.

As already discussed, Tomasello (2008) defends the idea that language grammar is formed by the interaction between individuals, from shared events and structures that are cognitively stored. Shared knowledge is used again when a similar situation occurs. Thus, when realizing, for example, that by saying *please*, a polite strategy, obtaining a desired item is facilitated, the next time the individual wants an item, it is possible that he or she may resort to a *please*, or a similar way to formulate the statement.

This notion is reaffirmed by Martelotta (2011), when the author argues that all cognitive relationships are perceptible in the interaction. Thus, not only linguistic, syntactic, and lexical choices are present, but also the knowledge of the culture stored and shared by individuals, arising from the mind's ability to store and access experiences.

These experiences, according to Langacker (2013), are responsible for how the individual builds the meanings they use linguistically. According to Goldberg (2006) and Bybee (2015), it is from enough similar experiences stored that the individual can re-elaborate existing patterns and generate new uses for existing forms.

Therefore, as also defended by Furtado da Cunha et al. (2013), language grammar emerges from the individual's use and is adapted by him to meet his communication needs. According to what the individual feels the need to communicate, he may resort to different communicative strategies, stored from the events experienced by this individual, as shown by Langacker (2013).

Given the above, the usage-based linguistics considers cognitive and social factors that go beyond linguistic materialization. Some of the cognitive factors that are expensive for the development of this study are present, as shown by Traugott and Trousdale (2013) and Bybee (2015), in activities that are not limited to the linguistic field – *categorization*, *prototypicality*⁹, *iconicity*, *perspective*, *analogy*, *rich memory and markedness* – they are present in all human activities.

⁹ We acknowledge the importance of *prototypes* while studying linguistic phenomena and its importance to categories formation; however, due to the nature of the object analyzed by this study, we will not consider *categorization* as an analysis parameter because the subject pronoun paradigm is, by default, a well-established category and we will not consider the insertion or loss of an element of this category either. In this same sense, we will not consider *prototypicality* as

These conceptions, associated with the notion of *informativity*, are necessary to show that the presence or absence of the subject have implications in the linguistic construction process, in the construction of linguistic meanings.

6.1 Iconicity

Iconicity within functional paradigms, according to Givón (2001), is related to the link between the expression planes and the content planes. This link is motivated by the being's relationships with the world, the way he conceives and conceptualizes all the events that occur around him. Linguistic structures adapt to human needs, thus establishing a strong connection between grammar and cognitive components linked to the way we linguistically represent the world.

As Wilson and Martelotta (2009) argue, linguistic representations are iconic, that is, they represent a natural relationship between linguistic elements and the meanings expressed by them. This relationship is so strong that it is possible to create onomatopoeias, words whose sound structure mimics the sound of the elements it designates. Onomatopoeia processes are good examples to show that language facts are not random, but a reflection of the experiences we live in the world.

Givón (2001) states that even the structure of a narrative is organized by the way events are experienced in the world. Narration of a bath, for example, is structured, prototypically, in the following order: we enter the bathroom, remove our clothes, and turn on the shower. However, if the water needs to be warmed up before we take a shower, it may be that, after entering the bathroom, the first thing we will do is turn on the shower. This example reinforces the notion that we structure language in line with our experiences of the world, however, the context of the structuring may strongly affect how the uttered results.

Givón (1984) still argues that there is a motivation that leads us to code the world as we code it. Thus, we adapted some life structures to linguistic structures, and, in this way, it is possible to state that language syntax is the representation of everything we live,

an indicator that a particular use of *null subject* or *expressed subject pronoun* as being a motivator of the use or not of a particular subject pronoun.

which goes against the idea that linguistic signs are simply arbitrary and confirms the assumptions of that the grammar of language emerges from our use of it.

In this sense, Givón (2001) defends three subprinciples of iconicity regarding iconic representations: *quantity*, *proximity*, and *linear ordering*.

The subprinciple of *quantity* is related to the linguistic material used in coding an event in the world. The greater the amount of content to be encoded, the greater the amount of linguistic material: the word *bye* has little content, and it is therefore encoded with little linguistic material; *I'll see you later*, that, functionally, it can be used at the same event as *bye*, in addition to the farewell, it comes imbued with the promise of a future meeting. Therefore, it has more content and a greater amount of linguistic material.

The *proximity* subprinciple is related to more specific contexts. Thus, the greater the conceptual integration, the greater its morphosyntactic integration. Givón (2001) defines, the more semantically integrated two events are, the greater the possibility that these two events are structured in a single sentence, for example, the verb *hear* in *I heard a bomb explode* informs two events that occur simultaneously and, therefore, are structured in a single sentence. *Hear* is a verb that encodes auditory perception; in *I heard that the bomb exploded*, the two events are not simultaneous and, therefore, are structured in a subordinate clause, *hear* encodes evidentiality (he learned from a third party that the event happened).

Linear ordering is related to the hierarchy of clause construction. Usually, according to Givón (2001), it is the least predictable or most important information that is placed in the foreground, a position normally occupied by the subject. Furthermore, Haiman (1983) discusses the idea that there is a strong tendency to organize narrative events in a chronological perspective. We narrate the facts, as we said before, as they happen in the world.

These three subprinciples reaffirm the idea that language is not arbitrary, but a reflection of how we relate to the world. Structurings, from the most basic to the most complex, are related to the cognitive principles of how we conceive the world, so much so that the more complex the experienced event, the more complex its linguistic representation will be.

Languages in which there is a possibility to omit the subjects, as defended by Chomsky (1981) and Rizzi (1998), will only make them explicit for a communicate purpose. This action of omitting and explicitating the subject can be considered iconic,

explicitating is linguistically more complex, and it can convey a more complex meaning, such as topicalization – which lays in agreement with the principle of quantity.

6.2 Perspective

The central ideas that guide this work are that there is a strong relationship between the way we conceive the world and conceptualize it and the way we organize and structure these conceptualizations. That is why it is necessary to discuss the different ways in which this relationship can be configured.

The perspective is related to the different points of view on a world event and the way this affects the linguistic organization for representing this event. Langacker (2008) asserts that the arrangement of visualization and dynamicity are two important cognitive mechanisms that affect the conceptualization of the world.

According to the Langacker's assumptions (2008), the visualization arrangement consists of the relationship between who sees and what is seen, which is configured in everyday conversational interactions, an event in which participants meet in the same place, observe the scene, and describe it. Therefore, they share the same point of view as they have the same perspective on the event.

In the example *João kissed Maria*, Langacker (2008) defines that, as it is a standard unmarked language arrangement, the inferences that can be made about this event go unnoticed by the speakers. It's as if we simply wanted to inform you that a boy in the world named John kissed a girl in the world named Mary; the implicatures and views about this event, also patterns, go unnoticed.

On the other hand, when we portray the same event in the imperative way, *kiss her*, it is possible to perceive what we did not perceive before: the linguistic configuration not only describes the event, but interferes in the development of the event, shows a desire for some situation to occur and, above all, translates a different perspective. Other aspects presented by Langacker (2008) which configure different perspectives, involve: word order, intonation, absence of subject or reordering of the subject's position, even the zero morpheme of the verb, which, in English, configures the imperative mode, reflects the speaker's intentions about the world event.

One factor that can change the perspective from which an event is described is whether the speaker is static or moving. Some linguistic constructions are only likely to

happen from the perspective of the speaker in motion, as shown by Langacker (2008), *when trees hurried past at 90 miles per hour*. In theory, such construction can only take place when the speaker observes the trees of a moving vehicle and never through a static speaker, since trees have no ability to move through the forest.

Another point that Langacker (2008) says is related to perspective are deictics. The separation of speakers in space or time and the attempt to locate the interlocutor, such as in *it's hot here*, reveals exactly the perspective of where one speaks, including whether this event occurs when the two people involved are in the same physical space. However, in a phone conversation, the perspectives can be different: *it's hot in here, but it must be cold where you are*. In this sense, perspective is determined only by the speaker.

Another example of differences, presented by Langacker (2008), between space, time and the perspective that linguistically constructs the event is the message of an answering machine. Normally, it would start with *I'm not here right now*; if the person is in the place, this denial is contradictory, since the person is in the place, but, when recording the message, the speaker alluded to a moment when he, possibly, would not be in the room.

The author also argues that one of the important factors for the arrangement of the point of view is the presupposed point of view. In a standard arrangement, the point of view is where both speaker and speaker are situated; the same situation can be observed and described from different points of view, which leads the speaker to formulate different constructions.

Langacker (2008) argues that some words and expressions already evoke the sense of advantage. The expressions in front and back are based on the location of the speaker and the event to build the vantage point:

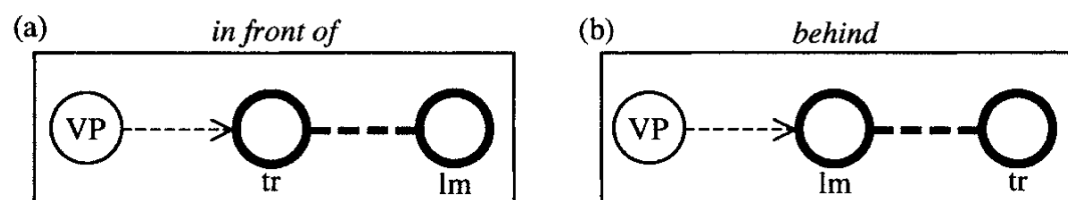
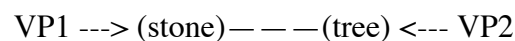


Figure 5 – Perspective based on spatial points (Source: LANGACKER, 2008, p. 76)

Langacker (2008) establishes that *VP* marks the *vantage point*, a chosen point of view, and the dashed arrow indicates the speaker's line of sight; *tr* marks the *trajectory*, the direction in which the perspective is constructed and, *lm*, the reference point,

landmark. In both cases presented by the author, in a and in b, the center of linguistic elaboration and what causes the semantic contrast to lie in the choice of *tr* and *lm*, but the event described is the same. To prove this perspective, Langacker (2008) proposes that we imagine a scene with a large rock and a tree. How we describe this scene linguistically depends on the point of view chosen. In this way, the same scene can be described from different points of view. Langacker (2008) uses the following scheme to illustrate this relationship:



From this scheme, it is possible to say that: 1) if the rock is chosen as the point of view, as the focal point of the line of sight, there are two construction possibilities: *the rock is in front of the tree* and *the tree is behind the rock*; 2) if the tree is chosen as the point of view, the constructions would be different: *the tree is in front of the rock* and *the rock is behind the tree*. Langacker (2008) points out that the point of view does not necessarily need to be the exact location where the speaker is, the location can be imagined. As abstract as it may be, the ability to adopt different points of view from fictitious places allows the speaker to elaborate constructions from other perspectives.

Another important point highlighted by the author is the constitution of the point of view based on the temporal position in which the speaker finds himself.

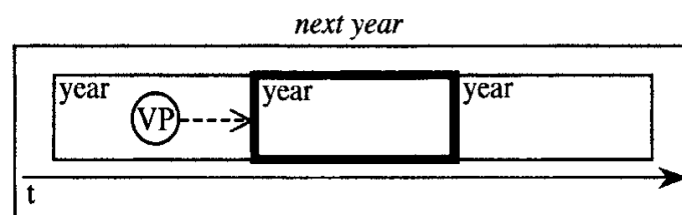


Figure 6 – Perspective based on time points (Source: LANGACKER, 2008, p. 77)

The expression *next year*, represented in the figure above, is based on the notion that there is a succession of years, with the following year being directly after the one that corresponds to the point of view. In this sense, in a standard construction, it corresponds to speaking time: *Next year will be full of surprises*. In other elaborations, even being the point of view, it may not correspond to the exact moment of speech. In *John believed that the next year would be full of surprises*, for example, next year corresponds to the year

following that when John came to believe and not the year in which the sentence was produced.

Langacker (2008) also highlights that it is the perspective adopted by the speaker that establishes the subject or object of a scene. Using the author's assumptions, when analyzing the structure *João took the lands* and *The lands were taken*, we have the same world event described in particular ways. In the first one, João occupies the central plane corresponding to the point of view, while in the second, the expression *the lands* takes over the central plane, even though the event is the same.

Thus, what one wants to inform the other is different and what determines the choice of how we structure this event is closely linked to what one wants to inform the interlocutor and, more than that, how one wants to reveal our perspective on the event.

eu falei "sabe que **eu** não sei." **eu** falei "bom, se fosse menino ia ser Rodrigo, agora, menina , **eu** não sei. ah! bota Carolina!" porque quando, assim, **eu** me juntei com a minha mulher, **eu** não estava numa boa em casa. (CLUL – BR – 80s)

In this occurrence¹⁰, for example, there seem to be only one of the times in which there is a specific reason for making the pronoun *eu* explicit, the last one, *eu não estava*, if we consider that the inflectional suffix can relate to both *eu*, *você*, *ele(a)*, and *a gente* and that there is in an immediate context a person that could be recovered by the verb *minha mulher – ela estava*, it would be reasonable that the subject need to be explicit to avoid ambiguity.

eu falei "sabe que **eu** não sei." **eu** falei "bom, se fosse menino ia ser Rodrigo, agora, menina , **eu** não sei. ah! bota Carolina!" porque quando, assim, **eu** me juntei com a minha mulher, **eu** não estava numa boa em casa. (M-BR)

However, the previous mentions of the subject could be avoided if we consider that we could recover the subject by looking into the verb inflectional suffix that could only indicate the 1st singular person *eu – falei, sei, juntei*. There's probably a syntactic pressure to make the subject explicit even when there's no necessity to do it, but we could also say, judging the context, that there is an informative need to reinforce that this

¹⁰ The occurrences presented during the development of this research are part of the *corpora Português Falado - Variedades Geográficas e Sociais* which belong to the University of Lisbon and can be accessed in the link: <https://clul.ulisboa.pt/recurso/portugues-falado-variedades-geograficas-e-sociais>

situation was happening to the locutor and, therefore, he feels necessity to reiterate the subject every possible time.

As the perspective of the event and the informative need of the locutor is what make him to express the subject, we decide to approach *informativity* as part of the perspective compound.

6.3 Informativity

As postulated in the previous section, the desire to show our point of view motivates us to organize sentences in specific ways, as postulated by informativity and, as Barros (2016) shows, these issues are beyond the structure of the utterance form. Informativeness is based on the relationship between those involved in the communicative act. In addition to showing our point of view about a particular event, we select ways to inform only what suits our interlocutor. Thus, informativity unites structure to pragmatic and discursive-functional functions and, according to García (1996), this determines, for example, the elements that configure *foreground* and *background*.

García (1996) conceives foreground as being the central part of the statement, the part that receives more focus, it is unique and obligatory; the background, on the other hand, is plural, has several elements that serve the foreground and configures an optional part for utterances. It is worth noting that, if one of the elements that make up the background is highlighted, it may cease to be a background and become a foreground.

According to Langacker (2013), depending on what the speaker wants to inform, some parts of the utterance can become more important and, normally, when this happens, this element needs to be, in some way, reinforced.

Going back to some examples introduced in the table 1, we can make this process explicit

- 1- EU não bati nele
- 2- Eu NÃO bati nele
- 3- Eu não BATI nele

Theoretically, the highlighted constituents, written in capital letters, also receive the contextual focus and they become the foreground of the scene. As it can happen to

the first component – the subject – when it is fully expressed, it can receive double focalization – one by being expressed when there was no syntactic need to and one by an extralinguistic force, such as the prosody.

6.4 Analogy

As defended by Bybee (2015), the term analogy has two meanings, a more specific and a more general one: the more specific is associated with changes in the morphological paradigm, while the more general applies to syntax and analyzes how new expressions arise from expressions already existing. A morphological change based on analogy analyzes how a word changes based on characteristics like those existing in other words. These changes tend to eliminate alternations or extend them to lexical items.

The author shows that changes that occur in sounds tend to affect all items likely to produce a certain sound; analogical change happens to one item at a time and does not necessarily affect all lexical items and morphological paradigms with the same production condition. As much as it may seem an irregular process, Bybee (2015) argues that analogical changes follow a unidirectionality.

The analogy process can take place in several ways. Bybee (2015) argues, for example, that the result of the alternative forms of *dreamed* and *dreamt* is a proportional analogy or four-part analogy, that is, two similar pairs are compared. In Portuguese, there are cases of creation of new verbs that normally use verbs from the first conjugation as a base, like this: *arma : armar :: mochila : mochilar*. From that same perspective, if it *seem : seemed* is true, probably *dream* is for *dreamed* and not *dreamt*:

seem : seemed :: dream : dreamed

However, the proportional analogy process does not consider the cognitive mechanisms involved in the changes. Also, hardly a single item will affect the system. The variation that occurs in *dream*, according to Bybee (2015), is based on a set of items that follow the same form: regular verbs, which receive the suffix -ed when conjugated in the past, a structure that is present in the language as a general form.

Bybee (2015) argues, then, that the process undergone by *dream* was analogical leveling. In this process, a new form eliminates the existing alternance on the old form.

The author gives an example of the change that occurred in the formation of some verbs in the past:

Base	Passado e Particípio Passado
[i:]	[ɛ]
<i>keep</i>	<i>kept</i>
<i>leave</i>	<i>left</i>
<i>sleep</i>	<i>slept</i>
<i>sweep</i>	<i>swept</i>
<i>feel</i>	<i>felt</i>
<i>kneel</i>	<i>knelt</i>
<i>mean</i>	<i>meant</i>
<i>dream</i>	<i>dreamt</i>
<i>creep</i>	<i>crept</i>
<i>leap</i>	<i>leapt</i>
<i>weep</i>	<i>wept</i>

This general change that occurs in the alternation of the sound of the verb base and the past participle is part of the *Great Vowel Shift* in which the extended sound of Middle English has been shortened when the end of the verb is *-t*.

Bybee (2015) also affirms that it is possible to perceive, in modern English, the *creeped* and *leaped* forms, although they cannot be seen as changes, from *leapt* to *leaped*, which characterizes as analogical leveling; what happens is that we have a new past form of the *leap* verb based on the regular verbs.

Bybee (2015) says that there is a strong link between the productivity of a pattern and the number of items affected by it: *type frequency*. In Modern English, 180 verbs have some productivity. Most of them are formed with the suffix *-ed* allomorphs ([d], [t] and [ɪd]) and, given the frequency, this pattern is usually applied to new verbs: *waltz/waltzed* (linguistic loan) / *hammer/hammered* (derivation – verb derived from a noun).

In Old English productivity was not *-ede* or *-ode*, the old form of *-ed*. The most frequent pattern was a complex system of alternating vowels in seven different classes, which made it difficult to create new verbs, and as there were some verbs (regular verbs) that formed the past by the suffix *-ede* and *-ode*, the formation of new derivation and borrowing verbs was more easily formed by this pattern, since it was simpler.

In addition to verbs, derivational affixes also compete for productivity: *-ness* is productive today for forming nouns in English. In Old English, there was the possibility

of creating nouns with *-ship*, *-hood*, and *-dom*, but over time they become less frequent and, consequently, *-ness* becomes more productive.

When questioning how to define which is the base form that can give rise to a new item, Bybee (2015, p.102) ensures that “high-frequency forms are resistant to change on the basis of the structure of other forms or patterns, and more likely to serve as the basis of such change in low-frequency forms.”

This is because each use of a word or construction strengthens its representation in memory and makes it more easily accessed in future uses. While more testing is needed to assert this, frequency is probably one of the strongest factors in predicting the direction of leveling.

Another process related to the analogy that Bybee (2015) refers to is the analog extension. In this process, the alternation starts to occur within a paradigm in which it did not occur before. There are circumstances for extension to occur: 1. one in which a base form serves as an alternation that carries a distinct meaning that spreads to other lexical items; 2. one that is an alternation introduced by the allomorph extension of an affix. In both cases, pattern frequency is a strong determinant of change.

The first order is those that represent the alternation of vowels in irregular verbs in Old English, for example, in *swim swam swum/spin spun*, by expansion. Other verbs are added to the list, and this prevents them from disappearing, for example in *ring rang rung/dig dug*. It is interesting to note that there are similar characteristics in these verbs: they are separated into two categories, those with three different forms and those with only two forms, with the past being replaced by the participle. Verbs that have three forms present regularity in the nasalization item, although the verbs that were added to the list by extension do not necessarily characterize the nasalization pattern, the phonological format being variable.

The second extension group involves alternations between plural marking suffixes in Brazilian Portuguese which are normally formed by *-s*, but when the singular ends in *-ão*; there are three ways to mark the plural: *s*, *-ões* and *ãos*. As the frequency of marking the plural with *-ões* is higher, some words that would receive the mark *-ãos* receive *-ões* (*cidadãos-cidadões*), which can be an extension process. Thus, the tag starts to exist in words that did not appear before or, as shown by Bybee (2015), it is possible to infer that there are three allomorphs of the plural affix that compete and co-occur with the *-s* tag. If we didn't make this consideration, we would be saying that there is a tendency to use *-s* and not the other forms, which is not the case.

Supplementation, another process presented by Bybee (2015), refers to any type of synchronous irregularity in the base forms of a paradigm, although its more restricted original meaning is used to refer to items whose base derives from other lexical items, like, for example, *go*, whose past was not *went*. *Went* was past *wend*, but today *go* past is *went*, while *wend* past is *wended*.

According to the author, this process normally occurs in inflectional languages that present few cases of supplementation. It is a frequent process and there are several possible generalizations about which categories can be expressed in supplementary forms.

Another process presented by Bybee (2015) is analogical reanalysis: the reformulation of an item based on other forms or patterns existing in the language, which happens when a meta-analysis takes place. Thus, some phonological material in a sequence is assigned a different morpheme or word: alternating the article *a* and *an* caused some words to lose initial *n* (naperon [French] – napron [ME] – apron) and others to acquire *n* (ekename – nickname).

Bybee (2015) also asserts that the change in sound occurs due to changes in articulatory habits. Although sound change is more likely to occur initially in high-frequency words, it is possible for change to occur in all words and is governed by phonetic factors.

Regarding BP, it is possible to say that some analogical process had occurred to the simplification of the pronominal paradigm. There is, as presented by Neves and Goulart (2017), an exclusive inflectional system for the pronoun *eu*, however, as we can see in Olbertz (2020), there is a syncretical form for almost all the other pronouns, basically, we would have an inflectional form for the *1st person of singular* and one form for all the pronouns.

To better understand this process of simplification, we primely must consider that, as stated by Duarte (1993), Portuguese underwent a change in its verbal inflectional suffixes its agreement process: first, there were six pronouns and six inflectional suffixes, one exclusive for each person, then, with the implementation of *você*, *a gente* and *os senhores* (as an alternative form to *vós*) it led us to six pronouns but only three inflectional suffixes, resulting in three syncretical forms:

six pronouns/six inflectional suffixes : six pronouns/three inflectional suffixes

The syncretical forms created an environment where the non-expressiveness of the subject could generate ambiguity since it was not possible to recover the subject directly from the verbal agreement process, i. e. *você/ele/a gente vai*. These situations where the subject expressiveness became mandatory, mainly to avoid ambiguity, and due to its frequency, may have enabled, via analogization, the expressiveness of the subject in contexts priorly thought as impossible, such as impersonal infinitive constructions and existential constructions.

Analogy can be, by default, used to explain how the process of pronouns simplification happened and how it is related to other changes in the language. It is only possible because analogy is related to a high level of cognition, it involves generalizations of word structures that are morphologically complex and affect one paradigm at a time, gradually enabling the change; however, some paradigms are unchanged, and when high-frequency paradigms resist change, it means they can still be easily accessed in memory.

6.5 Rich memory

Cognitive processes in general are associated with the human ability to store concepts and linguistic expressions in memory and, at appropriate times, trigger them. This process is called rich memory. As the individual interacts with the world, he stores informational data in memory so that, in a similar future event, this information can be activated.

Nader et al (2000) state that, in terms of language, this process involves how different sounds are combined to form words, the words themselves and their different meanings, and the situations, the context that allows those words to take on different meanings. Furthermore, as Goldberg (2006) shows, the individual also mentally compiles structures, complete linguistic patterns, and constructions.

It is important to emphasize, as Nader et al (2000) does, that all experiences, whether linguistic or not, affect the way we conceptualize and describe the world, as they structure cognitive representations and impact the neurological structure.

Bybee (2010), on the relationship between rich memory and language, argues that it is thanks to rich memory and the ability to form generalizations that we can deal with the complex and systematic form of communication. Memory is responsible for storing complex information, such as certain expressions that take on different meanings

depending on their context of use. Generalization is responsible for storing language abstractions that allow the generation of a functional scheme to formulate categories. Without these processes, it would not be possible to store all the elements necessary for human communication and interaction.

The definitions exposed so far allow us to state that all processes in the general cognitive domain are interconnected. The process of forming categories from a prototypical member, the frequency with which items/structures are triggered in memory and become favorable or resistant to the analogical process, the way in which we organize our utterance to interact with the other are associated with representations of similar situations stored in memory.

One key factor to the storage process to occur is the frequency of encounters one individual has with an input, some structures and patterns are so commonly used that they became easily accessed on the brain, being necessary little cognitive effort to access it. The frequency process is also responsible to determine, regarding the cognitive effort, which patterns will be marked, and which ones will be unmarked.

Thus, *rich memory* is a really important factor when analyzing the *expressiveness of the subject*, it is through our capacity to link cognitive processes and real-world live experiences to linguistic creativity and production, that we can understand the *subject* as a category and it being expressed via a pronoun as a subcategory and, in turn, that the subcategory is becoming the prototypical pattern which is iconic and shows the users' perspective of the event, then it also involves an *analogical thinking* that enabled the reduction of the pronominal paradigm to make it cognitively more accessible to finally result in a *non-marked* structure as we can see in the following section.

6.6 Markedness

According to Lakoff (1987), *markedness* is a process in which some morphological categories is entitled a “mark” while others are not. To the author, the singular form of nouns in English, for example, are not marked, while the morpheme-s that designates plural is a marked structure. In a scale of complexity, we could say that, cognitively and linguistically, the zero-marking form of plural is way less complex than the morpheme-s, it is also shorter and has less linguistic information, hence, it is simpler.

Lakoff (1987) also defends that we can also see the *markedness* process in phonology. The consonants that produce sound, that are voiced, are more complex than the ones that are voiceless, so we could say that the voiced consonants, the ones that includes vocal vibration in the production of sounds, are marked, while the voiceless are not.

In the semantics is no different, Lakoff (1987) proposes that meanings that are paired as contrasting pairs, like tall-short, may also have one of them that is marked. To prove his point, the author compares the question “How tall is Harry?” and “How short is Harry?” stating that just one of them would be elected in a current conversation, *tall*, it means that even among the contrasting pairs, one would be more likely accessed and, this way, unmarked.

Lakoff (1987) uses these examples to defend the idea that *markedness* is a process related to the asymmetries in the language, and this process can be present in different levels, basically, as it is said to be part of the prototype-effect, in all cognitive/linguistic category there will be a pattern that is simpler among the other members of the category and, therefore, unmarked.

Approaching this topic, Givón (1995) assumes that *markedness* is the same as *meta-iconicity*. To determine whether a pattern is marked or not it is necessary to analyze the context in which the pattern is used because markedness is highly dependent on the context, it is possible that a sentence be marked in on context and unmarked in another.

To establish a guide to check whether a linguistic phenomenon is marked or not, Givón (1995) proposes the following criteria: structural complexity; (ii) low frequency occurrence; (iii) cognitive complexity. The author ponders that a linguistic phenomenon does not need to have all three criteria, and that even though they may overlap one another, they must be considered separately.

The criterion of structural complexity is related to the syntactic elaboration, it can be considered a marked structure the ones that has a more complex elaboration; the low frequency occurrence is related to the frequency a pattern is accessed, the less frequent it is, the more marked it becomes; the cognitive complexity is related to the amount of cognitive effort one must use to process the information.

Thinking about the expressiveness of the subject, it is relevant to say that in PE, the omission of the subject seems to be more frequent than its explicitation, it is more complex and presents more linguistic material and therefore it requires more cognitive

effort to its interpretation, so, following Givón (1995) criteria, it is possible to say that sentences with an explicit subject have a marked structure.

Regarding BP, the expressed subject pronoun is being each time more frequent, mainly to the 2nd and 3rd person of singular, so, in this specific context, it is possible to consider the sentences lacking an explicit subject to have a marked structure.

To be sure about these assumptions, *markedness* will be one of the things to be considered, along the other cognitive process, while analyzing the data, mainly its production to whole discourse person panel. Other factors that will be considered lay on the guidelines present in the Construction Grammar, these guidelines are portrayed in the section *Methodology* – Part III of this thesis.

The next section will cover the phonological analysis theory, important feature of this work once it considers the prosody phenomena and its relation to the expressiveness of the subject.

7. Phonological Analysis Paramaters

This chapter presents one of the main aspects of the proposed analysis: the phonological parameters to analyze a linguistic phenomenon. Through the theoretical support of Dik (1989), Frota et al. (2015), Cagliari (1992), Massini-Cagliari (2003) Pietro and Roseano (2010) and Ramus, Nespor and Mehler (1999), it was possible to analyze the prosodic contour of our data, regarding the pitch accent pattern falling on the subject pronominally expressed.

Prosodic studies, according to Dik (1989), can be based on the analysis of the prosodic contour. For the theorist, this contour is equivalent to the “melody” in which linguistic expressions are produced and the main domains present in the melodic contour are tone, accent and intonation. Changes occurring in any of the domains can alter the melodic curve, the intonation, and the height of the curve, causing the curve to rise or fall, and these changes can come from different pragmatic functions, such as when at the peak of the curve melodic acts as a focuser.

According to Frota et al. (2015), it is important to understand how the intonation of a language operates, mainly because this system can operate on the meaning produced by usage or by the context where they happen. The authors also emphasize that a phonological analysis can provide essential information to studies across languages and

languages varieties to better comprehend how different languages can differ regarding features of the prosodic system such as *stress*, *rhythm*, and *intonation*.

Stress, in phonological studies, is understood as the part of a word or of a sentence which is emphasized given to its intensity resulting in a louder utterance. Generally, as posed by Pietro and Roseano (2010), Romance languages have the tendency of having the most prominent stress of a sentence to fall on the last content word of it, however, the stress position can change, as noted by the authors, due to emphatic cases or contrastive focus and Romance languages have shown flexibility on which part of a sentence can be stressed.

The category *rhythm* is related to the organization of timing of a speech, and a language can be classified into three different categories regarding this topic, it can be stress-timed, mora-timed or syllable-timed.

Ramus, Nespór and Mehler (1999) defines a stress-timed language as one in which the stressed syllable happens in regular patterns, intervals, and, generally, unstressed syllables tend to be shorter than the stressed ones so they can fit together in this rhythm. Syllable-timed languages are the ones in which the syllables, regardless of being stressed or not, have almost the same prominence and they usually do not present reduced vowels. Comparing both stress and syllable-timed languages, Ramus, Nespór and Mehler (1999) assumes that in both the syllables tend to have the same duration. Mora-timed, despite also having a pace, it is regulated by mora rather than syllables – such as Japanese.

Pietro and Roseano (2010) allude to the fact that Romance languages, such as Spanish and Portuguese, are syllable-timed; nevertheless, Frota et al. (2015), analyzing Brazilian Portuguese and European Portuguese, both Romance languages, affirm that they can display mixed rhythms, while BP can show syllable and mora-timing, EP can show syllable and stress-timing.

Another important category for prosodic analysis is the height of the curve. Casseb-Galvão (2001) defines it as being one of the components used to analyze the variation in the height of the melodic curve – also called intonation. According to the author, the function of this variation is to assign a relative value to a part of the information or to an information unit that establishes a link between the elements of the sentence.

The highest part of a melodic curve, according to Dik (1989), can vary in the high and low points, and it is in this variation and in the spacing between a higher and a lower tone that the different tones of a language are perceived. The analysis of the highest and

lowest points of the melodic curve is called intonation. Usually, it is clear from the study of intonation that the tone occurs within the syllables of a word, while the accent can be perceived in a word, sentence or even in a complete period, being directly linked to the prosodic contour.

Intonation, as quoted by Pietro and Roseano (2010), is the melody of the utterance and it has the modulation of F0 to measure the lowest harmonic voiced part of a speech. In addition, Cagliari (1992) defends that in relation to the *fundamental frequency parameter* (F0), property of the source of sound perceived by the ear as pitch, the highest part of the curve is responsible for representing the distance between the highest maximum F0 and the lowest minimum F0 of the melodic curve, generally ranging from 100hz to 300hz. The part of the statement that is at the apex of the melodic curve, according to Cagliari and Massini-Cagliari (2003), is the part that is in focus.

Pietro and Roseano (2010) also defends that there are two main reasons for a part of a speech to be at the apex of the melodic curve: (i) mark phrasing; (ii) to encode speech act distinctions, sentence modality, and belief state.

Some labels used as reference to the study of intonation are represented in the table below:










Monotonal pitch accents		
	L*	This pitch accent is phonetically realized as a low plateau at the minimum of the speaker's pitch range.
	H*	This accent is phonetically realized as a high plateau with no preceding F0 valley.
	jH*	This accent is phonetically realized as a rise from a high plateau to an extra-high level.
Bitonal pitch accents		
	L+H*	This accent is phonetically realized as a rising pitch movement during the stressed syllable with the F0 peak located at the end of this syllable.
	L+jH*	This pitch accent is phonetically realized as rise to a very high peak located in the accented syllable. It contrasts with L+H* in F0 scaling.
	L+<H*	This accent is phonetically realized as a rising pitch movement in the stressed syllable with the F0 peak in the post-accentual syllables.
	L*+H	This accent is phonetically realized as a F0 valley on the stressed syllable with a subsequent rise on the post-accentual syllable.
	H+L*	This accent is phonetically realized as a F0 fall from a high level within the stressed syllable.
Tritonal pitch accents		
	L+H*+L	This pitch accent displays a rising-falling pattern within the stressed syllable.

Table 7: Pitch accents patterns. Source: (PIETRO; ROSEANO, 2010 p.219)

The labeling for the *intonational phrasing* used in this study is based on Frota et al. (2015) – a study that searched for patterns of intonational phrasing and its correlation for its possible meanings in both EP and BP, for instance, the labeling H*+L L%¹¹ is used

¹¹ H corresponds to the highest part of the speech and L to the lowest; * marks when the speech is maintained in that position for some time, i.e., L* – the speech remains for a little while in the lowest part of the curve before the position changes, % marks the presence of a boundary. The Gray part is generally the part that is being analyzed.

for all varieties of EP as a focus context while in BP, $L^*+H L\%$ and $L+H^* !H\%$ can be used in the same context.

Also, for this study, it is highly important to analyze the implications of the cases when the most prominent part of the melodic contour, the highest part of the melodic curve, happens in prenuclear position and overlaps with the pronominal expressiveness of the subject, situations which can help us understand the speakers' motivations, in a pragmatic level, to alter the melodic curve to put the subject on its highest peak.

These assumptions are fully explored on the fourth part of this research, for the next part, however, we describe how this study was carried out and outline the analysis parameters used to reach our goals.

PART III – METHODOLOGY

8. Methodology

In this chapter we present the *corpora* we used and a few studies that were carried out with the same data base (YAO, 2020; FROTA, et al., 2021). We also present the PRAAT program developed with the guidance of Fant (1960) and Klatt and Klaat (1990) used in speech analysis. Then, we introduce our analysis criteria following the postulates of Tarallo, (1993); Duarte (1993; 1995) Duarte, Kato and Barbosa, (2001) and Kato to verify the morphological level; Neves (2018), and Lobo (2018) and Raposo (2018) to analyze the syntactic one; Olbertz (2020) to look into the semantic level and the four categories of subject; and, mainly, Gili-Fivela (2006; 2008) and Frota et al (2015) to analyze the phonologic level.

8.1 The *corpora*¹²

The *corpora* used in this research comes from the project *Português Falado - Variedades Geográficas e Sociais*¹³ developed by the Centro de Linguística da Universidade de Lisboa (CLUL). These *corpora* have been used to investigate the status of Portuguese, in all its different varieties, and its diverse linguistic phenomena, for example: Yao (2020) studied the resultative variation and its syntactic form under a Gerativism perspective and Frota, et al (2021) analyzed the prosody phenomena in patients with Parkinson's Disease.

Yao (2020) contrasted European Portuguese and Mandarin Chinese, regarding causative syntactic structure, and found out that the major difference in this phenomenon is what cause that state change, while in Portuguese a single verb has the power to express a complete caused-result event, Chinese express the same event in a complex syntactic elaboration. This difference may have its *roots*, as advocated by the author, in the semantic of the verbs itself, while, in Portuguese, it seems that the prime nature of the

¹² Our corpora are available on the following link: https://drive.google.com/drive/folders/1lwtIyzQ-XZrjaq0IMPrQFZwZbaE8ROWS?usp=share_link

¹³ The project can be accessed on the link: <https://www.clul.ulisboa.pt/recurso/portugues-falado-variedades-geograficas-e-sociais>

verbs has a *result* meaning, in Chinese they convey a pure activity meaning or a pure state of change meaning.

Frota, et al (2021) investigated the intonation and the prosodic system in patients that suffered from Parkinson's Disease to verify the relation between the use of a medication, dopaminomimetic, and the ability to produce structures of prosodic phrasing very similar to people without the disease. They discovered that patients with Parkinson had a difficult in using nuclear contours and the prosody, due to motor fluctuations, and the medication helped them to intonate better.

These studies are examples of how these corpora helped the development of research in different areas and with different purposes, hence, our choice of corpora to use. We also chose to work with *Português Falado* because it can give us the possibility to analyze our phenomenon with a constructional view, focusing on each level of the construction at a time. CLUL original project resulted in *corpora* of samples of varieties of Portuguese spoken in Portugal, Brazil, in the Portuguese-speaking African countries and in Macau, recent data collection includes samples of Portuguese spoken in Goa and Timor-Leste. These samples of oral speech, recorded in different places, dates and situations, are accompanied by the corresponding aligned orthographic transcripts. The two groups chosen to be analyzed were the ones from Brazil and Portugal.

From the *corpora* we select two regions in each country, from Brazil, we select *Rio de Janeiro* and *São Paulo*, and, from Portugal, we select *Porto* and *Lisbon*, one of the reasons for choosing these localities was because they were the most representatives of the *expressiveness of the subject* on the project's data, having the highest number of occurrences, and due to the geographic similarities between Rio de Janeiro and São Paulo and Porto and Lisbon, factor that could make a comparison fairer.

In total, there were 47 interviews analyzed, we read all interviews investigating the subject pronouns and we selected 5 from each region that contained the highest numbers of occurrences, totalizing 10 from Brazil, 10 from Portugal, as a social variable¹⁴, we only took into account the decade – 10 from the 90s, 10 from the 80s (Real-time methodology for data analysis). After selecting the interviews that would be analyzed, we reread them separating the occurrences (820 in total) and we exhaustively analyzed each

¹⁴ Sex, age and scholarship were not considered as relevant to our analysis because, during our reading, it didn't seem that these social variables played a considerable role in expressing the subject pronoun on the data analyzed.

and every one of them using the computer program PRAAT to determine the prosodical patterns.

8.2 PRAAT¹⁵

We opted to use the *PRAAT* program to analyze our data because it provides us with the necessary material to investigate pitch, format, and intensity of the prosody curve. The program was developed by Boersma and Weeninck (2017), has been used to study prosody: Gili Fivela (2002; 2008; 2018), and was used in Unisalento lab to analyze our data.

The program was developed following the research of Fant (1960) and Klatt and Klaat (1990) whose theories hypothesized the availability of turning acoustic speech into signals that contain information of the glottal source and the noises that are generated in the vocal tract constriction. This process made it possible to generate graphics as we can see below:

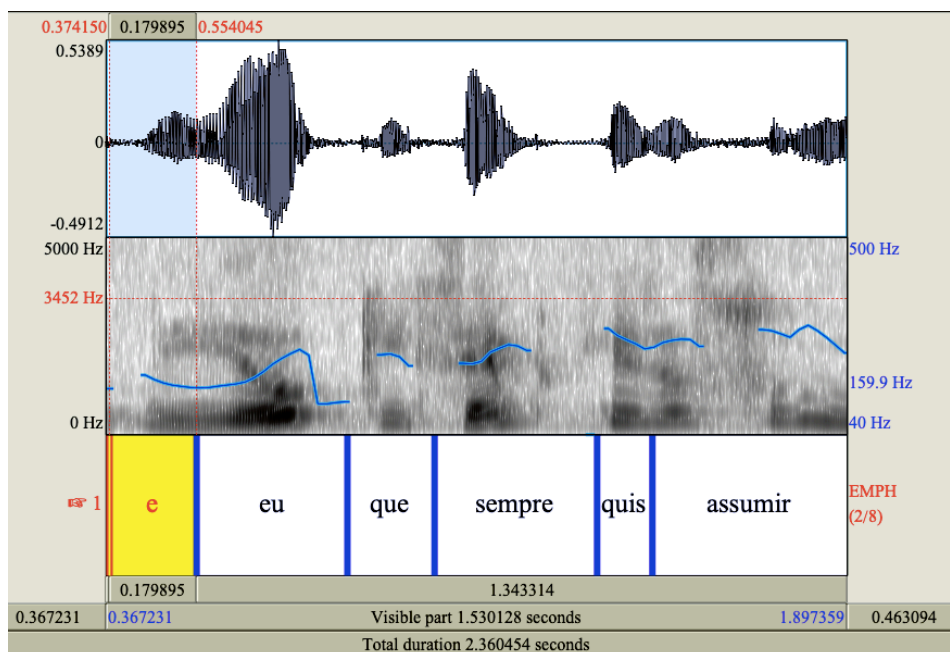


Figure 7– Example of graph generated by Praat¹⁶

¹⁵ The name of the program is reference of the work *listen/talk* in Dutch

¹⁶ The graph is multilayered: the first layer refers to the sound reading; the second, to the spectrogram and the intonational curve; and the third, to the grid where we can assign which words are being said.

The part of the graph represented by the blue lines is related to the curves of the speech, the highest pitch of the curve is considered to have prosodically emphasized an element and will be fully explored during our analyzes in chapter 9.

8.3 Analyses criteria

We will carry out a qualitative and quantitative approach to the expressiveness of the subject pronoun. Angouri (2018) defends that a study which considers both words and numbers can have a holistic and critical view over phenomena. Bybee (2015) will provide the alignment for the frequency and the numbers – the number of times the event happens may indicate the path of the linguistic change.

To analyze the morphological level, we considered, like Rubio (2012) did, the verbal agreement suffixes and the alternating pronoun subjects of 1st person of plural (*a gente* and *nós*), but we also considered the alternating pronoun of 2nd person of singular (*você* and *tu*) and the 2nd person of plural (*vós* and *vocês*). We relied on rely on the pronominal paradigm as it has been made by Tarallo, (1993); Duarte (1993; 1995) Duarte, Kato and Barbosa, (2001); Kato, (1999); and Neves and Goulart (2017). This part of the analyzes will focus on the shift in the pronominal paradigm in Brazilian Portuguese as it is stated by the authors that BP has gone to a process of simplification both in the subject pronouns and in its agreement while EP maintains a partially full pronominal system.

The morphological analysis as carried out as a starting point in our research, just as a matter of verification of the expressiveness of the subject pronoun in the data we chose to work with.

On the syntactic part, we investigated, the position of the subject, as proposed by Neves (2018), the usual position of the subject is pre-verbal, but it can also happen in other positions being always an extern argument. As proposed by Raposo (2018) and Lobo (2018), the position of the subject, its expressiveness and function can define whether the subject is a *grammatical subject* or not. This is an important definition regarding the expressiveness of the subject, mainly to verify with the subject pronoun is used with a function that is not on the syntactic level.

To analyze the semantic level, we considered the nature of the subject [+animated], [+human] and [-animated], [-human] and we also needed to add a few pragmatic considerations while typifying the function of the pronouns subject, we chose

to work, as Olbertz (2020) does, with four categories: *referential pronoun; topic pronoun; reactivation of topic; pronouns with no apparent motivation*.

The phonological analyzes resides on the description of the intonational pattern regarding the prosodic phrasing. For this purpose, the computer program PRAAT was used to verify which part of the statement is in the highest area of the prosodic contour and what the implications of this elevation are to produce different meanings. This analysis was carried out following the Gili-Fivela (2006; 2008) analysis standards with the theoretical contribution of Dik (1989); Scarpa and Fernandes-Svartman, (2012) and Pietro and Roseano (2010) and Frota et al (2015) was used for the analysis of the categories that are part of the prosody: prosodic contour, height of the curve, pitch, and melodic curve.

After considering the different levels of the construction on the expressiveness of the subject pronoun, we made a few considerations regarding the impact on one level of the construction to the others, on this part, we also make a few observations regarding the cognitive process involved in the act of expressing the subject via a pronoun.

While considering the analyze of the expressiveness of the subject pronoun as a construction, it was possible to propose schematic representation for this phenomenon in both BP and EP.

PART IV – CONSTRUCTIONAL ANALYSES

9. Expressiveness of the subject under constructional perspective

In this chapter, we discuss our findings about the expressiveness of the subject pronoun in both BP and EP. Firstly, we look at which pronouns are mostly used in both varieties and the verbal agreement, on this analysis, we are not considering the verbal tense and modes. In general, it was analyzed 820 occurrences of the subject pronominally expressed – 212 in EP, 28 in the 80s and 184 in the 90s, and 608 in BP, 214 in the 80s and 394 in the 90s – as it can be seen on the table below.

	EP	BP
80s	28 (3%)	214 (27%)
90s	184 (22%)	394 (48%)
Total	212 (25%)	608 (75%)

Table 8 – General Panorama of the Subject Pronoun Analyzed

Most of our data belong to BP, mainly from the 90s, it can indicate that in BP the subject has been used pronominally expressed with a higher frequency than in EP, however, if we check the differences between the decades, we can perceive that both varieties' frequency of expressing the subject via a pronoun has increased at a similar pace, EP with an increase of 23% and BP with 21%. It can indicate that both languages are developing toward a higher chance of expressing its subjects via pronouns.

After these considerations, on the next section, we check the alternating forms *tu/você* and *nós/a gente* and the verbal agreement, we also make some considerations regarding the pronouns that were not found in our data.

9.1 Expressiveness of the subject – Morphological Level

The following tables show the use of the pronouns, in EP, divided into discourse persons and the verbal inflectional suffixes:

Reference	Pronoun	Agreement	Quantity	Percentage
1 st singular person	<i>Eu</i>	1 _{SG}	14	50%
2 nd singular person	<i>Tu</i>	2 _{SG}	0	0%
2 nd singular person	<i>Você</i>	3 _{SG}	0	0%
3 rd singular person	<i>Ele/ela</i>	3 _{SG}	7	25%
1 st plural person	<i>Nós</i>	1 _{PL}	6	21%
1 st plural person	<i>A gente</i>	3 _{SG}	0	0%
2 nd plural person	<i>Vocês</i>	3 _{PL}	0	0%
2 nd plural person	<i>Vós</i>	2 _{PL}	0	0%
3 rd plural person	<i>Eles/Elas</i>	3 _{PL}	1	4%
		TOTAL	28	100%

Table 9 – Subject Pronouns in EP – 80s

Reference	Pronoun	Agreement	Quantity	Percentage
1 st singular person	<i>Eu</i>	1 _{SG}	121	66%
2 nd singular person	<i>Tu</i>	2 _{SG}	0	0%
2 nd singular person	<i>Você</i>	3 _{SG}	1	1%
3 rd singular person	<i>Ele/ela</i>	3 _{SG}	15	8%
1 st plural person	<i>Nós</i>	1 _{PL}	18	10%
1 st plural person	<i>A gente</i>	3 _{SG}	2	1%
2 nd plural person	<i>Vocês</i>	3 _{PL}	0	0%
2 nd plural person	<i>Vós</i>	2 _{PL}	0	0%
3 rd plural person	<i>Eles/Elas</i>	3 _{PL}	27	15%
		TOTAL	184	100%

Table 10 – Subject Pronouns in EP – 90s

Our first considerations are that the alternating pair *tu* and *você* do not occur in our data, there is just one occurrence of the pronoun *você* and its verbal agreement is related to the 3rd singular person. The alternating pair *vocês* and *vós* do not occur either.

While in the 80s there were no occurrence of the pronoun *a gente*, in the 90s, it happened twice and the verbal agreement is realized as it is supposed to be, according to the Grammar tradition, with the 3rd singular person.

In general, all the other pronouns occur with the verbal agreement postulated by the Traditional Grammar, the cases in which they are not are noted in the analysis at some point.

BP's occurrences are showcased in the next tables:

Reference	Pronoun	Agreement	Quantity	Percentage
1 st singular person	<i>Eu</i>	1 _{SG}	115	54%
2 nd singular person	<i>Tu</i>	2 _{SG}	1	0% ¹⁷
2 nd singular person	<i>Você</i>	3 _{SG}	26	12%
3 rd singular person	<i>Ele/ela</i>	3 _{SG}	43	20%
1 st plural person	<i>Nós</i>	1 _{PL}	8	4%
1 st plural person	<i>A gente</i>	1 _{PL}	4	2%
1 st plural person	<i>A gente</i>	3 _{SG}	2	1%
2 nd plural person	<i>Vocês</i>	3 _{PL}	0	0%
2 nd plural person	<i>Vós</i>	2 _{PL}	0	0%
3 rd plural person	<i>Eles/Elas</i>	3 _{PL}	15	7%
		TOTAL	214	100%

Table 11 – Subject Pronouns in BP – 80s

Reference	Pronoun	Agreement	Quantity	Percentage
1 st person of singular	<i>Eu</i>	1 _{SG}	181	46%
2 nd person of singular	<i>Tu</i>	2 _{SG}	0	0%
2 nd person of singular	<i>Você</i>	3 _{SG}	71	18%
3 rd person of singular	<i>Ele/ela</i>	3 _{SG}	72	18%
1 st person of plural	<i>Nós</i>	1 _{PL}	5	1%
1 st person of plural	<i>A gente</i>	3 _{SG}	34	9%
2 nd person of plural	<i>Vós</i>	2 _{PL}	0	0%
2 nd person of plural	<i>Vocês</i>	3 _{PL}	0	0%
3 rd person of plural	<i>Eles/Elas</i>	3 _{PL}	31	8%
		TOTAL	394	100%

Table 12 – Subject Pronouns in BP – 90s

¹⁷ The percentage of this particular use was littler than 0%, so it was not considered.

The pronoun *tu* occurs just once in the 80s, with the correct verbal suffix, against 26 times of the pronoun *você*, with the verb agreeing with the 3rd singular person. In the 90s, the reference to the second singular person of the discourse is realized only by the pronoun *você*. Thus, it can be inferred that the pronoun *tu*, in our data, is not the privileged one, like what has already been proved by Duarte (1993) in her data.

The pronoun *nós* has almost the same number of occurrences of *a gente* in the 80s, 8 against 6, but in the 90s, this number drastically change, 5 against 34, becoming, as advocated by Rubio (2012), the dominant form of expressing the first person of plural. About the verbal agreement, there is one interesting occurrence worth mentioning:

não sei o que eles fazem com tanto dinheiro lá fora, para apanhar dinheiro lá fora. e aqui, rapaz, **a gente temos** de tudo. de tudo! para ser um país, pó! rapaz! Sinceramente, **a gente não precisava** apanhar dinheiro com ninguém não, cara! É a mesma coisa com o petróleo. **a gente somos** tão burro. o que é que **a gente faz** ao petróleo, aqui? **a gente vende** o grosso, depois ele vai lá, refina, e vende para cá para gente mais caro. (CLUL – BR – 80s)

As it can be observed, there are two occurrences of the pronoun *a gente* being used with the verb of the 1st plural person (*nós*), *a gente temos* and *a gente somos*, while there are three occurrences in which the verbal agreement is with the 3rd singular person, *a gente não precisava*, *a gente faz* and *a gente vende*. It shows that although this informant wrongly agrees *a gente* with the 1st person of plural, there is an awareness that the correct correspondence is with the 3rd singular person.

Vocês and *Vós* do not appear in the occurrences and all the other discourse persons are expressed with the Grammarly traditionally expected verbal agreement.

It is also worth mentioning that, unanimously, the first singular person, *eu*, was the most expressed subject pronoun with, in total, 431 occurrences, 53% of the data.

Regardless of being the only discourse person that can only be referenced by a pronoun and not by a noun phrase, if we do not consider the verbal tenses in which the verbal agreement can generate a misunderstanding when the subject is not fully expressed, the past imperfect indicative, for example, *eu/você/ele/ela/a gente falava*, and considering that the 1st singular person has, generally, an almost exclusive verbal agreement suffix (NEVES; GOULART, 2017), it was expected that due to the morphology of the verb agreement suffix, there was no necessity of explicitating the subject pronominally and, even so, it is the most expressed one. Another factor that may

have contributed to the expressiveness of the 1st singular person is the model of the interviews, as the informants are guided to talk about their personal experiences on some topics, as they are positioning while telling their histories, it is reasonable to say that it motivated the use of this pronoun.

This necessity can be explained by both the principle of markedness and analogy. In terms of markedness, as proposed by De Rosa (2019), it seems that there is a growing number of cases in which Portuguese speakers, in general, feel the need to somehow express the subject and due to the frequency of occurrences, Portuguese pattern **SVO**¹⁸ with an expressed subject is becoming the new non-marked form and, although it has more linguistic material, which could lead to a more cognitive complexity, it is not necessarily the case.

When contrasting the occurrences of *a gente* and *nós*, mainly in the 90s, we have an expressive number of times in which the speakers opted to use *a gente* (34 occurrences) instead of *nós* (5 occurrences) and one of the explanations could lay on *a gente* verbal agreement (3_{SG}), while *nós* has an exclusive form (1_{PL}), so although it may seem that expressing the subject also means generating more linguistic material and more cognitive complexity, if we consider the verbal agreement suffixes and the possibility of reducing the desinential suffixes, via an analogy process, the cognitive complexity would be little since the speaker no longer needs to operate suffixation process on the verbal agreement level.

Then, only looking to the morphological aspect of expressing the subject via pronoun were not enough to understand it, that is why the next section is about the syntactical implications of expressing the subject.

9.2 Expressiveness of the subject – Syntactical Level

Regarding the purely syntactic part, the position of the pronoun subject is of almost 100% pre-verbal, being situated to left of the verb, there were only three occurrences of them being expressed in a non-prototypical position – once in BP and twice in EP, corresponding to 0,03% of the total data. Although it does not have a quantitative impact on the data, it is interesting to analyze these occurrences one by one.

¹⁸ Writing **SVO** with bold **S** indicates that there is no possibility of omitting it.

The occurrence in BP has a possible syntactic explanation to change the position of the subject, as it happens in the occurrence bellow:

ela fez a cesárea, foi onze horas, quando foi onze e vinte me deram a notícia. **estava eu e o meu cunhado** e eu só fazia chorar de nervoso, não é, porque nessa, eu sou muito bonzinho e tudo, mas nessas horas só faço chorar, meu nervoso ataca para chorar. (CLUL – BR – 80s)

Being located after the verb, it seems that the verb *estava* is mostly related to the pronoun *eu*, and the verbal agreement is, indeed, made with the 1st person of the singular, however, the subject is formed by the pronoun *eu* plus the noun phrase *o meu cunhado*, and the Grammar traditional general rule states that the verbal agreement in this case should have been with the 1st person of plural. Then, one plausible explanation the position change in this occurrence is to associate the verb and the expressed pronoun, the construction *eu e meu cunhado estávamos* (construction with the verb agreeing with the 1st person of plural) requires another verbal inflection and, as such, it is also more cognitively complex.

This occurrence shows a case in which the subject composed by a pronoun subject and a noun phrase cannot be considered a proper grammatical subject, since it does not trigger the verbal agreement process. All the other subjects expressed via a subject pronoun can be analyzed as a grammatical subject following Duarte (2018) proposal.

In EP, it does not seem the same case, we have a subject that is composed only by the subject pronoun, *eu*, in both occurrences:

visto que eu não, não tenho idade nem posição para me pôr em bicos dos pés e estar a dizer "**aqui estou eu**, aproveitem-me" (CLUL – PT – 90s)

In the use *aqui estou eu*, it seems that there is a necessity to highlight the place in which the speaker is located, the person is *here* and not in any other place. But other than that, it does not seem that there is a syntactic motivation to change the position of the pronoun.

e depois para se sair dali, tinha que se atravessar aquele bocado da areia e depois ia-se para a prisão... **suponho eu**. e, e de repente elas começaram-se a arranjar e já tinham saído muitas pessoas à nossa frente (CLUL – PT – 90s)

In *suponho eu*, it does not seem that there is a syntactic motivation to reposition the subject either. Doing quick research on google toolbar, we could find 18.100 results in which occurs *suponho eu* against roughly 118.000 cases of *eu suponho* – it rules out the idea that the construction *suponho eu* is used because of a particular scheme for the verb *supor*. We suppose, however, since its use seem rather discursive, and because it can be easily accessed through our *rich memory* and is available for the speaker, we found this kind of occurrence. If we consider that *suponho eu* and *eu suponho* take on different communicative purposes, such as a discursive function and a predicative function, for example, it could explain the motivation for using them both.

As it can be noted, properties other than the syntactic and morphologic must be considered while analyzing the subject, as it has been proposed by authors such as Casse-Galvão (2020), Neves (2018), and Raposo (2018). Thus, the next section is dedicated to investigating the semantic, pragmatic, and discursive aspects of the subject pronouns analyzed.

9.3 Expressiveness of the subject – Semantic-Pragmatic-Discursive Level¹⁹

Semantic Trait

The first thing that we needed to check while analyzing the semantic, pragmatic, and discursive level is its semantic trait, [+animated][+human]; then, we verified whether the pronoun had a specific reference or a generic one; and, lastly, we separated then into the categories: *referential pronoun*; *topic pronoun*; *reactivation of topic*; *pronouns with no apparent motivation*.

In BP, in total, there were six occurrences of subject pronoun referring to an entity that wasn't [+animated][+human], twice in the 80s and four times in the 90s. The pronoun used was from the 3rd person category in every time it happened:

¹⁹ Por que esse nível foi considerado como um todo? Porque para analisar o sujeito precisou-se analisar todos esses polos

Parades

– nesse, durante esse correr de ano, você viu alguma diferença, esses anos todos que você assistiu, alguma diferença, assim, gritante?

– não. cada, cad[...], de ano para ano, **elas [carnival parades]** vêm mais ricas, não é, mais ricas j, e vejo a diferença o seguinte: ano pa[...], há anos atrás, você via uma minoria de brancos desfilando, não é, hoje em dia, já está quase cinquenta por cento de brancos

Placenta

porque a, a placenta estava no[...], quase no colo do útero. se **ela** chega no colo do útero, a placenta sai na frente e a criança fica atrás, morre.

The law

dar um jeito nessa lei para **ela** caber dentro do que está acontecendo, sabe, então isso, quer dizer, Direito difícil, não é

ninguém pode assim dizer que não conhece a lei, porque a partir do momento que **ela** foi publicada, só aí **ela** vai começar a valer, sabe, então que são as leis assim... comuns.

Sovietic Union, Russia

aconteceu que em mil novecentos e dezessete, a União Soviética, a Rússia, **ela** tinha uma relação de produção eminentemente feudal, onde havia os czares e havia os servos

In EP, although the pronoun does not refer to a human entity, it has the trait [+animated]:

Larvae/eggs

vieram larvas ou ovos ou, ou qualquer coisa, vieram agarrados † minha pele, porque **eles**, depois quando eu me meti na banheira, na água, foi como se na água **eles** tivessem crescido.

In spite of it being worth mentioning, less than 1% of the occurrences have the trait [-human][-animated] and, as such, we cannot make further considerations in this regard. However, as it was shown by Soares da Silva (2006), De Rosa (2019) and Olbertz (2020), the possibility of a subject pronoun referring to entities, mainly 3rd persons pronouns, that are not human are becoming more evident.

Referenciality

Regarding the *referentiality* of the subject pronouns, as proposed by Neves (2011), there is a possibility that 1st and 2nd discourse persons to be used referring to a generic entity as it is usual to happen to the 3rd persons.

In BP, from the 608 occurrences analyzed, 549 showed pronouns with a specific reference, corresponding to 90% of the data, against 59 pronouns being used to refer to a generic entity.

There is not much of a difference if we separate this uses into decades. In the 80s, 8 percent of the data (18 occurrences) corresponded to generic reference and 10 percent (41 occurrences) in the 90s. However, the pronouns that were used to refer to a generic entity changed, as it can be seen in the next tables:

Pronoun	Quantity	Percentage
<i>Você</i>	3	17%
<i>Eles/Elas</i>	13	72%
<i>A gente</i>	2	11%
Total	18	100%

Table 13: Pronouns with generic reference in the 80s

Pronoun	Quantity	Percentage
<i>Você</i>	34	83%
<i>Eles</i>	4	10%
<i>Nós</i>	1	2,5%
<i>A gente</i>	2	4,5%
Total	41	100%

Table 14: Pronouns with generic reference in the 90s

While in the 80s the pronouns that were used to make a generic reference were the 3rd persons, which is more usual, according to Neves (2012), in the 90s the pronoun that is more used to accomplish this function was the 2nd singular person: *você*. The occurrence below is an example of how the 1st person *nós* and the 2nd person *você* were used with a generic referent:

como o pró[...], como a própria sociedade é estética, tanto é que **você**, fazendo uma análise histórica, **você** sabe que **nós** começamos numa, numa fase de produção escravista, passamos para uma fase de produção, eh, feudal onde existia o servo e o escravo e não se parou por aí (CLUL – BR – 90s)

In the uses for the pronoun *você* and *nós* above, the subject is not specified, in the first two, the referent is not the hearer, *você* is used as synonym of any person, while *nós* cannot indicate the participation of either of the persons involved in the communicational act in the event described and it acquires the meaning of *nation*.

Similar uses were found to the 1st pronoun *a gente*:

eu não acho legal **a gente** ter o filho com uma pessoa, o pai viver num lugar e a mãe viver em outro, eu acho que de acordo, tem que procurar contornar a situação, procurar viver bem, procurar se entender, conversar bastante, ter bastante diálogo. (CLUL – BR – 90s)

In this occurrence, *a gente* is used to refer to neither the speaker nor the hearer, but to refer to people in general and, as such, it has a generic reference.

The results from EP, regarding the *referentiality* of the subject were a little different, in the 80s, we could not find any expressed subject with the pronoun making a generic reference and in the 90s, there was only one occurrence that we present below:

é, pronto. **eles** podem pensar por que, pronto, ah, na cabeça deles, eh, não estão mentalizados, pronto, não estão muito mentalizados para isso, não é, eh, é tudo uma questão de educação, acho eu, e de princípios. (CLUL – PT – 90s)

Here, the pronoun *eles* refers to any person that identifies himself with the male gender.

It seems that while expressing the subject via a subject pronoun, BP has a higher tendency of using a generic reference than EP. There is an analogization force that could explain it, since there is growing number of contexts in which the subject can be expressed and the ones in which to not generate a misunderstanding, such as ambiguity, it is required to express the subject regardless of its generic or specific nature. First the subject starts being expressed to refer to specific entities, then, due the simplification of the verbal agreement suffixes, also to avoid ambiguity, then to refer to generic entities.

Even though it did not happen to EP yet, since BP and EP are in the same route of simplification of their pronominal system, it can be expected that in a near future more

subjects being expressed referring to a generic entity will be found in EP and even more in BP. To make more assumptions about the differences between these two varieties, we divided our data into the categories proposed by Olbertz (2020).

Olbertz's proposal

Olbertz (2020) acknowledges that most of the time, when the speaker changes the referent, a new topic is established, the establishment of a new topic has a pragmatic motivation and the change of referent has a semantic one, therefore, while establishing a new topic, at least two levels of the meaning pole are triggered.

We analyzed our data to find out if there was a subject pronoun which was used with a function purely semantic while changing the referent, but every time there was a change in the referent there was also a new topic. As a matter of following Olbertz's proposal, we did not rule out this category, yet none represent could be found neither in BP not in EP.

The subject pronominally expressed signaling a change in referent and establishing a new topic occupied a very expressive position in our data, in BP, from the 608 occurrences, 240 (39%) were *topic pronouns* – 86 in the 80s and 154 in the 90s. As it can be seen in the occurrence bellow:

ela viu as fotos que, **você** não, acho que não conhece[...], você não conheceu, mas **eu** trouxe um monte de fotos de lá, a metade está, metade está mal tirada e a outra... (CLUL – BR – 90s)

The use of *ela*, the first instance of *você* and *eu*, signalize the change of referent and the introduction of a new topic, not only it has a semantic motivation, but it also has a pragmatic one as well.

The reactivation of the topic, cases where after an interruption of the trail of thought there was a necessity of making the subject to become topic again, occupied the third position, 130 (21%) of the occurrences, 21 in the 80s and 79 in the 90s.

olha, **eu**... não, já assisti futebol, no Maracanã e, mas não vou sempre, **ele** vai sempre com o garoto, **eu** não vou. então, **ele** diz "Dora, assisti pela televisão é uma coisa e assistir lá, no campo, é outra." (CLUL – BR – 90s)

The first occurrence of the pronouns *eu* and *ele* indicate the change of referent and the introduction of a new topic, however, the second occurrence of both pronouns indicate the reactivation of the topic to keep the hearer following the trail of thought, avoiding misunderstandings, having, then, a semantic, pragmatic, and contextual motivation.

An intriguing category, *pronouns with no apparent motivation*, called our attention and motivated us to, indeed, find a reason why the subjects is being pronominally expressed. According to Olbertz (2020), the 3rd singular person pronoun is frequently used without a functionality such as emphasis, change of the referent or introduction/reactivation of the topic,

você fala em direito com um cara que é pobre, ele, para ele é um conceito abstrato de direito, **ele**¹ só viu falar nisso aí, mas **ele**² não tem a força, **ele**³ mesmo sabe que **ele**⁴ não tem a força de invocar o direito. (CLUL – BR – 90s)

The first *ele* highlighted is used as a new topic and there was a change in the referent, however, the other three uses the follows it indicate neither a topic nor a change in the referent not even it is a case where the omission of the pronoun could generate ambiguity; thus, apart from *ele*³ discussed in the end of this section, there were no real apparent necessity of expressing *ele*² and *ele*⁴.

But not only the 3rd singular person figures this category, actually, given its almost exclusive inflectional suffix, the 1st singular person is always used without being, as proposed by Olbertz (2020), fully functional:

Eu¹ estava em Bauru, **eu** senti qualquer coisa, **eu** senti uma preocupação com ela naquela semana, **eu** pensei nela de uma forma diferente, **eu** não sei como, **eu** não sei o quê, o que assim especificamente aconteceu, **eu** sei que alguma coisa assim me fez pensar mais fortemente nela naquela semana, naquele dia, e aí cheguei, ah, cheguei, entrei em casa assim no que **eu** abri a porta dei de cara com ela com a perna enfaixada para cima (CLUL – BR – 80s).

The first *eu* signalize a change of referent and a new topic, but there is no change of referent, and the topic remains the same during the rest of the speech. Through the inflectional suffix is also possible to retrieve the information from the subject so, its expressiveness was not necessary as well, it is the case, for example, when the speaker do not fully express the subject via a pronoun in *cheguei* and *entrei*, opting to omit it.

There is a higher frequency of use of pronouns with no apparent reason, being the second most frequent, there were 237 (39%) occurrences, 77 in the 80s and 160 in the 90s.

Summarizing this data, we have:

Pronoun	80s	90s	Total
<i>Change of referent</i>	0 (0%)	0 (0%)	0 (0%)
<i>Change of referent + new topic</i>	86 (40%)	154 (39%)	240 (39%)
<i>Reactivation of topic</i>	51 (24%)	79 (20%)	130 (21%)
<i>Pronoun + mesmo</i>	0 (0%)	1 (0,002%)	1 (0%)
<i>No apparent reason</i>	77 (36%)	160 (41%)	237 (39%)
Total	214	394	608

Table 15: Categories of subject pronouns in BP

Although there was a higher frequency of subject pronouns being not expressed with a tangible motivation in the 90s, overall, the pronouns that signalize the change of referent and introduction of a new topic and the pronouns with no apparent reason are used with the same proportion.

All discourse persons were used in the following categories – change of referent + new topic; reactivation of topic; no apparent reason without indicating numbers that could represent motivations for the phenomena analyzed, therefore, we opted to not subdivide these categories into the number of pronouns used in each of them.

In EP, the results were a little different:

Pronoun	80s	90s	Total
<i>Change of referent</i>	0 (0%)	0 (0%)	0 (0%)
<i>Change of referent + new topic</i>	19 (68%)	117 (64%)	136 (64%)
<i>Reactivation of topic</i>	3 (11%)	22 (12%)	25 (12%)
<i>Pronoun + mesmo</i>	0 (0%)	0 (0%)	(0%)
<i>No apparent reason</i>	6 (21%)	45 (21%)	51 (24%)
Total	28	184	212

Table 16: Categories of subject pronouns in EP

The occurrences in which there were a change of referent and the introduction of a new topic ranked first, with roughly two thirds of the occurrences, 136 (64%), 19 (68%) in the 80s and 117 (64%) in the 90s.

eles não têm, não conseguem viver nada que diga respeito à idade deles. eh, **eu** tenho uma filha, por exemplo, que está no décimo segundo ano, que eu ontem fui almoçar a casa de uns amigos, e **ela** ficou aqui a estudar. e tem sido isto sistematicamente, todos os fins-de-semana. (CLUL – PT – 90s).

The pronoun *eles* refers to the students in general and it is a change in the subject and a new topic; then, the speaker refers to herself by using the pronoun *eu*, changing the focus of the subject and establishing a new topic; in the end, the speaker uses the pronoun *she* to change the referent to her daughter, the last topic of the utterance.

The reactivation of the topic, in general, occurred with less frequency, compared to BP, just 12% against 21%, but it happens similarly in both languages.

eu comecei a gritar "Antônio anda cá ver, Antônio vem cá ver". e **ele** "o que é que foi, o que é que foi?" porque **eu**, entretanto, tinha, tinha ido lá para tirar o, o ralo da banheira, percebes (CLUL – PT – 90s).

There is the first pronoun *eu* as a new topic and the change of referent, then *ele* is inserted as a new topic to, in the sequence, be replaced by the reactivation of *eu*.

The ones with no apparent reason occurred in one fourth of the occurrences, 51 (24%). Although there were more occurrences in the 90s, 45 (21%), than in the 80s, 6 (21%), the percentage was the same since the use of pronouns scaled uniformly.

muitos candeeiros. gostam de luz. **eu** acho. é, é, é, estão, há, há uma fa[...], **eu** acho que para tudo há fases, **eu** acho que neste momento as pessoas estão, estão numa de... dar cor à casa (CLUL – PT – 90s).

Once again, the first *eu* represents a new topic and a new referent and, mainly because there is a specific inflectional suffix for the 1st person of singular, there was not a real necessity of expressing it again in an immediate context as it happened in the occurrence.

By contrasting the two varieties, it seems that the most feasible difference comes in the use with the syntactical motivation and the change of referent and introduction of

a new topic. While EP has a tendency of expressing the subject pronoun with a higher frequency when changing the referent, BP has a higher frequency of uses that lack functionality.

However, by looking at our data, it seems that there is, indeed, a reason for expressing the subject even when there is not a necessity to. It is becoming mandatory, in some contexts, to fully express a subject via a noun phrase or a pronoun, one of the contexts, as discussed by Paredes da Silva (1988), is that due the simplification of the verbal agreement system because of the change in the pronoun paradigm (syncretical forms), not expressing the subject may lead to misunderstandings: the singulars pronouns – 2nd (você) and 3rd (ele/ela) – and the plural – 1st (a gente) – have the same verbal agreement suffix – Você/ele/ela/a gente faz; in the past imperfect, we can added up the 1st singular person (eu) to this group of verbal agreement: Eu/você/ele/ela/a gente fazia.

This context of expressing the subject has been scaling through all the years, as proved by Tarallo (1993), Duarte (2003) and other authors already mentioned, and, according to Givón (1991) and Lakoff (1987), when a structure is used with a high frequency by the speakers of a language, it decreases its level of cognitive complexity, and it has the potential to become a non-marked structure. Therefore, given the frequency and the context that requires expressing the subject, it is possible to say that expressing the subject is increasingly becoming the non-marked structure.

It means that even though there are cases in which there is an exclusive inflection form and, this way, there would not be a need to express the subject, the existence of contexts where the non-marked structure is expressing the subject have an overall impact on *expressing/omitting* the subject and it can be proved by the process of analogization.

If the second and third person of singular and the first of plural need to be expressed and the structure **SVO** becomes the only option, the speaker can reanalyze contexts in which not expressing S was possible and, using **SVO** as a base model, starting to express it even when it had only a syntactical motivation.

So, *pronouns with no apparent motivation* are motivated both by the level word organization (syntactic level) with **SVO** becoming the non-marked structure and by cognitive process such as analogization.

The fourth category configures a particular combination *subject pronoun + emphatic modifier (mesmo)* had only one occurrence in BP and none in EP, accounted as 0,002% of the data, it is worth mentioning because it is a difference use, and it has an exclusively interpersonal function:

ele só viu falar nisso aí, mas **ele** não tem a força, **ele mesmo** sabe que **ele** não tem a força de invocar o direito.

The repetition of the subject may be an indicator that a different strategy must be used to emphasize a particular use, there are linguistic ways other than the grammatical combination above to make an element the emphasized one, the prosody is one of them and it is what we investigate on the next section.

9.4 Expressiveness of the subject – Phonological Level

Brazilian Portuguese

Regarding the analysis on the intonation contours, we analyzed the occurrences in which the highest part of the curve fell on the subject pronoun. In total, from the 608 occurrences of pronominally expressed subject pronouns in BP, there were 53 occurrences with the subject prosodically marked, corresponding to 9% of the data, while in EP, from the 212 occurrences, 24 (11%) were marked by prosody. Although the percentage from these two varieties were close, the way they mark they codify the prosody was slightly different.

The first thing to be noted is that, in all the occurrences, from both BP and EP, the subject had the trait [+human] and [+animated] and [+specified].

In BP, we found the following 8 patterns: H*; L + H; L+ H*; L* + H; H*+L; H + L*; H + H*; H* + H.

*The pattern H**

The *pattern H**, represented by the image bellow, happened when the whole pitch of the subject pronoun occurs in a high pitch frequency and it has a long duration, it was found in one fourth of the occurrences (14 – 26%) that, following our previous classification, were categorized into subject pronouns that had no apparent motivation to be expressed:

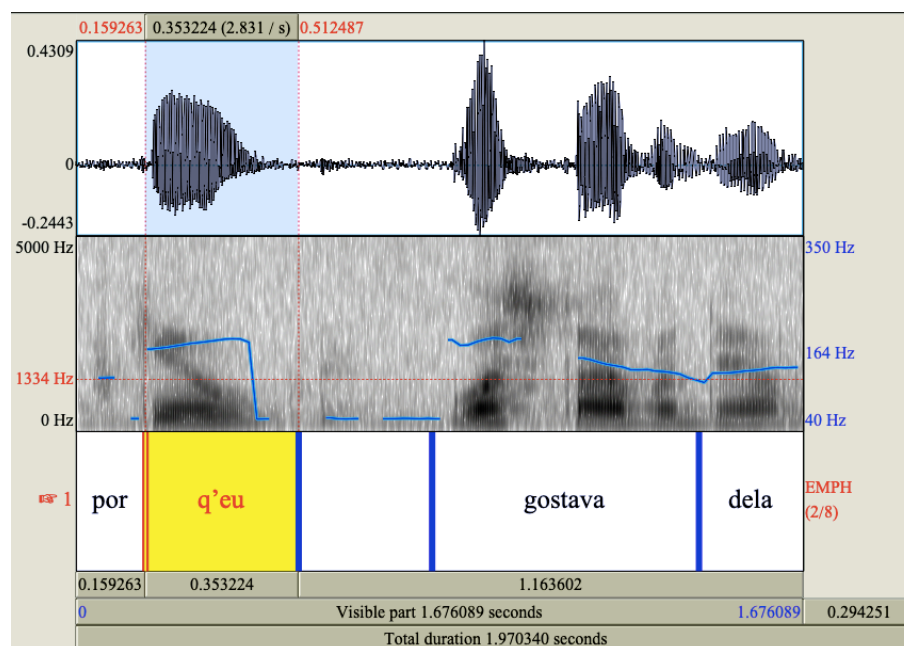


Figure 8: *Brazilian H* Pattern*

However, this pattern has a pragmatic implication of contrast of the kind X not Y, to emphasize that X, the subject, was the one to perform the action and not another entity. Below, we have the full transcription of scene pictured in the image above:

eu gosto dela demais, não é, porque eu tive muitas aventura na rua e nunca quis ter filho com ninguém. já com ela eu aceitei a criança porque **eu** gostava dela e, e senti nela (CLUL – BR – 90s).

There are four mentions of the pronoun *eu*, the first one would be enough to indicate who is the referent in the scene, the last one, though, is used to contrast the idea that it was him that liked her.

With a little less prominence, only 2 occurrences (4%), the same pattern was also used to signalize a change in the referent and the insertion of a new topic, the similar pragmatic function that was carried out to these uses, however, was the fact that both marked and contrast: X and not Y.

Interviewer: conta para gente como você faz a sua carne assada. dá receita para mim.

Speaker: aí **eu** boto umas cebola descascada, pimentão, umas batatinhas pe[...], miudinha, não é, (CLUL – BR – 90s).

The speaker has the informational need to set herself apart from the other cooks, then, she prosodically utters *eu* in a higher pitch to make this contrast.

L+H pattern*

Another pattern used to describe the subject we had classified as having no apparent motivation and with the contrast of the kind X and not Y, was the *L*+H pattern* (10 occurrences, 19%), this pattern describes a pitch that has a longer initial part with a higher pitch towards the end of the pronoun:

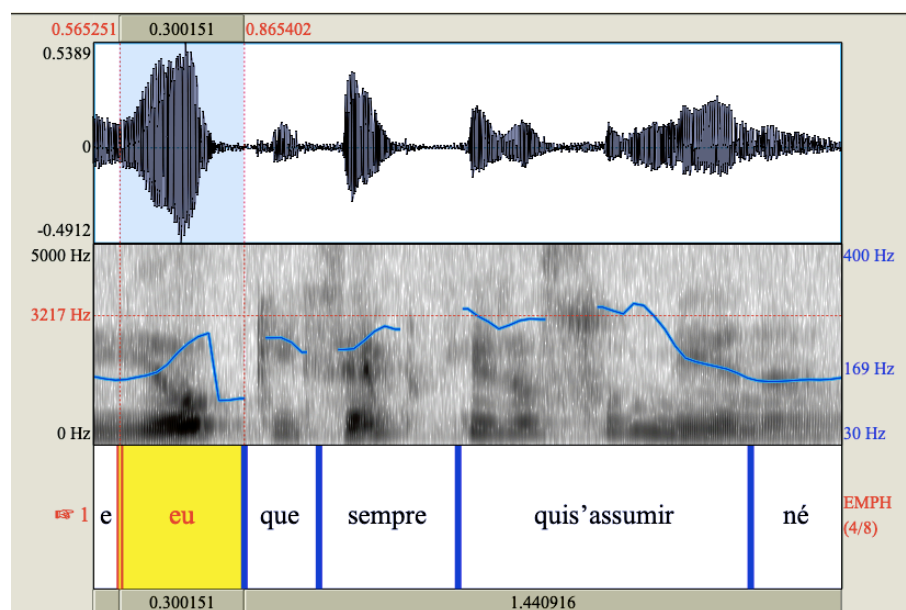


Figure 9: BP *L*+H Pattern*

eu, e eu quis[...], sempre quis assumir, não é, um compromisso com ela. (CLUL – BR – 90s).

Here, the speaker wants to reinforce that was him that always wanted to have a relationship with his spouse and not the other way around.

This pattern was also used to reactivate the subject, with the same idea of contrast (5 occurrences, 9%):

eu acho que a minha irmã, ela difere muito de mim em relação a preconceito, em relação a... muitas coisas que **eu** acho mito, eu acho tabu e ela acha que não (CLUL – BR – 90s).

The speaker, after making a little pause in her trail of thought and changing the topic to her sibling, reactive *eu* as a topic and emphasize the difference between her and her sister.

H+L Pattern*

In the same direction, the pattern *H*+L* also is used with subjects that marked a contrast and for which there were no motivation (3 occurrences, 6%), this pattern happens with a high pitch in the initial part, it has a long duration, and the intonation just slightly falls in the final part of the pronoun:

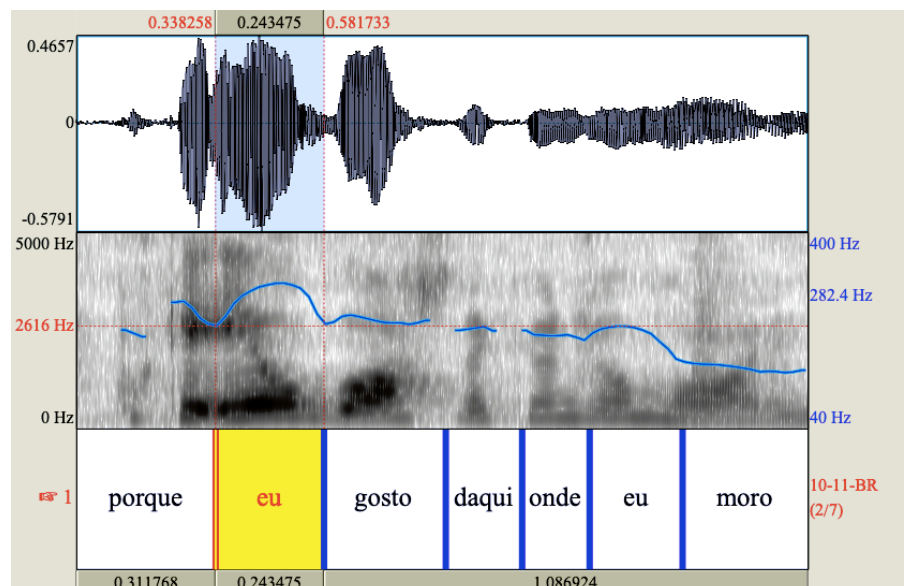


Figure 10: *BP H*+L Pattern*

This pattern is used to change the referent and introduce a new topic, pragmatically associated with contrast (4 occurrences, 8%):

meu marido não vai sair daqui mesmo, porque **eu** gosto daqui onde eu moro, entende, assim, dos vizinho. são muito bons vizinho (CLUL – BR – 80s).

In this occurrence, the meaning intend by the speaker is that *she* likes the neighborhood while her husband, despite not liking it, wouldn't like to move elsewhere, therefore, her opinion contrast with hers. Even though this contrast is not the same as in Y and not X, the contrast of ideas is strong enough to make it into the linguistic level.

In the next occurrence, however, the pattern shows a different purpose:

eu que cozinho mesmo. **ele** não sabe fazer nada, sabe nem fritar um ovo! . (CLUL –BR – 80s).

It seems that there is a double strategy to reinforce the contrast, the prosody *pattern H*+L* and the use of the determiner *emphatic modifier (mesmo)*. The speaker wants to make sure the hearer understand that it is she that cooks in her house, and not her husband.

L+ H Pattern*

The pattern *L+ H** differs from the other patterns described so far because it does not signalize the contrast of subjects or to contrast ideas, it was used in 3 occurrences (6%) to emphasize the new topic and in 4 (8%) with no apparent reason. It happens when the first part of the subject pronoun is lower with a fast growing increase in pitch and the highest part has a longer duration. This patter also shows a little difference because there is always a word before the subject that seem to motivate its emphatic context such as: *conjunctions and discursive markers*, or a negation, after the subject, as we can see bellow with *e* and mostly *que* are used to initiate the intonational pattern.

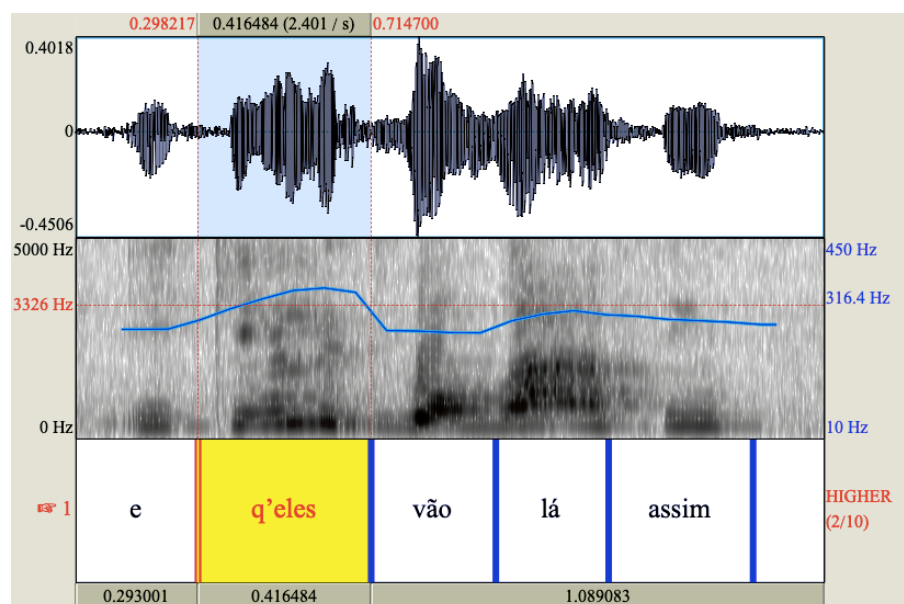


Figure 11: *BP L+ H* Pattern*

É lindinho e que vai aqueles pé de chinelo e tudo sujo de barro, sabe, ah, e que **eles** vão lá assim, e [...] o, as pe[...], eles têm muito medo do advogado. (CLUL – BR – 90s).

In this occurrence, the pronoun *eles* is used as a new topic followed by *que* and they both are also emphasized.

sou suspeita para fa[...], leiga **eu** não sou, nem um pouco (CLUL – BR – 90s).

In this occurrence, there was no need to express the subject given the exclusive inflectional form of the verb *ser* in *eu sou*, however, it is expressed, and it is prosodically emphasized, probably motivated, in both cases, by the negation.

L+H Pattern

The contexts prosodically marked by this pattern were only associated with a new topic and the change in the referent, just like the pattern L+ H*, but with a contrasting idea and, differently from the pattern L+H*, the final part of the subject pronoun, although in a higher pitch, does not have a long duration:

L+H pattern corresponded to 5 occurrences (9%):

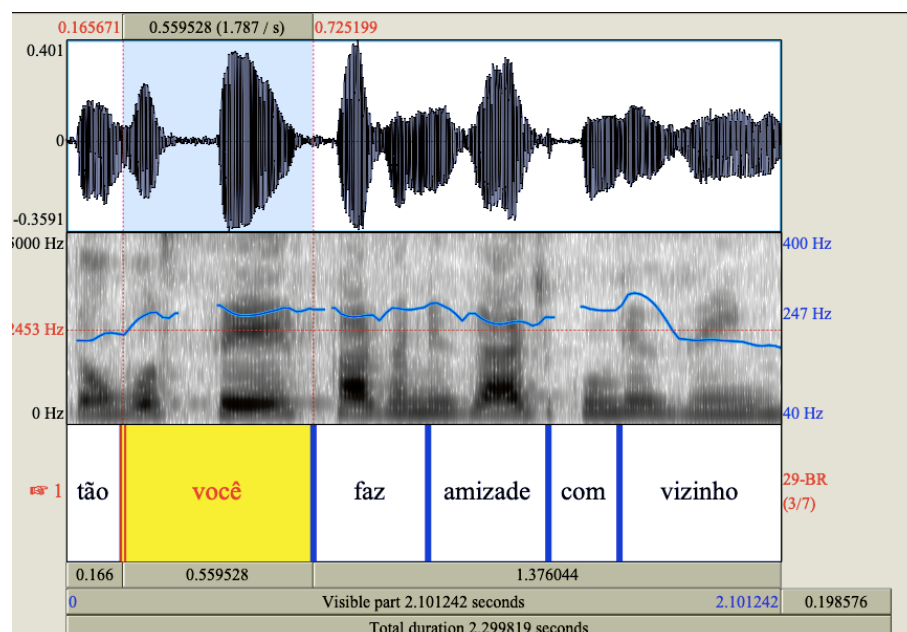


Figure 12: BP L+H Pattern

então, **vo**ç**ê** faz amizade com vizinho, já apartamento, não. (CLUL – BR – 90s).

The situation directed to the hearer is that in a flat there wasn't a possibility of making friends with the neighbors, while living in a house, this was a possibility. It is interesting that this pattern was used when the contrasting idea was not centralized on the subject.

H+H Pattern*

The *H*+H* pattern, with 5 occurrences (9%), also signaled the introduction of a new topic, but differently from the other patterns, the subject is always followed by a negation idea/word. This pattern can be described with two high pitches, the second being higher and with longer duration:

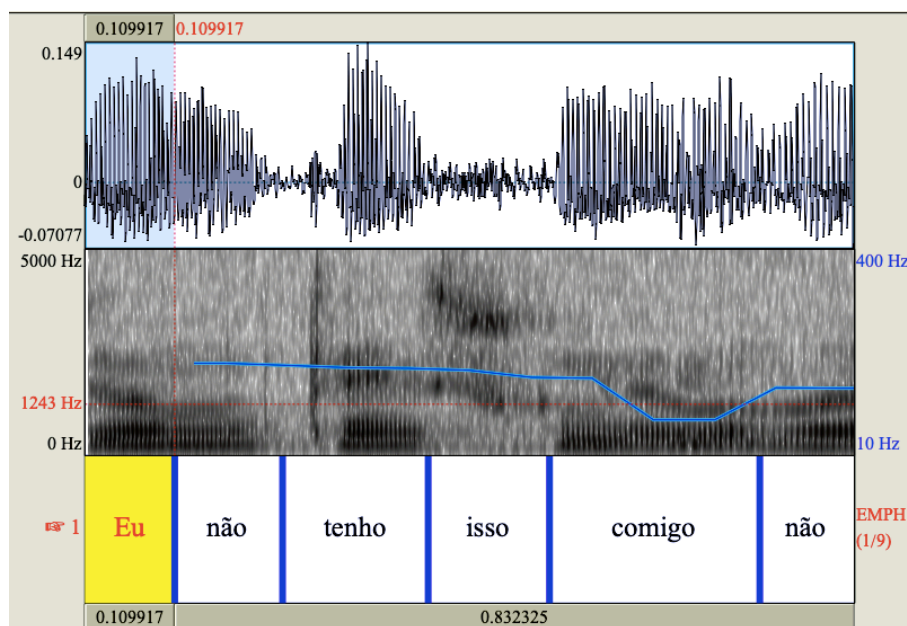


Figure 13: *BP H + H* Pattern*

eu não tenho isso comigo não, eu choro mesmo. aí, quando o médico chegou, não é, falou assim "quem é o papai?" eu só fiz assim. (CLUL – BR – 90s).

Besides the negation, there is a context of euphoria of becoming a parent and doing something that is not normally accepted by society, being man a crying. The other occurrences with this pattern had excitement, frustration, or indignation.

H + L Pattern*

There were only 1 occurrence with the pattern H+L*. This pattern starts in a high pitch, but towards the end, it has falling intonation with longer duration. It was similar to H*+H, when signaling a feeling of appreciation for the referent:

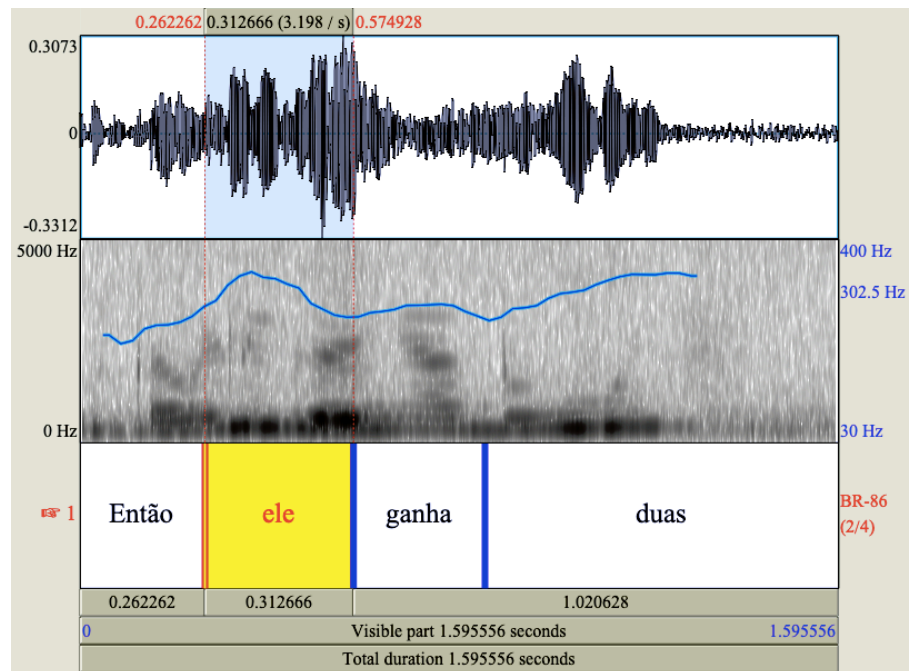


Figure 14: *BP H + L* Pattern*

é. o governador dá, né, a, o oficial superior tem direito. então, **ele** ganha duas, né (CLUL – BR – 80s).

In this occurrence, *ele* is the husband of the speaker, he has a prestigious position in the army and even has the admiration of the governor to the extent of receiving to tickets to the carnival parade, this admiration for the referent is what, possibly, motivated the speaker the emphasize the subject pronoun via prosody.

BP Patterns Summary

Based on the previous analysis, it is possible to formulate the following table for Brazilian prosody pattern when explicitating the subject via a subject pronoun:

Function / Subject	Pattern	Quantity
Contrast Subject: pronoun with no apparent reason	H*	14
	L* + H	10
	H* + L	3
Contrast Subject: Change of referent New topic	H*	2
	H* + L	4
	L + H	3
Reactivation of the Topic	L* + H	5
No contrast Subject: Change of referent New topic	L + H*	3
	H* + H	5
	H + L*	1
No motivation, word emphasize	L + H*	3
	Total	53

Table 17: Brazilian Prosody Pattern for the Subject Pronoun

It seems that BP doesn't have a clear pattern for the different subjects that are expressed, since the pattern H*, for example, can designate one function, contrast, but it can be activated by two kinds of subjects, with no apparent reason and as a change of referent/new topic.

However, it seems that there is a reason for why to emphasize phonologically the subject, from the 53 occurrences analyzed, 36, almost 70%, was used to signalize a contrast (X and not Y).

On our data, 30 occurrences with the subject pronoun in the highest part of the intonational contour matched with the pronouns with no apparent reason, another reason to believe that these uses are motivated, they needed to be expressed so that the prosody could also happen. This explanation becomes even more credible if we consider that from the 130 times that the pronoun was expressed with no apparent reason, 30 times it was emphasized through prosody.

The reactivation of the topic was the kind of subject that least were phonologically emphasized with only 5 occurrences and the ones that were emphasized with aid of a word, such as negation, had 3 occurrences, both these situations followed the same pattern: L + H*. And in both cases, the possible explanation is the same, there is already a process of emphasis happening, in the first case, we have the reactivation of the topic

that is, by default, an emphatic process, and the second case, the negation and words like conjunctions, already stress a part of the sentence.

European Portuguese

The first consideration that must be done regarding EP is that there were no occurrences in which the subject pronoun was emphasize through prosody with the pronouns having no apparent reason to be expressed. Another thing that can be easily perceived is the fact the patterns for the intonational contour in EP seems to be more solid.

Overall, in EP, we found 6 patterns for the intonational contour for our phenomenon: H^* ; H^*+L ; $L+H^*$; $H^*+H\%$; $L^*+H+H\%$; $L+H^*+L$.

H Pattern*

All the time H^* is the pattern for the intonational contour of the subject pronoun, 3 occurrences (13%), it signalizes a change of subject and a new topic, it also corresponds to a contrast system X and not Y. It happens when the pronoun has just one pitch, it is high and has a long duration.

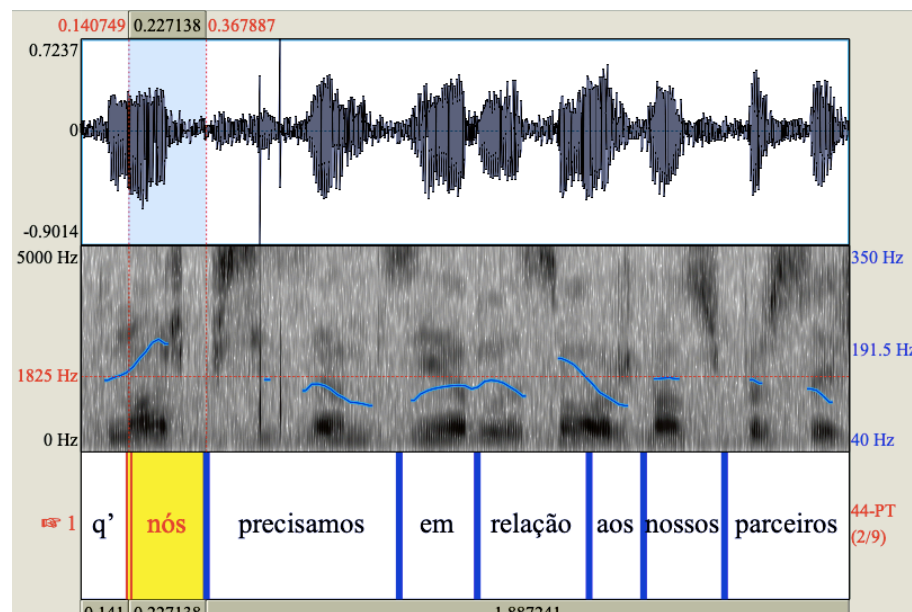


Figure 15: *EP H* Pattern*

isso eu reconheço, mas em relação àquilo que **nós** precisamos e em relação aos nossos parceiros comunitários (CLUL – PT – 90s).

In this occurrence, the pronoun *nós* corresponds to both new topic and referent and, in this case, it also symbolizes the speaker's opinion about what he and his company need in detriment of his business partner.

H+H% Pattern*

The only other pattern that was used with the same function and with the same kind of subject as the previous one was the pattern $H^*+H\%$, 2 occurrences, (8%). This pattern can be described with two high pitches, just like H^*+H , the difference is that just after the second pitch, there is a pause before initiating the next segment:

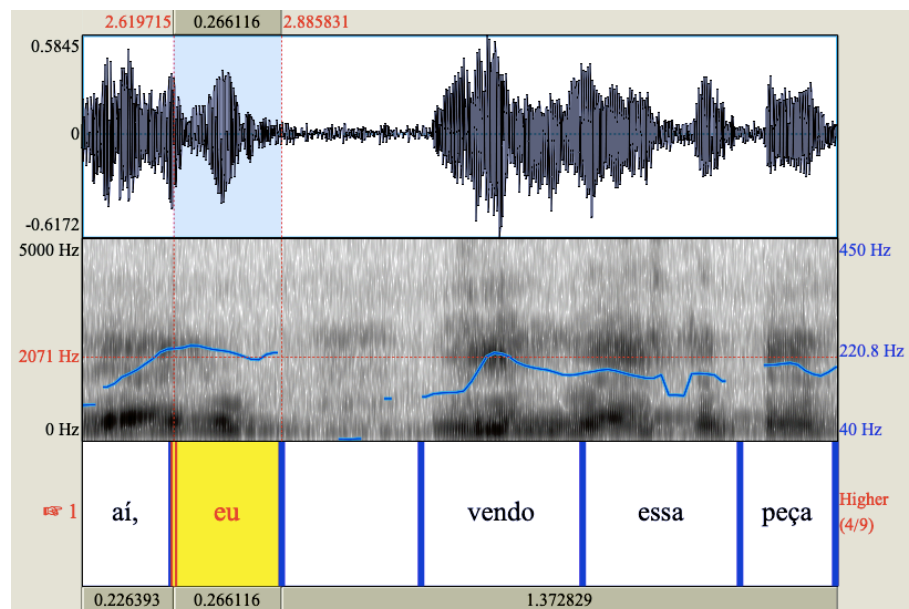


Figure 16: EP $H^*+H\%$ Pattern

às vezes as pessoas vão entregar uma peça e dizem "aí, **eu** vendo esta peça, mas não vendo por menos de xis". (CLUL – PT – 90s).

Even though the pronoun *eu* corresponds to the same exact kind of subject as H^* , the kind of contrast is slightly different, whereas in H^* the contrast occurs between the referents, in $H^*+H\%$, the contrast occurs only in the predicate. Another trait that makes the emphatic process more prominent is the fact that after expressing the subject there is a little pause marked by % in the transcription.

$L + H^*$

The pattern $L + H^*$ occurred for subjects that represents new topic and a referent change, however, in this context, there were no contrast, (5 occurrences, 21%):

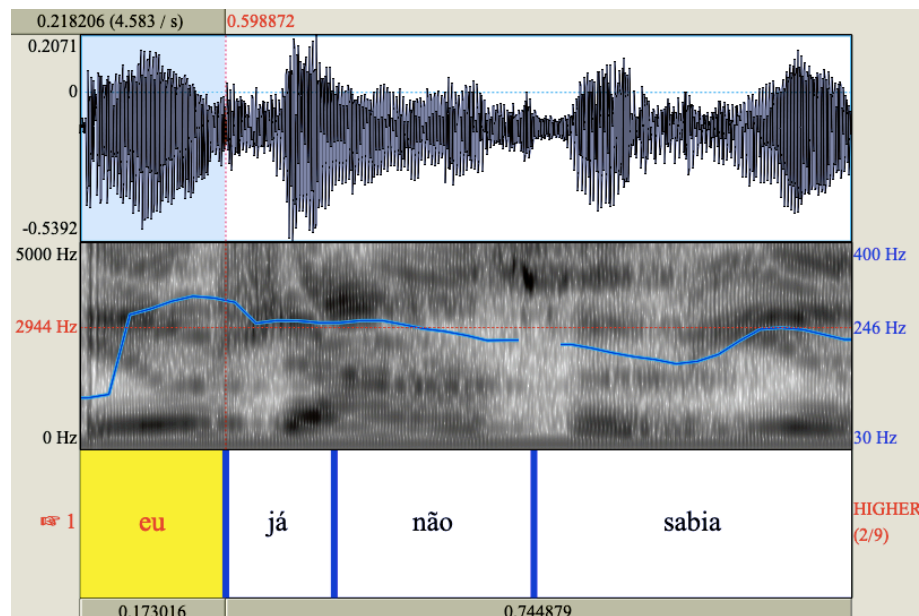


Figure 17: EP $L + H^*$ Pattern

quando a minha mãe me cha[...], chamava "Ô dona Adélia" **eu** já não sabia de "que virá por aí?" e nunca a tratei por tu, nunca na vida. (CLUL – PT – 90s).

Eu in this occurrence does not have an opposing idea, it just remarks the change of referent. Much like what happen in *BP H+H**, it seems that the occurrences with $L+H^*$ in EP were motivated by a strong feeling, in this case, fear/respect felt by the speaker toward her mother.

This same patter also happened when the subject was expressed differently from the marked structure SVO (2 occurrences, 8%):

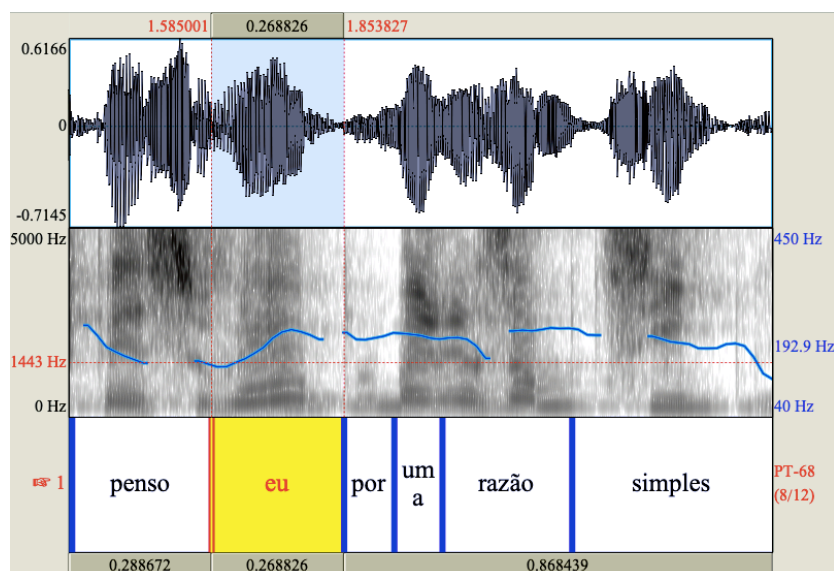


Figure 18: *EP L + H* Pattern 2*

eu penso, h· houve discussões variadas e até acesas, entre estes homens, de, para se chegar a uma conclusão, de como é que se fazia, uma carta, no século desaseis. não há motivo para discussões, penso **eu**, por uma razão muito simples (CLUL – PT – 80s).

In this case, there is three processes of emphasis on the second *eu*, the first one being the reactivation of the topic that is, by itself, and emphatic process, the second one marked by the intonational contour, the third one the dislocation to a position after the verb.

$L^*+H + H\%$

The pattern $L^*+H + H\%$ were also used when there was a reactivation of the topic, and the emphatic process happens by three different means in all the 4 occurrences (17%). This particular pattern describes a pitch that has three segments, a longer initial part with a higher pitch towards the end of the pronoun, followed by another higher pitch and a pause:

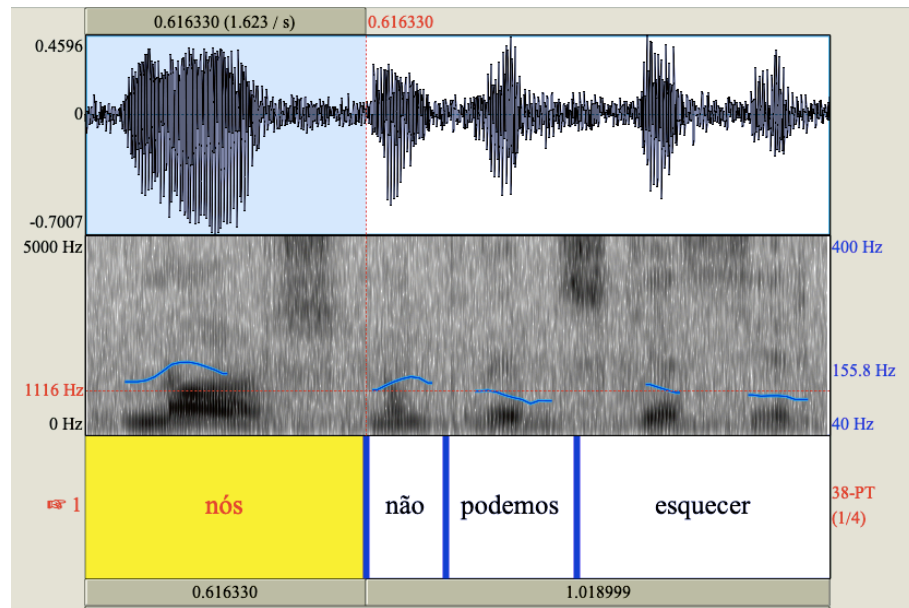


Figure 19: *EP L*+H + H% Pattern*

nós não podemos esquecer que o nosso país esteve parado no tempo durante cinquenta anos (CLUL – PT – 80s).

The first emphatic process is reactivating the topic, the second one comes from it being in the apex of the intonational curve, the last is the little pause that happens after the subject is expressed. In all the cases where this pattern happens, it seems that there are more than two emphatic processes going on.

H+L Pattern*

Another reactivation of the topic intonational contour pattern is *H*+L* (5 occurrences, 21%), this pattern, though, differs from the other reactivations patterns and is close to the change referent since it co-occurs with a contrast X and not Y, and it happens similarly to the BP *H*+L pattern*.

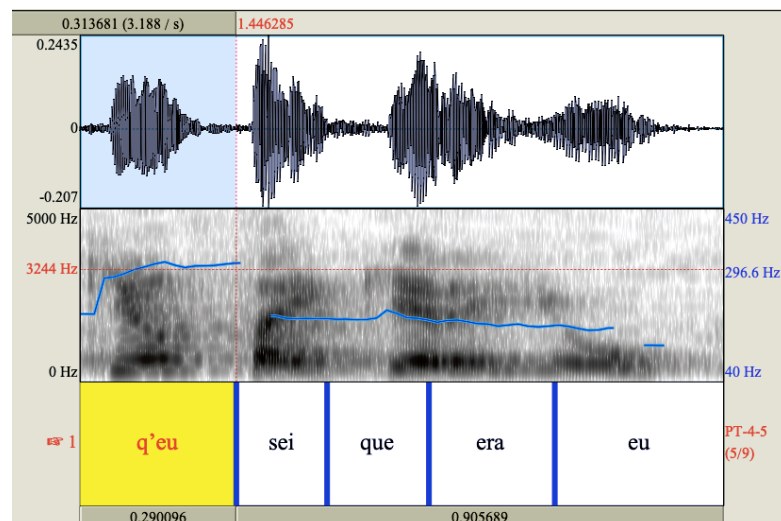


Figure 20: *EP H*+L Pattern*

era uma pessoa que **eu** sei que era eu, porque eu falava através daquela pessoa e sentia o que aquela pessoa sentia (CLUL – PT – 80s).

In the occurrence, there is a strong parallel between the speaker and the person in the speaker's dream, this contrast can be expressed like: I know it was me and not her – X and not Y.

$L + H^* + L$

The last pattern found was $L + H^* + L$, it is also a case of reactivating the subject and it happened 3 times (13%). This pattern has also three pitches, it starts lower, followed by a higher and long pitch that falls towards the ending:

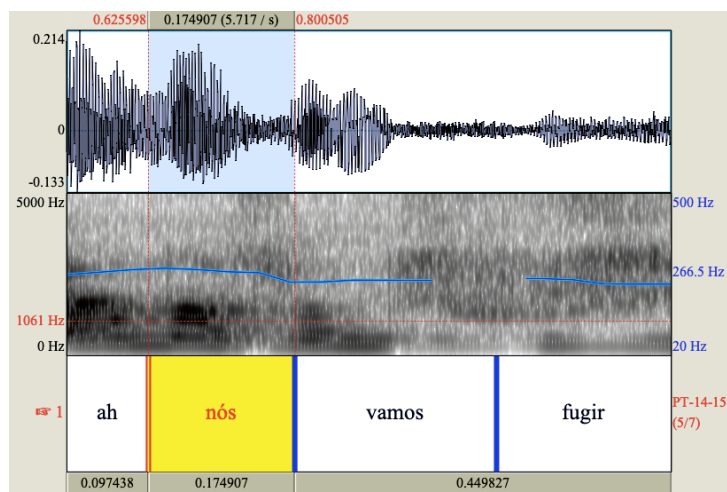


Figure 21: *EP L + H* + L Pattern*

porque elas em vez de irem para o lado da rede, estavam a i[...], estavam a ir para o outro lado. e elas disseram **ah nós** vamos fugir. (CLUL – PT – 80s).

In this occurrence, the pronoun *nós* is reactivated, *nós* are the people that were running away, besides the reactivation of the subject as an emphatic process, there is also an interjection before the pronoun showing emotion, exasperation, so, the reactivation of the subject and the expression of emotion is what possibly motivated this pattern.

EP Patterns Summary

The table below showcase a summary for EP prosody patterns and their functionality:

Function / Subject	Pattern	Quantity
Contrast Subject: Change of referent and new topic	H*	3
	H* + H%	2
No Contrast Subject: Change of referent and new topic	L + H*	5
		2
Reactivation of topic	L*+H + H%	4
	H*+ L	5
	L + H* + L	3
	Total	24

Table 18: EP Prosody Pattern for the Subject Pronoun

According to the table, EP generally uses the prosody when reactivating a topic – from the 24 occurrences, 11 where context of context reactivation. And has a littler tendency of marking contrast through prosody patter, only 5 occurrences were found.

Subject pronouns that did not have an apparent motivation to be expressed were not found in the apex of the intonational curve.

The prosody patterns seem to have a standard functionality, it means that to achieve a specific informational goal, a specific pattern will be triggered, i.e., with the speaker wants to make a contrast clear, a pattern H^* or $H^*+H\%$ would be used.

EP has a tendency of emphasizing the subject by different means in the same context: with a little pause, using the boundary ($\%$); reactivating the topic; putting the subject in the apex of the intonational curve; using words such as an interjection.

Contrasting BP and EP Patterns

With almost the same frequency in use, 11% of the data in EP corresponding to subjects pronominally expressed and prosodically emphasized, and 9% in BP, the main differences were found in the use the speakers make of the prosody.

EP uses prosody to emphasize the reactivation of a topic, it happened in 46% of the occurrences, while, in BP, only 9% of the occurrences were found in this context. Both varieties showed the same pattern for this purpose: $L^* + H$, being it the only pattern in BP, whereas in EP, it was found three more patterns: H^*+L ; $L^*+H + H\%$; $L + H^* + L$, the last two patterns were exclusively found in EP, and the first, H^*+L , in BP, has the function of marking an contrast with a pronoun with no apparent reason for its expressiveness.

In BP, the most frequent kind of pronoun was the one with no apparent reason, 51% of the data, with four different patterns – three showing contrast (H^* ; L^*+H ; H^*+L) and one using words to aid the emphasizing process ($L+H^*$). This kind of subject were not found in EP.

Contrast was another major difference between these two varieties. In BP, 68% of the time the subject was prosodically emphasized, it corresponded to contrast (X and not Y); in EP, it happened 21% of the time.

There are some differences in the patterns as well, even though the same pattern can be triggered when there is a change of referent and the introduction of a new topic while opposing to the previous referent (H^* pattern), the pattern H^*+L has this same the function in BP, but, in EP, it has no opposing meaning, and the subject is being reactivated.

Thus, despite having the same kind of subjects pronominally expressed, with the same emphatic process being available for the speaker's use, BP and EP trigger different processes to emphasize the subjects and these differences can be seen in the patterns the

intonational curve assume according to its functionality, and EP, in particular, has two three pitches pattern, while BP majorly has two pitches patterns.

10. Expressiveness of the subject: a constructional representation in Brazilian Portuguese and European Portuguese

In this final chapter, in accordance with our theoretical filiation, we are left with the task of representing the pronominal expressiveness of the subject with a constructional approach, following Goldberg (2006) proposal for the elaboration of a scheme and Traugott and Trousdale (2013) for the proposal of the network of our phenomena.

First, as general scheme, we have [(SUBJECT) V_{subject} (X)] as the basic representation of a clause, (SUBJECT) represents the possibility of expressing or omitting the subject; V_{subject} represents the verb, the predicator, followed by the inflectional suffix in agreement with the subject profiled; (X) for the option of having or not a noun phrase, a verbal phrase, or other constituent profiling the role of V complement.

Then, in a lower level, as a subschema, we have [V_{subject} (X)] for the null subject, omitting the subject element, and [(NP_{subject}) V_{subject} (X)] for the expressed subjects, (NP_{subject}) represents the possibility of filling the slot of the subject with a noun phrase or an equivalent SP – subject pronoun.

In a more detailed subschema, we introduce the kinds of subjects that can be profiled following the categories established in this work so far: *new topic; reactivation of topic; with emphatic modifier; syntactical motivation*²⁰. So, we have [(NP_{topic}) V_{subject} (X)], [(NP_{reactivation of topic}) V_{subject} (X)], [(NP_{topic}) EM V_{subject} (X)], EM being an emphatic modifier, [(NP_{syntatic motivation}) V_{subject} (X)].

It led us to several different microconstructions, here we separated than in groups:

²⁰ As we described the subjects with no apparent motivation as having, majorly, a syntactic motivation of filling the subject slot because of the new marked structure: SVO, here we labeled them according to its motivation to been expressed.

Group 1: [(NP_{topic}) V_{subject} (X)]:

Microconstruction 1:	[EU	V _{subject}	(X)]
Microconstruction 2:	[TU	V _{subject}	(X)]
Microconstruction 3:	[VOCÊ	V _{subject}	(X)]
Microconstruction 4:	[ELE/ELA	V _{subject}	(X)]
Microconstruction 5:	[NÓS	V _{subject}	(X)]
Microconstruction 6:	[A GENTE	V _{subject}	(X)]
Microconstruction 7:	[VOCÊS	V _{subject}	(X)]
Microconstruction 8:	[ELES/ELAS	V _{subject}	(X)]
Microconstruction N:	[...] ²¹		

Group 2: [(NP_{reactivation of topic}) V_{subject} (X)]

Microconstruction 1:	[EU	V _{subject}	(X)]
Microconstruction 2:	[TU	V _{subject}	(X)]
Microconstruction 3:	[VOCÊ	V _{subject}	(X)]
Microconstruction 4:	[ELE/ELA	V _{subject}	(X)]
Microconstruction 5:	[NÓS	V _{subject}	(X)]
Microconstruction 6:	[A GENTE	V _{subject}	(X)]
Microconstruction 7:	[VOCÊS	V _{subject}	(X)]
Microconstruction 8:	[ELES/ELAS	V _{subject}	(X)]
Microconstruction N:	[...]		

Group 3: [(NP_{topic}) EP V_{subject} (X)],

Microconstruction 1:	[EU	mesmo	V _{subject}	(X)]
Microconstruction 2:	[TU	mesmo	V _{subject}	(X)]
Microconstruction 3:	[VOCÊ	mesmo	V _{subject}	(X)]
Microconstruction 4:	[ELE/ELA	mesmo(a)	V _{subject}	(X)]
Microconstruction 5:	[NÓS	mesmos	V _{subject}	(X)]
Microconstruction 6:	[A GENTE	mesmo	V _{subject}	(X)]
Microconstruction 7:	[VOCÊS	mesmos	V _{subject}	(X)]
Microconstruction 8:	[ELES/ELAS	mesmos(as)	V _{subject}	(X)]
Microconstruction N:	[...]			

Group 4: [(NP_{syntactic motivation}) V_{subject} (X)].

Microconstruction 1:	[EU	V _{subject}	(X)]
Microconstruction 2:	[TU	V _{subject}	(X)]
Microconstruction 3:	[VOCÊ	V _{subject}	(X)]
Microconstruction 4:	[ELE/ELA	V _{subject}	(X)]
Microconstruction 5:	[NÓS	V _{subject}	(X)]
Microconstruction 6:	[A GENTE	V _{subject}	(X)]
Microconstruction 7:	[VOCÊS	V _{subject}	(X)]
Microconstruction 8:	[ELES/ELAS	V _{subject}	(X)]
Microconstruction N:	[...]		

²¹ The microconstructions N [...] – represents the possibility that the subject to be profiled by a full NP.

We can still add more layers to its representation if we consider, for example, $[(NP_{topic}) EP V_{subject}(X)]$, with bold (NP_{topic}) representing prosodically emphasized subject pronoun and, we can go even further with we consider that this same subject can be represented, in EP, as H^* but also as $H^* + H\%$, we would have a new set of groups, to be more easily visualized, we used just the first person as example:

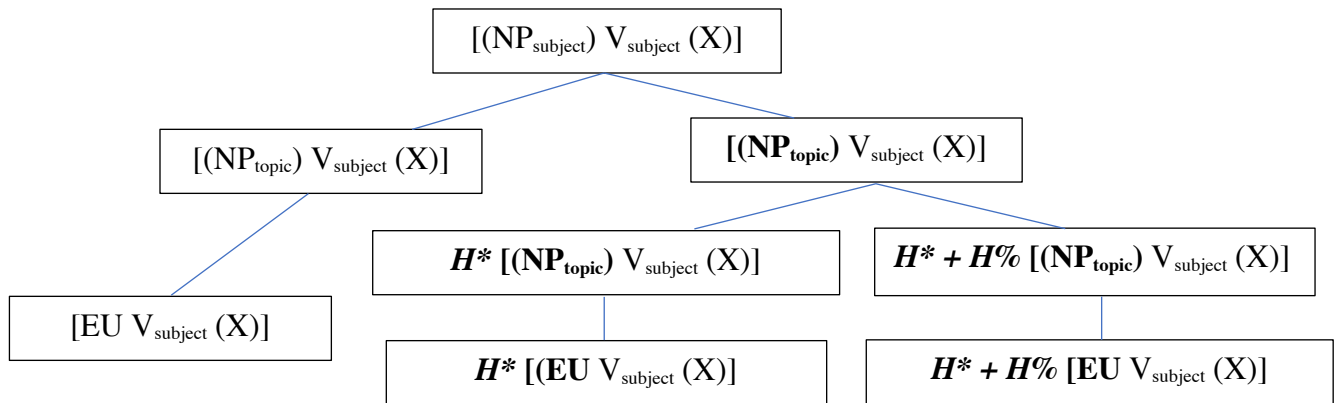


Figure 22: Expressiveness of the subject network

Still part of the schema would be the one used for the constructional representation of the null subject $[V_{subject}(X)]$, the $subject$ part having its inflections specified, roughly, something like:

- Microconstruction 1: $[V1sg(X)]$
- Microconstruction 2: $[V-2sg(X)]$
- Microconstruction 3: $[V-3sg(X)]$
- Microconstruction 4: $[V-1pl(X)]$
- Microconstruction 5: $[V-2pl(X)]$
- Microconstruction 6: $[V-3pl(X)]$

BP and EP have the same constructional scheme for the subject representation, being it expressed or not, we only ruled out the microconstruction $[V\acute{O}S V_{subject}(X)]$ because we did not find any occurrence like it in neither BP nor EP. The difference between these two varieties, in the constructional network, is only visible in the microconstructions with the subject prosodically emphasized since there is different intonational curves for each kind of subject.

CONSIDERATIONS AND FORTHCOMING WORKS

This thesis started with the will to investigate the status *pro-drop* in both Brazilian Portuguese and European Portuguese, but, by understanding a little bit better about the null subject, we realized that investigating how the speakers of both varieties are dealing with the necessity of expressing somehow the subject was more productive and could lead us to new findings, hence, we chose to investigate the subject expressed via a subject pronoun, referent grammar class by default.

Our primer goal was to verify in which point both varieties diverged regarding a syntactic phenomenon, however, throughout our analysis, we found out that the divergence was not only an issue of word order or omitting or explicating an element, expressing the subject pronoun held its roots deeper and linked to all six linguistic levels considered by Construction Grammar, since in its phonological level to its discursive one. And the divergences were between BP and EP were tangible in more than one of them, but mainly in the phonological one.

Thanks to the Constructional Grammar postulates, by proposing a schematic representation to the expressiveness of the subject pronoun, we understood that the main contrast between BP and EP is not whether the expresses it or not, but the motivations that led the speakers to do it.

Regarding the intonational curve, we realized that BP speakers generally have a tendency of emphasizing the subject pronouns that are motivated by the syntax to be expressed, the previously called, *subjects with no apparent motivation*; while EP speakers generally emphasize the subject pronouns that profiled the role of new topic. Also related to the intonational curve pattern, we discovered that the same intonational pattern is, sometimes, used in both varieties, but with different reasons, BP usually reactive a subject prosodically marked with the pattern $L^* + H$, in EP, there are four different patterns to represent the same kind of subject: $L + H^*$; $L^* + H + H\%$; $H^* + L$; $L + H^* + L$. Therefore, our first hypothesis was proved to be in the right direction.

We also expected to find more occurrences of pronominally expressed subjects in BP than in EP, BP indeed showed a higher frequency of occurrences, 608 against 212 in EP; the SVO structure with the subject having to be mandatory, or else the communicational event would be impaired, is becoming the non-marked structure for simple clauses, to an extent that even when there was no necessity to express the subject,

for instance, specific verbal inflectional suffix for the 1st singular person, the speakers express it anyway.

It explains, along with the impossibility of expressing the 1st person of singular via a NP, why the 1st singular person was the subject more frequently expressed via a pronoun, even though it maintains its inflectional system unaltered.

It was also possible to define what motivated the subject Olbertz (2020) calls *pronoun with no apparent motivation*. The organization of the language is, although chaotic, organized, and through an analogy process, to save cognitive energy, it tends to simplify its structures, therefore, if, in most contexts, there is a necessity to express the subject, via a syntactic process, for example, then, let it be the general rule.

Regarding the cognitive processes, we understand that although there is more linguistic material when expressing the subject, there is less cognitive complexity while doing so, it is related not only to the linguistic economy principle, reducing the verb agreement suffixes, but it also related to the iconicity principle, just explicitating the subject can no longer make it emphatic, it can still be considered iconic, since different strategies can be used to accomplish this function, such as prosody, repetition of the subject or even using the and emphatic modifier, the last two cases most related to the subprinciple of quantity.

Other factor that is motivating the expressiveness of the pronoun is related to the perspective and informativity, there is, due the simplification of the pronoun paradigm, both in BP and EP, an informational need to make the subject explicit, mainly to not impair the communicational process.

Nonetheless, the main factor that motivated the process of expressing the subject is certainly analogy, through the process of analogization, the pronoun paradigm underwent the simplification process, and this process can be furtherer in years to come making Portuguese more transparent, having only one exclusive desinential form for the 1st singular person and a syncretical form for all the other ones, which can lead to an even stronger motivation to express the subject via a NP, but mostly, via a pronoun.

These findings are important to understand that BP and EP are following the same routes in terms of expressing the subject via pronoun and how Portuguese, in general, is being impacted. Having, now, each time fewer contexts in which there is a possibility to use a *null subject* to express an event a new non-marked form SVO arising and being consolidated, how grammar is being taught needs to be revisited.

For future research, it is still necessary to investigate how other Portuguese varieties, such as the African variety or the Macao variety, (varieties also present in the CLUL project, Portuguese Spoken –, *Português Falado - Variedades Geográficas e Sociais*) have expressed the subject, via an inflectional suffix or with double marking, pronoun + inflection. The expansion of this research can ensure whether the phenomenon of expressiveness of the subject pronoun is a particularity of BP and EP or whether it is a path that Portuguese, in general, is going through.

REFERENCES

- ANGOURI, J. Quantitative, qualitative, mixed or holistic research? Combining methods in linguistic research. in LITOSSELITI, L. *Research Methods in Linguistics*, 2a. London: Bloomsbury, 2018.
- BARBOSA, P.; DUARTE, M. E. L. e KATO, M. A. Null subjects. in European and Brazilian Portuguese, *Journal of Portuguese Linguistics*, v. 4 (2), 2005, p. 11-52.
- BARROS, D. M. *Um estudo pancrônico da voz reflexiva em perspectiva construcional*. Tese de doutorado. 2016, 177. Programa de Pós-graduação em Letras e Linguística, Universidade Federal de Goiás, Goiânia, 2016.
- BOERSMA, P; WEENINK, D. *Praat: Doing phonetics by computer* [Computer program]. 2017, Version 6.1.54. <http://www.praat.org/>
- BYBEE, J. *Language, usage and cognition*. Cambridge: Cambridge University Press, 2010.
- BYBEE, J. *Language change*. Cambridge: Cambridge University Press, 2015.
- CAGLIARI, L. C. Prosódia: algumas funções dos supra-segmentos. in *Cad. Est. Ling.*, Campinas, Brasil, Jul/Dez, 1992, p. 135-150.
- CAGLIARI, L. C; MASSINI-CAGLIARI, G. O papel da tessitura dentro da prosódia portuguesa. in CASTRO, I; DUARTE, I (Orgs.), *Razões e emoções: miscelânea de estudos em homenagem a Maria Helena Mira Mateus*. Lisboa: Imprensa Nacional Casas da Moeda, 2003, p. 65-85.
- CASSEB-GALVÃO, V. C. *Evidencialidade e gramaticalização no português do Brasil: os usos da expressão diz que*. Tese de doutorado. 2001. p.231. Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara. 2001.
- CASSEB-GALVÃO, V. C. A função sujeito: definição, tratamento tradicional e funcionalidade. in CASSEB GALVÃO, V. C; et al (orgs.) *O sujeito gramatical no português brasileiro: expressão, concordância, ergatividade e afetamento*. 1ed. São Paulo : Parábola, 2020, p.47-59.

CARVALHO, G. A.; *A realização do sujeito na fala do português europeu. Tese de doutorado.* 2009. 259. Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara. 2009.

CARDINALETTI, A., & STARKE, M. The typology of structural deficiency: A case study of the three classes of pronouns. in *Clitics in the Languages of Europe*, H. van Riemsdijk (ed.), 1999, p.145-233.

CHOMSKY, N. *Lectures on government and binding*, Foris Publications, Dordrecht, 1981.

CRUSE, D. A. *A Glossary of Semantics and Pragmatics*. George Square, Edinburgh: Edinburgh University Press Ltd, 2004.

CROFT, W. *Radical Construction Grammar. Syntactic Theory in Typological Perspective* – Oxford Press, New York, 2001.

CROFT, W.; CRUSE, D. A. *Cognitive linguistics*. New York: Cambridge, 2004.

CYRINO S., DUARTE M.E.L. E KATO M.A., Visible subjects and invisible clitics in Brazilian Portuguese, in KATO M.A. E NEGRÃO E.V. (orgs.), *Brazilian Portuguese and the null subject parameter*, Vervuert/Iberoamericana, Frankfurt am Main/Madrid, 2000, p.55-105.

DE ROSA, G. L.; Os sujeitos de 3ª pessoa na fala fílmica brasileira. in: *Lingue Linguaggi*, vol. 32, 2019, p.225-245.

DIK, S. C. *The theory of Functional Grammar*. Dordrecht: Foris Publications, 1989.

DUARTE M.E.L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. in ROBERTS I. AND KATO M.A. (orgs.), *Português brasileiro: Uma viagem diacrônica. Homenagem a Fernando Tarallo*, Editora da Unicamp, Campinas, 1993, p. 107-128.

DUARTE M.E.L. *A perda do princípio Evite Pronome no português brasileiro* [Tese de Doutorado], IEL/UNICAMP, Campinas. 1995.

DUARTE M.E.L., The loss of the 'avoid pronoun' principle in Brazilian Portuguese. in KATO M.A. E NEGRÃO E.V. (orgs.), *Brazilian Portuguese and the null subject parameter*, Vervuert/Iberoamericana, Frankfurt am Main/Madrid, 2000, p. 17-36.

DUARTE M.E.L., On the 'embedding' of a syntactic change. in Gunnarsson B.L. et al, *Language Variation in Europe. Papers from ICLa VE2 - Second International Conference on Language Variation in English*, Universitetsstryckeriet, Uppsala, 2004, p. 145-155.

DUARTE M.E.L. Sujeito Nulo/Pleno e marcas de concordância. in VOTRE S. E RONCARATI C. (eds.), *Anthony Julius Naro e a Linguística no Brasil. Uma homenagem acadêmica*, 7, Letras, Rio de Janeiro, 2008, p. 265-277.

DUARTE M.E.L. Construções de topicalização. in RAPOSO, E. B. P. et al. *Gramática do português*. Lisboa: Fundação Calouste Gulbenkian, 2018. v. 1, p. 401-426.

DUARTE, M. E. L.; KATO, M. A.; BARBOSA, P. Sujeitos indeterminados em PE e PB. in *Anais do II Congresso Internacional da ABRAOLIN*, vol. I, p. 405-409, 2001.

FANT, G. (1960): *Acoustic theory of speech production*. Mouton, The Hague, 1960.

FILLIPPI, P. Emotional Voice Intonation: A Communication Code at the Origins of Speech Processing and Word-Meaning Associations?. in *J Nonverbal Behav* 44, 395-417 (2020). <https://doi.org/10.1007/s10919-020-00337-z>.

FRIED M. Principles of Constructional Change. In: HOFFMANN, T. TROUSDALE, G. (Eds) *The Oxford handbook of construction grammar*. Oxford Press, New York : 2013.

FROTA, S., et al. (Dys)Prosody in Parkinson's Disease: Effects of Medication and Disease Duration on Intonation and Prosodic Phrasing. in *Brain Sciences*, 11(8), 1100, 2021.

FROTA, S., ET al. Intonational variation in Portuguese: European and Brazilian varieties. in S. FROTA & P. PRIETO (eds.), *Intonation in Romance*. Oxford: Oxford University Press, 2015.

FURTADO DA CUNHA, M. A. et al. Linguística Funcional Centrada no Uso: conceitos básicos e categorias analíticas. in CEZARIO, M. C.; FURTADO DA CUNHA, M. A. *Linguística Centrada no Uso: uma homenagem a Mário Martelotta*. Rio de Janeiro: Mauad X FAPERJ, 2013m p. 13-40.

GARCÍA, A. L. *Gramática del español. II. La oración simple*. Arco/Libros: Madrid, 1996.

GALVES C.M.C. O enfraquecimento da concordância no português brasileiro. in ROBERTS I.; KATO M.A. (orgs.), *Português brasileiro: Uma viagem diacrônica. Homenagem a Fernando Tarallo*, Editora da Unicamp, Campinas, 1993. p. 387-408.

GILI FIVELA, B. Tonal alignment in two Pisa Italian peak accents. in *proc. Speech Prosody*, Aix-en-Pr., 2002, p. 339-342,

GILI FIVELA, B. The coding of target alignment and scaling in pitch accent transcription. in *Italian Journal of Linguistics*, vol. 18, 2006, p. 189-221.

GILI FIVELA, B. Broad focus vs contrastive focus: Is there categorical perception in Pisa Italian? in BARBOSA, P. A.; MADUREIRA, S.; REIS C. (eds.), *Atti del Convegno Internazionale Speech Prosody*, Campinas, Brasil, 2008, p.293-296.

GILI FIVELA, B. Multimodal analyses of audio-visual information: Some methods and issues in prosody research. in FELDHAUSEN, I; FLIESSBACH, J; VANRELL, M. M. (eds.), *Methods in prosody: A Romance language perspective*, Berlin: Language Science Press, 2018, p. 83–122

GIVÓN, T. *Syntax: a functional-typological introduction*. v. 1. Amsterdam/Philadelphia: John Benjamins, 1984.

GIVÓN, T. Distributional, Communicative and Cognitive Correlates of Syntactic Structure. in *Studies in Language*.v. 15:2, 1991, p. 335–370.

GIVÓN, T. *Functionalism and grammar*. Amsterdam/ Philadelphia: John Benjamins, 1995.

GIVÓN, T. *Syntax: a functional-typological introduction*. v. 2. Amsterdam/Philadelphia: John Benjamins, 2001.

GOLDBERG, A. E. *A construction grammar approach to argument structure*. Chicago: University of Chicago Press, 1995.

GOLDBERG, A. E. *Constructions at work: The nature of generalization in language*. New York: Oxford University Press, 2006.

GOLDBERG, A. E. *Explain me this: creativity, competition, and the partial productivity of constructions*. New Jersey: Princeton University Press, 2019.

GONÇALVES, S.C.L. et al. Tratado Geral sobre gramaticalização. in CASSEB-GALVÃO; LIMA-HERNADES, M.C.; GONÇALVES, S.C.L. *Introdução à gramaticalização*. São Paulo: Parábola, 2007, p. 15-66.

GONÇALVES, S. C. L.; FARIA, M. H. V. Mudança construcional pós-construcionalização de A GENTE e VOCÊ. In *Estudos Linguísticos* (São Paulo), v. 50, 2021, p. 207-226.

GUERRA, A. R. Variação diacrônica do grau de transparência linguística do português brasileiro. in *Cadernos de Estudos Linguísticos*, v. 63, 2021.

GUERRA, A. R. *Diacronia do grau de transparência do sistema de referência por expressão pronominal e desinencial do argumento-sujeito de 1ª e 2ª pessoas no português brasileiro*. Tese de doutorado. 2017, 170. Programa de Pós-Graduação em Estudos Linguísticos do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto. 2017.

HAIMAN, J. Iconic and economic motivation. in *Language*. 59, 1983, p. 781-819.

HENGEVELD, K; LEUFKENS, S. *Transparent and non-transparent languages*. in *Folia Linguistica* 52(1), 2018, p.139-175.

HOLMBERG, A.; NAYUDU, A. & SHEEHAN, M. Three partial null subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. in *Studia Linguistica*, v.63(1), p. 59-97. 2009.

KATO, M. A. Strong pronouns and weak pronominals in the null subject parameter. in *Probus*, vol. 11(1), 1999, p.1-37.

KATO, M. A. A evolução da noção de parâmetros. in *D.E.L.T.A.*, vol.18(1), p.300-337, 2002.

KATO M.A; DUARTE M.E.L. A Variação entre construções finitas pessoais e impessoais no português brasileiro, in *Sociodialeto*, vol. 4, p.153-177, 2014.

KAISER, G. A. Losing the null subject: a contrastive study of (Brazilian) Portuguese and (Medieval) French. In: Kaiser, G.A; Remberger, E.-M. (eds.), *Proceedings of the Workshop "Null-subjects, expletives, and locatives in Romance"*. Arbeitspapier 123. Fachbereich Sprachwissenschaft, Universität Konstanz, 2009, p.131-156.

KLATT, D.H; KLATT L.C. Analysis, synthesis and perception of voice quality variations among male and female talkers. in *Journal of the Acoustical Society of America* 87, 1990, p. 820–856.

LABOV, W. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell, 1994.

LAKOFF, G. *Women, Fire and Dangerous Things*. What Categories Reveal about the Mind. Chicago: University of Chicago Press, 1987.

LANGACKER, R. W. *Foundations of Cognitive Grammar*. v. 1. Theoretical prerequisites. Stanford, Cal.: Sanford University Press, 1987.

LANGACKER, R. W. *Cognitive Grammar: a basic introduction*. Oxford University Press, 2008

LANGACKER, R. W. *Investigations in Cognitive Grammar*. Mouton de Gruyter Berlin, New York, 2009

LANGACKER, R. W. *Essentials of Cognitive Grammar*. Oxford: Oxford University Press, 2013.

LEUFENS, S. The transparency of creoles. in *Journal of Pidgin and Creole Languages*, Amsterdam, v. 28, n. 2, 2013, p. 323-362.

LOBO, M. Sujeito nulo: sintaxe e interpretação. in RAPOSO, E. B. P. et al. *Gramática do português*. Lisboa: Fundação Calouste Gulbenkian, v. 2, 2018. p. 2309-2335.

MARTELOTTA, M.E.; VOTRE, S. CEZÁRIO, M.C. *Gramaticalização no português do Brasil: uma abordagem funcional*, 1996.

MARTELOTTA, M. E. *Mudança linguística: uma abordagem baseada no uso*. São Paulo: Cortez, 2011.

NADER, K., G.E. Schafe, and J.E. le Doux. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. In *Nature* 406: 722–6, 2000.

NEVES, M. H. M. *A gramática funcional*. São Paulo: Martins Fontes, 1997.

NEVES, M. H. M. *A gramática: história, teoria e análise, ensino*. São Paulo: Ed. UNESP, 2002.

NEVES, M. H. M. *Gramática de usos do português*. 2. ed. atualizada. São Paulo: Ed. UNESP, 2011.

NEVES, M. H. M. *A gramática do português revelada em textos*. São Paulo: Ed. Unesp, 2018.

NEVES, M. H. M; GOULART, F. V; A exclusividade desinencial e a realização do sujeito ‘eu’. in *Acta Scientiarum Language and Culture*. Maringá, v. 39, n. 4, 2017, p. 335-344.

OLBERTZ, H. A perda do sujeito nulo no português do Brasil: uma abordagem discursivo-funcional. in CASSEB GALVÃO, V. C; et al (orgs.) *O sujeito gramatical no português brasileiro: expressão, concordância, ergatividade e afetamento*. 1ed. São Paulo : Parábola, 2020, p.15-43.

OLIVEIRA, T. P. Subject expression in Brazilian Portuguese. in Keizer, E; Olbertz, H. (Orgs.). *Recent Developments in Functional Discourse Grammar*. Amsterdam: John Benjamins Publishing, 2018, p. 208-232.

PAREDES DA SILVA, V. L. *Cartas cariocas: a variação do pronome na escrita informal*. Tese de Doutorado, 1988, p.303, Rio de Janeiro, UFRJ. 1988.

PIETRO, P; ROSEANO, P. Prosody: Stress, rhythm, and intonation. in *The Cambridge Handbook of Spanish Linguistics*, Cambridge University Press, 2010, p.211-236.

PULVERMÜLLER, F. CAPPELLE, B; SHYROV Y. Brain Basis of Meaning, Words, Constructions, and Grammar. in HOFFMANN, T. TROUSDALE, G. (Eds) *The Oxford handbook of construction grammar*. Oxford Press, New York, 2013.

RAMUS, F; NESPOR, M; MEHLER, J. Correlates of linguistic rhythm in the speech signal. in *Cognition*. ed. 73, 1999 p.265-292.

RIZZI, L. *The new comparative syntax: principles and parameters of universal grammar*. ms. 1998.

ROBERTS. I. O português brasileiro no contexto das línguas românicas. in ROBERTS, I; KATO, M. A. (orgs.) *Português Brasileiro*. Campinas: Ed. da UNICAMP. 1993, p.409-42.

RAPOSO, E. B. P. Estrutura da frase. in RAPOSO, E. B. P. et al. *Gramática do português*. Lisboa: Fundação Calouste Gulbenkian, v.1, 2018. p. 303-394.

RAPOSO, E. B. P; MIGUEL, M. Introdução ao sintagma nominal. in RAPOSO, E. B. P. et al. *Gramática do português*. Lisboa: Fundação Calouste Gulbenkian, v.1, 2018. p. 703-731.

ROSÁRIO, I. C. do, Gramática, gramaticalização, construções e integração oracional: algumas reflexões, in OLIVEIRA, M. R; ROSÁRIO, I. C. (org). *Linguística centrada no uso: teoria e método*. Rio de Janeiro: Lamparina, 2015, p. 36-50.

RUBIO, C.F. *Padrões de concordância verbal e de alternância pronominal no português brasileiro e europeu: estudo sociolinguístico comparativo*. Tese de doutorado. 2012, 393.

Programa de Pós-Graduação em Estudos Linguísticos do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto. 2012.

SCARPA, E; FERNANDES-SVARTMAN, F. Entoação e léxico inicial. in *Veredas Online*, edição especial, 2002, p. 40-54.

SOARES DA SILVA, H. *O parâmetro do sujeito nulo: confronto entre o português e o espanhol*. 2006. 196f. Dissertação (Mestrado em Letras Vernáculas – área de Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

STEELS, L. Fluid Construction Grammar. in HOFFMANN, T. TROUSDALE, G. (Eds) *The Oxford handbook of construction grammar*. Oxford Press, New York : 2013.

TARALLO, F. Diagnosticando uma gramática brasileira: o português d’aquém e d’além-mar ao final do século XIX, In ROBERTS I; KATO M.A. (eds.), *Português Brasileiro: Uma viagem diacrônica – Homenagem a Fernando Tarallo*, Campinas: Editora da Unicamp, 1993, p.69-106.

TOMASELLO, M. *Origins of Human Communication* (Jean Nicod Lectures). Cambridge, MA: MIT Press, 2008

TOMASELLO, M. *Why we cooperate*. Cambridge: MIT Press, 2009.

TRAUGOTT, E. C; TROUSDALE, G. *Constructionalization and constructional changes*. Oxford: Oxford University Press, 2013.

TRAUGOTT, C.T. Toward a coherent account of grammatical Constructionalization. in BARÐDAL, J. (at. al), *Diacronic construction grammar*. Amesterdão: John Benjamins 2015. p. 51-79.

WILSON, V.; MARTELOTTA, M. E. Arbitrariedade e iconicidade. in MARTELOTTA, M. E. (org.). *Manual de linguística*. São Paulo: Contexto, 2009, p. 71-85.

YAO, J. A syntactic account of resultative variation. in *23rd Seoul International Conference on Generative Grammar*. Seoul: Hankook Munhwasa., 2021, p.217-229.