

Revista **Linguística**

Volume 17, número 3, set./dez. de 2021

Organizadores da edição:

Isabella Lopes Pederneira e
Thiago Laurentino de Oliveira

O presente número reúne artigos sobre fenômenos gramaticais sincrônicos e diacrônicos de diferentes níveis linguísticos, orientados por perspectivas teóricas variadas, e fornece um panorama representativo dos estudos linguísticos no Brasil.

ISSN: 2238-975X

UFRJ

COMISSÃO EDITORIAL

Editor-chefe

Aleria Lage, Universidade Federal do Rio de Janeiro, Brasil

Editores

Aniela Improta França, Universidade Federal do Rio de Janeiro, Brasil

Isabella Lopes Pederneira, Universidade Federal do Rio de Janeiro, Brasil

Marcus Maia, Universidade Federal do Rio de Janeiro, Brasil

Marije Soto, Universidade Estadual do Rio de Janeiro, Brasil

Editores Honorários

Maria Luiza Braga, Universidade Federal do Rio de Janeiro, Brasil

Editor Fundador

Lilian Ferrari, Universidade Federal do Rio de Janeiro, Brasil

Conselho Editorial

Anthony Naro, Universidade Federal do Rio de Janeiro, Brasil

Fernanda Ferreira Spoladore, Michigan State University, EUA

Gabriela Matos, Universidade de Lisboa, Portugal

Kees Hegenveld, Universidade de Amsterdam, Holanda

Leticia Sicuro Correa, Departamento de Letras - PUC/RJ

Leo Wetzels, Universidade Livre de Amsterdam, Holanda

Luiz Amaral, University of Massachusetts, USA

Maria Armanda Costa, Universidade de Lisboa, Portugal

Maria Cecília Mollica, Universidade Federal do Rio de Janeiro, Brasil

Roberto Almeida, Concordia University, Canada

Ruth Lopes - Universidade Estadual de Campinas, Brasil

Organização da Edição

Isabella Lopes Pederneira e Thiago Laurentino de Oliveira

Redação e Assinaturas

Programa de Pós-Graduação em Linguística - Faculdade de Letras da UFRJ

Av. Horácio de Macedo, 2151 - Sala F. 321. Ilha do Fundão Cidade Universitária

CEP 21941-917 - Rio de Janeiro - RJ. E-mail: ppglinguistica@letras.ufrj.br

Editor Operacional

Elir Ferrari, Editorarte /UERJ

Editores Eletrônica

Patricia Mabel Kelly Ramos, pattydesign/Brasil

Revista Linguística

RIO DE JANEIRO | VOLUME 17 | NÚMERO 3 | SET. - DEZ. DE 2021

ESTUDOS GRAMATICAIS SINCRÔNICOS E DIACRÔNICOS | *SYNCHRONIC AND DIACHRONIC GRAMMATICAL STUDIES*

UFRJ

Programa de Pós-Graduação em Linguística Faculdade de Letras
Universidade Federal do Rio de Janeiro

SUMÁRIO

Apresentação

- 6** **Estudos gramaticais sincrônicos e diacrônicos**
Isabella Lopes Pederneira e Thiago Laurentino de Oliveira

Entrevista

- 12** **Interview with Hagit Borer for Revista Linguística**
Isabella Lopes Pederneira

Artigos

- 38** **A estrutura argumental e a polissemia em nominalizações**
Rafaela do Nascimento Melo Aquino
- 60** **O comportamento dos pronomes possessivos seu(s); sua(s); dele(s) e dela(s) na recuperação de seus antecedentes**
Bruna Clara Santos de Almeida e Rafael Dias Minussi
- 82** **A cartografia dos pronomes resumptivos pessoais da construção de sujeito duplo do Português Brasileiro: uma proposta sincrônica**
Letícia Emília Kriek
- 100** **Contexts for null subjects in contemporary Brazilian Portuguese**
Mônica Rigo Ayres e Gabriel de Ávila Othero
- 125** **Restrições semânticas de animacidade do objeto nulo do Português Brasileiro**
Gabriel Guimarães Peixoto da Silva
- 150** **Pronomes plenos e clíticos pronominais no português brasileiro e no francês**
Luciano de Oliveira
- 174** **Se-passivo, se-impessoal e sujeitos nulos indefinidos: uma abordagem minimalista para a perda e o surgimento de construções na gramática da língua**
Humberto Borges
- 202** **Reconstruindo a história do Português do Brasil pelo Corpus Tycho Brahe Brasil: novos dados, novos olhares**
Paulo Ângelo Araújo-Adriano e Williane Silva Corôa

- 228** **Uma análise diacrônica das formas pronominais contigo e com você em cartas pessoais**
Thaissa Frota Teixeira de Araujo Silva
- 245** **Expressões de tempo decorrido em cartas pessoais do alagoano Graciliano Ramos**
Pedro Henrique Sousa dos Santos e Elyne Giselle de Santana Lima Aguiar Vitória
- 262** **Gramaticalização do verbo visar**
Geovana Portela de Moura, Célia Márcia Gonçalves Nunes Lôbo e Lúcia Regiane Lopes-Damasio
- 283** **Da emergência de formas compostas para a expressão do antepretérito: construtos morfossintáticos e semânticos**
Márluce Coan, Francisco José Gomes de Sousa e Laila Cavalcante Romualdo
- 300** **A construção X-ada: um caso de construcionalização**
João Carlos Tavares
- 320** **Homenagem à Nossa Miriam Lemle**
Isabella Lopes Pederneira

ESTUDOS GRAMATICAIS SINCRÔNICOS E DIACRÔNICOS
SYNCHRONIC AND DIACHRONIC GRAMMATICAL STUDIES

Organização: Isabella Lopes Pederneira¹

Thiago Laurentino de Oliveira²

Nesta edição da Revista *Linguística* (v. 17, n. 3), o nosso objetivo foi reunir artigos em que fossem discutidos fenômenos gramaticais relacionados aos níveis fonológico, morfológico, morfossintático e/ou sintático das línguas naturais, tanto do ponto de vista sincrônico quanto diacrônico. Outro ponto importante deste número foi a sua abrangência teórica. Graças a essa abertura quanto à perspectiva de análise, pudemos receber excelentes artigos cuja fundamentação teórica está baseada tanto em teorias formais quanto em teorias baseadas no uso, ambas a partir de diferentes vertentes de modelo gramatical. Esta característica nos fornece um panorama representativo sobre os estudos linguísticos no Brasil.

Este número procurou dar visibilidade à relação entre teorias linguísticas e suas respectivas explicações para os fenômenos linguísticos abordados, fossem eles sincrônicos ou diacrônicos. Os leitores observarão o quão diversificados são os fenômenos analisados e suas respectivas abordagens teóricas. Desse modo, colocamos em questão também, novamente, a clássica dicotomia saussureana, proposta no âmbito do estruturalismo linguístico mas constantemente retomada e redefinida por linguistas de orientações variadas, até os dias de hoje. Para Saussure (2021), a sincronia deveria ser preconizada, em detrimento do estudo diacrônico, devido ao fato de que é somente através da descrição do estado de língua que se pode chegar às diferentes relações pertinentes à gramática, que, no estruturalismo, era identificada como o sistema linguístico. No entanto, considerando o ponto de vista morfológico, morfossintático, sintático e suas interfaces, a análise culmina em alguns desafios e limites para explicações linguísticas. Deste modo, a linguística moderna tem se mostrado aberta também às análises diacrônicas para descrever, analisar e explicar mecanismos de mudança no âmbito das palavras, sintagmas e sentenças. Neste número, os leitores estão convidados ao debate democrático e respeitoso a partir de uma entrevista e treze artigos de alta qualidade. Esperamos que a leitura dos trabalhos aqui reunidos suscite muitos debates para o avanço das discussões e aprimoramento dos estudos linguísticos, imperativos para qualquer ciência.

¹ Universidade Federal do Rio de Janeiro (UFRJ), isbellapederneira@letras.ufrj.br, <https://orcid.org/0000-0002-5884-8472>.

² Universidade Federal do Rio de Janeiro (UFRJ), thiagolaurentino@letras.ufrj.br, <https://orcid.org/0000-0002-9537-5264>.

A entrevista realizada por uma das organizadoras deste número, Isabella Pederneira, e aqui apresentada foi com a renomada professora e pesquisadora Hagit Borer, atualmente vinculada à universidade de Londres - QMUL/Queen Mary University of London, ex-orientanda de Noam Chomsky no MIT e criadora do modelo Exoesqueletal de Gramática Gerativa (BORER, 2005a e b, 2013). A entrevista, como poderão apreciar, conta com mais de trinta páginas através das quais é possível fazer uma viagem pela Linguística Gerativa, com ênfase em modelos construcionistas de Gramática Gerativa, mais precisamente no modelo criado pela entrevistada, acima denominado. Hagit apresenta aos leitores não só seu conhecimento aprofundado sobre o Gerativismo, mas também sua generosidade em compartilhar conhecimento. Além de temas essenciais do cômputo teórico, Borer ainda dividiu com os leitores um pouco de seu posicionamento acerca do tema específico de nosso número - não só sobre sincronia, mas também sobre estudos diacrônicos.

No primeiro artigo deste número, a autora Rafaela do Nascimento Melo Aquino aborda a polissemia de nominalizações justamente a partir do modelo de gramática da Exoesqueletal, desenvolvido pela nossa ilustre entrevistada deste número. Assim, os leitores terão a oportunidade de observar o modelo sendo aplicado para analisar e explicar dados de nominalizações em que o mesmo item fonológico possui diferentes significados no português brasileiro. Na trilogia de livros desenvolvidos por Hagit Borer (2005a e b, 2013a), as nominalizações são o primeiro fenômeno abordado pela autora. Este trabalho da Rafaela possui ainda outra especificidade que é a de ser resultado da tese da autora, última a ser defendida sob orientação de nossa saudosa professora e pesquisadora Miriam Lemle.

O artigo “O comportamento dos pronomes possessivos seus(s); suas(s); dele(s) e dela(s) na recuperação de seus antecedentes”, de Bruna Clara Santos de Almeida e Rafael Dias Minussi, traz importantes contribuições para o tema sob a perspectiva da Morfologia Distribuída (HALLE; MARANTZ, 1993), abordagem que inaugura o construcionismo em Gramática Gerativa de maneira mais ampla e que também é bastante citada na entrevista concedida por Hagit Borer. Os dois modelos construcionistas possuem muito em comum, mas há algumas diferenças importantes, pontuadas na entrevista. É interessante observar que os dois modelos tenham sido fontes de fundamentações teóricas deste número, tendo os leitores a oportunidade de verificar os pontos de convergências e divergências, ainda que as análises tenham sido de fenômenos gramaticais diferentes da língua portuguesa.

Tivemos a sorte de termos uma diversidade imensa de abordagens dentro do conjunto das teorias formais. O artigo de Letícia Emília Kriek aborda, através de uma perspectiva sincrônica, os pronomes resumptivos pessoais da construção de sujeito duplo no português brasileiro sob a ótica da cartografia (RIZZI, 1997; CINQUE, 1999), modelo de Gramática Gerativa que tem como um de seus principais objetivos a descrição e explicação mais minuciosa da estrutura sintática. Desta maneira, concede-se aos leitores uma ampla diversidade teórica dentro do Gerativismo para analisar fenômenos gramaticais. Esse conjunto diversificado de modelos mostra a incansável tentativa de explicar os dados empíricos das línguas naturais da melhor maneira possível.

No artigo “Context for null subjects in contemporary Brazilian Portuguese”, os autores Mônica Rigo Ayres e Gabriel de Ávila Othero investigam sujeitos nulos no português brasileiro (PB) assumindo que sujeitos referenciais nulos são usados apenas em contextos ‘marcados’, específicos; enquanto sujeitos expressos são a estratégia ‘não marcada’ mais frequente. Para isso, os autores analisaram um corpus contemporâneo de PB falado em que encontraram 1.252 ocorrências de sujeitos nulos. Como poderão observar, um trabalho exaustivo de descrição e análise, seguindo modelos formais de análise de dados.

Se, no artigo precedente, os autores desenvolveram argumentos sobre o sujeito nulo, em “Restrições semânticas de animacidade do objeto nulo do Português Brasileiro”, o autor Gabriel Guimarães Peixoto da Silva oferece aos leitores uma análise sobre o objeto nulo no português brasileiro, seguindo referências teóricas e analíticas muito semelhantes ao artigo anterior, o que resulta em uma continuidade incomum e muito interessante. Há uma espécie de complementaridade casuística que resulta em uma sequência bastante didática.

Luciano de Oliveira, em seu texto “Pronomes plenos e clíticos pronominais no português brasileiro e no francês”, apresenta uma análise comparativa acerca do comportamento dos Pronomes plenos e clíticos pronominais e suas repercussões para o aprendizado do francês por brasileiros, tendo em vista algumas observações analíticas importantes. Por exemplo, segundo o autor, uma dessas diferenças é a perda dos pronomes clíticos acusativos de terceira pessoa por falantes do português brasileiro, ao passo que, em francês, os clíticos pronominais de terceira pessoa são muito produtivos, independentemente do caso sintático considerado. Outra divergência proveniente da comparação diz respeito ao fato de que o francês possui clíticos pronominais nominativos, locativos e partitivo/genitivo, enquanto o português não apresentaria, dentre outros aspectos considerados pelo autor.

A partir de uma abordagem minimalista de Gramática Gerativa, o autor Humberto Borges apresenta o artigo “Se-passivo, se-impessoal e sujeitos nulos indefinidos: uma abordagem minimalista para a perda e o surgimento de construções na gramática da língua”. O autor vai ao encontro da tradição gerativista que defende a perda das construções com o se-passivo no português brasileiro, resultante do enfraquecimento da concordância dos verbos transitivos. Em seu texto, o autor apresenta dados empíricos que mostram que a perda dessas construções, que perpassa a história do português, está associada à perda de uma projeção mais alta que T: FuncionalP. Este fenômeno seria, ainda, responsável por uma série de questões gramaticais apresentadas e argumentadas pelo autor.

No artigo “Reconstruindo a história do Português do Brasil pelo *Corpus* Tycho Brahe Brasil: novos dados, novos olhares”, produzido por Paulo Ângelo Araújo-Adriano e Williane Silva Corôa, os leitores são apresentados a dados históricos extraídos de novas amostras documentais do *corpus* Tycho Brahe, anotadas morfológica e morfossintaticamente. Com base nos pressupostos da linguística de *corpus* e da teoria gerativa, os pesquisadores discutem a questão da validade empírica e da representatividade dos dados diacrônicos. Ao explorar dois temas gramaticais relevantes para a história do português brasileiro – a colocação dos clíticos pronominais e a perífrase de progressivo –,

os autores demonstram a influência que o tipo textual e o modelo de periodização (baseada na data de nascimento dos indivíduos ou na data de produção do documento) podem exercer na análise diacrônica de fenômenos linguísticos.

Os últimos cinco artigos que compõem a sequência do presente número compartilham entre si dois aspectos: todos eles examinam fenômenos gramaticais a partir do ponto de vista diacrônico e fundamentam-se em modelos baseados no uso. Trata-se de análises de excelência que, amparadas em amostras documentais várias, exploram tópicos relativos à variação e/ou mudança linguística bem como os princípios e mecanismos que condicionam e determinam esses processos. Para além dos fatores morfossintáticos, encontramos, nestas análises, a observação de parâmetros semânticos, pragmáticos, discursivos e sociais.

O artigo “Uma análise diacrônica das formas pronominais contigo e com você em cartas pessoais”, assinado por Thaissa Frota Teixeira de Araujo Silva, traz outra interessante discussão baseada em dados de sincronias passadas. A autora toma como objeto de estudo as formas pronominais *contigo*, *com você*, *preposição+ti* e *preposição+você*, tratadas como variante da variável expressão da segunda pessoa do singular na relação gramatical oblíqua. Fundamentada nos princípios da sociolinguística histórica (HERNÁNDEZ-CAMPOY; CONDE SILVESTRE, 2012), a pesquisadora analisa o fenômeno em um vasto conjunto de cartas pessoais produzidas por indivíduos do Rio de Janeiro, entre as últimas décadas do século XIX e no decorrer do século XX. A partir do controle rigoroso de variáveis independentes linguísticas e extralinguísticas, a autora oferece aos leitores uma cuidadosa descrição do fenômeno na diacronia considerada, na qual identifica aspectos convergentes e divergentes no uso das variantes *contigo* e *preposição+ti*.

Os autores Pedro Henrique Sousa dos Santos e Elyne Giselle de Santana Lima Aguiar Vitória, no artigo “Expressões de tempo decorrido em cartas pessoais do alagoano Graciliano Ramos”, trazem aos leitores mais um estudo diacrônico interessante, ancorados na perspectiva da Sociolinguística Variacionista e da Sociolinguística Histórica. Os pesquisadores analisam as expressões de tempo decorrido (fenômeno morfossintático ainda pouco investigado no português brasileiro) em 112 cartas pessoais escritas pelo ilustre alagoano Graciliano Ramos na primeira metade do século XX. A partir de uma refinada metodologia de análise estatística, os resultados dos autores indicaram a predominância da variante “*haver*” e a influência de fatores linguísticos no condicionamento da variante “*fazer*”.

Enriquecendo ainda mais a pluralidade de abordagens e perspectivas de análise deste número, temos o artigo “Gramaticalização do verbo visar”, produzido por Geovana Portela de Moura, Célia Márcia Gonçalves Nunes Lôbo e Lúcia Regiane Lopes-Damasio. Fruto de uma pesquisa conduzida no âmbito do projeto Para a História do Português Brasileiro – Mato Grosso (PHPB-MT), o estudo investiga a trajetória de mudança linguística envolvendo as construções com o verbo visar. Fundamentando a análise nos pressupostos teóricos da Gramaticalização segundo a perspectiva funcionalista (TRAUGOTT; HEINE, 1991; HOPPER; TRAUGOTT, 1993), as autoras examinaram uma amostra pancrônica, extraída de jornais mato-grossenses publicados nos séculos XIX, XX e XXI.

Dentre os achados da investigação, as pesquisadoras observam que “visar” passa por um processo de gramaticalização, segundo o qual passa de verbo pleno com sentido de “olhar” para auxiliar, integrando perífrases visar+infinitivo com leitura de volição.

Também seguindo a perspectiva teórica da Gramaticalização, o artigo de Márluce Coan, Francisco José Gomes Sousa e Laila Cavalcante Romualdo, intitulado “Da emergência de formas compostas para a expressão do antepretérito: construtos morfossintáticos e semânticos”, transporta os leitores para séculos atrás, ao focalizar o português antigo e a formação dos tempos compostos envolvendo a construção haver/ter+particípio passado. Os pesquisadores analisaram 1.680 dados extraídos das cantigas medievais galego-portuguesas a fim de atestar motivações morfossintáticas e semânticas que elucidem o processo de mudança em questão. Os resultados gerais dos autores revelaram indícios do processo de gramaticalização nos textos medievais e apontaram contextos de favorecimento e de resistência à implementação das estruturas de tempo composto em português.

Fechando o presente número, temos mais uma contribuição aos estudos diacrônicos com o artigo “A construção X-ada: um caso de construcionalização”, assinado por João Carlos Tavares. O autor propõe uma articulação teórica entre os postulados centrais da Linguística Cognitiva (LANGACKER, 1987; 2008) e das abordagens construcionais baseadas no uso (BOOIJ, 2010; TRAUGOTT; TROUSDALE, 2013) para investigar a construção morfológica X-ada na história do português. Como hipótese, Tavares defende a existência de um processo de construcionalização, segundo o qual a construção derivacional teria emergido do particípio passado. Além disso, o pesquisador salienta também a importância da distinção entre mudança construcional e construcionalização, proposta por Traugott e Trousdale (2013), para embasar a coexistência de construções diferentes na língua que emergem de uma mesma construção diacronicamente anterior.

Por fim, os organizadores deste número da *Revista Linguística* agradecem aos autores pelo envio de artigos de altíssima qualidade e diversidade teórica, sem deixar de mencionar o cuidadoso trabalho de todos os pesquisadores que atuaram como pareceristas deste número. Cumpre registrar que esta edição foi organizada durante um momento bastante peculiar e difícil, no qual muitos atuavam no modelo de trabalho remoto (devido à pandemia da COVID-19), seguido do recesso e retorno às atividades presenciais. Esperamos que este número, portanto, contribua para o debate e desenvolvimento de outros temas e reflexões futuras em linguística nos mais diversos modelos teóricos de análise.

Referências

BOOIJ, Geert. *Construction Morphology*. Oxford: Oxford University Press, 2010.

BORER, H. *Structuring sense*. In Name Only. 1. ed. Nova Iorque: Oxford Univeristy Press, 2005a.

BORER, H. *Structuring sense*. The normal course of events. 1o ed. Nova Iorque: Oxford University Press, 2005b.

- BORER, H. *Structuring sense*. Taking Form. 1. ed. Nova Iorque: Oxford University Press, 2013a.
- CINQUE, G. Adverbs and functional heads: a cross-linguistic perspective. *Oxford studies in comparative Syntax*. Oxford: Oxford University Press, 1999.
- HALLE, Morris; Alec MARANTZ. Distributed Morphology and the Pieces of Inflection. In: HALE, Kenneth; KEYSER, Samuel Jay (Eds.) *The View from Building 20*. Cambridge: MIT Press, 1993, pp. 111-76.
- HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE SILVESTRE, Juan Camilo. *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, 2012.
- HOPPER, Paul; TRAUGOTT, Elizabeth Closs. *Grammaticalization*. Cambridge: Cambridge University Press, 1993.
- LANGACKER, Ronald Wayne. *Foundations of cognitive grammar*, v. 1: theoretical prerequisites. Stanford, Calif.: Stanford University Press, 1987.
- LANGACKER, Ronald Wayne. *Cognitive Grammar: a basic introduction*. Oxford: Oxford University Press, 2008.
- SAUSSURE, F. *Curso de Linguística Geral*. Parábola: São Paulo, 2021.
- RIZZI, L. The fine structure of left periphery. In: Haegman, L. (ed.). *Elements of Grammar*. Dordrecht: Kluwer Academic Publisher, 1997.
- TRAUGOTT, Elizabeth Closs; HEINE, Bernd. *Approches to Grammaticalization*. v. 1. John Benjamins Publishing Company. Amsterdam/ Filadelfia. 1991.
- TRAUGOTT, Elizabeth Closs; TROUSDALE, Graeme. *Constructionalization and constructional changes*. Oxford: Oxford University Press, 2013.

INTERVIEW WITH HAGIT BORER FOR REVISTA LINGUÍSTICA*ENTREVISTA COM HAGIT BORER PARA A REVISTA LINGUÍSTICA**Isabella Lopes Pederneira¹*

In this volume of *Revista Linguística*, which has articles on the topic “Synchronic and Diachronic



Grammatical Studies,” I had the honor of interviewing one of the most influential linguists of our time, Professor and researcher Hagit Borer.

Hagit Borer is a professor of linguistics at Queen Mary University of London. She develops research in Generative Grammar and has also developed her own Theoretical Model of Grammar - The Exo-Skeletal Model which shifts the computational load away from the lexical entry to the syntactic structure. She is the author of several books and articles in linguistics, including the trilogy – Structuring Sense – which details the Exo-Skeletal Model.

Borer got her PhD in linguistics in 1981 at the Massachusetts Institute of Technology (MIT), where she was a student of Noam Chomsky. She is a fellow of the British Academy and Linguistic Society of America. In her solid career, she has been developing studies mainly in three sub-areas of linguistics: comparative syntax, morphosyntax, and language acquisition.

Despite her extensive curriculum, she is an extremely attentive person. I also had the opportunity to get to know her welcoming side while I was in Queen Mary to develop part of my PhD under her supervision. This interview shares with all the readers of this volume her generosity in sharing her knowledge in view of her detailed and very enlightening answers.

REVISTA LINGUÍSTICA: You have developed a theoretical framework called the Exo-Skeletal Model (XSM). Researchers in the area frame your model as a constructionist approach. What are the main features that distinguish your theory from recent developments of the generative enterprise, in particular Distributed Morphology (DM)?

HAGIT BORER: It’s a family of approaches, not just mine and DM. The general term now mostly in use is Constructivism. At least one reasoning behind the term is to distinguish us from Construction Grammar (itself also in various executions) which is another one of your questions.

¹ Departamento de Letras Vernáculas e Programa de Pós-Graduação em Linguística - UFRJ, isabellapederneira@letras.ufrj.br, <https://orcid.org/0000-0002-5884-8472>.

As for the differences between my specific approach and Distributed Morphology, several things are important to stress. First, DM has now existed for some 25 years, and there are various diverging versions of it some of which are closer to mine and others more distant. Second, there are significant differences between DM with its various executions and my approach, but there are also some important similarities.

In discussing these differences and similarities, I set aside what I consider to be terminological differences that have no empirical consequences, for instance node labelling, and focus on more fundamental differences and fundamental similarities which may have empirical consequences. The fundamental similarity among all Constructivist approaches is the belief that grammatical structure should not be based on the properties of lexically-listed individual words. The idea in general is to minimize and to the degree possible altogether eliminate the syntactic role of the lexical component. Importantly, this concerns properties of substantive vocabulary, not properties of functional vocabulary or features. Because features or functional vocabulary, in most Constructivist approaches, do have grammatical properties. One could, of course, conceive of a model in which functional vocabulary as well does not have a formal role to play, and it does, indeed, play a minimal role in some Construction Grammar approaches, but in Constructivist models functional items, whether features or actual vocabulary items, do have properties. There are disagreements on how exactly to implement this. For instance, the relationship between features and functional vocabulary items is not agreed on. But that's not a distinction between XSM and DM, but a distinction between both XSM and DM on the one hand, and say Nano Syntax on the other hand. For both XSM and DM, there is a "lexicon", but it's a lexicon of grammatical functions and how they are pronounced. Within this view, substantive vocabulary either doesn't play a grammatical role altogether, or has a very limited grammatical role. Most importantly, and the idea all Constructivist approaches subscribe to, is that major syntactic categories (sometimes called Lexical Categories, such as N, V, A) are not listed with substantive vocabulary. To compare, in the lexicon as it comes to us from *Aspects of the Theory of Syntax* (Chomsky, 1965) and subsequent, you have a (substantive) word, let's say *kick*, and the word *kick* has various grammatical properties. First, it is pronounced as /KIK/. Second, it is listed as a verb (V). Third, it comes with a syntactic insertion frame (sometimes called *subcategorization*), which tells you, specifically, that it requires a complement, and that the complement needs to be the nominal (NP or DP) that gets interpreted as the entity that undergoes the *kick* action. Finally, it tells you what *kick* means, say KICK. So effectively, *kick* is a mini syntactic fragment, as it delivers not only sound-meaning relations, but also a piece of syntax, something like (1):

1. *Kick*: V /KIK/, KICK, [_{VP} *kick* NP]

Overwhelmingly, this view has been preserved to this day in both syntactic and semantic approaches, even if NP has become a DP, VP has acquired more complexity (e.g., vP or VoiceP) and insertion frames, or subcategorization frames, have been replaced by theta-roles assignment, or other

event roles associated specifically with listed entries, or by non-committal statements such as ‘*kick* takes a complement’. The original Constructivist agenda brought in the suggestion that this is not the right way to go, and that the right way to go is to incorporate these syntactic properties into the syntax and have both the complementation properties and the categorial properties emerge from the syntactic structure. These are theoretical desiderata that XSM and DM share. However, as I already noted, DM has now become a family of approaches, rather than just one, and differences relative to these desiderata have emerged. While all Constructivist approaches I am aware of do subscribe to the view that major categories are not listed properties and are syntactically derived, and that listed (non-functional) items, roots, are a-categorial, a lot of contemporary DM executions nonetheless ascribe to roots the ability to take syntactically projectable complements (most influentially HARLEY, 2009, 2014), and others go the extra step of assuming that these complements track the lexical semantics of the roots (EMBICK, 2004; ALEXIADOU *et al.*, 2006). I don’t subscribe to these views, and continue to maintain that roots have neither syntactic properties, nor syntactically consequential semantic ones. A similar take on the absence of any syntactic or semantic properties for the root is found in First Phase Syntax (RAMCHAND, 2008), although couched within a syntactic approach to event structure which differs from mine.

That’s one distinction. The second distinction which is quite fundamental is the interaction between roots and phonology, and by extension, the relationship between syntax and phonology. Relative to this question, XSM is not just different from DM, it is also different from (mainstream) Minimalism. As I have argued at some length, roots are no more than a package of phonological information, beyond which they have no properties. We could think of this phonological package as filling a hole at the bottom of every extended projection, where the phonological information does little beyond tracking the history of the root. In contrast, in many (although not all) DM-type approaches, roots have semantic or syntactic properties, but not phonological ones (see the debate on root suppletion in DORON, 2014).

Even more significant, however, is my view on the interaction between phonological representation and substantive word meaning, what I call *Content*. Within classical generative approaches, words are listed in the lexicon with their meaning. For the entry of *kick* in (1), KICK is in reference to the conceptual meaning of *kick*, and is a shorthand to some specific lexico-semantic representation. The lexical semantics of KICK or other listed substantive vocabulary, however, is only one part of our understanding of the semantic component available in natural language. The other part is that which is computed by formal semantic representations as constructed by the syntax and the properties of grammatical formatives, or grammatical *functors*. For instance, if you take something like “*the*” and you wish to characterize its relationship with the emergence of definite descriptions, we could (informally) say that *definite description* is a semantic meaning that is constructed through the semantic properties of the functor THE, but *definite descriptions* do not exist independently of the syntactic structure, in this specific case DP. By contrast, the listed lexico-semantic properties of

kick, KICK, are by assumption structure-independent. When you look at the classical “GB picture” as well as much of Minimalism, you see that although both lexical semantic meaning and formal meaning are integrated into the grammar, they emerge in distinct ways. One is associated with listed lexical entries and is unaffected by its syntactic environment, and the other is constructed through post-syntactic interpretation assigned to the syntactic output, in a component historically referred to as LF but nowadays mostly referred to as the ‘conceptual-intentional interface’.

Turning now to the Constructivist perspective and its anti-listing agenda, there are little immediate consequences for formal semantics from the diminished syntactic information associated with listed substantive items, as formal semantics depends primarily on functors and on structures. However, if we are to eliminate lexico-semantic and syntactic word-specific (or root-specific) meaning, leaving, effectively, only /KIK/ in the listed entry of *kick*, an important question immediately emerges concerning the way in which /KIK/ and KICK come together, as in most utterances, /KIK/ does mean KICK.

Because the overwhelming tendency within generative approaches from at least 1977 onwards has been to severely curtail the impact of phonological information on both the syntax and the interpretational component, the overwhelming Constructivist tendency has been to assume that the emergence of the interpretation which we refer to as KICK must be negotiated by the very same component in which formal semantics is negotiated, be it LF, or the Conceptual-Intentional interface, and as such, it crucially has no access to any phonological information. That, I think, is a mistake. Formal semantics, as negotiated through the syntactic and semantic features of functors and syntactic structures has little obvious use for phonological information (with some residual issues such as *focus* negotiable through a feature spellout system). Specifically, if we assume that ‘intentional’ is in reference to the formal-semantic computation, viewing it as a post-syntactic interface with no phonological properties appears to be on the right track. Not so, however, for what, broadly, we can assume to come under the heading ‘conceptual’, which, I assume, covers aspects of meaning, Content, related to the conceptual system, to world knowledge etc. The “conceptual interface”, and this goes back directly to Saussure, is the sound-meaning relation or more accurately the sign-meaning relations. If we take as our starting point the T-model as it now stands, this means that the ‘conceptual interface’ is not at the ‘end’ of the LF branch, but is linked to some well-defined representation on the PF branch. This is something that I have done some work on reported in Borer (2013, 2014), and this interface with PF is the topic in which I am presently mostly interested. This difference on how to model the emergence of conceptual Content, and how the phonology interacts with both syntax and Content is thus a significant difference between XSM and most Minimalist or Minimalist-oriented approaches, including DM.

There are additional differences. For instance, how exactly is categorization negotiated in a system with a-categorial roots, or what is (or isn’t) the distinction between inflectional morphology and derivational morphology, or whether there could exist roots in the structure that remain uncategorized. Although these are important issues, they are less intrinsic to these distinct approaches, but rather

concern issues of execution that ultimately should be decided on empirical grounds. I note one more important matter, concerning the comparison with DM. There are a lot more DM-practitioners than XSM-practitioners, and so there are, in consequence, many more well-developed accounts of phenomena that XSM at present is entirely silent on. One important example is the ongoing research on the licensing of non-direct arguments such as applicatives and others (see PYLKKANEN 2008 and much subsequent), a topic on which XSM approaches have been so-far silent. To the extent that much of this work licenses non-direct arguments through dedicated functional structure it is certainly compatible with fundamental XSM approaches. Not having studied the issue in detail, however, I cannot comment on the extent to which some assumptions or results of that research could be adopted, as such, with XSM.

REVISTA LINGÜÍSTICA: Many linguists compare your theory with Construction Grammar due to templates as syntactic structures. Could you clarify the differences between your approach and the one developed by Fillmore?

HAGIT BORER: Ok, this is not just me, this is the Constructivist family in general. And just a little historical note: when Constructivism, at the time it was called Neo Constructionism or Constructional grammar, first came on the scene in the late 90's or early 00's, we did try, collectively, as a group with a shared research agenda, to reach out to Construction Grammar communities. I was personally very enthusiastic to pursue the links, and so were, for a brief period, adherents of Construction Grammar. I remember a visit in 1999 to the University of Illinois at Urbana-Champaign, when Adele Goldberg was still there, and we were both very excited to discover shared theoretical interests. I still have her book with her dedication 'to a fellow constructionist' which she gave me at the time. That didn't last, however, as by the early 00's, Construction Grammarians were already seeking to distance themselves from Constructivism. I recall an article from about 20 years ago, possibly by Laura Wagner, which explicitly stated that the affinity Constructivists (as presently named) find with Construction Grammarians is on topics which are important to Constructivism, but which are peripheral to Construction Grammar, and on the issues that are important to Construction Grammar, there is a wide gap. The approaches are indeed fundamentally distinct in their underlying assumptions, but nonetheless, as there are significant commonalities in the empirical domains investigated, it would have been better had it been possible to pursue both similarities and differences in a cooperative fashion.

As for the differences, a major assumption that is built into Construction Grammar is that there are linguistic objects we can call *Constructions*, which are defined primarily through their communicative function or their discourse function. Take, for instance, *passive*. What does this term mean? In mainstream of Generative Grammar (from early transformational grammar, through Relational Grammar to GB onwards as well as in strictly lexicalist approaches such as LFG), *Passive* is not a linguistic object as such. Rather, the term refers to well-defined grammatical structures and/or

operations. Typically, some sort of a grammatical demotion of the logical subject and a grammatical promotion of the logical object. But if you look at natural language, you find that there are multiple ways of foregrounding the object and backgrounding the subject which do not involve grammatical demotion or promotion. Topicalization, emphatic stress, impersonal (active) constructions with an indefinite pronominal subject, *se* constructions in Romance that serve to communicatively demote the subject and promote the object, and so on. If you look at this picture from the perspective of Constructions as defined by communicative functions, the syntactic differences between these configurations are noted, of course, and may have some grammatical consequences, but at the end of the day, these syntactic differences are explicitly assumed NOT to be relevant to either the overall understanding of these constructions, their use, or their acquisition. Importantly, their *compositionality* is either denied or judged to be irrelevant. What matters is that a particular structural configuration is associated with a particular discourse function, call it Passive, and the relationship is arbitrary and listed. In one language, it may have the English syntactic structure; in another, it may have a completely different syntax, but its discourse function would be the same. This, for Construction Grammar is not only a legitimate result, but a desirable one, as the approach explicitly subscribes to the existence of a ‘Constructicon’, a dictionary of constructions with their discourse function and meaning as a whole, alongside the possibility of language-specific ‘Constructicons’. From that perspective, it is not surprising that the most important evidence for Construction Grammar comes from idioms, where meaning is, indeed, listed in conjunction with non-compositional structures. For *kick the bucket*, to take a classical example, a compositional approach is clearly either hopeless or circular, thus suggesting that at least sometimes, compositionality is not necessary. For that reason, many Construction Grammarians have worked on idioms or on idiomatic fragments (GOLDBERG, JACKENDOFF, and many others).

As an aside, it is really Chomsky (1965) who should be ‘blamed’ for the emergence this fundamentally non-compositional approach that associates meaning with whole phrases or even sentences, rather than composing it from the meaning of terminals. When one looks at insertion frames such as the one in (1), one sees that ‘lexical entries’ are in fact listed ‘mini constructions’ which are bigger than a terminal. But if one is allowed to list fragments of structure of a phrasal size, why stop with (1)? Why not push this approach to its logical limit, and say that the fundamental building block is exactly listed fragments of grammar of phrasal size or bigger, with their fully specified meaning, non-compositional meaning?

Returning to the comparison of Construction Grammar with Constructivism, this is a major difference, as Constructivist approaches are extremely compositional. More compositional than traditional lexicalist approaches, in fact, because so much is invested in the interpretation that emerges from composing functors, functional items. This is not just an issue of execution or formalism. Rather, this is a fundamentally different assumption about what the human linguistic ability consists of. For Construction Grammar, the human linguistic ability capitalizes on memory and the ability to store not just words, but chunks of structure with meaning. We do know human memory is rather good for

this kind of things and that is also the lexicalist assumption – that the role of rules is not as significant as the role of memorized lists. In contrast, Constructivist approaches adhere to the idea, which, I think, is fundamental to generative grammar and to structuralism in general, that the human linguistic capacity is first and foremost compositional, and rule driven. To be sure, *kick the bucket* is listed as such with its meaning. The question, however, is that of balance. For Constructivists, the question concerns the residue that is left once all compositional rule-driven operations are exhausted, and how to represent it. Conversely, for Construction Grammarian the question concerns the pervasiveness of compositional phenomena, and the danger of massive redundancy that non-compositional systems risk. In more concrete terms, one approach subscribes to the belief that we need to discover the rules, and that exceptions, by and large, signify the need to return to the drawing board and come back with better rules. Exceptions, in other words, define a further research agenda. The other approach highlights the fact that there are exceptions, possibly quite many, and because listing is inevitable for these, listing can be further extended making rules redundant.

REVISTA LINGUÍSTICA: Constructivist models have as main feature the fact that (certain aspects of) lexical content of lexical items is negotiated with the syntactic environment in which such items occur. In your theory, we can see that the morphological marking itself depends on the syntactic structure. How do you delimit and formalize the grammatical rules within this context?

HAGIT BORER: I am not sure I fully understand this question, as I no longer assume that lexical items have any syntactic properties?

REVISTA LINGUÍSTICA: The question is about the derivation of the sentences. Where, in the sentence can we say that a noun will be a noun, a verb will be a verb? Beyond that, when do these categories receive meanings, formal meanings and contextual meanings?

HAGIT BORER: I addressed at least some of it in the answer to the previous question, but to recap, I think that the essence of the root is some phonological properties, and these are not just phonological properties of the root in isolation, but also the root's ability to select particular allomorphs of functional elements that merge with it, and the ability of functional items that merge with the root to select a particular allomorph of the root. There is a lot of remarkably interesting discussion on such selection and the locality conditions that apply to it. That has not divided XSM from various DM proposals. There is a lingering puzzle here which, likewise, both approaches share, as it is not clear how to state local syntactic conditions on post-syntactic phonological insertion. This problem exists not only for DM with its late insertion, but for XSM as well. Although in XSM roots come with phonological properties, the phonological properties that are associated with the root are not a phonetic representation, but a reference to a phonological set of properties that are only fully articulated post-syntactically. For me, the association of roots with phonological properties is a device intended to ensure phonological faithfulness, which is to say, to ensure that the syntactic derivation

will *not* fundamentally change the phonological properties of the root. To illustrate, such faithfulness would altogether exclude or severely restrict a derivation in which e.g., *kill* is derived from the root which in other contexts would be realized as *die* (e.g., assuming that /KIL/ is the realization of [CAUSE [$\sqrt{\text{DIE}}$]], or in reference to specific arguments that I made in Borer (2003) and later, having the nominalization of a root such as $\sqrt{\text{WRITE}}$, say [NOM [$\sqrt{\text{WRITE}}$]] be realized phonologically as *essay*.

There is an ongoing debate on this issue, which centers on the availability of root suppletion, and the best place to consult this debate is Doron (2014). Focusing in particular on the *go-went* pair in English, the question is whether *go* and *went* share a phonologically unspecified root which is realized as *went* (or *wen*) in conjunction with PST, but as *go* elsewhere, or alternatively, these are different roots, and pairs such as *go-went* represent the merging of two defective paradigms – one associated with the root $\sqrt{\text{GO}}$ which is lacking past tense realization, and the other with $\sqrt{\text{WEN}}$ or possibly $\sqrt{\text{WENT}}$, which lacks realizations for all but past tense. This said, there is nothing in XSM that would prevent the phonological properties of the root from including suppletive realizations, as an option. The fundamental reason why I do not opt to integrate such suppletive pairs into the same phonological root package is because despite considerable research on the topic, we still do not have any insight as to why root suppletion (as opposed to root allomorphy) is so typologically rare, why, with few exceptions, most suppletive forms impact light verbs and auxiliaries, and how to precisely block such cases as the nominalized form of $\sqrt{\text{WRITE}}$ being realized as *essay*. In view of this, my perspective is that unless proven indispensable, root suppletion is to be excluded. This is the strongest position, and therefore, even if wrong, is the most productive starting point for the investigation.

Turning to the question of Content, where by Content I mean conceptual non-formal meaning, I believe the most fundamental question regarding roots and Content is whether the roots have Content, and beyond that, whether such Content, where it might exist, interacts with the grammar in any way. My answer to this question is an unequivocal “no”, and this is quite fundamental to the way in which XSM treats conceptual, so-called lexical Content. But there is no inherent reason to exclude this view within DM (although to the best of my knowledge there are no contemporary DM-grammarians who have fully endorsed it). The closest to this perspective comes from Harley (2014), who subscribes to the view that the root has no inherent meaning, but comprises a set of alternative meanings to be contextually chosen. (See also the alloseme account in MARANTZ, 2013 and subsequent). At the other end of the spectrum, and possibly not quite strictly within DM, we have the position taken by Malka Rappaport-Hovav (see RAPPAPORT-HOVAV, 2017 for representative argumentation), who argued in some detail that roots may incorporate significant lexical semantics erstwhile attributed to verbs, which in turn has a determining role in their emerging syntax (and see BEAVERS and KOONTZ-GARBODEN, 2012 for some perspectives and critique). In the middle of the spectrum, you find people such as Embick (2004) or Alexiadou *et al.* (2006), who classify roots into different lexico-semantic classes, allowing significant underspecification, but nonetheless assuming that such lexico-semantic classes for roots are syntactically informative in delimiting the occurrence of roots in syntactic structures.

A lot of scholars who attribute lexico-semantic properties to roots do so to solve the enduring overgeneration problem for root-based systems, including XSM. In the absence of syntactic and/or semantic properties for the root, the prediction is that any root could occur in any syntactic context, and while this is certainly more so than strict lexicalism would have it, it is not universally so. If roots, or lexically listed items, have insertion frames or lexical semantics that restricts their complementation possibilities, overgeneration vanishes. So, for instance, once we say that the meaning of $\sqrt{\text{THINK}}$ allows a propositional complement, but not so the meaning of $\sqrt{\text{EAT}}$, an infelicitous utterance such as *#I ate that Mary arrived late* is immediately excluded.

The cost for approaches that place syntactic restrictions on the distribution of roots, either through listed complementation structures, or through their lexical-semantics, comes from the massive proliferation of available verbal and nominal syntactic contexts for many, if not all mono-morphemic forms in English, by assumption co-extensive with roots. To accommodate such variable behaviour, a lexicalist approach, and one that seeks to derive the syntactic distribution from lexico-semantic considerations, would have to postulate multiple homophonous entries, each with its own insertion entry. This massive redundancy is eliminated by removing syntactic information from listed entries, roots. However, such elimination creates over-generation, as already noted. And XSM, indeed, over-generates.

This said, I have done at least a partial survey of the proposals currently under consideration within Constructivism that attempt to solve the over-generation issue, some mentioned above, and my sense is that they do not solve the problem. All of them exclude some cases of over-generation, but not others, and all of them, conversely, under-generate. To illustrate, Embick's (2004) proposes that roots are either eventive or stative, and only the former can take an event argument. This would suffice to correctly exclude (2a), under the assumption that *know* is stative. However, it would also exclude (2b), as *think* is equally stative, but (2b) *is* grammatical in English. It would also run into problem with verbs such as *touch*, *love* and others, which emerge as either 'stative' or 'eventive' depending on their syntax, as in (3)-(4):

2. #Jane knows to the store
Jane thought the book to Mary (following GLEITMAN, 1990)
3. The wall touches the fence (most saliently stative)
Kim loves Robin (most saliently stative)
4. The wall touched the fence twice (e.g., as a result of the wind) (most saliently eventive)
Kim loved Robin twice (most saliently eventive)

Alternatively, all such cases of variability, cases in which more than a single insertion frame is available for the same root, but potentially with a distinct overall interpretation emerging, would require postulating that e.g., *think* or *touch* each requires postulating two or more homophonous entries, each with its distinct syntax and semantics. As early as 1994 and in considerable subsequent

work, including work on acquisition, this is exactly what I opted NOT to do, much as I was from the very beginning aware of the over-generation issue. While I made some proposals to resolve at least some over-generation cases through the presence of covert affixation (notably in CHAPTER 10 of BORER, 2005b), I do believe that for Constructivism this remains an open question without a satisfactory solution, as of yet. My own belief is that a full answer must involve a better understanding of word complexity, on the one hand, and the conceptual system and its interaction with computational grammatical systems, on the other hand.

As a digression, I note that we have several thousand years of recorded philosophical discussions on the nature of concepts, to which we can add, in the last 150 years, insights by linguists and psychologists. This it's not exactly a new concern and yet very little progress has been made on understanding concepts, especially in context. That turns out to be a very tenacious problem, and that is before we get to the second tenacious problem, which is the characterization of the relationship between concepts and their labels, a massive philosophical discussion. If you assume that concepts could have a label, call it *word* or sound/sign, the relationship between that label and the conceptual meaning is extremely tricky. One commonly discussed problem has to do with polysemy, where a single label may be associated with two concepts which bear some relation to each other, but are clearly distinct concepts. An example which I give a lot has to do with words/labels such as English *pinch*. Depending on its context, it could mean 'to squeeze between two fingers', or 'to grab', or 'to steal'. It could also mean, as a noun, 'very quick' (*in a pinch*), or 'be in a tight spot', or 'very little' (*a pinch of salt*). Other meanings are available as well. But conceptually, the relationship between 'be in a tight spot' and 'steal' is quite simply non-existent, and between 'squeeze' and 'really quickly' at best yields itself to some vague metaphorical extension. Rather, it seems like the label *pinch* is in reference to what can be accomplished, literally or metaphorically, between two prongs, possibly but not necessarily, fingers, which are in proximity and may close and open. Whether the above description falls under our general understanding of *concept* is highly questionable. But even if it does, that so-called concept has zero value in informing the use of *pinch* in most of its occurrences. This illustrates quite clearly that labelling does not reliably track what we think of as *concepts*. And this is all before we integrate these labels into our grammatical representations, where phrasal factors play additional havoc in our understanding on how to associate these labels with meaning (*subsective and non-sective* readings being one example). It goes without saying that I no more have a solution to these issues than the thousands of scholars who precede me. It does, however, guide my avid conviction that concepts with or without their labels cannot reliably inform our grammatical abilities.

ISABELLA PEDERNEIRA: In Brazilian Portuguese we have many words like that, verbs, name and so on, in Brazilian Portuguese more than in European Portuguese.

HAGIT BORER: That's interesting. I offer the speculation that the distinction from European Portuguese is because Brazilian Portuguese is less canonized. To draw an analogy, it is well-known that children, when they are about 5 years old with nearly adult-like grammar, go through a very creative phase of language use, particularly in creative word formation, and allowing vocabulary items in non-canonical syntactic contexts. Once they go to school, however, they become much more conservative, because there is a lot of social pressure to conform to canonical, established vocabulary.

Returning to the issue of root meaning, and as I noted, the question is whether roots have meaning, and how root meaning (or lack thereof) is handled in various Constructivist approaches. I already noted that in my view roots do not have grammatically consequential meaning, and in the majority of cases, have no meaning of any kind. Substantive meaning, or Content, in XSM, is matched with units that are bigger than the root, and may include multiple derivational affixation (but not inflectional affixes). To be sure, the Content associated with complex forms that are bigger than the root must be listed somewhere, and I cannot see how a list of sign-Content pairs is avoidable in any approach to language that accepts the arbitrariness of the sign. This is not the issue under debate, however. The issue under debate is whether in the sign-Content listed pair, the 'sign' is a root. For XSM, the answer is *no*. The sign in the sign-meaning pair is always bigger than the root, and at the very minimum, includes a categorized root in contexts that do not include sign augmentation (e.g., in XSM, [_N √TABLE], or within an affixal approach to categorization, [_N ∅ [√TABLE]]).

As for DM, what is paradoxical to me is the frequent claim that the DM notion of what a *root* is, is modelled after the root in Semitic languages. Semitic words have roots, which have been argued to be an essential part of the morphological system (specifically of Arabic) at least since the philological investigations in the XVII-XVIII century. It's intuitively very clear that this is the case, especially for verbs, and it is hard to see how the morphological system can be explained without assuming such roots. These roots clearly do not have a category, which makes them an ideal starting point for theories such as DM or XSM that postulate an a-categorial listed item as the basic (non-functional) syntactic terminal.

However, and as has been pointed out by Constructivists and non-Constructivists alike (at the very least starting with ARONOFF, 1994), the Semitic root doesn't have a fixed Content. While some roots appear to give rise to a family of words that have a related Content (including metaphorical extensions), such as √KTB, whose derivatives tend to bear some relationship to writing, my own investigation revealed that when we exclude possible derivational relationship between distinct morphological verbal instantiations of the same root, 40% of cases exhibit some meaning affinity (sometimes quite loose and unpredictable), and in the other 60% of cases there is no meaning relatedness whatsoever. For the 60% of the roots where no meaning relatedness of any sort is attested, it is extremely difficult to see how an alloseme-like account (MARANTZ, 2013, HARLEY, 2014) could be explanatory, as it would require a circular listing of root meaning as based on the attested (unique) meaning of complex morphological forms, effectively boiling down to the (indirect) pairing of complex morphological signs with meaning. To illustrate some of this circularity, one only needs

to look at HARLEY (2014), who proposes that roots, including some that have no independent Content in English, are associated with a particular alloeme in their morphological/syntactic context, frequently unique. For example, the root $\sqrt{\text{CEIVE}}$ is associated with the alloeme RECEIVE in the context of the prefix *re-* but with the alloeme DECEIVE in the context of *de*. Not only does that effectively amount to listing meaning as associated with complex listed words, rather than roots, it sometimes requires choosing an allseme for the root in a non-local context. To illustrate, $\sqrt{\text{JECT}}$, an English root, must receive its interpretation in context, as it has none in isolation, and hence the claim would be that $\sqrt{\text{JECT}} \rightarrow \text{SUBJECT}$ in the context of the prefix *sub-*, but $\sqrt{\text{JECT}} \rightarrow \text{REJECT}$ in the context of *re-*. Consider however the complex English word *conjecture*, presumably with the structure in (5a), or, one might argue, in (5b):

5. a. [[[con [$\sqrt{\text{JECT}}$]] ure]
 b. [con [[$\sqrt{\text{JECT}}$] ure]]

As it turns out, not only is $\sqrt{\text{JECT}}$ devoid of independent meaning, so are *conject* and *jecture*, the potential analytic parts of *conjecture*. As meaning, by this system, must come from choosing the correct alloeme for the root, we are forced to the conclusion that $\sqrt{\text{JECT}}$ in *conjecture* is assigned the meaning CONJECTURE (either verb or noun) in one of the non-local domain in (5).

For proponents of root meaning, the other potential solution to the fact that a large percentage of roots do not share meaning across their instantiation (60% of Hebrew roots, as noted) would be to claim that e.g., in Hebrew, 60% of the roots have two or more homomorphs, each associated with its own meaning, and each, perforce, independently listed with its unique morphosyntactic insertion frame. Once again, such a system boils down, effectively, to the sign, in the sign-meaning pair, not being the root, but a bigger domain, which is inevitably more complex and categorized. To summarize, meaning, or Content, could not possibly be a grammatically significant part of the Semitic root. If, indeed, the Constructivist root is modelled after the properties of the Semitic root, one could not assume that it has Content without a significant loss of explanatory power. In turn, the only fail-safe generalization that applies to all instances of any one Semitic root is phonological. This is exactly the conclusion Aronoff (1976) and subsequent reaches as concerning not only the Semitic root, but also English non-functional morphemes such as $\sqrt{\text{JECT}}$, $\sqrt{\text{DUCE}}$, $\sqrt{\text{SCRIBE}}$ and similar.

This said, the occurrence of some form of Content relatedness in 40% of cases could not be a coincidence. To be sure, 40% is not sufficient to establish grammatically rigorous Content relations, even more so as the actual specific nature of the Content relatedness is idiosyncratic, ranging from synonymy to metaphorical extension, and grammatical properties such as argument structure realization cannot be predicted from it. It does suggest, however, that at least for these 40% of roots, some Content, however loosely captured, is cognitively encoded. A fuller understanding of that encoding requires a better understanding of the relationship between labelling and Content, a problem I already flagged as persistent and tenacious. My only claim, as relative to that encoding, however characterized, is that it is *not grammatically represented, nor does it play a grammatical role*.

And so, because the Content of roots is at best non-systematic and does not generalize, roots cannot be the basic building block in the construction of Content. Rather, my belief is that Content is matched with syntactic units of delimited size, which are always bigger than the root. Empirically, this claim is shared by DM. However, although the domain of Content is frequently the same in DM and XSM, the logic by which Content is associated with that domain is different. For DM, as noted, Content remains the prerogative of the root, but the specific root alloeme is chosen locally, in the domain of the immediate categorizer (but see *conjecture* and related discussion). In XSM, on the other hand, the domain of Content could extend to include all derivational affixation, thereby allowing the Content CONJECTURE to be associated with either structure in (5) in its entirety. This difference between DM and XSM is both empirical and theoretical. On the empirical level, it involves the existence (or lack thereof) of non-compositional Content that requires domains bigger than that of immediate categorization (beyond *conjecture*, also cases such as *civilization*, *existentialism* or *committal* with the meaning of BURIAL). On the theoretical side it involves claims about the nature of functional vocabulary, and whether categorial functors (typically in English derivational suffixes) are inherently similar or distinct from functors which are segments of extended projections (typically inflectional markers such as plural, tense, aspect, etc.). The other distinction is the one which I already discussed, and which is theoretical, and involves a fundamentally different view of the root. Is the root a Content set, or a set of alloemes, as Marantz (2013) and Harley (2014) contend, thereby putting forth a strong hypothesis of Content compositionality constructed from the root upwards, or alternatively, as I believe, roots have no Content, and Content, rather, is always associated with phrases of a prescribed size. For XSM, then, expressions such as *receive* and *deceive* are therefore assigned the Content RECEIVE and DECEIVE as complex formations, and the fact that they share a root translates into shared phonological properties, but neither syntactic ones, nor Content related.

You need to add to this the fact that under my approach, Content is not just assigned to phrases bigger than the root, but it does so based on partial phonological representation, and so what *re-* or *de-* are contributing here is first and foremost phonological, insofar as the internal syntax of *receive* and *deceive* may be identical, but the phonology is not. Similarly, while the internal syntax of *commitment* (DUTY) and *committal* (BURIAL) may be the same ($[_N [_V \text{con} [\sqrt{\text{MIT}}]] \text{NOM}]$), their phonology is different in allowing NOM to be realized either as *-al* or as *-ment*, and with a different Content emerging. This is a real significant difference between DM and XSM, in terms of understanding what Content is, and how it is modelled relative to roots, word-internal syntax (=morphological structure), phonological realization, and the relationship among these factors.

REVISTA LINGÜÍSTICA: The Exo-Skeletal approach proposes a deep split between the semantic-grammatical meaning conveyed by the functional structure and the grammar-independent Content conveyed by the roots. What is this grammar-independent Content carried by roots and how may the roots, which do not have any grammatical properties, interact with grammatical structures?

HAGIT BORER: I already covered much of this question in my previous answers. For me, as I already said, roots are just bundles of phonological properties. This is something I have been saying

for about 10 years, but I want to elaborate on this and clarify it, as it is frequently misunderstood. I will take as my starting point *Jabberwocky*, Lewis Carroll's poem from *Alice through the Looking Glass*, a brilliant linguistic experiment which lends huge support to the Constructivist agenda. In the first stanza of *Jabberwocky*, there are no recognizable Content words, and the only established English words are functional vocabulary items (both free and bound). Nonetheless it is clear that *Jabberwocky* is a well-formed chunk of English grammar, and it is also considerably more comprehensible than one would expect, given the absence of any recognizable Content. If, as lexicalism would have it, structure projects from lexical-semantic terminals, explaining *Jabberwocky* would require thrusting one into some sort of analogical loop, which, I submit, would be overlooking this striking result.

But what I want to do is to take the *Jabberwocky* experiment one step further. (Written) sound combinations in *Jabberwocky* that are place holders for substantive items, such as *brill(ig)*, *tove*, *slith(y)* etc. do obey English phonological restrictions, but they are not paired with any Content, and that is precisely the point of the Carroll experiment: labels without Content, sign without meaning. However, if these signs have no Content, why are they there, phonologically? Why don't we just have silence, where a substantive element would normally be? And so, the thought experiment I would suggest is to try to construct a sentence, a paragraph, in which only functional items have phonological representations. One could restrict oneself to free functional items, to avoid the problem of requiring a host, say, for the plural morpheme. A sample of such an experiment would be as in (6): The crossed-over items in (6a) are *Jabberwocky* nonce vocabulary items. In (6b), they have been eliminated. What is striking about the contrast between (6a) and (6b) is that (6a) is English. (6b) is not only not English, but, I submit, not a possible utterance in natural language:

6. a. I would (*gimble*) many a (*tove*), but you didn't (*outgribe*) every (*borogove*)
 b. I would many a , but you didn't every

And so, the *Jabberwocky* extension in (6a) is an utterance in natural language, and specifically in English, even though the substantive (non-functional) items have phonology but no Content, but if you eliminate the phonology, what you end up with is no longer natural language. It doesn't mean that a computer couldn't parse (6b) and assign it structure. To the contrary, I am sure it could, and that in fact, any linguist could assign it structure as well. But humans, including linguists, would still refuse to accept it as an utterance of natural language. It is worthwhile noting that our grammar does allow for phonologically silent spaces, produced through various forms of ellipsis or through the distribution of null elements with specific well-defined properties (null pronouns, null determiners, null complementizers, etc.). However, our licit silent elements are either designated functional elements, or alternatively, require a phonologically realized antecedent, thereby making them different from the Content-less silences in (6b). Another way of querying the oddity of (6b), then, would be to ask why we do not have null DPs with (non-functional) Content no matter how salient, or null 'gapping' without an antecedent, interpreted contextually within their syntactic environment, just like '*gimble*' would be, presumably.

The conclusion I draw from this thought experiment is that there must be some phonology at the bottom of the tree, or at the bottom of each extended projection. Don't ask me why, because I don't know why, but this seem to me to be an inevitable conclusion. Altogether, what it looks like, is that you start with some piece of unstructured sound, or sign, which the syntax then molds and shapes into something to which we can give Content. However, this is not a conclusion that our present-day syntactic theorizing is built to capture, and how to characterize this requirement without reevaluating radically the relationship between at least some aspects of phonology and syntax is not obvious. Methodologically, and at the very least from the mid-70's if not earlier, much theoretical care was taken to sever all aspects of phonology from the syntax, with the result that our present computational systems, and regardless of their merit, are specifically built to exclude statements such as 'first merge must involve a chunk of phonology', and perhaps it is time to reevaluate this.

XSM is by no means the only contemporary approach that is attempting to reintroduce phonology into the grammar. You see it quite prominently in recent work by Kayne, roughly from 2005 onwards, which is looking to create a phonology-syntax structural continuum. There are contact points between these agendas, the primary of which is the belief that altogether, the research methodology which generative syntax has followed at the very least since the mid 60's has overlooked the potentially essential part that sound or gesture play in the emergence of well-formed structures, including syntactic ones, in natural language. To summarize, the progression of generative research methodology, in principle, and into present day Minimalism and including DM, has involved an absolute separation between the phonology and the syntax and the phonology and the interpretation. While this may have been a beneficial theoretical move at its inception, there are more and more reasons, and more and more research, which indicate that it is high time to reconsider this perspective. I want to highlight the fact that such reevaluation is not an easy task. Because so much of our theorizing in the past 60 years has involved such strict separation, the formal tools for introducing phonological considerations into the syntax are very much in need of development, and I, myself, would very much like to see more research focused precisely on such formal questions.

Last comment on this topic - the problem I am highlighting here is *different* from the question raised, for instance, by Ramchand (2008), as to whether roots are needed in the syntactic representation altogether, alongside the possibility that what we call a 'root' is associated with a more elaborate syntactic structure, without the need for a particular terminal. Nor is it the same question as that put forth by De Belder (2011) and De Belder and Van Craenenbroeck (2015) who propose that roots are empty terminals which are formally forced to exist by the properties of first merge. The reason these are *not* the same question is that for Ramchand's system, and however stated, some phonological representation remains essential and obligatory, albeit in the context of a structure, rather than a terminal. Similarly, for De Belder (2011) and De Belder and Van Craenenbroeck (2015), the root may be a formal empty terminal in the syntax, but by the end of the derivation, one way or another it must be populated by some phonology. Some phonological representation for non-functional material

is thus obligatory in these systems as well, regardless of how we model its availability, and it is an explanation for this obligatoriness which remains missing.

REVISTA LINGÜÍSTICA: Your own work, at least until the mid - 90's, was guided by Lexicalism. What problems in lexicalist formulations made you propose a new syntactic approach to word structure?

HAGIT BORER: There were certainly important theoretical reasons, but I want to start from a personal perspective. We tend to set aside personal preferences when we do our research, and for a good reason, because reasons to adopt a theoretical perspective or an analysis must be grounded in rigorous reasoning and robust empirical evidence. But personal inclinations still matter in terms of which directions we choose to pursue. These may or may not pan out, of course, subject to more robust criteria of research methodology, but as this is an interview, not a linguistic article, I thought I should say that sometime in the early 90's, when I started considering the possibility that it is all just syntax, I became very excited, as such a wholly structural, top-down approach fits well with the general way I think. But it is important to note that my personal preferences aside, it is not always the case that syntactic structural systems are more coherent or explanatory than systems that are built, in our specific case, on lexical information. Take for instance a comparison between GB and LFG as they existed in the mid '80s. LFG was – and still is - a profoundly lexicalist model, in the sense that there were no longer any mapping operations on phrase structure, transformations. There are only mapping operations that take as their input information stored in lexical entries, and their output are modified lexical entries. By the mid to late 80's this was a much more theoretically coherent model than GB. GB during the period was a mixture of lexicalist assumption and grammatical non-lexicalist assumption, and the distribution of labor between the syntax and the lexicon was not entirely principled. There was also massive redundancy between lexical generalizations and syntactic phrasal ones. I will give you an example I frequently discuss. In (1986), Levin and Rappaport published their seminal adjectival passive article where they showed, to my mind conclusively, that the argument of an adjectival passive like “the window is (completely) broken” is syntactically external and hasn't moved there from the object position. They did not exclude the possibility that *window* may have started as the logical object of the verb *break*, but whatever mapping had transpired to land it in the subject position of *is broken* was not syntactic. The very same empirical tests that showed conclusively that *the window* is syntactically external in adjectival passive, however, also show that in “the window was broken by the children”, verbal passive, *the window* is internal, and *has* moved from the object position. And so, here is the theoretical conundrum: sometimes what looks like the logical object of the verb, say the *theme*, IS the object of the verb and needs to undergo syntactic movement to become subject, but at other times, what looks like the *theme* ostensibly associated with that very same verb has stopped being the object of that verb, if it ever was, well before the syntax, and rather, is mapped directly into the subject position without the need for syntactic movement. But

why should there be such duplicity of functions, with the promotion of the *theme* sometimes lexical and sometimes syntactic, and why should the lexical one give rise to an adjective, but the syntactic one to a verb? Why not the other way around? Why not both lexical? Why not both syntactic? GB, in the mid 80's, had no answer to these theoretical questions, and the result was a certain degree of arbitrariness, and considerable redundancy. By contrast, in LFG everything was an operation on lexical entries, and as a result, the redundancy and the need for two distinct promotion operations did not exist. It was, in other words, a cleaner and more coherent model. There was, however, a hitch – the mixed picture in GB, however theoretically sloppy, was empirically sound, insofar as the *theme* (if that is indeed what it is) is external in adjectival passive, but internal in verbal passive. This empirical result quite simply had no theoretical explanation in GB. The uniform, more rigorous LFG system, on the other hand, had an empirical problem, insofar as an identical lexical mapping operation did give rise to a distinct set of syntactic diagnostics. And so, LFG had an empirical problem, but GB had an in-principal theoretical issue – it had two parallel formal systems that effectively did the same thing, and which system would operate in which construction was a complete stipulation. If you look at much of the ensuing research in the following decade within both GB and LFG, you see attempts, within each one of these approaches, to address their respective problems. In GB (and subsequently early Minimalism) an attempt to eliminate argument structure mapping operations from the lexicon, and in LFG attempts to create a more complex representational structure that could in principle lend more nuanced relationship between operations of lexical mapping and the syntactic output.

You can see from this example at least some of the problems with GB that I and many within my cohort started to focus from the mid/late 80's onwards. On the one hand, GB was, to my generation, a huge structural playground, especially as it successfully integrated an incredible wealth of empirical results and structural observations inherited from Relational Grammar, and with multiple exciting empirical, rule-governed results all there ripe for the picking. At the very same time, and beginning with Pesetsky (1982), lexicalism acquired a lexico-semantic angle, as a result of which the role of lexical entries as, effectively, mini-constructions, was significantly strengthened. Reconciling these two trends was not necessarily an easy matter, and to some extent, you see different scholars making different choices as to how to distribute the labor between the syntax and the lexicon. For me, quite early on, and at least as early as (1984), and well before the emergence of the elaborate functional structures we have at our disposal today, the direction to go was to transfer to the syntax, as much as possible, what erstwhile was lexical, starting with inflection, which, following Halle (1973), came to be regarded as strictly lexical. In what was at the time a very controversial move (but most certainly no longer is), I argued in a NELS paper from 1984 that rules of inflection are syntactic. Particularly significant in advancing the cause of moving away from the lexicon in the 80's was Mark Baker's work (1985, 1988). Baker did not argue for syntactic rules of word formation then, or for that matter now, but his morphological structures were constrained by syntactic structures and interacted with them in such a way, that sooner or later, one had to ask why they are lexical, and not syntactic. Similar questions cropped up relative to the seminal work of Hale and Keyser (1993),

which proposed, effectively, a syntax of word formation lodged in the lexicon, but as subsequent work repeatedly showed, a failure to integrate it into the syntax proper may have helped with some lexical exceptions, but that was at the cost of major structural and theoretical redundancy. Work integrating Hale and Keyser (1993) into the syntax proper therefore started almost immediately, including my own earliest contribution from 1994, and within the budding DM framework and in particular, in Harley's (1995) dissertation.

It is worth noting that adjectival passive remained a tenacious problem. In the late 80's, with the emergence of an articulated analysis for verbal passive which took UTAH as its cornerstone (BAKER, JOHNSON and ROBERTS, 1989), it also became clear that this very approach absolutely cannot handle the argumental properties of adjectival passives, forcing adjectival passive, as a mapping operation, back into the lexicon. It is an appropriate illustration of the enduring force of lexicalism that it took Mark Baker some 15 years to concede (in his 2003 book) that adjectival passive indeed is a problem for UTAH, itself fundamentally entrenched in the existence of articulate lexical entries complete with syntactic information.

Problems internal to GB were definitely a reason I myself as well as others were moving away from the lexicon as a grammatical entity, but this tendency was very strongly enhanced by the fact that generative approaches to morphology have shown tremendous progress starting, roughly, at the mid 70's, with incredibly important contributions such as Aronoff (1976), Selkirk (1982), Lieber (1980), Williams (1981a,b), Pesetsky (1979), Kiparsky (1982) and many others. There is a bit of an ironic twist here. In Remarks on Nominalizations, Chomsky (1970) suggests that morphology is not a generative component. Lexical, to be sure, but without sufficient regularity that could serve as a foundation for a rigorous rule system. However, except for Jackendoff (1975), this was summarily rejected by the generative community, including Halle (1973). Rather, a flurry of work has emerged charting and documenting the myriad of ways in which morphology does lend itself to a generative, rigorous treatment. In turn, the more generative and rigorous the treatments of morphology, the more it all started to look an awful lot like syntax, with constituent structure and combinatorial constituent building operations, cyclical rule operations, and asymmetric structures involving heads and projections, all raising the distinct possibility that what is being constructed is a parallel structural system, with all the properties of the syntax, but, as it remains in the lexicon, still allowing exceptions to be stated more easily. To be sure, some syntactic structures as we perceived them at the time couldn't be accommodated within complex words, and most particularly, the X'-schemata of projection and obligatory complementation, but beyond that, it frequently appeared like the reason WF remains in the lexicon and is not integrated into the syntax is because of the existence of item-specific exceptions. By some rarely articulated consensus, then, morphology became a rigorous rule component (still distinct from the syntax) which nonetheless more easily accommodated 'exceptions', because, by virtue of being placed in the lexicon, it had access to item-specific information. The inherent contradiction between the existence of a rigorous rule system of any sort and in any component, and the license for that rule system, but no other, to access listed exceptions, is evident once spelled out.

We are now more or less at the early 90's, and several accounts start emerging which attempt to take on board the syntactic properties of word formation. Particularly noteworthy is Lieber (1992) and Ackema (1995) and subsequently Ackema and Neeleman (2004). These accounts developed along almost opposite lines. Lieber (1992) accommodated the difficulties WF had with X'-theory by integrating into her syntax phrase structure rules that were, in effect, specific to word formation, and not otherwise attested. Ackema (1995) and Ackema and Neeleman (2004), on the other hand, assumed that while word formation is syntactic in nature, it operates at the sub-zero level of the X'-theoretic projection, and therefore below the complementation level. I was myself developing a model at the time called Parallel Morphology. The idea was that rules of word formation are combinatorial but not syntactic, as such, and hence problems with both X'-theory or insertion frames do not emerge. However, these rules could apply to the output of syntactically generated structures, or alternatively, to lexically available terminals, pre-syntactically. The output would differ, because in one case the combination created by the word formation rule would be associated with a fully developed syntactic structure and its properties which it could not impact, while the 'lexical' output could be inserted into the syntactic structure as such without requiring additional syntactic structure. To illustrate, a verb, V, if it comes together with some nominalizing affix NOM would give rise to the structure $[_N \text{ NOM } [_V \text{ V }]]$. This structure could be inserted into the syntax as N, and have the syntactic properties of N. The V embedded within it would be syntactically inert. However, $[_N \text{ NOM } [_V \text{ V }]]$ could also be inserted to the output of a syntactic operation that would bring V and NOM together, say head movement, as in the simplified syntactic structure in (7):

$$7. \quad [_{NP} \text{ V+NOM } \dots [_{VP} (\text{subj}) \text{ V obj }]]$$

In (7) V is not inert, because regardless of its incorporation into NOM, it has a copy which projects syntactically, and which allows arguments. And so, the word formation rule and the morphological output may be identical, but given its interaction with different syntactic structures, the properties of the entire configuration may be different.

My original account, in 1991, of the differences between Grimshaw's (1990) complex event nominals (also labeled as process nominals, or argument structure nominals) which preserved the syntactic structure of the verb embedded within them, and so-called result nominals (also labelled R-nominals) where the verb embedded within the nominal appeared to be a non-projecting terminal with no associated VP or argument structure) was exactly based on this particular view in which Word Formation operations were parallel to the syntactic derivation.

Fundamentally, that has remained my position, in the sense that I still think Argument Structure nominals preserve the syntactic structure of the verbal constituent, while R-nominals only contain a bare, non-projecting verb. What has changed, however, is my view that word formation rules are sufficiently formally distinct from syntax to warrant a formal separation, and what cause that change were developments in syntax. In particular, X'-theory, a problem for Lieber (1992), and a major reason for postulating a distinct WF system both for Parallel Morphology and for Ackema and Ackema and

Neeleman, vanished from the syntactic landscape with the emergence of Bare Phrase Structure, and with its disappearance, a significant number of problems for integrating morphology into the syntax vanished as well. At the very same time, the emergence of models of argument structure that were less and less dependent on lexical entries solved the problem for syntactic word formation presented by the vanishing of ‘obligatory’ complements internal to words. I myself endorsed non-lexical models of argument structure at least as early as 1993, (in print in 1994), making the existence of *syntactic* structures without complements (especially for verbs) less problematic. And so, sometime in the mid ’90 I realized that I had run out of principled reasons to assume an independent Word Formation component, as syntax from the mid-90’s onwards could integrate the complexities of word structure with little need for word-specific compromises, especially once one endorses the view, as old as generative grammar itself, that full phonological realization is post-syntactic.

There remained, of course, the last tenacious reason to distinguish word formation from syntax, having to do with the prevalence of non-compositional Content for combined terminals, so-called morphology, when compared with syntactic phrasal outputs. While the generalization, to begin with, has been challenged, in particular by proponents of Construction Grammar and on the basis of the prevalence of phrasal and sentential idioms, I do think there is fundamental truth to the generalization. However, as I already discussed in some detail, there are at present a number of approaches to the modelling of non-compositional Content, all of which have considerable promise. As is the case, I believe, with other ‘exceptions’, whether they concern the non-compositionality of words, or of phrases (=idioms), I firmly believe these are not ‘problems’, but rather, these are names for research questions to pursue.

Perhaps some comments on my own personal linguistic development might be of interest here. I had this idea in summer of 1980 which I pursued in my dissertation, that all language variation stems from what at the time I referred to as grammatical formatives, and what now I would call functors (including both inflectional and derivational). As you know, the idea eventually became very well-known and universally endorsed, although it did take some 15 years for that to happen. From my perspective when I proposed that idea, and much as I continued to think it had to be right, my ability to pursue it was limited, because, as I realized almost immediately at the time, I didn’t really know anything about morphology. And so, I started making it my business to understand morphology. And understanding morphology, or to be a bit more realistic, understanding morphology better, is an extremely challenging task. On my way to my 2013 manuscript, which is probably as close as I am ever going to get to ‘understanding morphology’, I discovered that I had to understand lots of other things, so that I can see better whether they are morphology or not. And so, I had to understand, first and foremost, argument structure, which eventually allowed me to sever the verb from its arguments, a definite progress toward understanding morphology. I had to understand DP structure, not because it had anything inherently to do with morphology, necessarily, but because my analysis of DP structure is what laid the logical foundation for pursuing the radical stripping of insertion frames from lexical items. Linguists have always taken it for granted that e.g., the existential vs. generic meaning of *dogs*

does not come about because of dual lexical listing for *dogs*. But at the same time, the assumption that, say, *move* on its completive reading, *move* on its activity reading, and *move* on its transitive reading are distinct listings was explicitly or implicitly endorsed throughout the bigger part of the past 60 years. And so, I was hoping that once I showed that generics and existential, specific and non-specific, and most notably, mass and count polysemy does not entail double lexical listing, taking these conclusions to the verbal system would become more compelling.

Finally, I had to understand the nature of syntactic categories and the nature of the functional/substantial divide.

During this long journey, there were, to be sure, some digressions, and in particular my work on the maturation of syntax, my work on subjects and control, my work on verb movement in Hebrew and my work on inversion, some of which make for another coherent research complex, are not ‘morphology’. Just about everything else, and in particular the Exo-Skeletal trio of books, all came about as a result of following the morphological scent.

REVISTA LINGUÍSTICA: With the Exo-Skeletal approach you have developed a strong isomorphism between syntax and semantics.

HAGIT BORER: This is not entirely accurate. I don’t think I have an opinion about the relationship between syntax and semantics that’s different from mainstream Minimalism. Some people have explicitly challenged it, but overall, I would say that alongside many others, I believe that a syntactic structure *A* has a unique interpretation call it *I[A]*. Specifically, you take syntactic structure *A* of whatever relevant phrasal size, and you map it onto a semantic formula that would yield its interpretation, *I[A]*. And that’s the only possibility: $A \rightarrow I[A]$ via some rigorous semantic route. But not the other way around. There is no direct mapping between some specific well-defined semantic meaning and syntactic structure, and the same semantic meaning could correspond to distinct syntactic structures. For generative semantics there was full isomorphism, so that synonyms had to have the same syntactic structure. I don’t subscribe to this view, because I don’t think it is empirically or theoretically tenable, and I don’t believe my fellow formal syntacticians subscribe to this view either. A good example I can give you is generic expressions. So, if you look at the myriad of syntactic structures which express genericity within the same language, not to mention across languages, you see singulars both definite and indefinite *a lioness protects her cubs/the lioness protects her cubs*, some bare plurals *lionesses protect their cubs* and finally, even in English, a definite article plus plural in contexts that do not allow bare generics (*the extinction of the dinosaurs*). One could sit and tear out one’s hair to try to figure out whether there are semantic differences between these expressions of genericity, and sometimes there are, and sometimes there aren’t but at best they are very subtle. So here you go. Are there sufficient grounds for saying that we are dealing here with completely different type of genericity, or alternatively, as seems self-evident to me, we can conclude that generic meaning can be expressed in many syntactic ways. But if that’s so, there’s no isomorphism.

(Isabella Pederneira: Now, it’s clearer. Thank you.)

HAGIT BORER: I would say that's the belief of many, if not most formal syntacticians. There is a debate, however, concerning the relative strength of the uni-directional claim. Specifically, some would question the universality of the claim that $A \rightarrow I[A]$. For example, Chierchia (1998 and subsequent) and quite a few in his footsteps subscribe to the view that this uni-directional statement is too strong, and the interpretation of A , as a syntactic structure, may be parametrized. Thus, for Chierchia, in Italian, an NP structure could only map onto a predicative or property interpretation, but the very same NP structure in Chinese is interpreted as *kind*. The range of variations is of course delimited, but I still think that Chierchia, or, say Bošković (2008 and subsequent), would not agree that the same syntactic structure always gets the same interpretation. It follows, of course, that they cannot subscribe to the view that the same semantics would be translated into a unique syntactic structure.

REVISTA LINGÜÍSTICA: Do you think that diachronic linguistic studies can provide important answers to a generative analysis approach?

HAGIT BORER: Of course. Why not? What people spoke 500 years ago or 1,500 years ago is surely natural language? The question is not whether they can provide important answers. The question is how to study languages with no native speakers, and how to reason about diachronic development. The absence of native speakers is crucial, methodologically. I see with some sadness the capitulation, within theoretical syntax, to claims that grammatical judgements are insufficient as sources of evidence, to be replaced with corpus studies, or clumsy experiments to justify the simplest empirical generalizations. To be sure, some experiments are crucial to establish intricate nuances, particular contrasts and so on. But the overwhelming number of both grammatical and ungrammatical judgements reported in your average linguistic article, if we were to drag 40 undergraduate students to a lab to test them, would be at close to ceiling, and even in subtle cases well above the threshold of 'statistically significant' as set up in most published psychological work. It is precisely the ready availability of data, through judgements, which has allowed generative syntax to give rise to so many results, and so many novel observations about the structure of language that have been overlooked during the prior millennia. In the absence of native speakers, diachronic linguistics, just like corpus studies, faces two problems. The first has to do with the fact that the body of data is not 'clean', by which I mean it represents, indiscriminately, multiple speakers who may vary from each other. Synchronic data as obtained through grammaticality judgements is 'clean', because it reflects the single, coherent grammar in the mind the judgement giver. Although, of course, judgements may vary across the community, relative to any given speaker, it is possible to construct their individual grammar coherently and consistently, with subsequent accommodations made for inter-speaker variations. For corpus data and diachronic data, an outstanding problem is the fact that it is not possible to ascertain whether variation, whenever attested, resides in a single grammar, or is a generalization across a speech community with potentially diverse individual variations. Therefore, both diachronic data and corpus data give us evidence about what Chomsky has labelled E-language,

not I-language. The second problem concerns the absence, in both diachronic data and corpus data, of ungrammatical judgements, which have been crucial in delimiting the proper description of any grammatical system. All this does not mean that both diachronic data and corpus data could not be studied fruitfully, especially, in diachronic cases, when the historical development is well charted. Latin and its derivatives are particularly good example of the benefits of diachronic studies. It does, however, mean that conclusions drawn from both corpus data and diachronic data may only yield themselves to incomplete formalization.

REVISTA LINGÜÍSTICA: You have written a trilogy - *Structuring Sense* - and you have been working on morphosyntax, language acquisition and syntax-semantics interface within a grammatical theory you developed. What can we expect as future works?

HAGIT BORER: I should start with a brief comment on the emergence of trilogy. Most of what's in these three books, some 1,400 published pages, was all going to be one book. The oldest part, when I started working on the book in the year 2000 was the analysis of derived nominals, which I started working on in 1990, and that was going to be chapter 5 of the book. Chapter 2 (after the introduction) was going to be DP structure, and chapters 3-4 were to be the argument structure chapters. However, when I was done writing the draft of what was to be chapters 2-4, I already have well over 400 pages. At that point, I decided to set aside the derived nominals part, because this was getting too long, and furthermore, while I was working on the event structure part, I changed my mind about how argument structure works, and the derived nominals part, already written, needed some redoing. Even for what has already been written, OUP suggested 2 books, which with separate introductions and conclusions and some elaboration on some of the analyses made for two full size books. And so, the two books went to press, with the 3rd book to materialize shortly, or so I thought. The 3rd book, however, ended up taking very long time as my thinking went well beyond the derived nominals part to develop a much more articulated perspective on morphology, and in retrospect, that 3rd book, *Taking Form*, should have also been two books as well.

Anecdotally, I wanted to call the trio of books *The Exo-Skeletal Trilogy*, but OUP wouldn't let me as with that title, they claimed, the books would end up in the science fiction section of bookstores...

The next stop, in this XSM journey, I already spoke about, and that's the syntax-phonology interface. There is a lot of important work that's been going on for many years on focus, prosody, and intonation, but that's not what I want to do. I want to look at much smaller structures involving word-internal correlations between structure and phonological realization. To follow on things I said already, why do roots have to have phonology? What does it mean, exactly, that Content is matched with annotated phonological structures? These are all claims which are in my 2013 book, but which I need to understand better. I have work in progress which specifically asks whether the periphrastic divide that we find, phonologically, in participial passives in English, Romance, quite

a few other unrelated languages, corresponds to an important syntactic domain, and if so, can we find reflections of it in synthetic passive, such as that attested in Hebrew or Arabic. To argue that in such cases as well there is some sort of a boundary which, in English is periphrastically expressed but in Hebrew or Arabic is not immediately phonologically visible, requires delving into the details of the relevant phonological systems of these languages, and this would be new to me. I haven't done any serious phonology since 1979. But even if for whatever reason my own efforts do not pan out, I remain convinced that the syntax-phonology and the phonology-Content interfaces are the next linguistic frontier.

REVISTA LINGÜÍSTICA: I thank you immensely for participating in this interview. It is very important to me. It was very enriching, and the Linguistic Post Graduate Program also thanks you very much for your participation.

References

- ACKEMA, P. *Syntax below Zero*. Ph.D. dissertation, Utrecht University, 1995.
- ACKEMA, P.; & NEELEMAN, A. *Beyond Morphology*. Oxford: Oxford University Press, 2004.
- ALEXIADOU, A.; ANAGNOSTOPOULOU, E.; & SCHÄFER, F. The properties of anticausatives crosslinguistically. in Frascarelli, M. (ed.) *Phases of Interpretation*. Berlin: Mouton, 2006, pp. 187-212.
- ARONOFF, M. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press, 1976.
- ARONOFF, M. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press, 1994.
- BAKER, M. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry*, v. 16, pp. 373-416, 1985.
- BAKER, M. *Incorporation*. Chicago: University of Chicago Press, 1988.
- BAKER, M. *Lexical Categories*. Cambridge: Cambridge University Press, 2003.
- BAKER, M. K. Johnson, and I. Roberts. Passive arguments raised. *Linguistic Inquiry* v. 20, pp. 219-52, 1989.
- BEAVERS, J. & A. KOONTZ-GARBODEN. Manner and results in the roots of verbal meaning. *Linguistic Inquiry*, v. 43, n. 3, pp. 331-69, 2012.
- BORER, H. The Syntax and Morphology of Derived Nominals. MS. University of Massachusetts, 1991/1993.
- BORER, H. The projection of arguments. In: BENEDICTO, E.; and RUNNER, J. (eds.) *University of Massachusetts Occasional Papers in Linguistics*, v. 17, 1994.
- BORER, H. The syntactic domain of Content. In *Festschrift for Nina Hyams*. Amsterdam: John Benjamins, 2012.
- BORER, H. The projection principle and rules of morphology. *Northeast Linguistics Society*, v. 14, Article 3, 1984.

- BORER, H. Derived nominals and the domain of Content. *Lingua* v. 141, pp. 71-96, 2014.
- BORER, H. *Taking Form, Structuring Sense Volume III*. Oxford: Oxford University Press, 2013.
- BORER, H. Exo-skeletal vs. Endo-skeletal explanations: syntactic projections and the lexicon. Polinsky, M. and J. Moore (eds.) *The Nature of Explanation*. Chicago: Chicago University Press, (CSLI), 2003.
- BORER, H. *Structuring Sense*, Vols. I, II Oxford: Oxford University Press, 2005.
- BOŠKOVIĆ, Ž. On the Clausal and NP Structure of Serbo-Croatian. *Proceedings of FASL*
- CHIERCHIA, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics*, v. 6, pp. 339-405, 2008.
- CHOMSKY, N. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press, 1965.
- CHOMSKY, N. Remarks on nominalization. in Jacobs, R. A. and P.S. Rosenbaum (eds.) *Readings in English Transformational Grammar*. Waltham, MA: Ginn, 1970, pp. 184-221.
- DE BELDER, M. *Roots and Affixes: Eliminating Lexical Categories from Syntax*. LOT dissertation series. Landelijke: Netherlands Graduate School of Linguistics, 2011.
- De BELDER, Marijke, and Jeroen Van Craenenbroeck. How to merge a root. *Linguistic Inquiry*, v. 46, pp. 625-55, 2015.
- DORON, E. *On the identity of roots*. (ed.) *Theoretical Linguistics* v. 40, Issue 3-4, 2014.
- EMBICK, D. On the structure of resultative participles in English. *Linguistic Inquiry* v. 35, pp. 355-92, 2004.
- GLEITMAN, L. The structural sources of verb meanings. *Language Acquisition* v. 11, pp. 3-55, 1990.
- GRIMSHAW, J. *Argument Structure*. Cambridge, MA: MIT Press, 1990.
- HALE, K. & S. J. Keyser. On argument structure and the lexical expression of syntactic relations. in Hale, Kenneth and S. Jay Keyser. (eds.), 1993.
- HALLE, M. Prolegomena to a theory of word formation. *Linguistic Inquiry* v. 4, pp. 3-16, 1973.
- HARLEY, H. On the identity of roots. *Theoretical Linguistics*, v. 40, pp. 225-76, 2014.
- HARLEY, H. The morphology of nominalizations and the syntax of vP, in Rathert, M. and A. Giannakidou (eds.) *DP, QP and Nominalizations*. Oxford: Oxford University Press, 2009.
- HARLEY, H. *Subjects, Events and Licensing*, Ph.D. dissertation, MIT, 1995.
- JACKENDOFF, R. Morphological and semantic regularities in the lexicon, *Language* v. 51, pp. 639-71, 1975.
- KIPARSKY, P. Lexical morphology and phonology, in Linguistic Society of Korea (ed.) *Linguistics in the morning calm: Selected essays from SICOL-1981*. Seoul, South Korea: Hanshin, 1982a.
- LEVIN, B. & M. Rappaport. The formation of adjectival passive, *Linguistic Inquiry* v. 17, pp. 623-61, 1986.

- LIEBER, R. On the Organization of the Lexicon, Ph.D. dissertation, MIT, 1980.
- LIEBER, R. *Deconstructing Morphology*. Chicago/London: University of Chicago Press, 1992.
- MARANTZ, A. Locality domains for contextual allomorphy across the interfaces. In: *Distributed morphology today: Morphemes for Morris Halle*, eds. Ora Matushansky and Alec Marantz, pp. 95-115. Cambridge: MIT Press, 2013.
- PESETSKY, D. Russian morphology and lexical theory', MS., MIT., 1979.
- PESETSKY, D. Paths and Categories, PhD dissertation, MIT., 1982.
- PYLKKANEN, L. *Introducing Arguments*. Cambridge, MA: MIT Press, 2008.
- RAMCHAND, G. *Verb Meaning and the Lexicon: First Phase Syntax*. Cambridge: Cambridge University Press, 2008.
- RAPPAPORT HOVAV, M. Grammatically relevant ontological categories underlie manner/ result complementarity. In *Proceedings of IATL*, v. 32, pp. 77-98, 2017. Hebrew University of Jerusalem.
- WILLIAMS E. On the notions "Lexically Related" and "Head of a Word". *Linguistic Inquiry*, v. 12, pp. 45-74, 1981a.
- WILLIAMS. E. Argument structure and morphology. *The Linguistic Review*, v. 1, pp. 81-114, 1981b.

A ESTRUTURA ARGUMENTAL E A POLISSEMIA EM NOMINALIZAÇÕES

THE ARGUMENTAL STRUCTURE AND POLYSEMY IN NOMINALIZATIONS

Rafaela do Nascimento Melo Aquino¹

RESUMO

Este texto propõe-se a analisar nominalizações deverbais no português do Brasil (PB), tendo como recorte a relação entre a polissemia presente nesses itens e a sua estrutura sintática. O embasamento teórico desse estudo é o Modelo Exoesquelético (BORER, H. 2005a, 2005b, 2013a), abordagem teórica da Gramática Gerativa que analisa a formação de palavras de um ponto de vista construcionista derivacional. Isto é, as palavras são formadas pelo sistema computacional e apresentam estrutura sintática interna, assim como sintagmas e sentenças. A metodologia da pesquisa é o estudo formal do comportamento sintático-semântico das nominalizações deverbais, tendo como objeto de análise Nomes formados pelos sufixos *-ção* e *-mento* no PB. Os dados analisados foram selecionados do Google e dos corpora *corpus do português NOW* e *corpus brasileiro*. Em alguns casos, também foi utilizada a introspecção. A hipótese desta pesquisa é a de que as camadas funcionais na estrutura sintática do Nome deverbal influenciam diretamente em sua leitura, conforme propostas construcionistas da Gramática Gerativa defendem. O resultado das análises demonstra que a leitura de evento nos Nomes resulta da herança da estrutura de evento de sua contraparte verbal, tendo em sua estrutura interna todas as camadas funcionais formadoras de evento. Os nomes com leitura de resultado, por sua vez, apresentam somente as camadas aspectual e verbal em sua estrutura. Já nos nomes com leitura de entidade, o nominalizador é projetado acima da raiz e a tipifica como Verbo. Dessa maneira, a raiz ganha um status verbal, mas não faz parte de uma estrutura argumental.

PALAVRAS-CHAVE: Nominalizações. Interface sintaxe-semântica. Gramática gerativa. Estrutura argumental. Formação de palavras.

ABSTRACT

This paper analyzes deverbal nominalizations in Brazilian Portuguese (BP), focusing the relationship between the polysemy present in these items and their syntactic structure. The theoretical basis of this study is the Exoskeletal Model (BORER, H. 2005a, 2005b, 2013a), a theoretical approach to Generative Grammar that analyzes word formation from a derivational point of view. That is, words are formed by the computational system and present internal syntactic structure, as well as phrases and sentences. The research methodology is the formal study of the syntactic-semantic behavior of the deverbal nominalizations, having as object of analysis nominals formed by the suffixes *-ção* and *-mento* in BP. The analyzed data were selected from Google and from the corpora *corpus of the Portuguese NOW* and *Brazilian corpus*. In some cases, introspection was also used. The hypothesis of this research is that the functional layers in the syntactic structure of the deverbal nominals directly influence its reading, according to the different constructionist proposals of the Generative Grammar. The results of the analyzes demonstrate that the event reading in these nominals results from the inheritance of the event structure of its verbal counterpart, having in its internal structure all the event-forming functional layers. Nouns with result reading, in turn, present only the aspectual and verbal layers in their structure. In nominals with entity reading, the nominalizer is projected above the root and typifies it as a verb. In this way, the root gains a verbal status, but it does not present an argument structure.

¹ Pesquisadora de Pós-Doutorado do Programa de Pós-Graduação em Linguística da UFRJ, litteramelo@gmail.com, <https://orcid.org/0000-0002-2108-7550>

KEYWORDS: Nominalizations. Syntax-semantic interface. Generative grammar. Argument structure. Word formation.

Introdução

No estado da arte dos estudos formais sobre nominalizações deverbais, diversas pesquisas (e.g. ALEXIADOU & BORER, 2020; AQUINO, 2016, 2021; AQUINO, PEDERNEIRA & LEMLE, 2018; BORER, 2013; CHOMSKY, 1970; GRIMSHAW, 1990) têm tido interesse em compreender a relação morfossemântica complexa existente entre as nominalizações e os verbos dos quais são derivadas, bem como o modo pelo qual as propriedades nominais e verbais são geradas nesses nomes, uma vez que há diferenças sintático-semânticas entre as nominalizações e os verbos. Um exemplo dessa diferença, apresentado em (1), é a maneira como os argumentos são expressos nas nominalizações: ou por meio de um sintagma preposicionado, ou por meio de um genitivo.

- 1) a. Os bárbaros destruíram a cidade.
- b. A destruição da cidade pelos bárbaros.

Outro aspecto importante para o estudo das nominalizações é a polissemia presente nesses vocábulos, que podem apresentar leituras de evento, de resultado e de entidade, conforme os exemplos (2) e (3) demonstram.

- 2) a. Você pode adiantar o *carregamento* dos lotes para mim? (evento)
- b. Três *carregamentos* de madeira foram apreendidos pela polícia. (entidade)
- 3) a. A *abordagem* de pacientes alcoólatras pelos agentes de saúde deve ser respeitosa. (evento)
- b. Eu concordo com a sua *abordagem* do tema. (resultado)

(AQUINO, 2021, p. 104)

A diferentes leituras veiculadas pelas nominalizações foi estudada por Grimshaw (1990), que propôs três tipos de nominais: *Complex event nominals*, *Result nominals* e *Simple event nominals*. A principal diferença entre esses nomes, dentre muitas apontadas pela autora, é a que Nomes de evento complexo, diferentemente dos Nomes resultativos e Nomes de evento simples, mapeiam obrigatoriamente argumentos por apresentarem estrutura argumental. Desse modo, a autora defende que os sintagmas preposicionais iniciados pelas preposições *de* (*of-phrase*) e *por* (*by-phrase*) são argumentos desses Nomes. Essa divisão tem sido tomada como base de diferentes estudos sobre nominalizações, como os de Alexiadou e Borer (2020), Aquino, Lemle e Pederneira (2018), Aquino (2019, 2021) e Borer (2005, 2005b, 2013).

Dentro desse contexto, esse texto se propõe a estudar a relação entre a estrutura argumental e a polissemia das nominalizações com o intuito de compreender até que ponto a estrutura sintática dessas palavras contribui para as diferentes leituras nesses nominais, tendo como objeto de estudo nomes deverbais do Português do Brasil formados com o sufixo *-ção* e *-mento*². O embasamento teórico desta pesquisa é o modelo localista de gramática Exoesqueletal (BORER, 2005a, 2005b, 2013) da Teoria da Gramática Gerativa. Nesse modelo, assim como na Morfologia Distribuída, a formação de palavras, sintagmas e sentenças ocorrem no sistema computacional da gramática, tendo, como unidades básicas do processo de formação de palavras, as raízes (itens fonológicos) e funtores (sintáticos e semânticos).

A motivação principal para a escolha dessa abordagem, que será mais detalhada na parte de Fundamentação Teórica, está relacionada à assunção de que raízes são esvaziadas de conteúdo semântico (BORER, 2013a) e não introduzem argumentos (BASSANI & MINUSSI, 2015). Essa hipótese acerca das raízes possibilita uma maior abertura para o componente sintático gerar estruturas com informações bem claras a serem lidas pelos componentes fonológico e semântico da gramática. Isso significa, dentro do recorte de estudo deste artigo, que palavras complexas, dependendo de sua composição sintática, podem receber leitura não-composicional em camadas mais altas da derivação, conforme a proposta de Borer (2013c), pois são as relações entre os nós funcionais e lexicais presentes na estrutura sintática que auxiliarão ao componente enciclopédico a definir a leitura possível para a palavra, não a sua raiz. É importante notar que, nessa proposta teórica, o conteúdo enciclopédico das palavras é inserido após a derivação ser enviada para a Forma Fonológica, e a inserção de conteúdo é limitada pela primeira projeção estendida dominando um núcleo lexical (C-core) na derivação. (cf. BORER, 2013c, p. 455)

Diante disso, a hipótese dessa pesquisa é de que as estruturas sintáticas possíveis para os Nomes deverbais delimitarão as suas leituras possíveis. As questões base desse estudo são: (i) Quais são as camadas sintáticas presentes na estrutura morfossintática dos nomes deverbais? (ii) Como os nós funcionais presentes na estrutura dos nomes deverbais influenciam as diferentes leituras?

Para a análise dos dados, foi usada uma metodologia de análise formal das estruturas dos nomes deverbais, às quais foram aplicados alguns dos testes gramaticais propostos inicialmente por Grimshaw (1990) e ampliados por Borer (2013b) (cf. quadro 1), com o objetivo de verificar o comportamento dos dados do PB diante da tipologia proposta pelas autoras. Após essa fase de análise, foram observados os fatores, considerando a interface sintaxe-semântica, que contribuem para a diversidade de leituras nesses nomes.

² Esse texto não se propõe a discutir a relevante questão sobre as possíveis diferenças semânticas entre os sufixos *-ção* e *-mento*, mas tem como foco, principalmente, a estrutura sintática dos nomes formados por eles. Para saber mais sobre essa questão, veja Oliveira (2005) e Aquino (2021).

Quadro 1: Propriedades de Nomes deverbais

Nomes de evento complexo	Nomes de resultado
Não aceita determinante indefinido	Aceita determinante indefinido (um, uns)
Não aceita demonstrativos	Aceita demonstrativos (esse, essa)
Aceita determinante definido	Aceita determinante definido
Não podem pluralizar	Podem pluralizar
Leitura de evento	Sem leitura de evento
Argumentos obrigatórios	Argumentos não são obrigatórios
Compatível com modificação aspectual como <i>em três horas</i>	Não é compatível com modificação aspectual
<i>Constante e frequente</i> com o singular	<i>Constante e frequente</i> somente com o plural
<i>By phrase</i> é um argumento	<i>By phrase</i> não é argumento.

Fonte: Borer (2013c, p. 1, tradução da autora)

Os dados utilizados nesta pesquisa foram selecionados do *site Corpus do português – corpus NOW*³, do *Dicionário de Nomes deverbais*⁴ e de outras páginas diversas consultadas na internet com o auxílio do site de busca *Google*. Houve também o uso da introspecção para a escolha de alguns dados, principalmente em casos possíveis na língua que não foram encontrados nas buscas. A seleção dos dados foi feita seguindo os seguintes fatores: por sufixo, por tipo de leitura (evento, resultado, entidade) e tipo de estrutura da base verbal (inergativa, transitiva etc.).

Dito isso, o artigo seguirá a seguinte estrutura. Na seção a seguir, será feita uma breve descrição da fundamentação teórica, relacionando-a com o objeto de estudo. Na seção 3, veremos a análise dos dados, e, na seção 4, terá a discussão das análises e as considerações finais do artigo.

1. Fundamentação teórica

O Modelo Exoesquelético – XS (BORER, 2005a, 2005b, 2013a, 2013b, 2013c), abordagem teórica aqui adotada, assume uma visão fortemente computacional da gramática, na qual as propriedades sintáticas, como estrutura argumental e categoria sintática, são propriedades determinadas pela estrutura sintática e não pelo item em si. A teoria vai mais além ao propor que a estrutura também é responsável por oferecer instruções claras para o significado final das palavras.

Nesse modelo, assume-se, então, uma arquitetura de gramática em que há somente um componente computacional, a sintaxe, que opera na formação de hierarquia e na manipulação de constituintes na gramática. Nas operações realizadas por esse sistema, estão incluídas as operações que são responsáveis pela estrutura interna de palavras, isto é, as operações derivacionais que criam palavras complexas, como *globalização* e *guarda-roupa*. Assim como defendido pela MD, no XS as palavras não são a matéria-prima da sintaxe e não são atômicas. Por palavra entende-se um constituinte que contenha estrutura sintática e que corresponde a uma unidade fonológica de um determinado tamanho.

³ <https://www.corpusdoportugues.org/now/>

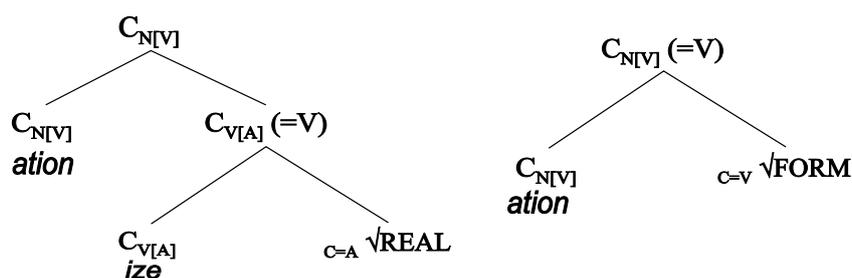
⁴ <http://www.portaldalinguaportuguesa.org/main.html?action=derdict>

Nesse sistema, a sintaxe opera com *raízes* (ou listemas) e *funtores gramaticais*. Na base do sistema há um reservatório de raízes, que não estão associadas com nenhuma informação gramatical. Em outras palavras, as raízes não têm categoria sintática definida, não apresentam um ambiente de inserção morfológica ou sintática específico e não contêm marcação morfológica específica, seja ela flexional ou derivacional. As raízes são índices fonológicos sem nenhum conteúdo conceitual. No processo de formação da estrutura sintática, as raízes se concatenam (*merge*) aos funtores e, em seu ambiente de concatenação, são categorizadas.

Em adição ao reservatório de raízes, há na gramática um *Léxico funcional* que contém *funtores gramaticais*. Esses funtores são definidos pela GU e fazem parte de um reservatório finito, sendo classificados em funtores categoriais (*C-functors*) e funtores semânticos (*S-functors*). A responsabilidade desses itens funcionais é designar as funções sintáticas e semânticas da estrutura através da derivação e dos contextos sintáticos. Isso quer dizer que os funtores são responsáveis por projetar e licenciar nós funcionais e lexicais. Alguns exemplos de funtores são os determinantes, os marcadores de tempo, os auxiliares, os marcadores de número, entre outros.

Os funtores categoriais (*C-functors*) projetam uma categoria e definem um espaço de complemento categorial (*Categorial Complement Space – CCS*) que equivale a uma outra categoria. Além disso, o funtor categorial também pode se concatenar à projeção de outro funtor, desde que essa projeção seja equivalente ao seu CCS. Por exemplo, um funtor do tipo $C_{N[V]}$, que pode ser representado fonologicamente como *-ation*, *-ance*, *-ment* etc. em inglês, ou como *-ção*, *-mento*, *-agem* etc. em português, projeta um N e define seu CCS como equivalente à classe de Verbo. Esse funtor pode ter em seu CCS uma raiz, que será equivalente a Verbo, ou um funtor que projeta um V, como por exemplo o funtor $C_{V[A]}$. Observe essas possibilidades de categorização nas estruturas das palavras *realization* e *formation* na figura 1.

Figura 1: Categorização no Modelo Exoesqueletal



Fonte: Borer (2013c, p. 13)

Os funtores semânticos (*S-functors*), por sua vez, têm como função atribuir valores semânticos a um nó funcional sem valor semântico marcado dentro de projeções estendidas (ExP). O modelo entende a projeção estendida como um conjunto de segmentos (ExP-segments) que recebem valores dos funtores semânticos e coletivamente definem um CCS particular. Como exemplo, considere a

notação, na qual D é um segmento ExP semanticamente valorado por um functor semântico como *the*, em inglês, ou *o*, em português. Nessa notação, {Ex[N]} é o conjunto de segmentos de uma projeção estendida que coletivamente definem seu CCS como equivalente a N. As expressões nominais do tipo *o menino forte* e *o céu azul* exemplificam esse caso. A leitura semântica atribuída por esses funtores é puramente formal. É importante frisar que a teoria assume uma diferença entre a semântica dos itens funcionais e o conteúdo conceitual que as palavras irão receber. É assumido que o sistema computacional não tem por si só a função de criar conteúdo conceitual e que também não é afetado gramaticalmente por ele.

Em relação às raízes, é importante ter em mente que, nessa teoria, esses itens não tomam argumentos e a sua categoria será definida pelo espaço categorial do functor que a domina. Assim como na MD, as raízes são acategoriais e precisam estar vinculadas a um functor para receber uma categoria. No entanto, diferente dessa teoria, no Modelo Exoesqueletal, um functor que gera um nome e define seu espaço categorial como verbo (e.g. $C_{V[N]}$) pode ser concatenado a uma raiz e defini-la como um verbo. Ou seja, não é necessário um nó verbalizador entre esse functor e a raiz para que ela seja verbalizada.

No que diz respeito às estruturas sintáticas geradas pelo sistema, Borer (1993, 2003, 2005b, 2013a) propõe a existência de estruturas sintáticas universais (i.e. *templatos*), representadas simplificada em (4). Nessa proposta, as diferenças aspectuais listadas pela *Aktionsart*, especialmente no que tange à telicidade, são realizadas sintaticamente por projeções sintáticas aspectuais (ASP_Q , F^{shl}), e os argumentos da estrutura sintática recebem sua leitura a partir da sua posição nessa estrutura.

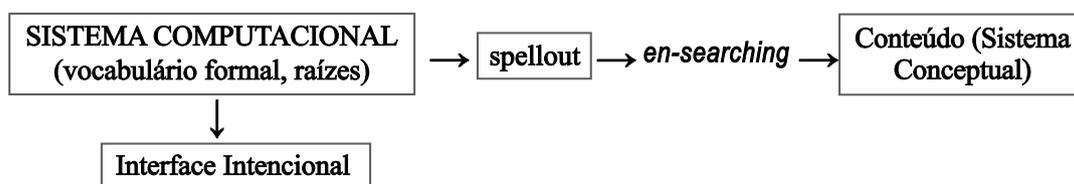
- 4) a. Transitivo, Télico:
 [EP DP1 <e>E [TP DP1 [ASPQ DP2 [VP V]]]] (em duas horas/*por duas horas)
 ORIGINADOR NOM ACC
- b. Transitivo, Atélico:
 [EP DP1 <e>E [TP DP1 [FP DP2 [VP V]]]] (*em duas horas/por duas horas)
 ORIGINADOR. NOM PRT
- c. Intransitivo, Télico:
 [EP DP1 [TP DP1 [ASPQ DP1 [VP V]]]] (em duas horas/*por duas horas)
 S-O-Q NOM S-O-Q
- d. Intransitivo, Atélico:
 [EP DP1 <e>E [TP DP1 [VP V]]]] (*em duas horas/por duas horas)
 NOM

(BORER, 2005b, p. 6, tradução da autora)

No decorrer da derivação, as estruturas sintáticas criadas a partir da concatenação destes elementos são enviadas para a Forma Fonológica – FF (*Phonological Form* – PF), a partir de *Spell-out*. Após o *Spell-out*, um mecanismo de busca enciclopédica é ativado e o conteúdo conceitual é combinado à estrutura criada. Diante dessas explicações, podemos compreender que o modelo de gramática proposto pelo XS tem esta representação de Borer (2014a, p. 94), da figura 2, no qual,

segundo Borer (2014a), não há uma interface conceitual-intencional unificada, de modo que, enquanto o input da semântica formal (Interface intencional), que é indiferente ao *spell-out*, é a estrutura sintática e as propriedades gramaticais dos funtores, o Conteúdo (Sistema Conceitual), que também acessa a estrutura sintática, é indiferente às propriedades formais dos funtores, mas inteiramente ligado a sua realização fonológica.

Figura 2: Modelo Exoesquelético de gramática



Fonte: Borer (2014a, p. 94, tradução da autora)

Além das propostas acima descritas, Borer (2014a), partindo da divisão proposta por Grimshaw (1990), propõe que a diferença entre os tipos de nomes deverbais, agora denominados *Argument Structure nominals* (AS- nominals) e *Referential nominals* (R- Nominals), é proveniente de estruturas sintáticas diferentes. A linguista parte da tipologia apresentada por Grimshaw (1990) para os Nomes deverbais e propõe propriedades adicionais a estes Nomes (cf. quadro 1).

Diante dessas propriedades, a pesquisadora analisa os casos de Nomes referenciais e os casos de Nomes com estrutura argumental. Ela propõe, então, que, no que diz respeito ao conteúdo semântico, os Nomes com leitura de evento (AS-nominals) sempre apresentarão conteúdo composicional, ou seja, o significado do todo será compreendido a partir do significado das partes. Essa proposta está pautada no fato de que esses Nomes apresentam, em sua estrutura sintática, uma projeção estendida verbal (VP) que bloqueia o mecanismo de busca enciclopédica, impossibilitando o surgimento de conteúdo não composicional.

No entanto, os Nomes referenciais (*R-nominals*) não apresentam essa projeção estendida, e, dessa forma, as leituras não composicionais são possíveis. Para estes Nomes, Borer assume que o item funcional (ou funtor categorial) concatenado à raiz da palavra torna essa raiz equivalente a Verbo, não um Verbo propriamente dito. Além disso, é importante ressaltar que, para Borer, não há diferença entre os nomes de resultado e de entidade apresentados no início deste texto. A autora considera os dois tipos como Nomes referenciais. No entanto, essa falta de distinção não é assumida nesta pesquisa e, como será visto na seção de análises, os dados de LP corroboram essa distinção.

No que diz respeito às camadas sintáticas presentes na estrutura dos AS-nominals, Borer (2013a) defende que uma estrutura de evento esteja encaixada sob o nominalizador (N). Desse modo, segundo a autora, esse item funcional é adjungido à projeção EP (Event Phrase – Sintagma de Evento). Essa projeção, por sua vez, tem como função licenciar o argumento do evento e é obrigatória para fazer emergir a estrutura de evento e a interpretação eventiva.

Conforme essa proposta, a estrutura dos Nomes argumentais contém todas as camadas funcionais e aspectuais presentes na sua contraparte verbal. Já para os Nomes referenciais, a pesquisadora propõe que esses Nomes não apresentem estrutura argumental, de modo que sua formação sintática se daria sem uma projeção estendida verbal (VP).

Com a fundamentação teórica descrita, na próxima seção, as análises das nominalizações em LP serão apresentadas.

2. Análise sintática de Nominalizações

2.1. Nominalizações que denotam evento

Conforme visto na introdução, os estudos de Grimshaw (1990) e Borer (2013a) apresentam diferentes evidências para defender a presença de estrutura argumental/eventiva nas nominalizações com leitura de evento. Duas delas, entre outras, são a modificação aspectual e a modificação com Adjetivos orientados a agentes. Em relação à presença de argumentos, no seu estudo comparativo do inglês e do hebraico, Borer (2013a) observa que há a possibilidade de o argumento externo agentivo realizado pelo *by-phrase* não ocorrer junto com o Nome, porém a recuperação desse sintagma é possível através de um controle implícito de argumentos. Por outro lado, Borer (2013a) defende que os AS-*nominals* têm o argumento interno obrigatoriamente realizado. Quanto à propriedade dos Nomes de evento selecionarem argumentos, observemos (5).

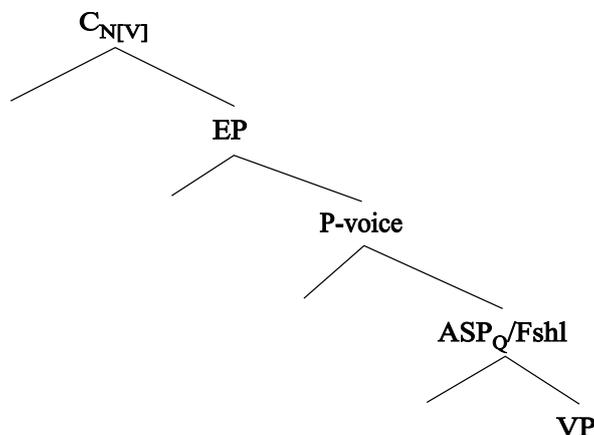
- 5) a. O *recrutamento* dos soldados pelo exército durou um mês.
 b. * O *recrutamento* pelo exército durou um mês.
 c. O *recrutamento* dos soldados durou um mês.
 d. O *recrutamento* durou um mês.

Em (5), o nome deverbal *recrutamento* denota um evento. Essa leitura é evidenciada pelo predicado *durou um mês*, que expressa a realização de um processo. Nas sentenças acima, o nominal é derivado de um verbo transitivo e, em (5a), mapeia os dois argumentos previstos pelos estudos de Grimshaw (1990) e Borer (2013a), isto é, o argumento interno (expresso como um *of-phrase*) e o argumento externo (expresso como um *by-phrase*).

Em (5b), porém, o argumento externo ocorre sozinho, sem o argumento interno expresso por *of-phrase*. Nesse caso, a sentença torna-se agramatical. Além disso, o que ocorre nesta sentença parece corroborar a defesa feita por Borer de que os argumentos internos dos nomes com estrutura argumental são obrigatoriamente expressos, principalmente quando o argumento externo está realizado. Esse fato também é observado por Lebeaux (1986 *apud* GRIMSHAW, 1990), Grimshaw (1990) e Oliveira (2014). Esses autores afirmam que, se estiver presente o argumento externo realizado por um *by-phrase*, o argumento interno deve ocorrer obrigatoriamente. Nesse sentido, podemos supor que, na estrutura do nome com leitura de evento, o nominalizador é projetado acima de uma estrutura

em que há camadas funcionais responsáveis pela projeção do argumento externo (e.g EP, P-voice) e do argumento interno (ASPQ ou Fshl) dominando o VP, conforme a representação na figura 3.

Figura 3: Proposta de estrutura para Nomes deverbais com leitura de evento



Fonte: Aquino (2021, p. 126)

Em (5c) há o caso inverso ao de (5b), pois em (5c) o Nome seleciona apenas o argumento interno, e a aceitabilidade da sentença não é afetada. Esse dado indica que, de fato, o argumento externo pode não ser expresso fonologicamente, embora seja possível recuperá-lo.

Já em (5d), o Nome ocorre sem os argumentos e ainda apresenta leitura de evento. Neste caso, além da leitura de evento ser mantida, a sentença mantém ainda a gramaticalidade e a aceitabilidade. Esse exemplo demonstra que também é possível não expressar o argumento interno do Nome, sem alterar a sua gramaticalidade. Na literatura, esse fenômeno também foi apontado no estudo de Sleeman e Brito (2007). Diante disso, é possível questionar se a leitura eventiva do Nome está de fato diretamente relacionada à propriedade do nominal de selecionar argumentos. Na tentativa de compreender esse fato, observemos dados de estruturas verbais no PB, nas quais não há a expressão fonológica do objeto.

- 6) a. A Joana viu ____ na TV ontem.
 b. Eu informei à polícia da possibilidade de o Manuel ter guardado ____ no cofre da sala de jantar.⁵

(CYRINO, 1994, p. 1)

O fenômeno de objeto nulo, representado em (6), tem sido estudado por diferentes autores (cf. KATO, 1993, 2011; CYRINO, 1994, 2000; JANSEN, 2016), que analisam o objeto nulo no PB como um elemento pronominal de caráter fórico. Assim como o objeto nulo, o argumento interno não realizado dos Nomes deverbais pode ser recuperado pelo contexto. Diante disso, surge a questão de se a possibilidade de ocorrência do objeto nulo no PB estaria influenciando a não obrigatoriedade

⁵ Segundo Cyrino (1994), esta sentença em português europeu (PE) é agramatical.

da realização fonológica do argumento interno das nominalizações em PB. Além disso, segundo Grimshaw (1990, p. 49), “*obligatoriedade* deve significar o mesmo para Nomes e Verbos: capazes, em princípio, de serem obrigatórios, mas talvez sujeitos à variação lexical. Afinal, até mesmo os objetos diretos dos Verbos podem ser opcionais.”⁶ Podemos, então, concluir que a falta de expressão do argumento em certos contextos não implica a falta da capacidade de Nomes deverbais que denotam evento selecionarem argumento, um fato evidenciado por (5a).

Em relação ao argumento externo, sabemos, com evidências de outras línguas e de (5b), que a falta de realização desse argumento é possível. Além de (5), podemos ver que em (7a) há a possibilidade de o argumento externo do Nome ser realizado com um *of-phrase* agentivo. A comparação entre os exemplos (5) e (7) mostra que o *sujeito* pode ser expresso tanto por um *by-phrase* como por um *of-phrase*.

- 7) a. A *latição dos cachorros* durou o dia todo.
 b. *A *latição pelos cachorros* durou o dia todo.
 c. A *latição* durou o dia todo.

É relevante notar que o Nome *latição* acima apresenta somente a leitura eventiva, pois faz parte do grupo de nomes deverbais que apresentam leitura iterativa (cf. Aquino, 2021). No entanto, quando um Nome deverbal apresenta polissemia, como os Nomes *avaliação* e *entrada*, representados em (8), há a possibilidade de o sintagma iniciado pela preposição *de* comportar-se como um modificador genitivo ou como um sintagma agentivo relacionado ao sujeito verbal (i.e., o argumento externo).

- 8) a. A *avaliação de João* estava longa. (genitivo)
 b. A *avaliação de João* durou o dia todo. (agentivo)
 c. A *entrada da escola* fica na rua Manoel Reis. (genitivo)
 d. A *entrada dos alunos* durou uma hora. (agentivo)

Os dados apresentados até aqui corroboram a proposta de que os Nomes eventivos mapeiam argumentos, ainda que em alguns casos eles não sejam expressos. Esse fato nos mostra que, no processo de derivação sintática de Nomes com leitura de evento, a estrutura argumental é projetada de maneira semelhante à estrutura argumental verbal. Ou seja, em Nomes derivados de Verbos com estrutura transitiva, há o mapeamento dos argumentos interno e externo, como em (8); em Nomes derivados de Verbos inergativos, também como em (8), há o mapeamento do argumento externo; e, em Nomes derivados de Verbos inacusativos, representado no exemplo (9) a seguir, há o mapeamento do argumento interno.

- 9) A síndrome do olho seco pode evoluir para a *ulceração* da córnea.

⁶ Of course, obligatory must mean the same for nouns as for verbs: capable in principle of being obligatory but perhaps subject to lexical variation. After all, even direct objects of verbs are sometimes optional. (GRIMSHAW, 1990, p. 49)

A partir da discussão acima, já é possível supor que há evidências de que os Nomes que denotam eventos no PB são formados pelo mesmo caminho derivacional de sua contraparte verbal. Desse modo, é possível assumir para a sua estrutura a presença da projeção estendida verbal (VP). Outros estudos na literatura (Alexiadou, 2001, 2020) também propõem essa camada para as nominalizações deverbais do inglês, do hebraico, do holandês, do catalão e do russo. Além da projeção estendida verbal, assumindo a proposta de estrutura de evento de Borer (2013a, 2013b), outras camadas funcionais podem ser consideradas para a estrutura interna dos Nomes eventivos. Estas são: ASP_Q / F^{shl} , *P-voice*, EP.

Tomemos outra evidência para a presença dessas camadas nos Nomes eventivos, a modificação aspectual. Essa modificação é exercida por expressões temporais do tipo *por X tempo*, *em X tempo*, *às X tempo* e *durante X tempo* (cf. VLENDER, 1967; DOWTY, 1979). Observe os exemplos a seguir.

- 10) a. O planejamento do seminário pela comissão (*em dois meses*).
b. A reativação do site (*às sete horas*).

(AQUINO, 2021, p. 120)

Como os dados acima demonstram, a inserção das expressões temporais como modificadores da nominalização comprova a sua leitura de evento, pois essa modificação pode delimitar o evento descrito pelo nominal. Esse fato nos fornece ainda mais evidência de que esses Nomes apresentam as camadas funcionais presentes na derivação de sua contraparte verbal. Desse modo, é possível assumir que os Nomes com leitura de evento apresentam uma estrutura sintática em que o nominalizador ou funtor categorial é projetado acima da estrutura de evento de sua base verbal. Isso indica que abaixo da camada nominal estarão as projeções EP, ASP_Q / F^{SHL} e V, propostas por Borer (2013a) como componentes da estrutura de evento dos Verbos.

Outro ponto importante a ser considerado na análise das nominalizações é a presença do argumento externo *by-phrase*. Autores como Roeper e van Hout (1999) e Borer (2013a) apresentaram propostas para a posição sintática responsável por liberar o argumento externo nas nominalizações. Roeper e van Hout (1999) propõem que o agente da nominalização é projetado em *VoiceP*, uma projeção mais alta que o VP⁷. Borer (2013a), no entanto, com dados do hebraico e do inglês, propõe que o agente expresso como *by-phrase* é projetado em um subevento encaixado no evento principal, um *eventinho* (*e*). Esse sub-evento é responsável por licenciar a voz passiva analítica, ou algumas formas participiais. Esse *eventinho* é dominado por um operador chamado de *Passive Voice* (*P-Voice*). Nessa proposta, a leitura de *originador* emerge em *e*. Desse modo, as nominalizações que contêm *by-phrase* são formadas de modo semelhante às sentenças passivas.

⁷ Para os dados do inglês, os autores afirmam que o agente do *Spec* do *VoiceP* é alçado para *Spec* de DP, podendo gerar o Nome *the enemy's destruction of the city*. Nesse texto, os autores também analisam as nominalizações de Adjetivos deverbais formados em *-able*, como *learnability*. O adjetivador *-able*, segundo os autores, se comporta como um passivizador e, quando o Adjetivo formado por ele é nominalizado, não é possível o alçamento do sujeito verbal: **John's learnability of French*. Porém, é possível o alçamento do objeto: *grammar's learnability*. Esse fenômeno ocorre, segundo os autores, porque *-able* projeta um traço [+tema] para o seu *Spec*, impedindo o alçamento do *Spec* do *VoiceP*, que é [+ agente].

Quanto à presença de um *VoiceP* ou *P-Voice* na estrutura das nominalizações do PB, vamos analisar os seguintes exemplos.

- 11) a. A prefeitura construiu o campo de futebol.
 b. O campo de futebol foi construído pela Prefeitura.
 c. O campo de futebol foi construído (pela Prefeitura) no mês passado.
- 12) a. A construção do campo de futebol.
 b. A construção do campo de futebol pela prefeitura.
 c. A construção do campo de futebol (feita pela prefeitura).
 d. ?? A construção da prefeitura do campo de futebol.
- 13) a. A pintura de Tarsila do Amaral.
 b. *A pintura por Tarsila do Amaral.
 c. A pintura feita por Tarsila do Amaral.

(AQUINO, 2021, p.132)

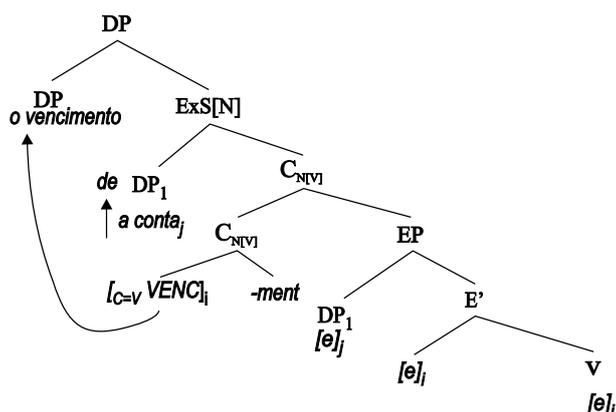
Em (11) e (12), vemos que o agente da passiva e o agente da nominalização transitiva apresentam a mesma realização fonológica: *por* + DP. No entanto, em (13) podemos ver que essa estrutura não é possível quando se refere a uma autoria, diferentemente do inglês, em que se pode dizer *the painting by Tarsila do Amaral*. Além disso, em (11c) observa-se que o PP agentivo pode ser encaixado em um subevento expresso pelo sintagma [feito por DP]; assim como em (12c), em que a autoria pode ser expressa por essa construção. Diante disso, pode-se assumir, assim como Borer (2013a), que o agente expresso por um *by-phrase* é gerado em *e*. Outra evidência para essa hipótese é a estranheza causada na sentença (12d), em que o agente é expresso por um *of-phrase*⁸.

A análise feita até o momento demonstra que os Nomes deverbais formados por derivação sufixal no PB que têm leitura de evento são formados a partir da estrutura argumental do verbo que está em sua base, concordando com outras análises disponíveis na literatura. Dessa maneira, tomando os templates de Borer (2005, 2005a, 2013), em (14), os nomes *vencimento* e *escovação* são formados pelas estruturas representadas nas figuras 4 e 5.

- 14) a. O vencimento da conta será na próxima semana.
 b. A escovação do pelo do cavalo pelo treinador.

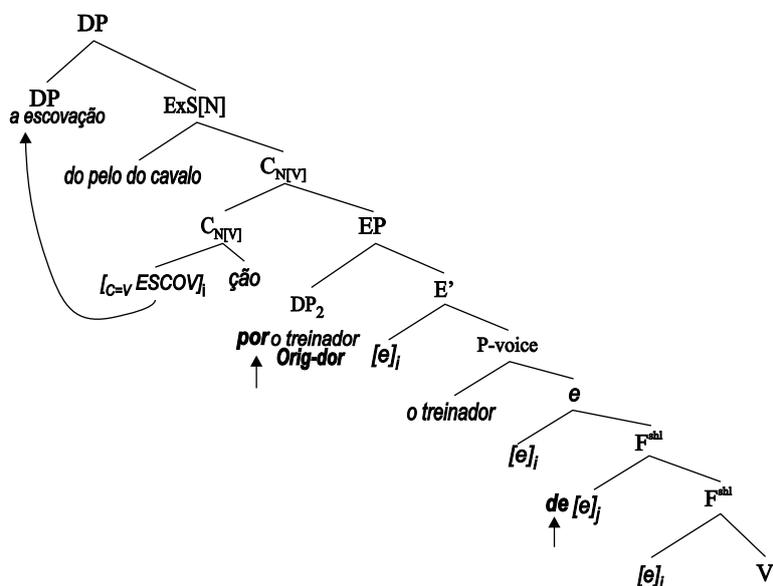
⁸ No inglês e no hebraico, a ordem de palavras Nome Agente Objeto é possível com nominalizações transitivas. Nesse caso, a proposta de Borer (2013a) é que o agente emerge em EP e é alçado para *Spec* de uma projeção estendida no domínio nominal. Em PB, essa ordem de palavras nos parece estranha, até agramatical. Nesse sentido, optamos por considerar que em PB há somente a opção em que o agente da nominalização transitiva emerge em *P-Voice*. Esse seria então um caso de variação paramétrica.

Figura 4: Nome derivado de estrutura inergativa



Fonte: Aquino (2021, p.133)

Figura 5: Nome derivado de estrutura transitiva atética



Fonte: Aquino (2021, p.136)

2.2. Nomes que denotam resultado

Para iniciar a análise dos Nomes com leitura de resultado, é importante relembrar que a noção de *resultado* é entendida como “o *output* de um processo ou um elemento associado com o processo” (GRIMSHAW, 1990, p. 49). Vamos analisar os dados a seguir.

- 15) a. *O planejamento da festa pelos pais foi arquivado.
- b. *O planejamento pelos pais foi arquivado.
- c. O planejamento dos pais foi arquivado.
- d. O planejamento da festa foi arquivado.
- e. O planejamento foi arquivado.

Como evidenciado pelo predicado *foi arquivado*, que seleciona semanticamente um nominal não eventivo, o Nome *planejamento* denota resultado. Em (15a) há um exemplo em que o Nome com leitura de resultado não pode selecionar dois argumentos (i.e., interno e externo). Nesse caso, a sentença é agramatical.

Ainda em relação ao argumento externo expresso como um *by-phrase*, é possível ver em (15b) que o Nome também não seleciona somente esse argumento, uma vez que a sentença na presença desse único argumento se torna agramatical. O exemplo (15c), por sua vez, nos mostra que o DP *os pais*, o sujeito da sentença, é realizado como um modificador *of-phrase* genitivo. Esse caso é diferente dos Nomes com leituras de evento, em que o sintagma preposicional é agentivo. Em (15d) há um caso em que o Nome de resultado pode mapear o argumento interno e ainda manter sua leitura. E em (15e) há um caso em que o Nome aparece sem os argumentos.

Diante desses exemplos e com a discussão feita para os Nomes de evento, surge a questão: os Nomes que denotam resultado apresentam a mesma estrutura dos Nomes que denotam evento? Para responder a essa pergunta, é necessário focalizar alguns dados. Observe as sentenças a seguir.

- 16) a. O planejamento *da festa* foi arquivado.
b. A gravação *da palestra* está na biblioteca.

Em (16) vemos que é possível, em PB, o Nome com leitura de resultado selecionar argumento interno. Conforme defendido por Alexiadou (2001, 2010), o sintagma preposicional que ocorre com o Nome de resultado (no caso dos exemplos acima, os sintagmas *da festa* e *da palestra*) apresenta o mesmo papel temático do complemento de sua contraparte verbal (i.e., tema), indicando que o complemento verbal é herdado pela nominalização. Neste trabalho, assumo a proposta de Alexiadou (2001, 2010) e analiso aqui os sintagmas *da festa* e *da palestra* como argumento interno do Nome, uma vez que o seu papel temático é semelhante ao dos complementos dos verbos *planejar* e *gravar*. Assim, é possível assumir que a camada vP está presente também na estrutura de Nomes com leitura de resultado no PB. Esse fato também já foi observado por Harley (2006) e Sleeman e Brito (2007, 2010) para as línguas inglesa e holandesa.

Além disso, segundo Jorge (1986), Nomes de resultado ainda apresentam algum traço eventivo na estrutura, uma vez que podemos recuperar a ação que produz o resultado. Desse modo, assumindo a proposta de Borer (2005b) para a estrutura de evento, proponho que além da camada verbal há, na estrutura dos Nomes resultativos, a camada ASP_Q/F^{shl} . Essas camadas, como visto no início do texto, são responsáveis por medir o evento e por contribuir para a formação do próprio evento.

Em relação à camada aspectual em Nomes resultativos, diferentes autores propõem essa camada na estrutura desses Nomes. Uma evidência de que há essa projeção na estrutura nominal de Nomes de resultado é a possibilidade de modificação aspectual pela expressão *durante X tempo*. No diagnóstico de Grimshaw (1990), endossado por Borer (2013a), essa propriedade não faz parte dos Nomes de resultado. Porém, observe o seguinte dado apresentado por Reis e Figueiredo (2018).

- 17) O porta-voz[...] afirmou hoje que a conturbação *durante duas semanas* e dificuldades internas do Syriza não justificam a invenção de histórias nem de desculpas.

(REIS, FIGUEIREDO, 2018, p. 192)

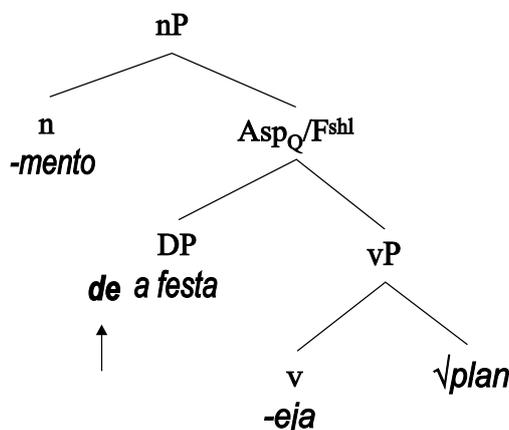
Segundo Reis e Figueiredo (2018, p. 192), no que se refere a Nomes eventivos, “essa expressão indica o desenvolvimento de um processo”, enquanto para Nomes de resultado “indica a duração de um estado”. Esses fatos, demonstrados pelas autoras, contribuem para a hipótese de que a camada aspectual também está presente nas nominalizações. Desse modo, temos evidência, com dados do PB, de que as camadas vP e AspP estão na estrutura de Nomes com leitura de resultado.

Em relação às camadas EP e P-voice presentes nos Nomes eventivos, os exemplos (15a), (15b) e (15c), rerepresentados em (18), demonstram que os PPs agentivos (i.e *of-phrase* e *by-phrase* agentivos) não ocorrem com Nomes de resultado.

- 18) a. *O planejamento da festa pelos pais foi arquivado.
 b. *O planejamento pelos pais foi arquivado.
 c. *O planejamento dos pais (*agente*) foi arquivado.

Esse fato fornece evidência de que as camadas funcionais EP e P-voice não fazem partes desses Nomes. Nesse sentido, os dados apontam para uma proposta de estrutura em que o nominalizador é concatenado acima das camadas aspectuais ASP_Q/F^{shl} e VP, conforme a representação arbórea na figura 6 a seguir demonstra. Essa é a proposta aqui defendida para Nomes resultativos.

Figura 6: Estrutura de Nomes que denotam resultado



Fonte: Aquino (2021, p. 139)

Conforme visto até o momento, os Nomes com leitura de evento e os Nomes resultativos apresentam estruturas sintáticas complexas nas quais diferentes camadas funcionais contribuem para a sua leitura final. No primeiro grupo de nomes, a nominalização contém estrutura argumental, e, no segundo grupo, o Nome é formado com as camadas aspectual e verbal, sem herdar a estrutura argumental completa de sua contraparte verbal. Dito isso, passemos à análise dos Nomes de entidade.

2.3. Nomes que denotam entidade

Além das leituras de evento e de resultado, os nomes deverbais podem apresentar leitura de entidade. Na literatura, a diferença entre os nomes que denotam resultado e os nomes que denotam entidade foi notada inicialmente por Brito e Oliveira (1997). Segundo as autoras, essas nominalizações denotam *produto* ou *coisas* e podem apresentar diferentes valores semânticos: (i) entidade concreta, (ii) entidade concreta coletiva, (iii) entidade concreta de lugar e (iv) entidade abstrata, conforme os exemplos a seguir demonstram, respectivamente.

19)

- a. Segundo a porta-voz, na precária *embarcação* viajam quatro bebês e três mulheres grávidas⁹
- b. O bioma abriga o último remanescente carioca deste tipo de *vegetação*¹⁰
- c. Hoje o *assentamento* conta com cerca de 100 famílias, que trabalham basicamente na produção de frutas e hortaliças de pouca renda¹¹
- d. A novidade apresentada pela Amifest [...]garante ao visitante com vestimenta em conformidade com as regras duas opções de *bonificação*: (1) quatro tickets de água/refrigerante ou (2) dois tickets de chopp¹²

De acordo com Brito e Oliveira (1997), os nomes que denotam entidade, além do comportamento semântico diverso, apresentam comportamento sintático diferente dos nomes com leitura de evento e de resultado. Isto é, os nomes que denotam entidade, segundo as autoras, não podem vir acompanhados com um sintagma preposicional agentivo (tanto *by-phrase* quanto *of-phrase*), não podem selecionar um *of-phrase* como argumento, somente como adjunto, e não podem ser modificados por expressões de tempo do tipo *em X tempo*, *durante X tempo* etc. Esse diagnóstico pode ser confirmado com os exemplos que se seguem.

20)

- a. Hoje o *assentamento* ***pela população** conta com cerca de cem famílias, que trabalham basicamente na produção de frutas e hortaliças de pouca renda.
- b. Hoje o *assentamento* ***da cidade** conta com cerca de cem famílias, que trabalham basicamente na produção de frutas e hortaliças de pouca renda.

⁹ Disponível em: <https://noticias.uol.com.br/ultimas-noticias/efe/2019/06/30/navio-de-resgate-acha-bote-com-40-pessoas-a-deriva-e-se-dirige-a-lampedusa.htm>. Acesso em: 04 ago. 2019.

¹⁰ Disponível em: <https://www.otempo.com.br/superfc/local-onde-ser%C3%A1-aut%C3%B3dromo-do-rio-abriga-quadrilhas-de-ladr%C3%B5es-de-cargas-1.2202515>. Acesso em: 04 ago. 2019.

¹¹ Disponível em: <http://odocumento.com.br/governo-encaminha-projeto-de-lei-que-visa-obrigar-o-presos-a-custear-tornozeleira/>. Acesso em: 07 jul. 2019.

¹² Disponível em: <https://www.promoview.com.br/promocao/traje-tipico-na-oktoberfest-de-igrejinha-vale-bebidas.html>. Acesso em: 04 ago. 2019.

- c. Hoje o *assentamento* ***durante cinco minutos/*em cinco minutos/*por cinco minutos** conta com cerca de cem famílias, que trabalham basicamente na produção de frutas e hortaliças de pouca renda.

21)

- a. A novidade apresentada pela Amifest [...]garante ao visitante com vestimenta em conformidade com as regras duas opções de *bonificação* ***pela empresa**: (1) quatro tickets de água/refrigerante ou (2) dois tickets de chopp.
- b. A novidade apresentada pela Amifest [...]garante ao visitante com vestimenta em conformidade com as regras duas opções de *bonificação* ***da empresa**: (1) quatro tickets de água/refrigerante ou (2) dois tickets de chopp.
- c. A novidade apresentada pela Amifest [...]garante ao visitante com vestimenta em conformidade com as regras duas opções de *bonificação* ***durante cinco minutos/*em cinco minutos/*por cinco minutos**: (1) quatro tickets de água/refrigerante ou (2) dois tickets de chopp.

Diante desse diagnóstico, Brito e Oliveira (1997) e outros autores, como Reis e Figueiredo (2018), assumem que, diferentemente dos Nomes eventivos e de resultado, os Nomes de entidade não são Nomes derivados de Verbos, mas Nomes em que a raiz é diretamente dominada pelo nominalizador. Desse modo, as palavras *embarcação*, *vegetação*, *assentamento* e *bonificação* dos exemplos analisados aqui teriam, respectivamente, as estruturas [$\sqrt{\text{embarca}} + \text{ção}$]_n [$\sqrt{\text{vegeta}} + \text{ção}$]_n, [$\sqrt{\text{assenta}} + \text{mento}$]_n e [$\sqrt{\text{bonifica}} + \text{ção}$]_n.

Essa é uma abordagem possível na MD, pois considera que há uma reanálise das raízes. Porém, há um problema quando observamos a realização morfofonológica de alguns Nomes de entidade e o seu significado. No exemplo (22), a seguir, estão destacadas em negrito as marcações morfológicas verbais dos Nomes listados.

- 22) *bonificação* (*recompensa*), *assentamento* (*lugar*), *vegetação* (*coletivo*), *contribuição* (*imposto*), **acampamento** (*lugar*), **alojamento** (*lugar*), **hospedagem** (*lugar*), **pousada** (*lugar*), **pagamento** (*salário*), **engrenagem** (*conjunto de peças*), **embarcação** (*barco*), **rolamento** (*peça de carro*)

Com o conjunto de dados em (22), podemos ver que, assim como as suas contrapartes *eventiva* e *resultativa*, os Nomes de entidade também podem apresentar complexidade morfológica. Além disso, ainda que tenhamos leitura de *produto* ou *coisa*, esses Nomes não podem ser considerados idiomáticos, o que evidenciaria a reanálise das raízes. Isto é, esses itens não perderam totalmente a relação semântica com a sua raiz, uma vez que podemos resgatá-la, se fizermos uma paráfrase. Por exemplo, os vocábulos *acampamento* e *rolamento* podem ser parafraseados como *lugar onde se acampa* e *peça que faz rolar*. Esse é um caso distinto do que ocorre com os vocábulos *refrigerar* - *refrigerante* (bebida gasosa) e *ramo* - *derramar* (tirar ramo) - *derramar* (líquido), nos quais a

semântica da palavra derivada não emerge por meios derivacionais, mas é arbitrária¹³.

Nesse sentido, considerando que, no arcabouço da MD e do Modelo Exoesqueletal, as peças morfológicas representam nós sintáticos, nos deparamos, então, com a seguinte questão: como analisarmos esses Nomes, sem desconsiderar a complexidade morfológica, já que eles não se comportam sintaticamente como os Nomes eventivos ou de resultado?

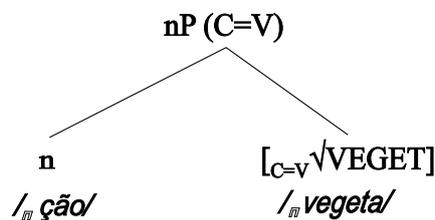
Como vimos no início desta seção, os Nomes com leitura de entidade apresentam um comportamento sintático que indica a falta de estrutura argumental e de evento em sua derivação, motivando as análises em que esses Nomes, em sua formação, não passam pelas camadas verbal, aspectual e outras que formem estrutura de evento. Por outro lado, no exemplo (22), temos evidência morfológica de que esses Nomes podem ser complexos.

Diante disso, proponho resgatar a proposta de Borer (2013a) para Nomes referenciais. Na proposta da autora, esses Nomes são os que Grimshaw (1990) denomina de Nomes de resultado. Porém, considero que a análise feita por Borer (2013a) pode ser aproveitada de melhor maneira para os Nomes que denotam entidade. Nessa análise, os Nomes não herdam estrutura de evento de sua contraparte verbal, mas apresentam uma raiz tipificada em Verbo pelo funtor categorial nominalizador.

Para a melhor compreensão dessa proposta, é importante lembrar como os funtores categoriais funcionam no Modelo Exoesqueletal. Conforme discutido na fundamentação teórica, os funtores categoriais são responsáveis por atribuir uma classe gramatical a uma palavra e criar um espaço complementar categorial (CCS). Esse CCS torna o complemento do funtor equivalente a uma categoria lexical. Nesse sentido, um funtor do tipo $C_{N[V]}$, como os sufixos *-ção* e *-mento*, forma um Nome e transforma o seu complemento equivalente a um Verbo, de modo que, quando esse funtor domina diretamente uma raiz, essa raiz passa a ser equivalente a Verbo, possibilitando à sua realização fonológica ser semelhante à expressão de um Verbo. Relembremos também que as raízes são apenas índices fonológicos, tendo a sua realização fonológica definida em *Spell-Out*.

De acordo com essa proposta, os Nomes de entidade seriam formados com uma estrutura semelhante à representada na figura 7.

Figura 7: Estrutura de Nomes que denotam entidade



Fonte: AQUINO (2021, p. 143)

¹³ Para saber mais sobre a análise formal de palavras idiomáticas, confira Pederneira e Lemle (2009), Pederneira (2010), Lemle e Pederneira (2012), Aquino (2016), Aquino *et al.* (2018).

Nessa estrutura, o nominalizador seleciona a raiz $\sqrt{\text{VEGET}}$ como complemento, tornando-a equivalente a um Verbo. Quando a estrutura é enviada para *Spell-Out*, a raiz é realizada como /*vegeta*/, ganhando uma realização fonológica verbal possível na língua, e a estrutura toda é realizada como *vegetação*. Em relação ao recorte semântico dessa estrutura, a falta de uma projeção estendida verbal (cf. BORER, 2013b) possibilitaria o surgimento de uma nova leitura semântica, gerando as noções de lugar, coletividade etc. Diante disso, essa proposta se apresenta como um caminho possível para compreender a complexidade morfofonológica desses Nomes e o seu comportamento sintático diferenciado.

Considerações finais

Com base em uma análise sintática para os processos de formação de palavras da Gramática Gerativa, adotando principalmente os conceitos do modelo Exoesqueletal de Borer (2005a, 2005b, 2013a), este texto teve como objetivo principal estudar o fenômeno da polissemia em nominalizações deverbais, tendo como foco os dados do português do Brasil. Para isso, foi utilizada uma metodologia de análise linguística formal e foram aplicados testes sintáticos variados para compreendermos como a estrutura interna das nominalizações (i.e dentro do escopo do nó nominalizador) é formada e como ela exerce influência nas diferentes leituras. Os testes utilizados para análise foram: *presença/ausência de argumentos*, *modificação adverbial* e *modificação aspectual*.

Diante da análise realizada, propõe-se, com base nos templates de Borer (2013a), que as camadas sintáticas formadoras de uma estrutura de argumento/evento podem fazer parte da estrutura interna das nominalizações. Nesse sentido, a leitura de evento emerge quando todas essas camadas são projetadas abaixo do nominalizador, gerando uma estrutura de evento. A leitura de resultado, por sua vez, emerge quando a estrutura do Nome apresenta somente as camadas vP e ASPQ/Fshl, licenciando um argumento interno e a possibilidade de modificação aspectual, mas limitando a formação de uma estrutura de evento completa. Por fim, propõe-se que a leitura de entidade é resultado de uma estrutura em que não há a projeção dessas camadas, explicando assim o comportamento sintático diferenciado desses Nomes. Mas, para dar conta da realidade de complexidade morfofonológica nesses Nomes, que não podem ser analisados como idiomáticos, é assumido, retomando a proposta de categorização de palavras de Borer (2013a, 2013c), que o funtor categorial nominalizador tipifica a raiz por ele selecionada como equivalente a um verbo, possibilitando a realização fonológica dessa raiz de maneira semelhante a verbos da língua.

Diante disso, as análises aqui expostas corroboram a assunção de diferentes propostas na literatura de que a estrutura de evento não está presente somente em verbos, podendo ser herdada pelos seus nomes derivados. Por fim, é importante reiterar que este texto não encerra a discussão a respeito do processo de formação das nominalizações e que as propostas aqui defendidas suscitam novas perguntas. Uma delas, por exemplo, é como analisar a formação de nominalizações do infinitivo.

Referências

- ALEXIADOU, A. *Functional Structure in Nominals*. Nominalization and Ergativity. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2001.
- ALEXIADOU, A. On the role of syntactic locality in morphological processes: the case of (Greek) derived nominals. In: GIANNAKIDOU, A.; RATHER, M. (org.). *Quantification, Definiteness and Nominalization*. Oxford: Oxford University Press, 2009. pp. 253-80.
- ALEXIADOU, A.; BORER, H. Introduction. In: ALEXIADOU, A.; BORER, H. (eds.) *Nominalization: 50 years on from Chomsky's Remarks*. Oxford: Oxford University Press, pp. 1-24, 2020.
- AQUINO, R. N. M. *Encontros e desencontros semânticos entre palavras cognatas das línguas portuguesa e espanhola*. 2016. 110 f. Dissertação (Mestrado em Linguística). Programa de Pós-Graduação em Linguística. Faculdade de Letras - Universidade Federal do Rio de Janeiro. 2016
- AQUINO, R. N. M. Nomes deverbais em português: um estudo na interface sintaxe- semântica. In: *Caderno de Resumos SEPLA 2019 – Seminário de Pesquisas Linguísticas em Andamento*. Rio de Janeiro: Revista Linguística. Rio, v. 5. pp. 78-80, 2019.
- AQUINO, R. N. M. *Nomes deverbais em português do Brasil*. 2021. 145 f. Tese (Doutorado em Linguística). Programa de Pós-Graduação em Linguística. Faculdade de Letras - Universidade Federal do Rio de Janeiro. 2021
- AQUINO, R. N. M.; PEDERNEIRA, I. L.; LEMLE, M. A relação raiz e estrutura sintática na semântica da formação de palavras no português brasileiro e espanhol. *Revista da ANPOLL*. v. 1 n. 45, pp. 90-105, 2018.
- BASSANI, I; MINUSSI, R. Contra a seleção de argumentos pelas raízes: nominalizações e verbos complexos. *ReVEL*, v. 13, n. 24, 2015.
- BORER, H. Morphology and Syntax. In SPENCER, A; ZWICKY, A (eds.). *The Handbook of Morphology*. Oxford, Blackwell Publishers Ltd. pp. 149-190, 1998.
- BORER, H. Exo-Skeletal vs. Endo-Skeletal Explanations: Syntactic Projections and the Lexicon. In. POLINSKY, M.; MOORE, J. (eds.) *The nature of explanation in linguistic theory*. Standford: CSLI Publication, 2003.
- BORER, H. *Structuring sense*. In Name Only. 1. ed. Nova Iorque: Oxford Univeristy Press, 2005a.
- BORER, H. *Structuring sense*. The normal course of events. 1. ed. Nova Iorque: Oxford Univeristy Press, 2005b.
- BORER, H. *Roots and Categories*. Circle of Generative Grammar. University of the Basque Country, 2009. Handout. Disponível em: http://hborer.sllf.qmul.ac.uk/Downloads/Borer%202009%20roots_and_categories.pdf.
- BORER, H. *Structuring sense*. Taking Form. 1.ed. Nova Iorque: Oxford Univeristy Press, 2013a.
- BORER, H. The syntactic domain of content. In: BECKER, M.; GRINSTEAD, J.; ROTHMAN, J. (eds.). *Generative Linguistics and Acquisition*. Studies in honor of Nina M. Hyams. John Benjamins Publishing Company, 2013b.
- BORER, H. Derived nominals and the domain of content. *Lingua*, pp. 71-96, 2014a.

- BORER, H. The category of roots. In: ALEXIADOU, A.; BORER, H.; SCHÄFER, F. (eds.) *The Syntax of Roots and the Roots of Syntax*. Oxford, 2014b.
- BRITO, A. M.; OLIVEIRA, F. Nominalization, aspect and argument structure. In: MATOS, G. et al. (eds.), *Interfaces in Linguistic Theory*. Lisbon: APL/Colibri. 1997, pp. 57-80.
- CINQUE, G. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford, 1999.
- CYRINO, S. M. L. *O objeto nulo no Português do Brasil: um estudo sintático-diacrônico*. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística - Unicamp, Campinas, 1994.
- CYRINO, S. M. L. O objeto nulo no português brasileiro. In: GÄRTNER; HUNDTE; SCHÖNBERGER (orgs), *Estudos de gramática portuguesa*, v. III, pp. 61-73. Frankfurt am Main. 2000
- CHOMSKY, N. *Aspects of the theory of syntax*. Cambridge. MA: MIT Press, 1965.
- CHOMSKY, N. Remarks on nominalization. In: JACOBS, R. A.; ROSENBAUM, P.S. (eds.) *Readings in English Transformational Grammar*. Waltham, MA: Ginn, pp. 184-221, 1970.
- CHOMSKY, N. *The Minimalist Program*. Current Studies in Linguistics. MIT Press, Cambridge, Massachusetts. 1995
- CHOMSKY, N. Minimalist inquiries: The framework. In: MARTIN, R.; David MICHAELS, D.; URIAGEREKA, J. (eds.). *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, pp. 89-156. Cambridge, MA: MIT Press, 2000.
- CHOMSKY, N. Derivation by phase. In: KENSTOWICZ, M. (ed.). *Ken Hale: A Life in Language*, pp. 1-52. Cambridge, MA: MIT Press, 2001.
- DOWTY, D. Thematic Proto-Roles and Argument Selection. *Language*, v. 67, n. 3, pp. 547-619, 1991.
- DOWTY, D. *Word meaning and Montague grammar*. Dordrecht, Holland, Kluwer Academic Publishers. 1979
- FU, J; ROEPER, T.; BORER, H. The VP within Process Nominals: Evidence from Adverbs and the VP Anaphor Do-So. *Natural Language & Linguistic Theory*, v. 19, n. 3, pp. 549-82, 2001.
- JORGE, L. T. Complementação do Nome, relações semânticas e estruturas sintáticas: uma proposta de revisão da análise tradicional. Dissertação (Mestrado em Linguística). Universidade de Brasília, Brasília, 1986.
- GRIMSHAW, J.; MESTER, A. Light verbs and theta-marking. *Linguistic Inquiry*, v. 19, pp. 205-32, 1988.
- GRIMSHAW, J. *Argument Structure*. Linguistic Inquiry Monographs. 1990.
- OLIVEIRA, D. C. M. *Nominalizações no Português do Brasil: Aspectos Morfossintáticos e Semânticos*. Dissertação (Mestrado em Linguística). Universidade de Brasília, Brasília, 2005.
- OLIVEIRA, D. C. M. *Nominalizações no Português brasileiro: Estrutura argumental, formação e morfossintaxe*, 2014. Tese (Doutorado em Linguística). Universidade de Brasília, Brasília, 2014.
- REIS, R. *Os traços de aspecto lexical e as nominalizações em -ção e -mento*. Dissertação. f. 280. (Mestrado em Língua e Cultura) Instituto de Letras, Universidade Federal da Bahia, Salvador. Salvador: BA. 2016.

REIS, R; FIGUEIREDO, C. Os traços de aspecto lexical e as nominações em -ÇÃO no Português Brasileiro. *Revista da ANPOLL*, v. 1, n. 46, 2018.

ROEPER, T.; VAN HOUT, A. The impact of nominalization on passive, -able and middle: Burzio's generalization and feature-movement in the lexicon'. *MITWPL*, v. 35, pp. 185-211, 1999.

SLEEMAN, P.; BRITO, A. M. Nominalization, event, aspect, and argument structure: a syntactic approach. *In: DUGUINE, M.; HUIDOBRO, S.; MADARIAGA, N. (eds). Argument Structure and Syntactic Relations. A cross-linguistic perspective. Linguistik Aktuell/Linguistics Today, John Benjamins B.V. 2010.*

SLEEMAN, P.; BRITO, A. M. Aspect and argument structure of deverbal nominalizations: a split vP analysis. *In: ALEXIADOU, A.; RATHERT, M. (eds.). The Syntax of Nominalizations across Languages and Frameworks*, pp. 199-217. Berlin: De Gruyter, 2010.

O COMPORTAMENTO DOS PRONOMES POSSESSIVOS SEU(S); SUA(S); DELE(S) E DELA(S) NA RECUPERAÇÃO DE SEUS ANTECEDENTES

THE BEHAVIOR OF POSSESSIVE PRONOUNS SEU(S); SUA(S); DELE(S) E DELA(S) IN THE RECOVERY OF THEIR ANTECEDENTS

Bruna Clara Santos de Almeida¹

Rafael Dias Minussi²

RESUMO

Este artigo faz uma descrição e análise dos pronomes possessivos seu(s), sua(s), dele(s) e dela(s), tomando como base o Princípio B da teoria de Regência e Ligação (*Government and Binding*) (CHOMSKY, 1981), segundo o qual o pronome tem que estar livre em sua categoria de regência, isso significa que os pronomes não precisam de antecedentes, mas quando os possuem, esses não podem c-comandá-los dentro de sua categoria de regência. Entretanto, em alguns casos os pronomes possessivos seu(s); sua(s); dele(s) e dela(s) parecem ser ligados, como nas sentenças: [Pedro_i considera seu_i óculos] o mais bonito; Felipe acha [Gabriela_i orgulhosa dela_i]. Por meio de dados retirados do corpus do projeto SP2010, transcrito pelo Grupo de Estudos e Pesquisa em Sociolinguística da USP, coordenado pelo Prof. Dr. Ronald Beline, identificamos que os pronomes possessivos seu(s), sua(s), dele(s) e dela(s) não atuam como prevê o Princípio B da Teoria da Ligação. A partir da perspectiva da Morfologia Distribuída, sugerimos que tais pronomes são subespecificados, porém os pronomes seu(s) e sua(s) parecem ser mais subespecificados do que dele(s) e dela(s), já que possuem vários tipos de antecedentes: 3ª pessoa, 2ª pessoa, 2G e 3G ocasionando o fenômeno de sincretismo, pois esses antecedentes possuem mais de um traço sintático-semântico, sendo retomado pelo mesmo conteúdo fonológico. Os pronomes seu(s), sua(s), dele(s) e dela(s) parecem possuir os traços [+anafórico, +pronominal], pois se comportam tanto como pronomes quanto como anáforas, sendo o seu(s) e sua(s) aparentemente mais anafóricos e dele(s) e dela(s) mais pronominais.

PALAVRAS-CHAVE: Pronomes possessivos. Princípio B. Morfologia Distribuída. Traços.

ABSTRACT

This article makes a description and analysis of possessive pronouns *seu(s)*, *sua(s)*, *dele(s)*, and *dela(s)*, based on Principle B of the theory of Government and Binding (CHOMSKY, 1981), according to which the pronoun must be free in its regency category. This means that pronouns do not need antecedents, but when they have them, antecedents can't c-command them within their regency category. However, in some cases the possessive pronouns *seu(s)*; *sua(s)*; *dele(s)* and *dela(s)* seem to be linked, as in the sentences: (1) [*Pedro_i considera seu_i óculos*] *o mais bonito*; (2) *Felipe acha [Gabriela_i orgulhosa dela_i]*. Through data taken from the SP2010 project corpus, transcribed by the Study and Research Group in Sociolinguistics at USP, coordinated by Prof. Dr. Ronald Beline, we identify that the possessive pronouns *seu(s)*, *sua(s)*, *dele(s)* and *dela(s)* do not act as predicted by Principle B of the Binding Theory. From the perspective of Distributed Morphology, we suggest

¹ Mestranda do Programa de Pós-Graduação em Letras (PPGL) da Universidade Federal de São Paulo (UNIFESP), bruna.clara02@unifesp.br, <https://orcid.org/0000-0001-8171-7860>.

² Professor Doutor do Departamento de Letras da Universidade Federal de São Paulo (UNIFESP), rafael.minussi@unifesp.br, <https://orcid.org/0000-0003-4103-8796>.

that such pronouns are underspecified, but the pronouns *seu(s)* and *sua(s)* seem to be more underspecified than *dele(s)* and *dela(s)*, since they have several types of background: 3rd person, 2nd person, 2G and 3G causing the phenomenon of syncretism, because these antecedents have more than one syntactic-semantic feature, being retaken by the same phonological content. The pronouns *seu(s)*, *sua(s)*, *dele(s)*, and *dela(s)* seem to have features [+anaphoric, +pronominal] since they behave both as pronouns and as anaphors, with *seu(s)* and *sua(s)* more anaphoric and *dele(s)* and *dela(s)* more pronomial.

KEYWORDS: Possessive pronouns. Principle B. Distributed Morphology. Features.

Introdução

O objetivo principal deste trabalho é realizar uma descrição e análise do comportamento dos pronomes possessivos *seu(s)*, *sua(s)*, *dele* e *dela(s)*, a partir de um corpus³ formado por 709 enunciados⁴.

O ponto de partida para a descrição e análise foi a Teoria da Ligação (CHOMSKY, 1981). Assim sendo, a fim de buscar um entendimento dessa teoria em relação aos dados do português brasileiro (PB), vamos nos apoiar na descrição encontrada em Miotto; Silva e Lopes (2016). Segundo os autores, o Princípio A determina que uma anáfora tem de estar ligada em sua categoria de regência (CR), que é o domínio onde se encontra a anáfora e seu regente: “A categoria de regência de α é o XP mínimo que contém α , o regente de α e (i) um sujeito que é distinto de α e que não contém α ; ou (ii) a flexão que atribui Caso Nominativo para α ” (MIOTTO *et al.*, 2016, p. 215). A anáfora exige a presença de um antecedente, em que se identifica referencialmente, como na sentença:

(1) [A Diana_i, se_i ama]

Em (1), a anáfora ‘se’ retoma o sujeito Diana. Tal retomada ocorre dentro de sua categoria de regência e, portanto, está de acordo com o Princípio A da teoria.

Por sua vez, o Princípio B ressalta que o pronome tem que estar livre em sua categoria de regência, de modo que os pronomes não precisam de antecedentes, mas quando os possuem, esses não podem c-comandar os pronomes dentro de sua categoria de regência, mas fora deste domínio o antecedente poderá c-comandá-los⁵. No exemplo (2), apresentado a seguir, percebemos que, diferente das anáforas, o pronome ‘ele’ retoma o sujeito Pedro. Contudo, ‘ele’ e ‘Pedro’ não estão na mesma categoria de regência.

(2) O Pedro_i disse que [a Patrícia_j ama ele_i]

O Princípio C descreve que uma expressão-R tem que estar livre, como na sentença (3):

³ O corpus foi formado a partir de entrevistas do Projeto SP2010. Todo o Corpus pode ser acessado gratuitamente através de um cadastro no site <http://projetosp2010.fflch.usp.br/>.

⁴ Consistem em entrevistas retiradas do corpus, realizadas com homens e mulheres paulistanas, com a faixa etária de 19 a 34 anos, selecionados independentes da renda, com Ensino Médio completo e ou com Ensino Superior.

⁵ Miotto *et al.* (2016, p. 56) explicam a relação de c-comando: “ α c-comanda β se e somente se β é o irmão de α ou se β é dominado pelo irmão de α ”.

(3) Luiz gosta de chocolate.

No exemplo em (3), vemos que o sujeito ‘Luiz’ está livre, pois não é retomado nem dentro e nem fora do seu domínio de regência. No PB, os pronomes possessivos seu(s); sua(s); dele(s) e dela(s), em alguns contextos, parecem ser ligados dentro de sua categoria de regência, confrontando o Princípio B da Teoria da Ligação:

(4) ? [Pedro_i considera o seu_i hamburguer o melhor da cidade]

(5) ? O João acha [a Maria_i orgulhosa dela_i]

Em (4), descrevemos uma das interpretações possíveis para a sentença “Pedro considera o seu hamburguer o melhor da cidade”. Nessa interpretação, o pronome ‘seu’ está retomando o sujeito ‘Pedro’ dentro do seu domínio de regência. Da mesma forma, em (5), a interpretação ressaltada é a de que o pronome ‘dela’ está retomando ‘Maria’ no mesmo domínio.

Diante desses e de outros dados coletados, o interesse deste trabalho é descrever e verificar o comportamento desses pronomes. Em um primeiro momento, apresentaremos as principais propriedades da Teoria da ligação (CHOMSKY, 1981), para, assim, abordar alguns aspectos sobre os pronomes possessivos seu(s), sua(s), dele(s) e dela(s) trazidos pela literatura linguística sobre o assunto. Em seguida, descreveremos e analisaremos os dados retirados do corpus do projeto SP2010, transcrito pelo Grupo de Estudos e Pesquisa em Sociolinguística da USP, à luz do modelo da Morfologia Distribuída (doravante MD) (HALLE; MARANTZ, 1993). A partir desse estudo, pretendemos responder as seguintes questões: 1. Os pronomes seu(s); sua(s); dele(s) e dela(s) se comportam como prevê o Princípio B da Teoria da Ligação? 2. Há desenvolvimento dentro da MD capaz de explicar esses dados?

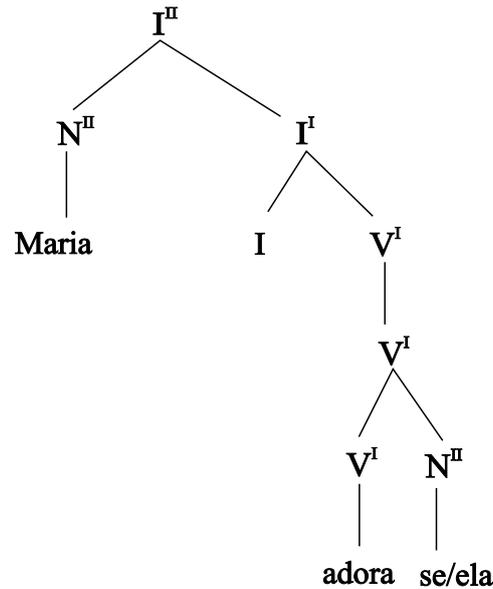
1. A teoria da ligação e seus desdobramentos

Chomsky (1981), como já dissemos, apresenta três princípios que delimitam o modo de ordenação de diferentes DPs⁶. O primeiro, o princípio A, refere-se ao funcionamento das anáforas; o princípio B refere-se aos pronomes e o princípio C trata da função das expressões referenciais.

Mioto *et al.* (2016) explicam tais princípios com base nos dados do PB. Esses autores mostram que as anáforas impõem algumas condições com relação à estrutura sintática em que aparecem, a começar pela exigência da presença do elemento que referencialmente a identificam, os chamados antecedentes. A partir da noção de c-comando, os autores demonstram que: “ α c-comanda β se e somente se β é o irmão de α ou se β é denominado pelo irmão de α ” MIOTO *et al.* (2016, p. 210). Então, a segunda exigência feita pela anáfora consiste na ideia de que apenas um DP que c-comande a anáfora pode ser o seu antecedente. Essa combinação de elementos, segundo os autores, chama-se “ligação”. A noção de c-comando é demonstrada através da estrutura sintática formulada por Müller (1997, p. 36):

⁶ Do inglês Determiner Phrase (Sintagma Determinante).

(6)



Müller (1997, p. 36)

Na estrutura (6) está representada a sentença “Maria_i adora se_i/ela_k”, na qual o sujeito ‘Maria’ c-comanda o objeto ‘se’ ou ‘ela’. Assim sendo, tanto o ‘se’ quanto o ‘ela’ pertencem à mesma categoria de regência. Existe, portanto, um sujeito ‘Maria’ disponível ao objeto e o verbo adorar atua como regente. Em conformidade ao princípio A, a anáfora ‘se’ deve estar ligada, ou seja, coindexada a um antecedente que a c-comande neste mesmo domínio. Já o pronome ‘ela’ deve ser livre, pois não pode ser coindexado a um antecedente que o c-comande no mesmo domínio.

Ademais, Mioto *et al.* (2016) explicam que é necessário assegurar que a anáfora procure seu antecedente na própria sentença encaixada e, quando o preceito (i) não é satisfeito, a anáfora deve apurar se o preceito (ii) é satisfeito, antes de ir procurar o seu antecedente na sentença matriz. Dessa forma, os autores resumem com a seguinte constatação: “as anáforas devem ser ligadas em sua categoria de regência” (MIOTO *et al.*, 2016, p. 216).

Ao tratar dos pronomes, Mioto *et al.* (2016) explicam que, diferente das anáforas, esses não podem estar ligados na sua categoria de regência, ou seja, devem ser livres. Então, os pronomes não precisam de antecedentes, mas quando os antecedentes existem, estes não podem estar muito próximos aos pronomes, caso contrário, a sentença pode ser agramatical. Os estudiosos observam que os contextos nos quais as anáforas se dão fornecem resultados de gramaticalidade inversos das sentenças com pronomes. Com base nesses argumentos, resume-se o comportamento dos pronomes a partir da seguinte confirmação: “Os pronomes não devem ser ligados em sua categoria de regência” (MIOTO *et al.*, 2016, p. 218).

Por sua vez, as expressões referenciais, também conhecidas como expressões-R se referem aos DPs que possuem autonomia referencial. Assim como os pronomes, as expressões-R não precisam de um antecedente, de modo que são mal reformuladas em situações em que as anáforas são gramaticais, bem como, não são possíveis em contextos sintáticos adequados aos pronomes. Desse modo, diferente

dos pronomes que não podem ocorrer dentro de sua categoria de regência, as expressões-R não poderão nunca ser ligadas. Resumindo: “As expressões-R não devem ser ligadas” (MIOTO *et al.*, 2016, p.220).

2. Relações anafóricas

Passando a falar das relações anafóricas, destacamos o trabalho de Müller (2001), que explica as anáforas pronominais. Nesse estudo, a autora discute a seguinte questão: o que são pronomes e como se dá a interpretação desses? O trabalho é fundamentado em duas teses. A primeira tese é a de que um pronome pode determinar dois tipos de relações com seu antecedente: (i) a relação de correferência ou (ii) a relação de ligação. A segunda tese é a de que um pronome é sempre uma variável, em que a denotação não é fixa no léxico, ou seja, varia de acordo com uma atribuição de valores que é feita a ele, quer por sua ligação a um operador, quer por uma função contextual. Em seu trabalho, a autora defende a tese de que os pronomes devem ser sempre tratados como variáveis, de modo que a denotação não é fixa lexicalmente:

- (7) *Você* está cansada;
- (8) João brigou com Arnold Schwazenegger. *Ele se* machucou bastante;
- (9) *Ele* é o cara que *eu* conheci na festa;
- (10) Qualquer aluno gosta quando *seu* trabalho é julgado melhor que o trabalho dos outros.⁷

A partir dos exemplos (7-10), Müller (2001) explica que os itens em itálico são tradicionalmente chamados de pronomes, todos eles possuem a mesma distribuição sintática dos sintagmas nominais e não possuem uma denotação que possa ser determinada por seu significado lexical, de modo que é o significado lexical que diz que a sua referência deverá ser solucionada todas as vezes que um pronome surgir em um discurso.

A autora assume, assim como Heim e Kratzer (1998), que um pronome é uma variável, ou seja, é uma expressão em que a denotação não é fixa somente pelo seu significado lexical e varia de acordo com uma atribuição de valores. Assim sendo, Müller esclarece que nas sentenças apresentadas anteriormente, os pronomes ‘eu’ e ‘ele’, sendo variáveis num sentido semântico, irão ter sua denotação estabelecida por algum valor que pode ser atribuído em função do contexto linguístico e/ou extralinguístico. Já em sentenças como (8) e (10), a denotação do pronome ‘se’ e do pronome possessivo ‘seu’ covaria com valor de seu antecedente, podendo tomar com antecedente o nome próprio ‘Arnold Schwazenegger’ ou o sintagma quantificado ‘qualquer aluno’.

Ao falar de correferência *versus* ligação de variáveis, Müller (2001) explica que a diferença entre o uso anafórico do uso dêitico dos pronomes é que no uso dêitico, a referência de um pronome é definida pelo contexto extralinguístico, já em seu uso anafórico, a referência é definida pelo discurso, seja ele anterior ou posterior, como explicado pela autora nas seguintes sentenças:

⁷ Sentenças retiradas de Müller (2001, pp. 260-61).

(11) “João_i brigou com Arnold Schwazenegger. Ele_i se machucou bastante.”

(12) “Ninguém parecia gostar dele_i. No entanto, João_i insistia em comparecer a todas as festas.”

No exemplo (11) o ‘ele’ se refere ao nome ‘João’ mencionado anteriormente, enquanto que em (12) o pronome ‘dele’ se refere ao ‘João’ que é mencionado posteriormente.

Porém, a autora esclarece que em Teoria Gramatical assume-se que os usos anafóricos e dêiticos podem ser vistos como um fenômeno em que a referência do pronome é definida pelo contexto, linguístico ou extralinguístico. Nas duas ocorrências, o pronome se refere a um indivíduo que está eminentemente evidente no instante em que sua referência é decidida, ou seja, tanto o uso dêitico quanto o uso anafórico de um pronome são casos especiais da mesma ocorrência. Assim, os pronomes, quando variáveis, recebem tanto uma interpretação de variável presa quanto uma interpretação de (co)referência, a qual dependerá de como seu valor é definido em cada uso.

Em Müller (2000), a autora explica que o modo como os Princípios de Ligação são colocados provoca problemas para a interpretação semântica. O primeiro problema apontado pela autora é: como interpretar a coindexação? Segundo a autora, uma coindexação pode representar tanto a correferência como a ligação de variáveis.

Com o intuito de excluir da Teoria da Ligação os problemas da interpretação semântica, Reinhart (1983) apud Müller (2000, p. 15) reinterpreta as condições dos Princípio A, B e C, como um mecanismo sintático que dá às formas pronominais o status de anáfora ligada:

O conceito de anáfora ligada é ao mesmo tempo sintático e semântico. Sintaticamente temos uma anáfora ligada quando um sintagma nominal está ligado a seu antecedente. Semanticamente, temos uma anáfora ligada quando uma forma pronominal é interpretada como uma variável presa por um operador (MÜLLER 2000, pp. 15-6).

Heim & Kratzer (1998, apud. Müller, 2000) reformulam a proposta de Reinhart em termos de um princípio que admite um vínculo entre a Semântica e a Sintaxe, no qual eles chamam de “Princípio de Ligação”, tal princípio determina que uma ligação semântica respeite os princípios sintáticos, bem como uma ligação sintática deve refletir uma ligação semântica. O “Princípio de Ligação” apresentado pelos autores é demonstrado do seguinte modo: “Princípio de Ligação: Sejam α e β sintagmas de determinantes, onde β não é foneticamente vazio. Então α liga β sintaticamente na estrutura superficial se e somente se α liga β semanticamente na forma lógica” (MÜLLER 2000, p. 16).

Menuzzi (1996) também colabora com os estudos sobre os possessivos seu(s), sua(s), dele(s) e dela(s). Alguns pontos levantados pelo autor são considerados em nosso trabalho. Para o autor, bem como para Müller (1997), ‘seu’ e ‘dele’ não são análogos, ou seja, não existe uma substituição de um pronome por outro, o que ocorre é uma especialização que é imposta pela característica do antecedente. O pronome ‘seu’ se comporta como variável ligada em retomadas de antecedentes que não são referenciais, os chamados genéricos e quantificadores, enquanto ‘dele’ é o possessivo que

retoma antecedentes referenciais e por isso manifestam uma correferência. Abaixo apresentamos alguns exemplos retirados de Menuzzi (1996, p. 4). Esses exemplos exemplificam a relação de distribuição, considerando a natureza do antecedente da forma ‘seu’ vs. ‘dele’:

- (13) *Quem* esqueceu {*seu* livro/(?)*o livro *dele*} em casa?
- (14) *Cada um* deve fazer {*seu* trabalho/(?)*o trabalho *dele*} sozinho.
- (15) *Qualquer linguista* quer ter {*seus* artigos/os artigos *dele*} lidos por Chomsky.
- (16) *O João* já viu um retrato {(?)?seu/dele} no jornal?

A partir das sentenças apresentadas (13), (14), (15) e (16) o autor observou que em contextos de quantificadores gerais como: ‘quem’ e ‘cada’, a forma possessiva ‘seu’ é mais produtiva em oposição a forma ‘dele’ (13 e 14). Já quando o antecedente é indefinido (15), as formas ‘seu’ e ‘dele’ podem intercalar. No entanto, quando os antecedentes são referenciais/definidos, o possessivo ‘seu’ parece o menos adequado, como apresentado na sentença (16).

Segundo Menuzzi (1996), o sistema pronominal do PB está se modificando, de modo que a forma ‘seu’ está se tornando improdutiva como forma anafórica para antecedentes de 3ª pessoa referenciais. Tal aspecto faz com que o PB se diferencie de outras línguas românicas. Com base nessa observação, o autor levanta as seguintes questões: (i) Por que a forma ‘seu’ se tornou desfavorecida na retomada do antecedente referencial? (ii) Por que essa restrição é suspensa em situações discursivas apropriadas?

Para chegar a algumas conclusões, o autor, através de um corpus retirados da narrativa “Agosto” de Rubem Fonseca, investiga um padrão em contextos que aparecem a forma ‘seu’. Menuzzi (1996) separa os trechos que apresentam discurso direto nas falas das personagens, dos trechos narrativos e discurso indireto. Os resultados mostram que nas passagens em que ocorrem discurso indireto o uso da forma possessiva ‘seu’ aparece em 94,5% das situações, já nas passagens que foram reproduzidos os discursos diretos a forma ‘dele’ foi a favorita, atingindo 84,4% dos casos, ao mesmo tempo que a forma ‘seu’ refere-se, na maioria dos casos, à 2ª pessoa em 90,8%. A partir dos dados observados em “Agosto”, Menuzzi (1996) explica que o possessivo ‘seu’, quando possui um antecedente de 3ª pessoa, é licenciado através de circunstâncias especiais do discurso, delimitadas como Princípios de Acessibilidade⁸, esses fatores são organizados entre o anafórico e o antecedente, através dos seguintes pontos:

- (i) Distância entre eles, ou seja, o anafórico está na mesma sentença que o antecedente;
- (ii) Concorrência entre os prováveis antecedentes;
- (iii) Se o antecedente é tópico ou não tópico;
- (iv) O anafórico e o antecedente estão no interior da mesma unidade do discurso.

⁸ Para um melhor entendimento sobre a Teoria da Acessibilidade ver: Ariel, M. (2001). Accessibility theory: *an overview*. In: Sanders, T., Schilperoord, J. & Spooren, W. (eds.). Text representation: linguistic and psycholinguistic aspects. Amsterdam: J. Benjamins Publishing Co.

Desse modo, o autor explica a atuação do discurso da Escala de Marcação de Acessibilidade, que tem o seguinte efeito: quanto mais acessível o antecedente, menos marcado o item anafórico deve ser, isto é, quanto menos traços gramaticais o pronome expressar do seu antecedente, este será mais acessível em termos funcionais.

Menuzzi (1996) também esclarece que no PB falado, o número de vezes que ‘seu’ aparece como antecedentes referenciais de 3ª pessoa é muito inferior, ele propõe que tais contextos não promovem condições adequadas para esse possessivo no PB falado.

Em resumo, Müller (1997) e Menuzzi (1996) parecem encontrar uma solução para resolver as oposições apresentadas pelos Princípios da Teoria da Ligação. Neste estudo, pretendemos também trazer luz a essas oposições, apoiados nas propostas de Müller (1997) e Menuzzi (1996); sobretudo, ancorados no modelo teórico de Morfologia Distribuída, para compreender melhor o comportamento das formas possessivas canônicas – seu(s)/sua(s) e não canônicas dele(s)/dela(s) em contextos de ligação. Na próxima seção iremos apresentar o modelo teórico de Morfologia Distribuída que será usado no desenvolvimento da nossa análise.

3. O modelo teórico de Morfologia Distribuída

De acordo com Scher, Bassani e Minussi (2013), a MD assume a separação entre traços morfo-sintáticossemânticos e a realização fonológica desses traços, isto é, os traços são abstratos, sem conteúdo fonológico, o qual será inserido depois da derivação sintática. Entretanto, diferente da Morfologia Baseada no Léxico – MBL, a MD rejeita o aspecto não afixal e transforma a noção de morfema, para que seja possível reparar algumas violações da relação entre a forma e significados dos morfemas flexionais.

Scher, Bassani e Minussi (2013, p. 18) explicam que:

o modelo da MD foi formalmente introduzido em Halle e Marantz (1993) e se configura como um modelo de gramática que compartilha com a Teoria de Princípios e Parâmetros de uma série de ideias sobre a organização da gramática, mas apresenta um conjunto de propriedades particulares que os diferenciam. O objetivo desse distanciamento é a busca de um tratamento teórico que seja mais satisfatório para a formação de palavras.

Segundo os estudiosos, a MD propõe que qualquer processo de formação de palavras ou de constituintes maiores (sintagmas e sentenças) ocorre no Sistema Computacional, em outras palavras, no componente sintático. Assim sendo, as mesmas operações que formam as sentenças (concatenar/mover) estão na base dos processos de formação de palavras. Consequentemente, essas operações vão manipular elementos básicos, traços e raízes, e é através deles que palavras e sentenças vão ser geradas.

Os autores ressaltam ainda que, para a MD, não existem elementos linguísticos exclusivos que são dedicados a formações, sejam derivacionais ou flexionais ao passo que, não há operações singulares para os tipos de formações, isto é, toda formação é sintática, sendo consequência de operações combinatórias que ocorrem no Sistema Computacional. Essa propriedade do modelo ficou conhecida como Estrutura Sintática por toda a derivação⁹.

⁹ Do inglês *Syntactic Structure all the way down* (HARLEY; NOYER, 1999).

Scher, Bassani e Minussi (2013) apontam também que, com a supressão do modelo gerativo de Léxico, é necessário admitir outras formas de armazenamento de informações. Para eles, a distribuição da morfologia retoma a ideia antiga dos estudos linguísticos, em que a morfologia não deve ser tratada isoladamente; sobretudo, deve-se considerar a integração que mantém com a fonologia, sintaxe e semântica.

Retomando a discussão dos elementos primitivos que vão alimentar a derivação sintática, na MD, a sintaxe opera com um conjunto de traços abstratos. Os pesquisadores destacam que esses traços podem ser gramaticais e/ou semânticos:

[singular], [plural], [dual], [paucal], [1], [2], [3], [passado], [presente], [futuro], [perfectivo], [imperfectivo], [nominativo], [acusativo], [dativo], [causa], [reflexivo], [determinante], [ativa], [passiva], entre muitos outros a depender da proposta de análise, e núcleos categorizadores n (nome), v (verbo), a (adjetivo), além de funcionais, tais como T (Tempo), C (Complementizador), Asp (Aspecto), Voice. (SCHER, BASSANI e MINUSSI, 2013, p. 23).

Tais traços são armazenados em um local denominado de Lista 1 e são manipulados pelo Sistema Computacional, que forma conjuntos de traços abstratos, cujo conteúdo fonológico final, como já dissemos, será definido somente após a atuação de operações morfológicas, após o término da derivação sintática e envio da estrutura sintática formada para PF (do inglês *Phonological Form* ‘Forma Fonológica’). Essa propriedade ficou conhecida como Inserção Tardia (*Late Insertion*) de conteúdo fonológico (cf. HALLE; MARANTZ, 1993).

Os autores explicam que o processo de Inserção Tardia é ordenado pelos Itens de Vocabulários (IVs), que são regras que caracterizam uma relação que ocorre entre uma expressão fonológica e um traço gramatical ou semântico e, ocasionalmente, um contexto de inserção. Tais itens estão armazenados na Lista 2.

Os autores também explicam a propriedade de Subespecificação¹⁰ dos IVs. Essa propriedade é comandada pelo Princípio do Subconjunto, segundo o qual os IVs não precisam ser integralmente especificados para a aplicação de uma determinada posição sintática, precisam apenas possuir traços que não existem no nó terminal. Essa Subespecificação age sobre os IVs e não sobre esses nós, que continuam plenamente especificados. Tais IVs competem para a inserção em um nó-terminal, revelam os autores, ao contrário do que acontece nas teorias lexicalistas.

4. Metodologia

Através de um corpus do projeto SP2010, coletado e transcrito pelo Grupo de Estudos e Pesquisa em Sociolinguística da USP, o qual foi coordenado pelo Prof. Dr. Ronald Beline, selecionamos alguns enunciados de entrevistas, realizados com um grupo de vinte pessoas, dez homens e dez mulheres, com a faixa etária de 19 a 34 anos, com o Ensino Médio e ou Ensino Superior completo. O perfil socioeconômico não foi considerado para a seleção do corpus utilizado.

¹⁰ Do Inglês Underspecification.

Para selecionar os enunciados que continham os pronomes possessivos seu(s); sua(s); dele(s) e dela(s) em contextos anafóricos, utilizamos o programa AntConc. Esse programa é um *freeware* (programa de computador gratuito à disposição dos usuários) desenvolvido pelo Lawrence Anthony e está disponível para download¹¹.

A fim de organizar os enunciados para uma melhor análise, utilizamos o aplicativo – Excel. Os enunciados foram organizados da seguinte maneira: coluna A – contendo os pronomes; B – Gênero do Entrevistado que não foi levado em consideração; coluna C – Pessoa, ou seja, a pessoa a qual o pronome está retomando: 2ª; 3ª ou se é 2ª, ou 3ª pessoa generalizante. Neste estudo a sigla 3G foi criada para especificar os casos em que a pessoa do antecedente é uma 3ª pessoa, mas é generaliza, como por exemplo: todo mundo, o povo, as mulheres, cada um, etc. Já a sigla 2G será utilizada para designar também uma pessoa generalizada, mas em casos em que a 2ª pessoa ‘você’ se refere a qualquer pessoa, se comportando como um pronome de 2ª pessoa indeterminada. Utilizamos a coluna D para especificar se o enunciado pertence ao entrevistado ou ao entrevistador, em que o entrevistado é representado por F (falante) e o entrevistador por E; na coluna E, inserimos o item anafórico, para descrever se a retomada do pronome ocorre dentro ou fora da Categoria de Regência (CR) – RD (Retomada dentro de CR), RF (Retomada fora de CR).

Neste estudo, consideramos apenas os enunciados do entrevistado, em que será visualizado várias situações nas quais os pronomes vão agir - retomando um sujeito ou não e não apenas em contextos de perguntas, como ocorre nos enunciados do entrevistador, que são direcionadas a um interlocutor que está ali presente, evitando um enviesamento dos dados.

5. Análise e descrição do fenômeno em contextos anafóricos

O corpus completo possui 709 enunciados. Nele há situações em que os pronomes aparecem antes do sujeito, por exemplo: “e com **seus** amigos você, **vocês** saiam pra outros bairros”. Nesses contextos, não consideramos que ocorreu uma retomada. Além disso, escolhemos demarcar o tamanho do enunciado com dez palavras a esquerda e dez a direita do pronome, então, desconsideramos também os contextos em que o sujeito que o pronome está retomando não aparece. Delimitamos a análise, focando apenas nos enunciados do entrevistado e os contextos RD e RF, que são os contextos que nos interessam. Assim, a quantidade de enunciados foi reduzida de 709 para 242 enunciados.

Entre os pronomes possessivos que serão analisados, o mais recorrente nos enunciados é o “dele”. O quadro 1, apresentado a seguir, demonstra quantas vezes cada pronome aparece. Também observamos a quantidade de vezes que cada pronome aparece nos contextos RD e RF.

¹¹ <http://www.laurenceanthony.net/software.html>.

Quadro 1: Pronomes seu(s)/sua(s)/dele(s) e dela(s) em cada contexto

seu(s)/sua(s)							
Ocorrências				Percentual			
Pronome	RD	RF	Total Geral	Pronome	RD	RF	Total Geral
seu	22	15	37	seu	59,4	40,6	100
seus	6	3	9	seus	66,6	33,4	100
sua	27	13	40	sua	67,5	32,5	100
suas	4	2	6	suas	66,6	33,4	100
Total	59	33	92	Percentual	64,1	35,9	100
dele(s)/dela(s)							
Ocorrências				Percentual			
Pronome	RD	RF	Total Geral	Pronome	RD	RF	Total Geral
dela	19	31	50	dela	38	62	100
delas	2	5	7	delas	28,6	71,4	100
dele	20	36	56	dele	35,7	64,3	100
deles	15	22	37	deles	40,5	59,5	100
Total	56	94	150	Percentual	37,3	62,7	100
Total Geral	115	127	242	Percentual Total	47,6	52,4	100

Com base no quadro 1, observamos que os pronomes seu(s) e sua(s) aparecem em 92 sentenças. Em 64,1% dos casos, esses pronomes ocorreram no contexto de RD, enquanto os pronomes dele(s) e dela(s), das 150 vezes que aparecem, 62,7% foi no contexto de RF. Ou seja, os pronomes seu(s) e sua(s) ocorrem mais em contexto de RD e os pronomes dele(s) e dela(s) ocorrem mais em contexto de RF.

Verificamos nos quadros 2 e 3 apresentados a seguir, os tipos de antecedentes. O quadro 2 possui a quantidade de ocorrências e o quadro 3 possui o percentual das ocorrências. A partir dessa descrição, percebemos que os pronomes seu(s) e sua(s) possuem diversos antecedentes: 2ª, 3ª pessoa, 3G e 2G:

Quadro 2: ocorrências do tipo de pessoa que os pronomes seu(s)/sua(s) retomam

Ocorrências					
Pronome/Pessoa	2	3	2G	3G	Total Geral
seu	2	3	14	18	37
seus	0	1	3	5	9
sua	3	3	17	17	40
suas	0	0	2	4	6
Total Geral	5	7	36	44	92

Quadro 3: Percentual do tipo de pessoa que os pronomes seu(s)/sua(s) retomam

Percentual					
Pronome/Pessoa	2	2G	3	3G	Total Geral
seu	5,4	37,8	8,1	48,7	100
seus	0	33,3	11,1	55,6	100
sua	7,5	42,5	7,5	42,5	100
suas	0	33,3	0	66,7	100
Percentual Total	5,4	39,2	7,6	47,8	100

O antecedente na pessoa 3G parece ser escolha preferida dos falantes, já que, de acordo com o quadro 2 e 3, das 92 sentenças em que os pronomes seu(s) e sua(s) aparecem, 44 casos são utilizados para retomar uma 3G, isso equivale a 47,8% dos casos. Em seguida está a pessoa 2G, utilizada em 36 situações, isto é, 39,2% dos casos.

Nos quadros 4 e 5 a seguir, descrevemos as formas dele(e) e dela(s), que possuem apenas antecedentes na 3ª pessoa e 3G, porém a 3ª pessoa é muito mais recorrente, cerca de 95,3% das vezes:

Quadro 4: ocorrências do tipo de pessoa que os pronomes dele(s) e dela(s) retomam

Ocorrências					
Pronome/Pessoa	2	2G	3	3G	Total Geral
dela	0	0	48	2	50
delas	0	0	5	2	7
dele	0	0	55	1	56
deles	0	0	35	2	37
Total Geral	0	0	143	7	150

Quadro 5: Percentual do tipo de pessoa que os pronomes dele(s) e dela(s) retomam

Percentual					
Pronome/Pessoa	2	2G	3	3G	Total Geral
dela	0	0	96	4	100
delas	0	0	71,4	28,6	100
dele	0	0	98,2	1,8	100
deles	0	0	94,6	5,4	100
Percentual	0	0	95,3	4,7	100

De acordo com o quadro 4 e 5 as formas dele(s) e dela(s) possuem apenas dois tipos de antecedentes (3G e 3ª pessoa); porém, das 150 sentenças nas quais tais elementos aparecem, os antecedentes são 3G em apenas em 4,7% dos casos. Isso significa que em contextos de ligação os pronomes dele(s) e dela(s) são muito mais produtivos quando possuem antecedentes na 3ª pessoa.

5.1. Pronomes possessivos canônicos: seu(s) e sua(s)

Nesta seção, descreveremos algumas características dos pronomes seu(s) e sua(s) em contextos anafóricos, a partir de algumas sentenças retiradas do corpus.

Vejam algumas sentenças em que aparecem os pronomes seu(s) e sua(s):

- (17) [eu acho que as mulheres_i também têm que ter os seu_i direito iguais] assim com o homem
- (18) [cada um_i tem o seu_i momento certo] e eu acho que está faltando isso um pouco
- (19) [está todo mundo_i olhando pro seu_i umbigo]
- (20) [o povo_i sai do do dos seus_i estados né?] e vem tentar a vida em São Paulo né?
- (21) [você_i não ganha um trabalho pela sua_i capacidade]
- (22) [a gente_i respeita o que cada um quer da sua_i vida]
- (23) [você_i via então confissões de adolescente na sua_i época né?]
- (24) [você_i não pode sair fazendo qualquer coisa na rua] uhum [tem que ter a sua_i a sua privacidade] e eu não não sou contra não eu sou sou a favor.
- (25) [cada umi vai pro seu_i lado] [cria a sua_i vida] eu já não eu já não penso assim
- (26) ahn e cresceu... lá... [eu_i veio pra São Paulo] [acho que já com seus_i vinte e poucos anos] ne começou a trabalhar na metalúrgica.

Utilizando um dos pressupostos da MD, descreveremos as características dos pronomes nessas sentenças por meio de traços. Detalharemos os traços, levando em conta a anaforicidade, a posição pré/pós nominal e a natureza do antecedente.

O primeiro aspecto elucidado a partir das sentenças e a partir dos dados apresentados na seção anterior é que os pronomes seu(s) e sua(s) aparecem mais em contextos de RD, tendo um comportamento semelhante às anáforas, que devem estar ligadas dentro de sua CR. Desse modo, descreveremos essa característica por meio dos traços [+anafóricos, -pronominais]. Embora os pronomes seu(s) e sua(s) apareçam mais nos contextos de RD, percebemos que aparecem também em contextos de RF, como nos enunciados (24-26).

Observamos que a maioria dos antecedentes estão na 3ª pessoa, aparentemente generalizadas. Nos enunciados, a 3ª pessoa generalizante (3G) é usada para fazer referência mais gerais como, por exemplo: ‘o povo’; ‘a gente’; ‘todo mundo’, ‘cada um’, bem como o ‘você’ indeterminado (2G) na sentença (21). Na sentença (23), o pronome ‘sua’ retoma uma 2ª pessoa, o ‘você’.

A partir das sentenças, observamos que as formas possessivas ‘seu(s)’ e ‘sua(s)’ podem recuperar referentes na 2ª pessoa, 3ª pessoa, 2G e 3G, possuindo então os seguintes traços para a pessoa do antecedente: [+2ªpessoa, +3ªpessoa, +2G, +3G]. No entanto, como vimos a partir dos dados apresentados nos quadros 2 e 3, os pronomes seu(s) e sua(s) são mais produtivos com antecedentes na 2G e 3G, com isso, já é possível verificar que pelo menos em contextos orais os pronomes seu(s) e sua(s) são utilizados na maior parte das vezes para retomar antecedentes generalizados. Portanto,

assim como descrito em Menuzzi (1996), percebemos que no PB falado, o número de vezes que ‘seu’ aparece com antecedentes referenciais de 3ª pessoa é muito inferior à quantidade de vezes em que ele aparece com antecedentes generalizantes.

Embora os pronomes seu(s) e sua(s) possam possuir tanto antecedentes específicos e definidos quanto antecedentes não-específicos e não-definidos nas sentenças, eles retomam mais as 2ª pessoas e 3ª pessoas generalizantes (2G e 3G). Desse modo, vamos considerá-los [–específico] e [–definido]. Para exemplificarmos o comportamento desses traços, podemos substituir o antecedente do exemplo (22) por um antecedente [+específico] e [+definido]. Após essa substituição, teríamos:

(27) [João_i respeita o que ele_i quer da sua_i vida]

O exemplo em (27) também passa a ter uma nova interpretação, pois o pronome ‘sua’ pode se referir não apenas a ‘João’, mas também a alguém fora da CR. Desse modo, nesta análise vamos considerar que os pronomes seu(s) e sua(s) em contextos anafóricos parecem não ser ambíguos, quando possuem um antecedente [–específico] e [–definido].

Em relação à posição, as formas possessivas ‘seu(s)’ e sua(s)’ aparecem apenas em posição pré-nominal: em (17) ‘seu direito iguais’; (18) ‘seu momento certo’; (19) ‘pro seu umbigo’; (20) ‘dos seus estados’; (21) ‘pela sua capacidade’; (22) ‘da sua vida’; (23) ‘na sua época’; (24) ‘sua privacidade’; (25) ‘seu lado’/ ‘sua vida’ e em (26) ‘seus vinte e poucos anos’. Apesar das formas canônicas funcionarem também em posição pós-nominal, dentro das relações anafóricas, a posição pré-nominal parece ser a preferida para os falantes. Se alternarmos a posição dos pronomes em uma dessas sentenças, a sentença pode parecer estranha para alguns falantes. Vejamos no exemplo a seguir, quando alternamos a posição do pronome ‘seu’ da sentença (19), reescrita em (28):

(28) [?está todo mundo_i olhando pro umbigo seu_i]

Quando a posição do pronome é alternada, como em (28), a sentença parece se tornar ambígua, pois ‘seu’ pode se referir a ‘todo mundo’, mas também parece se referir a um sujeito que não está saliente no discurso. Neste contexto, o pronome ‘seu’ não é interpretado como uma anáfora, de modo que temos um exemplo do seu uso como dêitico, em que a referência do pronome é definida pelo contexto extralinguístico, ou seja, alguma situação que o pronome é mencionado após um gesto em que se aponta para uma determinada pessoa/objeto. É extralinguístico, pois não depende do discurso. Sendo assim, vamos considerar que os traços que expressam a posição das formas possessivas canônicas, dentro das relações anafóricas, são [+pré-nominal, –pós-nominal].

Através dos dados levantados na seção anterior é possível perceber que os pronomes seu(s) e sua(s) são os únicos que retomam 2G. No enunciado (21), notamos que o ‘você’ não está se referindo a um indivíduo específico, mas parece se referir a qualquer indivíduo, por isso é generalizado.

Souza (2015) faz um estudo aprofundado sobre o uso do ‘você’ em contextos de referência indeterminada. A autora explica que:

[...] não devemos fazer análises isoladas, levando em consideração apenas o nível sintático, mas também os níveis semântico, pragmático e discursivo das situações concretas de uso. Como também, chamamos atenção que é aí que a estrutura da língua é adquirida, é onde a forma se ajusta na função que passou a servir, como a forma pronominal *você* , utilizada frequentemente para contextos anafóricos e dêiticos, mas que passou a adquirir referência indeterminada em dadas situações comunicativas (SOUZA, 2015, p. 111).

Quando olhamos para os contextos nos quais o ‘ *você* ’ indeterminado aparece nas entrevistas, percebemos que ele é sempre retomado pelas formas *seu(s)* e *sua(s)* e nunca por *dele(s)* e *dela(s)* . De acordo com Müller (2000), o pronome muitas vezes não se refere a qualquer entidade, é o caso dos pronomes com antecedentes quantificados, pronomes relativos e interrogativos. Diante disso, vemos que esse aspecto trazido pelo estudo de Müller (2000) sobre a natureza do antecedente, pode se referir também aos casos de 2G.

Como vimos anteriormente, existem dois tipos de relações anafóricas, segundo Müller (2000):

correferência e ligação de variáveis. Como assumimos que todos os pronomes são variáveis, o fato de estes receberem uma interpretação (co-)referencial ou de variável presa, irá depender de como seu valor é determinado em cada caso. Um pronome referencial tem sua interpretação determinada por uma função que atribui um indivíduo enquanto valor de uma variável. Já a interpretação de variável ligada significa que o valor do pronome é determinado pelo seu antecedente e independe de qualquer função (con)textual (MÜLLER, 2000, p. 8).

Analisando os contextos em que os pronomes *seu(s)* e *sua(s)* aparecem, identificamos que as relações anafóricas ocorrem como ligação de variáveis, ou seja, a relação anafórica em que os pronomes *seu(s)* e *sua(s)* aparecem mais são [+ligação de variáveis, -correferência]. Müller (1997) argumenta que ocorre uma especialização das formas possessivas ‘ *seu* ’ e ‘ *dele* ’, de modo que, ‘ *seu* ’ é escolhido para atuar como variável presa e ‘ *dele* ’ para expressar uma relação semântica de correferência. Tais formas, segundo a autora, estão sujeita às imposições sintáticas decorrentes deste modo de funcionamento.

Em (25), ‘ *cada um* ’, que é interpretado como 3ª pessoa generalizante (3G), é retomado por ‘ *seu* ’ dentro de CR, mas em seguida é retomado por ‘ *sua* ’ fora de CR. Já em (26), embora o referente não esteja evidente, a ‘ *ec* ’, que é uma categoria vazia, pode se referir a um ‘ *ele* ’ implícito, ou seja uma 3ª pessoa, que é retomada por ‘ *seus* ’ fora da CR. As ‘ *ecs* ’, segundo Miotto *et al.* (2016), correspondem a DPs que podem ser classificadas a partir da Teoria da Ligação. Os autores explicam que esse tipo de ‘ *ec* ’ possui uma versão sem matriz fonética de um verdadeiro pronome, com propriedades de distribuição complementar semelhantes. Esse tipo de ‘ *ec* ’ é chamada de ‘ *pro* ’ (*prozinho*).

5.2. Pronomes possessivos não-canônicos: *dele(s)* e *dela(s)*

Os pronomes *dele(s)* e *dela(s)* aparecem nos dois contextos, RD e RF, porém são mais utilizados em RF. Vejamos algumas sentenças retirada do corpus, em que tais pronomes aparecem em contexto RF:

- (29) [tem uma moça_i que trabalha comigo] [que ela_i é da Bahia] [e eu acho lindo o sotaque dela_i]
- (30) [ele resolveu registrar ela_i] [mas porque ele voltou a namorar com a mãe dela_i]
- (31) [um sítio inclusive que o meu pai_i morou quando criança]... né que legal [os amigos dele_i estavam lá] receberam a gente
- (32) [aqui aonde eu moro são cinco sobrados_i]... quatro deles_i já foram assaltados
- (33) [o dinheiro_i hoje você tem amanhã não tem] [mas a gente... não é que a gente é... escravo dele_i]
- (34) [uma pessoa_i que] [acabou de... perder tudo na casa dela_i... né?]

Nesses enunciados, observamos que os pronomes agem dentro do que espera o Princípio B da Teoria da Ligação, ou seja, eles não estão ligados em sua CR. Em (29) o pronome ‘dela’ está retomando o referente ‘moça’ e ‘ela’ ao mesmo tempo, mas essa ligação ocorre fora do domínio de regência. O mesmo ocorre nos enunciados subsequentes. No entanto, esses pronomes também podem aparecer ligados dentro de sua CR, como nos seguintes enunciados:

- (35) [meu namorado_i por exemplo eh [ele_i é super próximo da família dele_i]
- (36) [porque a minha mãe_i tinha a mãe dela_i viva] a minha vó era viva ainda
- (37) [eles_i sempre eles_i tem associações deles_i]
- (38) [meu ex-namorado_i roubaram duas vezes o carro dele_i em Pinheiros]

Diante dos enunciados apresentados acima, vemos que as formas possessivas dele(s) e dela(s), ora aparecem fora de CR, como nos exemplos de (29-34), ora dentro de CR como nas sentenças (35-38). Ao atribuir alguns traços para esses pronomes, o primeiro aspecto que pode ser elucidado é que os pronomes dele(s) e dela(s) aparecem mais em contextos de RF, comportando-se como pronomes de fato. Desse modo, esses pronomes podem ser descritos como [+pronominais, -anafóricos].

Através dos enunciados e dos dados apresentados anteriormente nos quadros 4 e 5, percebemos que as formas dele(s) e dela(s) podem retomar também antecedentes de 3G, como demonstrado em (34). Nesta sentença, o sujeito ‘uma pessoa’ parece se comportar como sintagma quantificado. Então, em relação ao antecedente, diferente dos pronomes seu(s) e sua(s) que possuem antecedentes diversos, as formas não canônicas retomam apenas a 3ª pessoa e a 3G; contudo, é muito mais recorrente quando usado para retomar um antecedente na 3ª pessoa. Portanto, os antecedentes dos pronomes dele(s) e dela(s) são [+3ª pessoa, -2ª pessoa, -2G, -3G].

Diferente do que vimos com as formas canônicas, os antecedentes dos pronomes dele(s) e dela(s) serão, na maioria dos casos, especificados e definidos. Para ilustrar tal proposição, vamos

tomar a sentença (22) apresentada na seção anterior, a qual contém uma forma possessiva canônica, e, no lugar dessa sentença, vamos inserir a forma não-canônica ‘dele’. Vejamos a sentença (39):

(39) ?[a gente_i respeita o que cada um_i quer da vida dele_i]

Nota-se que o pronome ‘dele’ parece não poder ser ligado ao antecedente ‘a gente’ e ‘cada um’, pois a sentença fica agramatical. Porém, se considerarmos que o pronome ‘dele’ não está retomando nenhum dos dois antecedentes, mas está se referindo a outro sujeito não saliente no enunciado, a sentença é considerada gramatical. Por este motivo, em contextos que as formas possessivas não-canônicas são utilizadas em uma relação anafórica, espera-se encontrar antecedentes [+definido] e [+específico].

No que se refere à posição em que o pronome aparece nas sentenças, as formas possessivas dele(s) e dela(s) aparecem apenas em posição pós-nominal: (29) ‘sotaque dela’; (30) ‘mãe dela’; (31) ‘amigos dele’; (32) ‘quatro dele’; (33) ‘escravo dele’; (34) ‘casa dela’; (35) ‘família dele’; (36) ‘mãe dela’; (37) ‘associações deles’; (38) ‘carro dele’. Quando alternamos a posição desses elementos nas sentenças, para pré-nominal, as sentenças tornam-se agramaticais. No exemplo (40) a seguir, alternamos a posição do pronome ‘dele’ da sentença (38):

(40) *[meu ex-namorado_i roubaram duas vezes o dele_i carro em Pinheiros]

Desse modo, a posição das formas possessivas não-canônicas, dentro das relações anafóricas são [–pré-nominal, +pós-nominal].

Em relação aos contextos em que os pronomes dele(s) e dela(s) aparecem, a relação anafórica em que esses pronomes aparecem são ambas: ligação de variáveis e correferência; porém, como tais formas retomam na grande maioria dos casos antecedentes referenciais, manifestando uma correferência, podemos considerar que são descritos pelos traços [–ligação de variáveis + correferência].

5.3. Proposta de análise

A utilização da MD em nossa análise se justifica pelo fato de que será possível realizar uma análise mais elegante e satisfatória sobre o porquê de os pronomes possessivos seu(s), sua(s), dele(s) e dela(s) poderem ser interpretados ora como anáforas, ora como pronomes. Sobretudo, a partir desse modelo teórico será possível discutir a categoria desses elementos e compreender se são uma categoria especial de pronome.

Retomando algumas ideias, para a MD, na Lista 1 estão os elementos primitivos, traços morfossintáticos e semânticos, além de categorizadores e núcleos funcionais, que serão manipulados pela sintaxe. Todos esses elementos não possuem material fonológicos (BASSANI; MINUSSI, 2020).

Ressaltamos ainda um fenômeno presente nas línguas e que está ligado aos nossos dados. Notamos que os pronomes seu(s) e sua(s) podem possuir diversos tipos de antecedentes, nas sentenças em que tais pronomes aparecem, eles retomam antecedentes de 3ª pessoa, 2ª pessoa, 2G e 3G. O

sincretismo, segundo Lazzarini-Cyrino (2015) é um dos fenômenos mais instigantes nos estudos da Morfologia, pois ocorre quando traços sintático-semânticos diferentes são realizados pelo mesmo conteúdo fonológico. Assim sendo, os pronomes seu(s) e sua (s) estão realizando diferentes traços no PB, o que pode caracterizar o fenômeno de sincretismo nessas formas.

A seguir retomamos as sentenças (21), (22) e (23) que possuem os antecedentes 2G, 3G e 2ª pessoa respectivamente:

Quadro 6: Pronome possessivo canônico em relação anafórica com seu antecedente

Antecedente	Pronome possessivo canônico em relação anafórica
2G	(41) [você _i não ganha um trabalho pela sua _i capacidade]
3G	(42) [a gente _i respeita o que cada um _i quer da sua _i vida]
2ª pessoa	(43) [você _i via então confissões de adolescente na sua _i época né?]

A partir do quadro 6, sugerimos que o item /seu(s)/sua(s)/ é subespecificado para os traços da pessoa do antecedente, uma vez que não precisa ser plenamente especificado para a inserção em uma determinada posição sintática. Assim, um item pode possuir menos traços que os traços existentes no nó terminal. Os traços que são manipulados pelo Sistema Computacional recebem a sua forma fonológica tardiamente, através do processo de Inserção Tardia de material fonológico que ocorre através da operação de *spell-out* fonológico.

A fim de analisar os pronomes por meio de traços, apresentaremos o quadro 7 a seguir, para explicar cada traço sugerido:

Quadro 7: Traços de Anaforicidade, pessoa do antecedente, posição do pronome, relação anafórica e definitude e especificidade do antecedente

Anaforicidade	[+/-anafórico, +/-pronominal]	Quando o pronome apresenta mais características de anáfora será [+anafórico], ou seja, casos em que são ligados em sua CR. São [+pronominal] quando apresentarem mais características de pronomes, casos em que os pronomes são ligados somente fora de sua CR.
Pessoa do antecedente	[+/-2, +/- 3, +/- 2G, +/- 3G]	Quando o antecedente for, por exemplo, 'você' será [+2], bem como quando possuir um antecedente que seja um sujeito determinado, como: 'João' será [+3]. A pessoa [+2G] foi criada nesse estudo para designar o antecedente 'você' em referência indeterminada. Casos em que o sujeito for indeterminado, ex.: 'cada um' será [+3G].
Posição do pronome	[+/-pré-nominal, +/-pós-nominal]	A posição se refere ao lugar onde o pronome está posicionado, ou seja, antes ou depois do nome, de modo que quando estiver antes do nome será [+pré-nominal] e quando estiver após o nome será [+pós-nominal].

Relação anafórica	[+/-ligação de variáveis, +/- correferência]	Como vimos em Müller (2000), a interpretação de ligação de variáveis significa que o valor do pronome é determinado pelo seu antecedente e independe de qualquer função (con)textual [+ligação de variável]. Já a interpretação de correferência se dá quando um pronome possui sua interpretação determinada por uma função que atribui a um indivíduo enquanto valor de uma variável [+correferência].
Definitude e especificidade do Antecedente	[+/-definido, +/- específico]	Em casos que possuem antecedente genérico, o pronome receberá o traço [-definido, -específico], por exemplo: 'todos'. Enquanto nos casos com antecedentes determinados, o pronome será [+definido, +específico], por exemplo: 'João'.

No quadro 8, sugerimos os seguintes traços para descrevermos os pronomes seu(s) e sua(s):

Quadro 8: Criação de traços para os pronomes seu(s) e sua(s) com foco na natureza dos antecedentes

Antecedente	Pronome	Traços
[2ª pessoa]	seu(s)/sua(s) ↔	[+anafórico, -pronominal, +2ª pessoa, -3ª pessoa, -2G, -3G, +ligação de variáveis, -correferência, -específico, -definido, +pré-nominal, -pós-nominal]
[3ª pessoa]	seu(s)/sua(s) ↔	[+anafórico, -pronominal, -2ª pessoa, +3ª pessoa, -2G, -3G, +ligação de variáveis, -correferência, -específico, -definido, +pré-nominal, -pós-nominal]
[2G]	seu(s)/sua(s) ↔	[+anafórico, -pronominal, -2ª pessoa, -3ª pessoa, +2G, -3G, +ligação de variáveis, -correferência, -específico, -definido, +pré-nominal, -pós-nominal]
[3G]	seu(s)/sua(s) ↔	[+anafórico, -pronominal, -2ª pessoa, -3ª pessoa, -2G, +3G, +ligação de variáveis, -correferência, -específico, -definido, +pré-nominal, -pós-nominal]

Segundo Embick (2015), os IVs subespecificados permitem que o sincretismo seja verificado como não acidentais, já que fornecem um mecanismo pelo mesmo IV, ou seja, o mesmo expoente fonológico. A partir do quadro 8, podemos concluir que o IV /seu(s)/sua(s)/ possui os seguintes traços:

- (44) /seu(s)/sua(s)/ ↔ [+anafórico, -pronominal, +ligação de variáveis, -correferência, -específico, -definido, +pré-nominal, -pós-nominal]

De acordo com Embick (2015), a Inserção Tardia permite que algumas generalizações sejam maximizadas pela redução do número de IVs, de modo que, um sincretismo é encontrado, quando traços sintáticos-semânticos distintos são realizados fonologicamente do mesmo modo.

Na perspectiva do autor, podemos dizer que a realização dos morfemas abstratos de 2ª pessoa, 3ª pessoa, 2G e 3G como antecedentes dos pronomes seu(s) e sua(s) é sistemática, não é acidental,

isto é, essa sistematicidade deve ser codificada na gramática, de maneira a desconsiderar o uso de morfemas tradicionais. Vejamos no Quadro 9 abaixo, como podemos derivar os traços para os pronomes dele(s) e dela(s):

Quadro 9: Criação de traços dos pronomes dele(s) e dela(s) com foco na natureza dos antecedentes

Antecedente	Pronome	Traços
[3ª pessoa]	dele(s)/dela(s) ↔	[-anafórico, +pronominal, -2ª pessoa, +3ª pessoa, -ligação de variáveis, +correferência, +específico, +definido, -pré-nominal, +pós-nominal]

A partir do quadro temos os seguintes traços para o IV /dele(s)/dela(s):

- (45) /dele(s)/dela(s)/ ↔ [-anafórico, +pronominal, -2ª pessoa, +3ª pessoa, -ligação de variáveis, +correferência, +específico, +definido, -pré-nominal, +pós-nominal]

Bertolino e Grolla (2012) fazem um estudo sobre o pronome ‘ele’ e trazem algumas reflexões interessantes, que podem ser consideradas aqui. As autoras dizem que no PB o pronome ‘ele’ pode ter a leitura correferencial e a leitura localmente ligada. As autoras propõem que o ‘ele’ é uma expressão não especificada para ligação, uma UBE (*Unspecified Bindable Expression*). Os pronomes seu(s), sua(s), dele(s) e dela(s) parecem também ser subespecificados para ligação.

Considerações finais

Por meio do arcabouço teórico da MD mostramos que os pronomes possessivos seu(s), sua(s), dele(s) e dela(s) não atuam como prevê o Princípio B da Teoria da Ligação, já que esses elementos apareceram tanto em contextos de RD, como em contextos de RF. Tais pronomes são subespecificados, ou seja, os IVs não são plenamente especificados para a inserção em uma determinada posição sintática, de modo que, esses pronomes parecem possuir menos traços que os traços existentes no nó terminal. A subespecificação atua sobre os IVs e não sobre os nós, que continuam integralmente subespecificados.

No entanto, os pronomes seu(s) e sua(s) parecem ser mais subespecificados do que dele(s) e dela(s), pois possuem diversos tipos de antecedentes: 3ª pessoa, 2ª pessoa, 2G e 3G ocasionando o fenômeno de sincretismo, já que tais antecedentes possuem mais de um traço sintático-semânticos, sendo retomado pelo mesmo conteúdo fonológico. Além disso, os pronomes seu(s) e sua(s) são mais encontrados em contextos RD do que em contexto de RF, possuindo então mais traços de anáforas do que pronomes, enquanto dele(s) e dela(s) apresentam mais traços de pronomes. Desse modo, consideramos os traços para os IVs /seu/sua(s)/ e /dele(s)/dela(s)/ respectivamente demonstrados a seguir:

- (46) /seu(s)/sua(s)/ ↔ [+ligação de variáveis, -correferência, -específico, -definido, +pré-nominal, -pós-nominal]

(47) /dele(s)/dela(s)/ ↔ [-2ª pessoa, +3ª pessoa, -ligação de variáveis, +correferência, +específico, +definido, -pré-nominal, +pós-nominal]

Mioto *et al.* (2016), quando apresentam a ‘ec’ PRO¹², a classificam como [+anafórico, +pronominal], os autores explicam através do Teorema de PRO, que este não tem categoria de regência. Embora, os pronomes seu(s), sua(s), dele(s) e dela(s) sejam, na maioria dos casos regidos, por receberem Caso, esses parecem possuir os mesmos traços que PRO: [+anafórico, +pronominal], pois se comportam tanto como pronome, quanto como anáfora, sendo o seu(s) e sua(s) aparentemente mais anafóricos e dele(s) e dela(s) mais pronominal.

Referências

- ARIEL, Mira. Accessibility theory: *an overview*. In: SANDERS, T.; SCHILPEROORD, J.; & SPOOREN, W. (eds.). *Text representation: linguistic and psycholinguistic aspects*. Amsterdam: J. Benjamins Publishing Co, 2001.
- BASSANI, I. DE S.; MINUSSI, R. D. Sobre o conteúdo fonológico das raízes: raízes supletivas, fonologias genéricas e erros de fala. *Revista do GELNE*, v. 22, n. 2, pp. 267-79, 6 nov. 2020.
- BERTOLINO, Karina G.; GROLLA, Elaine. *O pronome “ele” está sujeito ao princípio B? Uma discussão sobre os resultados experimentais*. *Revista Linguística / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro*, v. 8, n. 2, dez. de 2012.
- CHOMSKY, Noam. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.
- LAZZARINI-CYRINO, João Paulo. *O sincretismo passivo-reflexivo: um estudo translinguístico*. Tese (Doutorado) - Faculdade de Filosofia, Letras e Ciências Humanas. Universidade de São Paulo (USP). São Paulo. (11/21973-1)
- EMBICK, D. *The Morpheme: A Theoretical Introduction*, volume 31. Mouton de Gruyter, 2015.
- HALLE, Morris; Alec MARANTZ. Distributed Morphology and the Pieces of Inflection. In: HALE, Kenneth; KEYSER, Samuel Jay (eds.) *The View from Building 20*. Cambridge: MIT Press, 1993, pp. 111-76.
- HALLE, Morris. Distributed Morphology: Impoverishment and Fission. In: BRUENING, B.; KANG, Y.; MCGINNIS, Martha (eds.). *MIT Working Papers in Linguistics 30: Papers at the Interface*, Cambridge, MA, MIT Press, 1997, pp. 425-49.
- HEIM, Irene; KRATZER, Angelika. *Semantics in Generative Grammar*. Malden, Mass, Oxford, U.K.: Blackwell, 1998.
- MENUZZI, Sérgio. 3rd person possessives in Brazilian Portuguese: *On the syntax-discourse relation*. University Centre for Computing Corpus Research on Language (UCREL), v. 8, pp. 191-210, 1996.
- MIOTO, Carlos; SILVA, Maria Cristina Figueiredo; LOPES, Ruth. Teoria da Ligação. In: *Novo manual de sintaxe*. São Paulo: Contexto, 2016. pp. 207-33.

¹² Prozão.

MÜLLER, Ana Lúcia de Paula. *A gramática das formas possessivas no português do Brasil*. São Paulo: Campinas, 1997.

MÜLLER, Ana Lúcia de Paula. *Anáfora Pronominal*. Revista Letras, Curitiba: Editora da UFPR, jul./dez. 2001 n. 56, pp. 259-75.

MÜLLER, Ana Lúcia de Paula. *Pronomes E Anáfora - O Estado Da Arte*. São Paulo: USP, 2000.

REINHARDT, Tanya. *The Syntactic Domain of Anaphora*. Ph. D. Dissertation. Cambridge, Mass., USA, 1976.

SCHER, Ana Paula; BASSANI, Indaiá de Santana; MINUSSI, Rafael Dias. Morfologia em Morfologia Distribuída: *Morphology In Distributed Morphology*. Salvador: Estudos Linguísticos e Literários, jan.-jun. 2013, n. 47. pp. 9-29.

SOUZA, Mônica dos Santos. *Estudo da Referência Indeterminada do Pronome “Você” na Perspectiva Funcionalista da Linguagem*. 2015. Dissertação (Mestrado em Linguística) – Centro de Ciências Humanas e Naturais, Universidade Federal do Espírito Santo, Vitória, 2015.

A CARTOGRAFIA DOS PRONOMES RESUMPTIVOS PESSOAIS DA CONSTRUÇÃO DE SUJEITO DUPLO DO PORTUGUÊS BRASILEIRO: UMA PROPOSTA SINCRÔNICA

THE CARTOGRAPHY OF PERSONAL RESUMPTIVE PRONOUNS OF THE BRAZILIAN PORTUGUESE'S SUBJECT DOUBLING CONSTRUCTION: A SYNCHRONIC PROPOSAL

Letícia Emília Krieck¹

RESUMO

Este artigo articula uma proposta de análise cartográfica e de caráter sincrônico para as sentenças formadas por um sujeito pré-verbal duplicado por um pronome pessoal resumptivo (SD) do português brasileiro (PB). Parte-se da tese já abordada na literatura (cf. COSTA; DUARTE; SILVA, 2004; QUAREZEMIN, 2019, 2020; entre outros trabalhos) de que, nessa língua, os constituintes iniciais de tais sentenças não são necessariamente topicalizados, podendo, ainda, se situar no domínio argumental. Assume-se, em concordância com Quarezemin (2019, 2020) e Krieck (2022), que esses constituintes não topicalizados ocupem Spec, SubjP, configurando-se, portanto, como sujeitos sentenciais *default*. A partir de tal premissa, é assumido que o pronome resumptivo manifesta comportamento heterogêneo ao longo da sorte de sentenças de SD do PB, dependendo da posição ocupada pelo elemento inicial e de outras condições gramaticais envolvidas. As peculiaridades que caracterizam os pronomes resumptivos permitem que sejam assim categorizados, em acordo com a tripartição pronominal proposta em Cardinaletti e Starke (1994): como pronomes fortes ou fracos, em face de um DP inicial à esquerda, e como pronomes fracos ou clíticos, quando esse elemento ocupa uma posição interna a IP. Sob tal concepção, este trabalho propõe que a classe em que o elemento resumptivo se enquadra define a posição ocupada por esse na estrutura, de tal forma que: (i) pronomes resumptivos fortes ocupem Spec, SubjP, podendo, muito restritamente, se situar à esquerda; (ii) pronomes resumptivos clíticos ocupem a posição nuclear Subj^o; e (iii) pronomes resumptivos fracos figurem em Spec, TP. Com base nessa análise, são delineados, ainda, os distintos processos de formação estrutural que subjazem a produção de tais sentenças.

PALAVRAS-CHAVE: Português brasileiro. Sintaxe cartográfica. Sujeito duplo. Pronomes.

ABSTRACT

This article articulates a cartographic and synchronic proposal to the Brazilian Portuguese's (PB) sentences formed by a preverbal subject duplicated by a resumptive personal pronoun (SD). It is understood by academic writing on the theme that, in this language, the initial phrase of these sentences is not necessarily topicalized; it might occupy a position within the argumental domain. It is assumed, in accordance with Quarezemin (2019, 2020) and Krieck (2022), that these non-topicalized phrases can fill Spec, SubjP, as default sentence subjects. Under this supposition, it is expected that the resumptive pronoun manifests heterogeneous behavior across the variety of SD sentences in PB, depending on the position occupied by the initial phrase and other grammatical conditions involved. The singularities that characterize the resumptive pronouns allow these items to be categorized as follows, under the Cardinaletti & Starke's (1994) pronominal tripartition: strong or weak ones, when the initial DP is a topic, and clitic or weak ones, when this element occupies a position within IP. On this assumption, this analysis proposes that the category of the resumptive item defines the position it occupies in structure, so that: (i) strong resumptive pronouns occupy Spec, SubjP and very narrowly a position inside CP domain; (ii) clitic resumptive pronouns occupy Subj^o; and (iii) weak resumptive pronouns is at Spec,

¹ Mestra e Doutoranda em Linguística pela Universidade Federal de Santa Catarina, leticiakrieck@gmail.com, <https://orcid.org/0000-0003-4392-4193>.

TP. Based on this examination, the different processes of structural formation that underlie these sentences' production are also described.

KEYWORDS: Brazilian Portuguese. Cartographic Syntax. Double subject. Pronouns.

Introdução

O português brasileiro (PB) compartilha com línguas como o inglês, o português caboverdiano, o francês e o árabe tunisiano as instigantes construções formadas por um sujeito pré-verbal duplicado por um pronome pessoal (SD). Tais construções são formadas por um DP em posição inicial seguido direta (cf. (1)) ou indiretamente por uma forma pronominal pessoal correferente, que atua como resumptiva (PR). Quando da quebra de adjacência entre ambos os elementos, a interpolação pode ser realizada por um ou mais sintagmas (cf. (2a)) ou por quebra entoacional saliente, representada neste trabalho pela adição de vírgulas (cf. (2b)):

- (1) O meu irmão_i ele_i é um cara que viaja muito.
- (2) (a) Essa radiação_i, obviamente, ela_i pode ser transmitida pelo vácuo.
(b) Esse cheirinho do ácido_i, ele_i pode ficar no cabelo até as próximas lavagens.

(KRIECK, 2022)

Gradualmente mais comum entre os grupos de fala mais jovens (cf. DUARTE, 1998), no PB, as sentenças de SD encontram paulatinamente menos restrições à sua ocorrência, frente a um cenário de alta produtividade nessa língua. Esse fato, ao longo das últimas décadas, vem motivando uma ampla gama de análises linguísticas, que se dedicam, entre outras questões, a compreender as propriedades gramaticais dessa construção. Este trabalho se insere nesse contexto, ao analisar o comportamento de tal estrutura a partir, sobretudo, de um de seus elementos constitutivos: o PR. Essa investigação leva em conta tal objeto sob a perspectiva sincrônica, de forma a considerar, como fonte principal para as análises a serem realizadas, os dados de fala do *corpus* de Krieck (2022), formado por 550 sentenças de SD coletadas em materiais videográficos provenientes da plataforma virtual *YouTube*, resguardados por direitos autorais segundo a Lei nº. 9.610/1998 (BRASIL, 1998). A seleção dos dados foi realizada a partir de conteúdos gravados e disponibilizados necessariamente entre os anos de 2018 e 2020.

O estudo realizado neste trabalho segue os direcionamentos fornecidos pela vertente cartográfica de análise sintática (cf. CINQUE, 1999; RIZZI, 1997; CINQUE; RIZZI, 2010), que pode fornecer contribuições significativas e importantes a respeito desse item, bem como do fenômeno de SD em sua totalidade. São aproveitados, sobretudo, os direcionamentos fornecidos pelo princípio *one feature, one head* (“um traço, um núcleo”) (cf. KAYNE, 2005), a partir do qual “cada traço morfossintático corresponderia a um núcleo sintático independente com um espaço específico na hierarquia funcional” (CINQUE; RIZZI, 2010, p. 54, tradução da autora²). Assumi-lo lega a esta análise a decisiva

² Tradução do original: “each morphosyntactic feature would correspond to an independent syntactic head with a specific slot in the functional hierarchy”.

consequência de que constituintes que dispõem de um traço x serão categorizados em uma mesma categoria destinada a hospedá-los, fatalmente distinta de outra que incorpore um traço y .

A partir dessa orientação, são analisadas as propriedades gramaticais que caracterizam os PRs da estrutura de SD no PB, de sorte que sejam encontradas explicações a estas questões: (i) como se comportam tais elementos?; (ii) eles manifestam, uniformemente, as mesmas características em toda sorte de sentenças de SD?; (iii) de que forma(s) podem ser caracterizados?; e (iv) qual(is) é(são) a(s) posição(ões) destinada(s) a comportá-los? Os direcionamentos fornecidos por esta investigação podem prover subsídios a uma caracterização mais precisa dos PRs da estrutura de SD, além de contribuir aos estudos sobre a sintaxe dessas sentenças.

Aos objetivos a que se propõe, este artigo compreende, na seção 2, análises voltadas às posições ocupadas pelos constituintes iniciais da construção de SD no PB, de modo a definir a concepção adotada neste trabalho. Por sua vez, a seção 3 apresenta uma análise das propriedades gramaticais dos PRs da estrutura em questão, que permita classificá-los em acordo com a tipologia de Cardinaletti e Starke (1994) (C&S) para os pronomes das línguas naturais. A seção 4, por fim, caracteriza as posições ocupadas por tais itens, bem como aborda uma proposta estrutural para esses elementos em acordo com o comportamento gramatical que manifestam.

1. As posições estruturais dos DPs iniciais

Ao longo das últimas quatro décadas, a literatura vem se dedicando a compreender as propriedades da estrutura de SD no PB e por que ela é paulatinamente mais frequente entre os falantes dessa língua. Ainda, são objetos de análise os aspectos estruturais circunscritos a tal construção, especialmente em relação ao constituinte que figura em posição inicial (cf. PONTES, 1987; DUARTE, 1995; BRITTO, 1998; VASCO, 2006; ORSINI; VASCO, 2007; PIRES, 2007; KATO, 1993; COSTA; DUARTE; SILVA, 2004; SILVA; ZIOBER, 2017; QUAREZEMIN, 2019, 2020; CYRINO, 2021; e numerosos outros estudos). São algumas questões que motivam tais investigações: esse elemento ocupa invariavelmente a mesma posição? Seria tal um caso obrigatório de topicalização, ou existiram mais posições disponíveis para comportá-lo? Nesta seção, tais discussões são retomadas, com a finalidade de definir o posicionamento adotado por este estudo no que concerne aos aspectos estruturais circunscritos a esse sintagma. Essa posição, por sua vez, subsidia a análise a ser realizada sequentemente neste trabalho.

Através da literatura, há certa orientação que considera que a estrutura de SD envolva, obrigatoriamente, a periferação do constituinte inicial. Destaca-se, nesse sentido, o trabalho pioneiro e notório de Pontes (1987), que descreveu a forma como as construções de tópico acontecem no PB, incluindo o fenômeno em questão como uma das mais frequentes. A autora caracterizou a estrutura como envolvendo movimento de algum sintagma ao domínio à esquerda da sentença, seguido pela retomada desse elemento por meio de uma forma pronominal na sentença-comentário (cf. (3¹), a partir de (3)):

(3) O carro, ele está na oficina mecânica.

(PONTES, 1987)

(3') [_{CP} O carro_i [_{IP} t_i ele está [_{VP} ...

Posicionamento semelhante é adotado por Duarte (1995) e Barbosa, Duarte e Kato (2005), análises que consideram que a SD seja resultado da perda do sujeito nulo no PB. Segundo tais, a língua passa a dispor de um pronome lexicalizado fraco que figura na zona argumental, duplicando o constituinte situado no domínio à esquerda (cf. (4')), que estrutura (4)):

(4) A Clarinha ela cozinha que é uma maravilha.

(DUARTE, 1995)

(4') [A Clarinha] [[_{IP} ela_i cozinha [_{VP} t_i ...]]

(BARBOSA; DUARTE; KATO, 2005, p. 21)

Pires (2007), em complementaridade, argumenta que uma sentença de SD envolveria ambas as posições dedicadas a comportar sujeitos no PB: o constituinte inicial, à esquerda, em Spec, TopP, e o PR internamente à zona flexional, em Spec, TP (cf. (5b), a partir de (5a)). Entre os argumentos utilizados pelo autor em favor de considerar que os constituintes iniciais são tópicos sentenciais, está o fato de que essas sentenças não figurariam licitamente em contextos *out of the blue*, que bloqueiam tópicos (cf. (6b)). Por outro lado, quando o sujeito não se encontra ocupando uma posição no domínio periférico – ou seja, não há SD, de acordo com a análise –, há possibilidade plena de se situar nesse mesmo cenário (cf. (6a)). Outro indício abordado pelo trabalho se centra no fato de que o pronome reduzido de segunda pessoa do singular *cê* não pode ocupar a posição inicial (*i.e.*, topicalizada, segundo tal abordagem) em sentenças de SD, já que, por ser um pronome fraco, é impossibilitado de ser deslocado à esquerda, segundo as definições de C&S (cf. (7-8)).

(5) (a) A Silvia ela saiu cedo.

(b) [_{TopP} A Silvia_i, [_{TP} ela_i [_T saiu cedo]]]

(6) P: O que aconteceu?

R: (a) O Paulo me deu o anel de noivado no parque.

R: (b) ?? O Paulo, ele me deu o anel de noivado no parque.

(7) (a) * Cê, o seu pai cê pode convidar.

(b) Você, o seu pai cê pode convidar.

(8) (a) * Cê, cê pode convidar o seu pai.

(b) Você, (vo)cê pode convidar o seu pai.

(PIRES, 2007, p. 132)

Já trabalhos como Costa, Duarte e Silva (2004) oferecem outro ponto de vista a respeito dos aspectos circunscritos à estrutura das sentenças de SD do PB. Em tal análise, especificamente, os autores observam algumas diferenças entre esse fenômeno nessa língua em relação ao francês, em que a SD ocorre necessariamente com os DPs iniciais figurando como tópicos. Em suma, as distinções se fundamentam no fato de que, diferentemente do que ocorre no francês, no PB, tais sentenças: (i) podem conter um DP inicial quantificado (cf. (9)); (ii) podem manter esse elemento e PR adjacentes, sem haver quebra entoacional entre ambos (cf. (10)); (iii) não têm relevância estatística em dados de aquisição (cf. GONÇALVES, 2004; GROLLA, 2000); (iv) não podem ter adjacência rompida por XPs se o PR for fraco (cf. (11)); (v) não incorporam constituintes nominais indefinidos específicos ou nus figurando como iniciais (cf. (12)); e (vi) ocorrem sobretudo envolvendo 2ª ou 3ª pessoas.

(9) Cada criança, ela leva seu livro para a escola.

(DUARTE, 2003)

(10) Pedro ele telefonou faz horas.

(COSTA; DUARTE; SILVA, 2004, p. 140)

(11) (a) * Você intencionalmente/sempe cê lê livros românticos.

(b) Você cê intencionalmente/sempe lê livros românticos.

(c) * Você, a mulher de você, cê deve tratar ela bem.

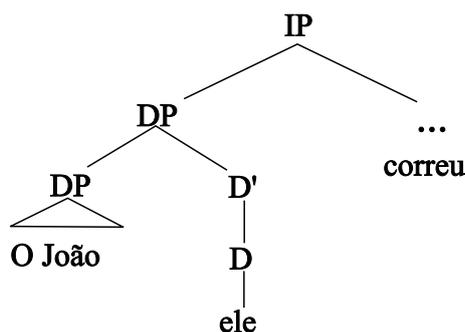
(12) (a) ?? Um menino ele chegou.

(b) * Gatos eles são felinos.

(c) * Brasileiro ele vive grudado no samba.

(COSTA; DUARTE; SILVA, 2004, p. 140)

Essas propriedades se apresentam enquanto evidências para Costa, Duarte e Silva (2004) definirem que a SD no PB se estrutura diferentemente daquela no francês, já que o PB parece não exigir a topicalização do constituinte inicial – fato que, categoricamente, distingue essa análise daquelas precedentemente apresentadas, segundo as quais o fenômeno do SD depende da periferização. Para os autores, o DP inicial é duplicado por um PR fraco que hospeda traços de pessoa e preenche a posição de núcleo da mesma projeção ocupada por esse constituinte. Tal raciocínio é representado na figura 1:

Figura 1: A proposta de Costa, Duarte e Silva (2004) para a SD no PB

Fonte: Costa, Duarte e Silva (2004)

Os estudos de Quarezemin (2019, 2020) se associam a tal proposta à medida que notam questões pertinentes acerca do fenômeno de SD no PB: (i) não admite rompimento de adjacência entre um DP inicial e um PR reduzido tal qual *ei* (cf. (13)); (ii) quanto à extração, existem dissemelhanças referentes a tal estrutura (cf. (14)); e (iii) essa construção pode ser lícita em contextos *out of the blue*³, ao passo que, na presença de um XP ocupando, necessariamente, posição de tópico, a formação se torna sistematicamente agramatical (cf. (15)):

- (13) (a) Ele *ei* sempre trabalha até tarde
 (b) * Ele sempre *ei* trabalha até tarde.

(QUAREZEMIN, 2020, p. 119)

- (14) (a) Onde o Pedro acha que a Ana *ela* encontrou o João?
 (b) * Onde o Pedro acha que o João a Ana (*ela*) encontrou?

(QUAREZEMIN, 2020, p. 126)

- (15) P: O que aconteceu?
 R: (a) A Ana *ela* comprou um carro.
 R: (b) # Um carro, Ana comprou *ele*.

(QUAREZEMIN, 2020, p. 126)

Para a autora, os contrastes em (13-15) são satisfatoriamente justificados se considerado que a SD dispõe de uma dupla possibilidade de articulação: para além da viabilidade de o DP inicial figurar como um tópico, no domínio à esquerda da sentença, ele pode, ainda, se situar em Spec, SubjP,

³ Nota-se um evidente contraste entre tal concepção e Pires (2007), análise sob a qual sentenças de SD não se comportam licitamente em contextos *out of the blue*. Diferentemente dessa perspectiva, contudo, Quarezemin (2019, 2020) insere a SD na condição em tese para defender que os DPs iniciais de tal construção, nesse contexto, se comportam como itens situados em uma posição no domínio IP. Essa circunstância, em específico, não é considerada em Pires (2007), investigação que considera que tais DPs são, necessariamente, elementos topicalizados – e que, por esse motivo, são lacunares em condições *out of the blue*.

tal qual um sujeito sentencial canônico. Como uma consequência dessa análise, Quarezemin (2019, 2020) argumenta que os PRs também possuem comportamento dissemelhante, a depender da posição ocupada pelo constituinte inicial: os DPs iniciais topicalizados seriam duplicados por pronomes fortes, e aqueles situados na zona argumental, por pronomes fracos, que, nesse caso assumiriam *status* de clíticos. Apoiando-se na articulação proposta em Cardinaletti (1997, 2004), Rizzi (2005, 2015, 2018) e Rizzi e Shlonsky (2007) para as posições de sujeito, é proposta a dupla articulação em (16) para as estruturas das sentenças de SD do PB:

- (16) (a) $[_{TopP} DP [_{SubjP} pron_{forte} [_{Subj} [_{TP} t_i [_{T'} V (...)]]]]]$
 (b) $[_{SubjP} DP_i [_{Subj} pron_{clítico} [_{TP} t_i [_{T'} V (...)]]]]$

(adaptado de QUAREZEMIN, 2020, p. 127)

Krieck (2022) acrescenta a tais análises outras evidências favoráveis ao fato de que, no PB, a SD não necessariamente se trata de casos de topicalização do constituinte inicial. O primeiro ponto sinalizado pelo trabalho é a produtividade, pautada em seu *corpus* de dados de fala, de sentenças cujo constituinte inicial é indefinido (cf. (17)). Apoiando-se em Rizzi (2005), a autora argumenta que, já que incorpora informação antiga, um XP topicalizado encontra severas restrições para ser indefinido; diferentemente, sentenças de SD são lícitas dispendo de tal traço, mesmo em contextos *out of the blue* (cf. (18)). Por outro lado, em face do rompimento da adjacência entre DP inicial e PR por um XP que figura em posição de tópico – o que demanda que esse DP ocupe, também, uma posição à esquerda –, a sentença se torna sistematicamente agramatical (cf. (19)).

- (17) (a) Um dia solar ele possui 24 horas. Porém, um dia sideral ele possui 23 horas.
 (b) Note que, sempre, um projeto ele é gerenciado por um gerente.
 (c) Algumas ações políticas do governo elas ocasionaram revoltas populares.

- (18) P: O que aconteceu?
 R: (a) Um carro ele bateu no poste.
 R: (b) Pelo que eu supus, um carro ele bateu no poste.

- (19) ??/? Um carro, pelo que eu supus, ele bateu no poste.

(KRIECK, 2022)

A análise, ainda, sustenta que a expansão de CP proposta em Rizzi e Bocci (2017) pode fornecer contribuições à compreensão dos aspectos estruturais das sentenças de SD. Segundo os autores, a zona baixa desse domínio comporta uma projeção $Q_{emb}P$, que aloja elementos interrogativos que introduzem orações encaixadas. Conforme expressa a articulação em (20), que apresenta a formação hierárquica proposta pela análise, essa categoria se encontra em uma posição estruturalmente inferior a quaisquer tópicos:

(20) $[_{ForceP} [_{TopP*} [_{IntP} [_{TopP*} [_{FocP} [_{TopP*} [_{ModP} [_{TopP*} [_{QembP} [_{FinP} [_{IP} \dots$

(adaptado de RIZZI; BOCCI, 2017, p. 9, grifo da autora)

A partir de (20), Krieck (2022) aponta para o fato de que DPs iniciais de construções de SD podem ocorrer adjacentes a PRs em posição mais baixa em relação a sintagmas interrogativos encaixados (cf. (21a)) – indicando, assim, que o constituinte em posição inicial na SD não figure à esquerda, mas em uma posição argumental, conforme representa (21a’). Presumivelmente, o rompimento da adjacência entre ambos os elementos por um XP parentético acarreta uma formação lacunar (cf. (21b)), uma vez que não há posição disponível para comportar tal constituinte inferiormente a $Q_{emb}P$, consideradas as propriedades gramaticais desse sintagma⁴:

(21) (a) Eu fico me perguntando que coisa (que) o Bruno ele cozinhou pra Carla.

(a’) $[_{QembP}$ **que coisa** $[_{Fin}$ (que) $[_{IP}$ **o Bruno ele...**

(b) ?? Eu fico me perguntando que coisa (que) o Bruno, pra Carla, ele cozinhou.

(KRIECK, 2022)

Krieck (2022) admite, em associação às análises anteriores, que existam duas posições que possam comportar os elementos iniciais das sentenças de SD do PB: uma na zona CP, e a outra na zona IP. À esteira de Quarezemin (2019, 2020), o estudo propõe que, além de poderem se situar em Spec,TopP, como tópicos – tais quais quaisquer outros constituintes sentenciais –, esses DPs podem figurar em Spec,SubjP, *i.e.*, como sujeitos *default* da oração (cf. (22)):

(22) (a) $[_{TopP}$ **DP inicial** $[_{SubjP} [_{TP} \dots$

(b) $[_{TopP} [_{SubjP}$ **DP inicial** $[_{TP} \dots$

(KRIECK, 2022)

Essa proposição é assumida neste trabalho, que considera a dupla possibilidade estrutural às análises doravante realizadas. Certos padrões de SD são assumidos pela autora como necessariamente envolvendo o constituinte inicial em uma posição de tópico: (i) tópicos-anacolutos, sob a compreensão de Orsini (2003) (cf. (23a-a’)); (ii) contextos de duplicação do PR, referidos por Krieck (2022) como “triplicação do sujeito” (cf. (23b-b’)); e (iii) quebra de adjacência entre DP inicial e PR por XP(s) interveniente(s) à esquerda ou por contorno característico de tópico, no sentido de Rizzi (1997) (cf. (23c-c’). Por outro lado, o DP inicial pode ser atestado como ocupando Spec,SubjP sob circunstâncias

⁴ A hierarquia apresentada em (20) fornece a disponibilidade da posição Spec,FinP para alojar constituintes, que, dessa forma, estariam dispostos no domínio à esquerda. Sob certo ponto de vista, poder-se-ia conjecturar a alocação do DP inicial de uma sentença de SD nessa posição; contudo, essa opção se mostra incongruente com as análises de natureza cartográfica, que são adotadas por este estudo. Neste ponto, toma-se a concepção de Rizzi (1997), segundo a qual compete ao núcleo Fin^o a alocação das informações sentenciais relacionadas à finitude/forma. Uma consequência dessa análise aponta para o fato de que um sintagma sujeito oracional é impossibilitado de ocupar tal projeção, em face da incompatibilidade das propriedades desse elemento com o núcleo em foco.

tais como quando o contorno de tópico é inexistente e a pausa entre DP inicial e PR é breve (cf. (24), a partir de (1)).

- (23) (a) O catarinense, a gente tá acostumado a receber migalha.
 (a') [_{TopP} o catarinense [_{IP} a gente [...
 (b) O fast-food ele, em algum momento, ele termina com a ideia da comensalidade.
 (b') [_{TopP} o fast-food [_{TopP} ele [_{XP} em algum momento [_{IP} ele [...
 (c) Essas incertezas, ao longo do tempo, elas vão decaindo.
 (c') [_{TopP} essas incertezas [_{XP} ao longo do tempo [_{IP} elas [...

(KRIECK, 2022)

- (24) [_{SubjP} o meu irmão [_{X(P)} ele [...

2. O comportamento dos pronomes resumptivos

Assim como a literatura vem apontando para a importância de serem averiguados os aspectos circunscritos aos sintagmas iniciais do fenômeno de SD, um exame atento das propriedades inerentes aos PRs se torna imprescindível à análise global do fenômeno. Esta seção traz uma proposta de cunho cartográfico a tal questão, abordando propriedades gramaticais desses itens que possam fornecer sugestões sobre o seu comportamento no PB. A esta altura, é relevante o destaque para o fato de que as sentenças de SD, em sua ampla variedade, podem envolver pronomes polissêmicos, *i.e.*, casos em que uma mesma forma morfofonológica manifesta propriedades distintas, a depender das condições nas quais se enquadra. Formações de SD cujos elementos são apenas pronomes (quase) homófonos – como no exemplo de Silva, Carvalho e Ziober (2017), em (25), a seguir – são exemplares de tal circunstância:

- (25) Ele, ele é um pouco ciclotímico.

(SILVA; CARVALHO; ZIOBER, 2017, p. 96)

A esta etapa, são utilizados os postulados do notório trabalho de C&S à classe pronominal ao longo das línguas naturais, análise que compreende a existência – em princípio, grosso modo – de duas classes de pronomes, fortes e deficientes, que se distinguem entre si por meio de um conjunto de características gramaticais. Em uma abordagem sintética da investigação, são considerados os seguintes aspectos, cruciais a este estudo: (i) apenas pronomes fortes podem ocupar alguma posição no domínio periférico; (ii) também exclusivamente essa classe pode figurar em estruturas de coordenação; (iii) referentes [- animados] são exclusivos à classe deficiente; (iv) tão somente as formas fortes podem portar modificadores categoriais; e (v) a redução morfofonológica se aplica exclusivamente a pronomes deficientes.

Em primeiro lugar, volta-se aos pronomes coordenados, que podem funcionar como resumptivos em sentenças de SD. Essa configuração é satisfatoriamente atestada diante de quebra de adjacência entre DP inicial e PR, além de em construções de tópico-anacoluto, conforme os modelos em (26). Por outro lado, PRs em coordenação se mostram ilícitos caso não existam indícios de que o constituinte inicial ocupe uma posição à esquerda, tal qual sugere (27). Sendo os pronomes coordenados encarados como formas fortes, sob a concepção de C&S, é possível sugerir, ao menos parcialmente, que constituintes iniciais situados no domínio IP sejam incompatíveis com tal classe como resumptiva em sentenças de SD.

- (26) (a) O Pedro e a Laurinha, ele e ela brigam todo dia.
 (b) O Pedro e a Laurinha, se bobear, ele e ela brigam todo dia.
 (c) Os meus filhos, ele e ela brigam todo dia.
- (27) */?? O Pedro e a Laurinha ele e ela brigam todo dia.

A validade de tal asserção pode ser atestada em se considerando as formas pronominais c-modificadas; conforme estabelecido por C&S, somente pronomes de classe forte admitem modificadores categoriais. Se tais pronomes forem compatíveis com sentenças de SD somente se o DP inicial figurar à esquerda, é esperado que, assim como no primeiro teste, sejam aceitáveis somente as construções cujo DP inicial funciona como tópico; diferentemente, quando esse elemento se situa em Spec,SubjP, a formação resultante se torna lacunar. Os testes em (28-29) confirmam tal defesa, apontando que pronomes fortes são, de fato, impossibilitados de ocorrer em face de DPs iniciais que ocupam uma posição no domínio argumental:

- (28) (a) O catarinense, só a gente tá acostumado a receber migalha.
 (b) A Clara, de todos os irmãos, só ela quis fazer faculdade.
 (c) A Clara, só ela quis fazer faculdade.
- (29) * A Clara só ela quis fazer faculdade.

Resta, ainda, atentar aos pronomes deficientes. Em acordo com C&S, pertencem a essa classe pronomes cujos referentes são dotados de traços [- animados]. O *corpus* de Kriek (2022) aponta que tais formas se inserem produtivamente como PRs em formações de SD reportando constituintes que figuram tanto à esquerda quanto em posições argumentais (cf. (30-31)). Esse fato, em específico, indica que a posição ocupada pelo DP inicial na sentença de SD não se mostra sensível à aceitação de formas deficientes como resumptivas, diferentemente do que acontece em relação àquelas fortes.

- (30) O talco, por exemplo, ele é facilmente riscado com a unha.
- (31) A tristeza ela faz parte da vida.

(KRIECK, 2022)

É notável, ainda, a incompatibilidade de DPs iniciais [- animados] em relação à retomada por PRs coordenados ou c-modificados, que só podem ser fortes, em acordo com C&S (cf. (32)). Essa má-formação se deve, precisamente, à necessidade de uma forma deficiente recuperar um constituinte dotado do traço em questão, independentemente, assim, da posição ocupada por tal:

- (32) (a) */?? A tristeza e o medo, ela e ele fazem parte da vida.
 (b) ??/? O talco, por exemplo, só ele é facilmente riscado com a unha.

Uma segunda análise possível à finalidade de atestar a tese em foco – de que PRs deficientes são lícitos para recuperar DPs iniciais independentemente do domínio em que tais constituintes figuram – abarca formas pronominais morfofonologicamente reduzidas. O processo de redução, em acordo com a proposta de C&S, afeta tão somente pronomes deficientes. Neste ponto, torna-se relevante a nota de que, segundo Kato e Duarte (2014), o PB dispõe de um produtivo paradigma de pronomes reduzidos (cf. (33)):

- | | | | | |
|------|------|------|-------|--------|
| (33) | EU | [ô] | VOCÊS | [ceis] |
| | VOCÊ | [cê] | ELES | [eis] |
| | ELE | [ei] | | |

(KATO; DUARTE, 2014, p. 16)

A literatura sobre a construção de SD aponta que PRs reduzidos são satisfatoriamente possíveis nessas sentenças caso não constatado o rompimento da adjacência entre o DP inicial e esse elemento (cf. (34), em que (34a) recupera (13a)). Quarezemin (2020), contudo, alerta para o fato de que, caso exista um XP interveniente entre um constituinte sujeito e um PR reduzido como *ei*, verifica-se má-formação sentencial (cf. (35a), que reinterpreta (13b)). Por outro lado, em Duarte (1995), atesta-se que esse fato não se aplica à forma reduzida *cê*, cuja adjacência com o DP inicial pode ser violada (cf. (35b)) sem haver qualquer prejuízo à aceitabilidade da construção:

- (34) (a) Ele *ei* sempre trabalha até tarde.

(QUAREZEMIN, 2020, p. 119)

- (b) Você *cê* não fala pra ele o que lhe contei.

(COSTA; DUARTE; SILVA, 2004, p. 141)

- (35) (a) * Ele sempre *ei* trabalha até tarde.

(QUAREZEMIN, 2020, p. 119)

- (b) Você, quando viaja, *cê* passa a ser turista.

(DUARTE, 1995, p. 138)

Apartir de (30), (31), (34) e (35b), é plausível a asserção de que pronomes deficientes são possíveis como elementos resumptivos independentemente do domínio em que o DP inicial se situa. Contudo, a inaceitabilidade de (35a) fornece um entrave a essa análise. Levando em conta essa formação, podem ser colocadas algumas questões, como: quais os aspectos que diferenciam a aceitabilidade de (34a)

em comparação a (35a), ambas que envolvem o mesmo elemento resumptivo? Por que a inserção de um sintagma interveniente em (35a) torna a sentença lacunar, enquanto a parentetização em (35b) não afeta a aceitabilidade da formação? Tais indagações revelam a necessidade de uma investigação mais complexa e particularizada dos pronomes deficientes, como um todo, e de formas pronominais reduzidas do PB, em particular.

Ao discorrer acerca das particularidades que definem os pronomes ao longo das línguas, C&S notam que a divisão da classe entre *fortes vs. deficientes* não atende plenamente à sua complexidade. Envolvendo-se em tal questão, os autores propõem que os pronomes deficientes, por manifestarem comportamentos divergentes entre si, devem ser divididos entre *fracos e clíticos*. A diferença entre ambas as categorias se assenta em aspectos como estes: (i) pronomes fracos, assim como os fortes, são constituintes que ocupam Spec da categoria correspondente, enquanto clíticos são núcleos, situando-se, portanto, em X⁰; e (ii) tão somente pronomes fracos podem portar acento lexical.

Torna-se imprescindível, neste ponto, caracterizar de que forma os pronomes reduzidos envolvidos em (34-35) se enquadram na divisão estabelecida para a classe deficiente. Em primeiro lugar, são recuperados os aspectos inerentes ao pronome de segunda pessoa do singular *cê*. A teorização linguística acerca dessa forma (cf. COELHO, 1999; BARBOSA, 2005; PEREZ, 2006; PETERSEN, 2008) aponta que a sua ocorrência mostra produtividade em contextos nos quais clíticos são vetados, manifestando comportamento análogo a um pronome fraco, nos moldes de C&S. Barbosa (2005) destaca os seguintes pontos em favor dessa classificação: (i) que um clítico mostra dependência em relação ao acento de outra palavra, fato que se aplica ao pronome *se* enclítico, mas não à forma *cê* (cf. (36)); e (ii) que um clítico não pode ser substituído, o que não acontece com esse pronome (cf. (37-37')). Mesmo assim, a autora sinaliza que existem registros muito restritos de que esse elemento possa ocorrer em contextos em que formas fracas são impossíveis, como c-modificação e coordenação (cf. (38)), que são admitidas somente com pronomes de classe forte.

- (36) (a) Foi-se o culpado de tudo.
(b) Foi *cê* o culpado de tudo.

(37) Duas horas e *cê* só falou isso agora!

(37') Duas horas e tu só falou isso agora!

- (38) (a) De todos, *cê* só é o mais sem-vergonha.
(b) *Cê* e ele serão felizes.

(BARBOSA, 2005)

Versando acerca das constatações de Barbosa (2005), Perez (2006) demonstra que o emprego de *cê* nas condições descritas em (38) se mostra marcado – possivelmente, sugerindo um curso lento de mudança dessa forma em direção ao comportamento de pronomes como *você* e *ocê*, *i.e.*, formas fortes. Torna-se, então, plausível considerar *cê* enquanto um pronome que predominantemente

manifesta propriedades análogas à classe fraco, em acordo com a tripartição de C&S. A partir de tal asserção e de exemplares como os dispostos em (34b) e (35b), depreende-se que pronomes dessa categoria são lícitos como resumptivos em sentenças de SD do PB independentemente da posição ocupada pelo constituinte inicial.

Levar em conta que um pronome reduzido como *cê* seja fraco lega a esta análise a inevitável necessidade de esse item ocupar certa posição Spec na estrutura. Esse fato, por si só, respalda a viabilidade de ser rompida a adjacência entre esse item e o DP inicial de uma sentença de SD, como acontece em (35b). Diferentemente, a interveniência entre um constituinte sujeito e uma forma reduzida tal qual *ei* provoca má-formação – circunstância essa que revela que as formas reduzidas em foco não podem ser homogeneamente categorizadas, uma vez que manifestam comportamento incongruente entre si. Assim, de modo dissemelhante a *cê*, assume-se, com Quarezemin (2019, 2020), que *ei* seja uma forma clítica, situando-se, portanto, em uma posição nuclear (X⁰). Com a autora, esta análise defende que o núcleo ocupado por esse item atrai o DP inicial da sentença de SD à posição Spec correspondente, de forma que a contiguidade de ambos os elementos não possa ser rompida por quaisquer materiais, tais como sintagmas ou, mesmo, entonação parentética.

Em suma, esta seção argumentou em favor de que os PRs das sentenças de SD do PB manifestam comportamento dissonante, a depender de condições discursivas e da posição ocupada pelo DP inicial. Com base na tripartição da classe pronominal de C&S, averiguou-se que, quando esse constituinte ocupa uma posição no domínio CP, o PR é, necessariamente, ou forte, ou fraco; por outro lado, quando o DP inicial figura na zona argumental, o PR é impossibilitado de ser forte, devendo ser ou fraco, ou clítico. A seguir, o quadro 1 sintetiza tais ponderações:

Quadro 1: As classificações dos PRs das sentenças de SD do PB

	DP inicial em CP	DP inicial em IP
Pronome resumptivo forte	✓	*
Pronome resumptivo fraco	✓	✓
Pronome resumptivo clítico	*	✓

Fonte: Elaboração da autora

3. As posições estruturais dos pronomes resumptivos e as articulações das sentenças de sujeito duplo

O diagnóstico do comportamento heterogêneo manifestado pelos PRs das sentenças de SD do PB, além de fornecer informações importantes acerca das questões gramaticais circunscritas a tal construção, são favoráveis a contribuir a uma caracterização precisa de seus aspectos estruturais. A esse ponto, em específico, os direcionamentos fornecidos pela Cartografia Sintática podem desempenhar papel significativo: sob escopo do princípio cartográfico *one feature, one head* (cf. KAYNE, 2005), compreende-se que elementos que compartilham de um traço específico serão alocados em uma mesma categoria, que, por sua vez, é necessariamente distinta daquela que comportará um traço que

diverge em relação ao primeiro. Aplicadamente aos PRs, é esperado que as três classes admitidas nas sentenças de SD do PB, por apresentarem particularidades distintivas, sejam alocadas, também, em categorias particulares. Assumindo tal proposta, esta seção traça uma proposta às posições ocupadas por esses itens, bem como aos processos de formação das sentenças de SD a partir de tais.

Em primeiro plano, são considerados os PRs pertencentes à classe forte. Conforme delineou a seção 3, formas resumptivas desse grupo são possíveis tão somente em face de DPs iniciais que figurem à esquerda, *i.e.*, na posição Spec,TopP. Plausivelmente, encontra-se nesse fato um recurso para definir o ponto em que se situa esse item na estrutura: é esperado que, por os PRs fortes não poderem co-ocorrer com DPs iniciais que ocupam uma posição na zona argumental, esses itens estejam em distribuição complementar. Assumindo as propostas de Quarezemin (2019, 2020) e Krieck (2022), segundo as quais os constituintes iniciais que se situam no domínio IP ocupam Spec,SubjP, define-se que PRs fortes estejam situados nessa posição, sendo tais os argumentos s-selecionados pelo verbo na estrutura profunda (cf. (39)):

$$(39) \quad [_{\text{TopP}} \mathbf{DP}_{\text{INICIAL}} [_{\text{SubjP}} \mathbf{PR}_{\text{FORTEi}} [_{\text{TP}} \mathbf{t}_i [_{\text{VP}} \mathbf{t}_i \dots$$

Em se tratando de tais formas pronominais, torna-se impreterível a menção ao fato de que existe a possibilidade de esses elementos figurarem no domínio-A', como tópicos. Trata-se das restritas ocorrências de triplicação do sujeito, referidas no *corpus* sincrônico de Krieck (2022). A obrigatoriedade de o constituinte parentético ocupar uma posição à esquerda lega a inevitabilidade de os constituintes hierarquicamente mais altos se situarem, também, no domínio CP. Por consequência, PRs fortes – os únicos que admitem deslocamento a uma posição A', em acordo com C&S – podem ocupar a posição Spec,TopP (cf. (40), a partir de (23b)):

$$(40) \quad [_{\text{TopP}} \mathbf{o\ fast-food} [_{\text{TopP}} \mathbf{ele}_{\text{FORTE}} [_{\text{XP}} \mathbf{em\ algum\ momento} [_{\text{IP}} \mathbf{ele} \dots$$

Situação oposta cabe aos PRs clíticos, por suas vezes, que são admitidos em sentenças de SD tão somente caso o DP inicial figure em Spec,SubjP, *i.e.*, na zona argumental. Tais itens figuram em posição nuclear, atraindo o constituinte inicial do SD ao seu Spec. Nesse caso, esse DP e o PR se encontram em relação Spec-head na projeção SubjP. Assumindo a análise de Quarezemin (2019, 2020), considera-se que é o elemento inicial que cumpre a seleção argumental do verbo (cf. (41)):

$$(41) \quad [_{\text{SubjP}} \mathbf{DP}_{\text{INICIALi}} [_{\text{Subj}} \mathbf{PR}_{\text{CLÍTICO}} [_{\text{TP}} \mathbf{t}_i [_{\text{VP}} \mathbf{t}_i \dots$$

Uma atenção específica deve ser voltada aos pronomes fracos, elementos que podem funcionar como PRs independentemente do domínio sentencial em que se situa o constituinte inicial. Neste aspecto, associa-se à análise Cardinaletti (2004), segundo a qual pronomes fracos, que não dispõem da mesma autonomia em relação a pronomes fortes, são impossibilitados de subir a Spec,SubjP, devendo estacionar na posição hierarquicamente inferior Spec,TP. Transpondo a asserção da autora a

esta análise, assume-se que a situação se estenda a quando essa forma assume o *status* de resumptiva nas estruturas de SD (cf. (42)):

$$(42) \quad [_{\text{TopP}} [_{\text{SubjP}} [_{\text{TP}} \mathbf{PR}_{\text{FRACO}} [_{\text{VP}} \dots$$

Considerada a inviabilidade de os PRs fracos figurarem em SubjP, cabe ao DP inicial a satisfação do traço relativo a sujeito da predicação (cf. CARDINALETTI, 2004) nessa categoria. É esse elemento, então, o selecionado pelo verbo e que se move a Spec,SubjP; nessa posição, o constituinte se congela (cf. (43a)) ou se desloca ao domínio-A', atendendo a necessidades discursivas (cf. (43b)):

$$(43) \quad (a) \quad [_{\text{TopP}} [_{\text{SubjP}} \mathbf{DP}_{\text{INICIALi}} [_{\text{TP}} \mathbf{PR}_{\text{FRACO}} [_{\text{VP}} t_i \dots$$

$$(b) \quad [_{\text{TopP}} \mathbf{DP}_{\text{INICIALi}} [_{\text{SubjP}} t_i [_{\text{TP}} \mathbf{PR}_{\text{FRACO}} [_{\text{VP}} t_i \dots$$

Em síntese, esta etapa delineou uma proposta cartográfica aos PRs da construção de SD do PB, de acordo com as asserções assumidas na seção precedente. Advogou-se em função de que tais elementos, cujas classes dispõem de propriedades particulares e distintivas, assumem diferentes posições estruturais, a depender do grupo em que se incluem: (i) pronomes fortes, em Spec,SubjP ou, restritamente, em Spec,TopP; (ii) pronomes clíticos, em Subj°; e (iii) pronomes fracos, em Spec,TP. Foi, ainda, articulada a estruturação das sentenças de SD a partir das categorizações possíveis aos PRs.

Considerações finais

Este artigo trouxe contribuições à compreensão dos aspectos sintáticos circunscritos ao fenômeno de SD do PB a partir de seus PRs, traçando uma proposta de caráter sincrônico, que tomou, como alicerce, o *corpus* de Krieck (2022), formado por sentenças produzidas por falantes do PB entre os anos de 2018 e 2020. A análise realizada aponta para o fato de que tais elementos resumptivos manifestam comportamento heterogêneo, a depender, também, das propriedades estruturais correspondentes aos constituintes em posição inicial. As particularidades circunscritas a tais elementos permitem categorizá-los desta forma: (i) enquanto fortes ou fracos, caso o DP inicial ocupe posição no domínio periférico; e (ii) como fracos ou clíticos, se esse constituinte se situar na zona argumental.

Ainda, a análise desenvolvida seguiu direcionamentos de base cartográfica para advogar em função da tese de que, a depender das propriedades dos PRs que permitam classificá-los em acordo com a tripartição de C&S, esses elementos ocupam posições distintas na estrutura. Em primeiro plano, formas fortes ocupam Spec, SubjP, exercendo a função de sujeitos *default* da oração; sob circunstâncias restritas, tais elementos podem figurar em Spec, TopP, *i.e.*, como tópicos oracionais. PRs clíticos, por suas vezes, ocupam a posição nuclear Subj°, atraindo os constituintes iniciais da SD à posição Spec correspondente. Já as formas resumptivas fracos devem, necessariamente, figurar em Spec, TP. Por fim, este trabalho concebeu propostas de estruturação das sentenças de SD do PB, fundamentadas na análise desenvolvida ao longo do trabalho.

Referências

- BARBOSA, L. P. *Estatuto da forma cê: clítico ou palavra?*. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal de Minas Gerais. Belo Horizonte, 2005.
- BARBOSA, P; DUARTE, M. E. L; KATO, M. A. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, v. 4, n. 2, 2005.
- BRASIL. *Lei nº 9.610, de 19 de fevereiro de 1998*. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Brasília, 1998. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/19610.htm. Acesso em: 24 ago. 2021.
- BRITTO, H. Pronomes fracos nulos e lexicalizados: das línguas verdadeiramente *pro-drop* ao português do Brasil. *Cad. Est. Ling.*, Campinas, n. 34, jan./jun. 1998.
- CARDINALETTI, A. Toward a cartography of subject positions. In: RIZZI, L. (ed.). *The structure of CP and IP: the cartography of syntactic structures*. New York: Oxford University, v. 2, 2004.
- CARDINALETTI, A. Subjects and clause structure. In: HAEGEMAN, L. (ed.). *The new comparative Syntax*. London: Longman, 1997.
- CARDINALETTI, A; STARKE, M. The typology of structural deficiency on the three grammatical classes. *Working paper in Linguistics*. v. 4, n. 2. University of Venice, 1994.
- CINQUE, G. Adverbs and functional heads: a cross-linguistic perspective. *Oxford studies in comparative Syntax*. Oxford: Oxford University Press, 1999.
- CINQUE, G; RIZZI, L. The cartography of syntactic structures. In: BERND, H; HEIKO; N. (ed.). *The Oxford handbook of linguistic analysis*. Oxford: Oxford University Press, 2010.
- COELHO, M. S. V. *Uma abordagem variacionista do uso de formas de tratamento no norte de Minas*. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal de Minas Gerais. Belo Horizonte, 1999.
- COSTA, J; DUARTE, I; SILVA, C. R. T. Construções de redobro em português brasileiro: sujeitos tópicos vs. soletração do traço de pessoa. *Estudos em Sintaxe Comparativa*, n. 33, jan./jun. 2004.
- CYRINO, S. A “restrição do nome nu” e o português brasileiro. In: MARINS, J. E; ORSINI, M. T; CAVALCANTE, S. R. O. (org.). *Contribuições à descrição e ao ensino do português brasileiro: da fonética ao discurso, com parada obrigatória na sintaxe – uma homenagem a Maria Eugênia Lammoglia Duarte*. São Paulo: Pimenta Cultural, 2021.
- DUARTE, M. E. L. *Reestruturação nas construções existenciais e de alçamento*. Comunicação apresentada no Centro de Linguística da Universidade de Lisboa, 2003.
- DUARTE, M. E. L. Left-dislocated subjects and parametric change in Brazilian Portuguese. In: CARON, B. (org.). *Proceedings of the XVIth Congress of Linguists*. Pergamon: Paris, 1998.
- DUARTE, M. E. L. *A perda do princípio “evite pronome” no português brasileiro*. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. Campinas, 1995.

- GONÇALVES, F. M. R. *Riqueza morfológica e aquisição da Sintaxe em português europeu e brasileiro*. Tese (Doutoramento em Linguística Portuguesa) – Universidade de Évora. Évora, 2004.
- GROLLA, E. *A aquisição da periferia esquerda da sentença em português brasileiro*. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. Campinas, 2000.
- KATO, M. A. Recontando a história das relativas em uma perspectiva paramétrica. In: ROBERTS, I; KATO, M. (org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993.
- KATO, M. A; DUARTE, M. E. L. Restrições na distribuição de sujeitos nulos no português brasileiro. *Veredas*. v. 18, n. 1, 2014.
- KAYNE, R. S. Pronouns and their antecedents. In: KAYNE, R. S. *Movement and silence*. New York: Oxford University Press, 2005.
- KRIECK, L. E. *As sentenças com duplicação do sujeito no português brasileiro: uma análise cartográfica*. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina. Florianópolis, 2022.
- ORSINI, M. T. *As construções de tópico no português do Brasil: uma análise sintático-discursiva e prosódica*. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, 2003.
- ORSINI, M; VASCO, S. L. Português do Brasil: língua de tópico e de sujeito. *Revista Diadorim*, v. 2, 2007.
- PEREZ, E. P. *O uso de você, ocê e cê em Belo Horizonte: um estudo em tempo aparente e em tempo real*. Tese (Doutorado em Letras – Linguística) – Faculdade de Letras, Universidade Federal de Minas Gerais. Belo Horizonte, 2006.
- PETERSEN, C. A tripartição pronominal e o estatuto das proformas cê, ocê e você. *D.E.L.T.A.*, v. 24, n. 2, 2008.
- PIRES, A. The subject, it is here!: the varying structural positions of preverbal subjects. *D.E.L.T.A.*, v. 23, 2007.
- PONTES, E. *O tópico no português brasileiro*. Campinas: Pontes, 1987.
- QUAREZEMIN, S. Brazilian double subjects and the sentence structure. In: PIRES DE OLIVEIRA, R; EMMEL, I; QUAREZEMIN, S. *Brazilian Portuguese, Syntax and Semantics: 20 years of Núcleo de Estudos Gramaticais*. John Benjamins Publishing Company, 2020.
- QUAREZEMIN, S. Um novo olhar sobre as sentenças com redobro em português brasileiro. *Revista da Anpoll*, v. 1, n. 48, Florianópolis, jan./jun. 2019.
- RIZZI, L. Subjects, topics and the interpretation of *pro*. In: PETROSINO, R; CERRONE, P; HULST, H. (ed.). *From sounds to structures: beyond the veil of Maya* [Studies in Generative Grammar, n. 135]. Berlin: Mouton de Gruyter, 2018.
- RIZZI, L. Notes on labeling and subject positions. In: DI DOMENICO, E; HAMMAN, C; MATTEINI, S. (ed.). *Structures, strategies and beyond: studies in honour of Adriana Belletti*. Amsterdã: John Benjamins Publishing Company, 2015.
- RIZZI, L. On some properties of subjects and topics. In: BRUGÉ, L. et al. (ed.). *Proceedings of the XXX Incontro di Grammatica Generativa*. Venezia: Cafoscarina, 2005.

RIZZI, L. The fine structure of left periphery. In: Haegman, L. (ed.). *Elements of Grammar*. Dordrecht: Kluwer Academic Publisher, 1997.

RIZZI, L; BOCCI, G. Left periphery of the clause: primarily illustrated for Italian. In: EVERAERT, M; RIEMSDIJK, H. C. (ed.). *The Wiley Blackwell Companion to Syntax*. 2. ed. Hoboken: John Wiley & Sons, 2017.

RIZZI, L; SHLONSKY, U. Strategies of subject extraction. In: GÄRTNER, H. M; SAUERLAND, U. (ed.). *Interfaces + recursion = language?: Chomsky's minimalism and the view from syntax-semantics*. Berlin: Mouton de Gruyter, 2007.

SILVA, C. R. T; CARVALHO, D. S; ZIOBER, F. M. Licenciamento de duplos sujeitos em variedades do português: pessoa, definitude e estrutura de traços. *Letras escreve*, Macapá, v. 7, n. 2, 2017.

VASCO, S. L. *Construções de tópico na fala popular*. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2006.

CONTEXTS FOR NULL SUBJECTS IN CONTEMPORARY BRAZILIAN PORTUGUESE*CONTEXTOS PARA SUJEITOS NULOS EM PORTUGUÊS BRASILEIRO CONTEMPORÂNEO*Mônica Rigo Ayres¹Gabriel de Ávila Othero²**ABSTRACT**

Null subjects have been extensively investigated in natural languages. Brazilian Portuguese (BP), in particular, raises a lot of interest in the subject, for it is apparently in a process of grammatical change, from a *pro-drop* language to a *partial pro-drop language* (cf. BIBERAUER *et al.*, 2010; HOLMBERG *et al.*, 2009; DUARTE; MARINS, 2021; SOARES *et al.*, 2019). In this article, we investigate null subjects in BP assuming that null referential subjects are only used in ‘marked’, specific contexts, whereas overt subjects are the ‘unmarked’, more frequent strategy. We have analyzed a contemporary corpus of spoken BP and found 1,252 occurrences of null subjects. We then present four factors that are traditionally known for playing a role in determining the optimal contexts for null subjects in BP (and several other languages). We argue here that all these four factors are both necessary and sufficient to explain all null subject occurrences in any given corpus of contemporary spoken BP. We analyzed all the 1,252 occurrences of null subjects in the corpus and could explain 99.8% of the data, leading us to a better understanding of the null subject phenomenon in BP, at least when it comes to identifying its licensing contexts.

KEYWORDS: Null subject. Corpus analysis. Brazilian Portuguese.

RESUMO

Sujeitos nulos têm sido largamente investigados em línguas naturais. O português brasileiro (PB), em particular, desperta muito interesse pelo assunto, pois aparentemente o PB está em processo de mudança, de uma língua *pro-drop* para uma língua *pro-drop* parcial (cf. BIBERAUER *et al.*, 2010; HOLMBERG *et al.*, 2009; DUARTE; MARINS, 2021; SOARES *et al.*, 2019). Neste artigo, investigamos sujeitos nulos em PB assumindo que sujeitos referenciais nulos são usados apenas em contextos “marcados”, específicos, enquanto sujeitos expressos são a estratégia “não marcada”, mais frequente. Analisamos um *corpus* contemporâneo de PB falado e encontramos 1.252 ocorrências de sujeitos nulos. Em seguida, apresentamos quatro fatores que são tradicionalmente conhecidos por desempenhar algum papel na determinação dos contextos ótimos para sujeitos nulos em PB (e várias outras línguas). Argumentamos aqui que todos esses quatro fatores são necessários e suficientes para explicar todas as ocorrências de sujeito nulo em qualquer *corpus* de PB falado contemporâneo. Analisamos todas as 1.252 ocorrências de sujeitos nulos no *corpus* e conseguimos explicar 99,8% dos dados, o que nos leva a uma melhor compreensão do fenômeno do sujeito nulo no PB, pelo menos que tange à identificação de seus contextos de licenciamento.

PALAVRAS-CHAVE: Sujeito nulo. Análise de *corpus*. Português Brasileiro.

¹ Universidade Federal do Ceará, monica.ayres@ufrgs.br, <https://orcid.org/0000-0003-0972-5633>.

² Universidade Federal do Rio Grande do Sul, gabriel.othero@ufrgs.br, <https://orcid.org/0000-0002-2060-6312>.

Introduction³

Null subjects have already been extensively investigated in natural languages. Brazilian Portuguese (BP), in particular, raises a lot of interest in the subject, for it is apparently changing from a pro-drop language to a partial pro-drop language (BIBERAUER *et al.*, 2010; HOLMBERG *et al.*, 2009, DUARTE; MARINS, 2021; SOARES *et al.*, 2019). Here, we will investigate null subjects in BP assuming that null referential subjects are only used in specific contexts, whereas overt subjects are the ‘unmarked’, more frequent strategy.

According to Widera; Kaiser (2019, p. 144), “Portuguese is considered to be a null subject language. This means that, in this language, the subject’s omission represents the normal case and the use of the subject pronoun is limited to restricted contexts” – that is true for European Portuguese; in Brazilian Portuguese, we find the exact opposite scenario (DUARTE, 1995; DUARTE, 2012; AYRES, 2021), i.e. in BP the subject’s omission represents the ‘marked’, less common case, and it is restricted to some specific contexts – which we intend to cover here.

Literature on BP null subjects (and on null subjects in general, across the languages) usually points to four factors that favor null subjects: (i) rich verbal AGR morphology (DUARTE, 1993); (ii) semantic features of the referent DP [human, referential] (CYRINO; DUARTE; KATO, 2000) or [semantic gender] (OTHERO; SPINELLI, 2019b); (iii) *V1 linear order (KATO, 2020); and (iv) optimal discourse connection (or topic chain) (PAREDES SILVA, 2003).

Here, in our investigation, we intend to analyze a BP spoken language corpus to verify if the null subjects we encounter can be explained by one or more of these four factors. In other words, we aim to check if null subjects are licensed in contexts where

- i. verbal AGR morphology is rich enough to explicit the subject;
- ii. the referent of the null subject is a [–human, –referential] DP (or a DP with the feature [–semantic gender]);
- iii. the verb is not the first element in the clause;
- iv. the null subject occurs in an optimal connection context, i.e. it is a prominent element in a topic chain.

In the next sections, we will explain in detail each of these contexts and features and their relations to null subject occurrences in BP. After that, we will present the data we have investigated and analyze the occurrences according to these four factors.

1. Null subjects in Brazilian Portuguese

The investigation of null subjects in natural languages, in the generative framework, dates back,

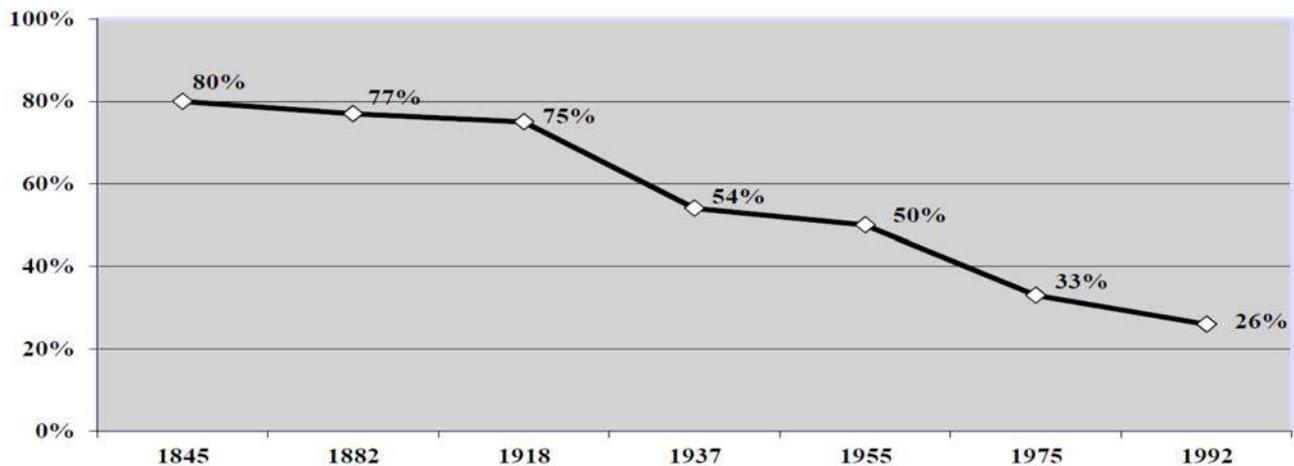
³ We would like to thank Aline Gravina, Camila Ulrich, Georg Kaiser, Karoline Gasque, Melissa Lazzari and Sergio Menuzzi for debating with us many of the ideas we present here. We would also like to thank Elisa Battisti, for sharing the LínguaPOA corpus with the academic community. Any inconsistencies in our text are of our own responsibility.

at least, to Perlmutter (1971), who pointed out the distinction between languages that allow null subjects and languages that do not. The topic received more attention after Chomsky (1981), when the pro-drop parameter was formalized. This formalization took place in the Principles and Parameters framework, which predicts that there are linguistic universals (principles), with parameters being the major responsible for differences among languages. For example, think about the Extended Projection Principle (EPP): this principle postulates that, in any natural language, every clause will have a subject. The value that the null subject parameter receives in each language will define whether or not this subject needs to be overt, i.e., morphologically and phonetically expressed. If the subject does not necessarily need to be expressed, we are facing a common phenomenon in natural languages: the null subject (1). If a language does not allow the omission of the subject, i.e., the phonetically expressed subject is mandatory, we say that language is a –pro-drop language (2).

- 1) (Io) Sono uno studente (Italian)
(Yo) Soy estudiante (Spanish)
- 2) *(I) am a student (English)
*(Je) suis un étudiant (French)

The picture, though, is not so clear as one might think, for there are languages that have some +pro-drop characteristics and –pro-drop as well. Brazilian Portuguese is such a language (cf. BIBERAUER *et al.*, 2010, DUARTE; FIGUEIREDO SILVA, 2016, HOLMBERG *et al.*, 2009). Tarallo (1983) was probably the first to investigate a possible change in the parameter setting in BP. Duarte (1993, 1995), observing this change, followed Tarallo’s suggestions and pointed out that BP goes through a transition phase from a +pro-drop to a –pro-drop language stating that “the increasingly less frequent occurrences of first person singular null subjects in oral language, still licensed by the verbal agreement morphology, should continue to manifest themselves indefinitely as residues of the pro-drop property” (DUARTE 1995, p. 56). Still in the mid-1990s, Monteiro (1994, p. 123), in a comprehensive investigation of pronouns in BP, questions if it is valid to consider BP as a null subject (+pro-drop) language, as it is the case for European Portuguese.

This question is well founded, as several studies have shown that the subject position/function is preferably overt in BP, as opposed to what used to happen in earlier periods of time – the change over time can be seen in the following graph:

Graph 1: Null subjects across time in Brazilian Portuguese

Source: DUARTE, 2018, p. 88

As we can see, null subjects went through a decrease: they reached 80% in the first period analyzed by Duarte (1845) and dropped to only 26% in the most recent period of Duarte (2018)' analyses. As Duarte (1995, p. 4) states: "although it cannot be said that we have lost the possibility of omitting the subject, there is a clear preference for using the full pronominal form".

When comparing BP to other +pro-drop Romance languages (European Portuguese, Spanish and Italian, cf. KATO, 1999), we see that the contexts in which null subjects occur are more restricted in BP than in these other languages. This is why BP has been classified by some as a partial pro-drop language (RODRIGUES, 2004; HOLMBERG, 2005; GRAVINA, 2014). According to Gravina (2014, p. 33), BP is a partial null subject language, since "it still licenses null subject, but with context restriction". Soares; Miller; Hemforth (2019, p. 3581) point to two different approaches:

Within the Generative tradition, there are two general approaches to so-called 'pro-drop' in BP, one diachronic, the other synchronic. These can be taken to be complementary. According to the diachronic approach, there is an ongoing change in the pro-drop parameter towards obligatory phonological realization of subject pronouns (TARALLO, 1983, among others); according to the synchronic approach, present-day BP is a partial pro-drop language (HOLMBERG *et al.*, 2009; BIBERAUER *et al.*, 2010, *inter alia*).

In this same text, a reservation is made regarding these two positions:

As Sérgio Menuzzi (p. c.) has pointed out, though these two positions do not exclude each other, they were not put forward as compatible in the literature and no specific framework within which they might form a single combined diachronic and synchronic hypothesis has been proposed. Specifically, researchers defending the 'partial pro-drop' analysis have mainly argued for this position against the idea that BP was becoming a non-null subject language of the English type. (SOARES; MILLER; HEMFORTH, 2019, p. 3581)

Partial null subject languages, though, do not behave exactly in the same way. Duarte (2020, p. 95) emphasizes that “the inclusion of BP among the partial null subject languages needs to be seen carefully. According to Biberauer (2010), this ‘partial’ label may contain very different systems, which require a rigorous analysis of the contexts that license null subjects.” Framing languages into categories is not a trivial task, for they do not behave in the same way in licensing null subjects.

In any case, we will not discuss this typology here. We agree to Gravina’s observation: BP still licenses null subject, but with context restriction. Our goal here is precisely to investigate which contexts are these. In other words, we will investigate the contexts that allow and favor null subjects in BP, showing that null subjects are less frequent and depend on these contexts to occur. In the next section, we start with verbal morphology and its relation to null subject in BP.

1.1. Verbal morphology

In BP, there are several cases where verbal agreement (AGR) morphology can mark the subject (1st, 2nd, 3rd persons, singular or plural). The relationship between verbal morphology and null subjects can be seen in the sentences below: in (3), verbal morphology marks the subject, i.e., the verb has an exclusive inflectional suffix for the grammatical person; in (4), the verb has syncretic morphology; as a result, the subject cannot be identified by AGR morphology alone in (4), only in (3):

3) Adoro pizza.
Love.1st.person.sg pizza
‘I love pizza’

4) Adora pizza.
Love. $\left. \begin{array}{l} 2^{\text{nd}}.\text{person.sg} \\ 3^{\text{rd}}.\text{person.sg} \\ 1^{\text{st}}.\text{person.pl} \end{array} \right\}$ pizza.

‘You\he\she\it\we love(s) pizza’.

In the past tense, the syncretism is even bigger than in the present, as we show below:

	Present tense:	Past tense:
Eu	ador-o	ador-ava
Você ⁴	ador-a	ador-ava
Ele\ela	ador-a	ador-ava
A gente ⁵	ador-a	ador-ava
Vocês	ador-am	ador-avam
Eles\elas	ador-am	ador-avam

⁴ We also find the pronoun *tu* as a 2nd person singular form. The verb, nevertheless, does not inflect any different in non-monitored speech (see table 1, below).

⁵ We also find the pronoun *nós* as a 1st person plural form. In the present form of the verb, the inflection is different (“*nós ador-ávamos*”); in the past tense, it remains the same (“*nós ador-avam*”) in non-monitored speech.

Until the work of Huang (1984), it seemed clear that the licensing of null subjects in a language would be related to verbal AGR morphology in the following sense: if a language has rich AGR morphology, null subjects could occur (as in Spanish, European Portuguese or Italian); if a language has ‘poor’ verbal AGR morphology, null subjects would not occur (as in English or French) – cf. also Jaeggli; Safir (1989) and Meillet (1918). However, Huang showed that there are languages, such as Chinese and Japanese, for example, that, despite not having rich AGR morphology, license null subjects. Thus, morphology proved to be incapable of explaining the null subject phenomenon, at least independently. As Duarte (1995, p. 2) points out, “the relationship between rich inflection and null subject is no longer exclusive in the processes of licensing and recovery of the null subject’s content.”

In BP, however, there is a relationship between the loss of verbal AGR morphology and the gradual loss of null subjects. In the table below (adapted from DUARTE, 1993, p. 109), we can see, through three periods of time, how the inflectional paradigm of BP has become impoverished:

Table 1: Verbal inflection in BP, present tense

	Paradigm 1	Paradigm 2	Paradigm 3
1st person singular	falo ⁶ (<i>speak+o</i>)	falo	falo
2nd person singular	falas (<i>speak+s</i>) fala (<i>speak+Ø</i>)	fala	fala
3rd person singular	fala (<i>speak+Ø</i>)	fala	fala
1st person plural	falamos (<i>speak+mos</i>)	falamos	(falamos) fala
2nd person plural	falais (<i>speak+is</i>)	falam	falam
3rd person plural	falam (<i>speak+m</i>)	falam	falam

Adapted from DUARTE, 1993, p. 109

In table 1, we see the present tense of the verb ‘falar’ (to speak) in the present tense. Paradigm 3 corresponds to current spoken BP. In this paradigm there are three main forms: 1st p. sing. (*falo*), 2nd p. sing./3rd p. sing./1st p. pl. (*fala*) and 2nd p. pl./3rd p. pl. (*falam*) – and also the less used 1st p. pl. (*falamos*), much less used than ‘fala’. BP’s verbal system is even more syncretic in the past tense (‘pretérito imperfeito’), as we have just shown.

Duarte (1993, p. 110) states that “with a paradigm so impoverished or weakened, nothing is more natural than to expect profound changes in the representation of the pronominal subject.” In view of this shift in the inflection paradigm (from paradigm 1 to 3), morphology is no longer enough a criterion to identify the subjects. Soares; Miller; Hemforth (2019) carried out a quantitative study regarding the effect of verbal agreement on the use of null and pronominal subjects, in relation to the first person in BP. In this study, the authors make a reanalysis of the corpus used by Duarte (1995). According to them,

⁶ The verbal form *falo* is actually the verb stem *fal-* (which is the infinitive form *falar* without the infinitive morpheme *-r* and the thematic vowel *-a-*) with the morpheme *-o* added to its end.

the crucial case for evaluating the effects of the impoverishment of verbal morphology on the choice between overt and null subjects in present-day BP is the first-person singular inflectional marking system. Contrary to second-person markers, which systematically converge with the third person, leading to the precipitous increase in the number of overt subjects [...], first-person subjects differ according to Tense Type [...]. Some tenses have an exclusive marking for first-person singular, others are ambiguous. If verbal inflectional marking is a significant factor in the choice between an overt or a null subject, the Tense Type of the verb should be a significant factor in the choice of null vs. overt subjects for the first-person singular. But it should not interact with the third discourse person singular and plural, since the forms for these are systematically ambiguous with those of the second discourse singular and plural, irrespective of the Tense Type of the verb. The results obtained in this study partially support this prediction (SOARES; MILLER; HEMFORTH, 2019, p. 3586).

Verbal morphology plays a role favoring null subjects in BP, but that is certainly not the only grammatical factor involved in the phenomenon of null subjects. In the next section, we discuss semantic features of the referents of anaphoric subjects and their relation to null and pronominal subjects.

1.2. Semantic features

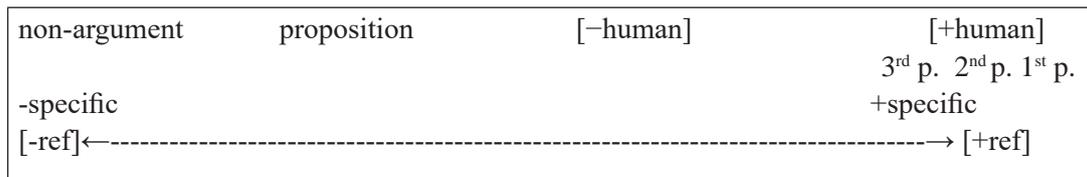
1.2.1 [\pm human] and [\pm specific]

Instigated by the work of Tarallo (1983) – which presented the decrease in the use of null subjects and the increase in the use of pronominal direct objects in BP –, Cyrino; Duarte; Kato (2000) investigated the two phenomena and sought a unified explanation for them. According to them,

If null subjects are licensed by ‘rich’ morphology (TARALDSEN 1978, JAEGGLI; SAFIR 1989, ROBERTS 1993), and if BP’s agreement inflection has been undergoing a change in the direction of a poorer morphology, the decrease of null subjects is justified. But, it is a mystery why the null object should have undergone the opposite trend over time.

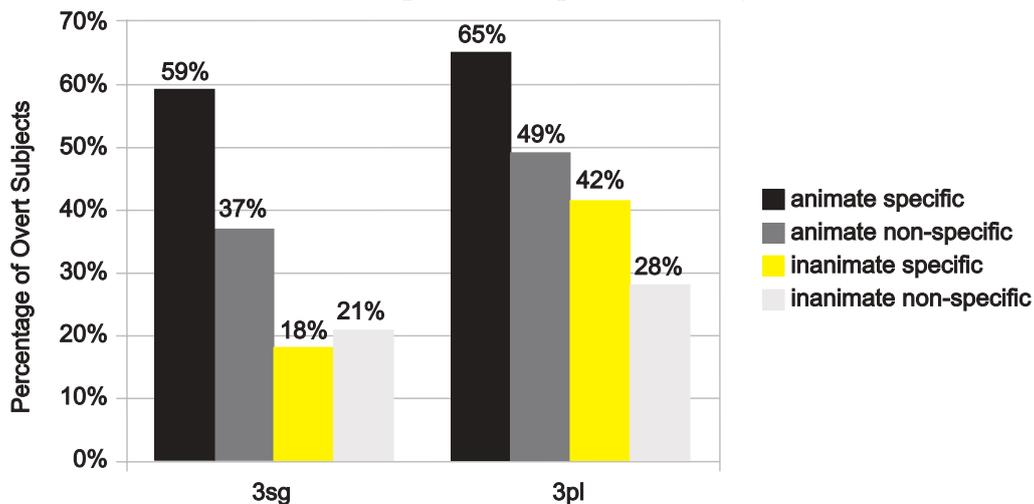
If, on the other hand, null subjects and objects in BP are like those in Chinese (see HUANG, 1984), in that they are not identified by inflection, but rather ‘controlled’ by a co-referent [...], one might expect that, for functional reasons, one or the other, or both, could be null. (CYRINO; DUARTE; KATO, 2000, p. 55).

Cyrino; Duarte; Kato (2000) proposed a referential hierarchy, in which elements at the highest point of the scale would favor pronominal forms, while elements at the lowest point of the scale would favor null forms. The authors present this hierarchy based on the idea that “cross-linguistically, referentiality is highly relevant for pronominalization” (CYRINO; DUARTE; KATO, 2000, p. 58).

Graph 2: Referentiality Hierarchy

Source: adapted from CYRINO; DUARTE; KATO, 2000, p. 59

The farther to the right, the higher the point in the hierarchy. Elements that have the [+human] feature stay at the top, and, therefore, tend to be referred to by pronominal forms. At the opposite end, the referents are preferably referred to by null elements. Cyrino, Duarte; Kato (2000, p. 55) state that “facts found in two independent studies working specifically with the diachronic distribution of null and full pronouns in subject position (Duarte 1993, 1995) and in object position (CYRINO, 1993, 1994[7])” can sustain this analysis – to which we agree. We find difficult, though, to present generalizations based on the analysis considering only the features [\pm human] and [\pm specific]. Soares (2017), for example, analyzes the corpus used by Duarte (1995, the NURC-RJ) in order to investigate the role of [animacy] and [specificity] in null subjects. He then presents the following results:

Graph 3: 3rd person overt pronominal subjects

Source: SOARES, 2017, p. 53

Soares states that

the features Animacy and Specificity take over the purported role of the impoverishment of verbal morphology in BP and explain straightforwardly the imbalanced number of null subjects across discourse persons. However [...] **these two features are not sufficient on their own**” (our emphasis added) (SOARES, 2017, p. 12).

Investigating the same corpus, Soares; Miller; Hemforth, (2020, p. 18) state that

there is a clear tendency: animate and specific antecedents are mostly retrieved by overt subjects while inanimate and non-specific antecedents are preferably recovered by null subjects. The third person singular subject is more frequently overt when its antecedent is animate and specific, followed by the animate non-specific antecedent, and inanimate specific and non-specific antecedents are more or less at the same level; in the third person plural, animate specific overt subjects are close in frequency to other discourse persons, followed by animate non-specific subjects, then by inanimate specific subjects with slightly lower frequency and, at the bottom of the scale, inanimate non-specific subjects.

In other words, we can again verify trends and tendencies, but not definitive generalizations to explain the null subject phenomenon in BP.

1.2.2 Semantic gender of the referent

In addition to Cyrino; Duarte; Kato (2000), other researchers have been working with a different, but related, semantic feature: the semantic gender of the referent. Spinelli (2018), Otero; Spinelli (2019a, b) and Otero; Goldnadel (2020) investigated the relationship of the semantic gender of referents with null subjects in BP. The authors base their hypothesis on Creus; Menuzzi (2004), who demonstrate that natural gender agreement is responsible for 3rd person anaphoric direct objects in BP. BP has a masculine pronoun that refers to male beings and referents ('ele', he), a feminine pronoun that refers to female beings ('ela', she) and an empty category (possibly, little pro) that refers to neutral beings and objects (\emptyset , it), a parallel to the English pronominal system, for example. This idea was then, generalized to cover both null object and null subject phenomena in BP.

Semantic gender is not equivalent to grammatical gender, although the two may overlap. In the words of Ayres (2016, p. 22)

The semantic gender feature differs from the grammatical gender feature: the first is "natural", it is about the semantic classification of beings denoted by nouns, which can have natural sex identifiable and recognizable by the speaker; the second refers to the morphosyntactic classification of nouns, which will determine their grammatical agreement relationships. Therefore, it can be said that all nouns in Portuguese have grammatical gender (...). Among the nouns, some will have an inherent semantic gender and others will not. To exemplify: among the nouns *girl*, *cow*, *victim*, and *bottle*, only the first two have inherent semantic gender, but all four have grammatical gender [in Portuguese, all these four nouns are feminine: *a menina*, *a vaca*, *a vítima* and *a garrafa*]. That is, although they are all feminine (grammatical gender), only the first two belong to the class of nouns with the feature [+semantic gender]; the other two nouns ("victim" and "bottle") are marked as [-semantic gender] nouns.

Creus; Menuzzi (2004, p. 16), as we said, propose the semantic gender hypothesis to explain 3rd person anaphoric direct objects, based on the premise that the basic difference between the pronominal forms 'ele/ela' (he/she) and the empty category is that pronouns carry gender features, whereas null forms are not gender-specified. Thus, the choice for an overt pronoun or an empty category, in BP, would be the result of, in the words of the authors (2004, p. 7), a "process of agreement between the antecedent and the anaphoric form." Antecedents with semantic gender would then favor the

use of overt pronouns, since pronouns have gender specification. In contrast, antecedents without semantic gender would favor null objects, since null forms do not have specification for gender. This hypothesis has been pursued with some success (cf. OTHERO; LAZZARI, 2022, for a survey of the latest advances), but it does not have full empirical coverage.

1.3. Avoid V1

Berlinck (1989, 1995), based on diachronic analysis, showed that, in past synchronies, BP had a high use of constructions of the V1 type (VSO, more specifically). That changed over time (cf. also MARTINS; CAVALCANTE; COELHO, 2020), leading to the SVO order in contemporary BP.

In European Portuguese, a null subject language, the first element of the sentence is often a verb (cf. GUILHERME, 2016). In BP, however, there is a great tendency to filling in the initial position of the clause, which is usually the subject position. According to Guilherme (2016, p. 95),

European Portuguese, being a language that maintains characteristics of a null subject language, exhibits a high occurrence of V1 clauses. BP, in turn, seems to have characteristics of a V2 language, as it presents a finite verb in the second position of the sentence, preceded by only one constituent.

Kato; Duarte (2018, pp. 621-2), following Kato (2000)'s suggestion, also point to the fact that

BP rejects the verb in sentence-initial position, filling it with an adjunct or a discursive element. The phenomenon is analyzed as distinct from the syntactic V2 structures, as any category can fill this position: a syntactic adjunct (whether a head or an XP) or a discursive element. Kato's hypothesis is that the constraint here is rhythmic or prosodic.

We will argue here that, in BP, null subjects are allowed when some other constituent occupies the position in the leftmost part of the clause, 'performing the function' of phonetically filling the left periphery of the verb, be it an adjunct or a discourse marker (e.g. *Ontem Ø vi a Maria* [Yesterday (I) saw Maria] should be better/more frequent than *Ø Vi a Maria ontem* [(I) saw Maria yesterday]).

1.4. Discursive Connection

In addition to the generative tradition, there are functionalist studies on null subject in BP. Paredes Silva (1988, 2003), for example, investigated the pronominal subject in BP, bringing together functionalism and Labovian sociolinguistics. From this perspective, she demonstrated that the variation we find between null subjects and overt pronouns has discursive motivations, specially related to topic chain and the 'activation' of the referent in discourse. This idea dates at least to Givón (1979, 1983) and Ariel (1988), i.e., the idea of 'predictability of information': the more predictable a referent (a subject, for example), lesser the need to make it (morphologically\phonetically) explicit. According to Paredes Silva (1998, p. 125),

our conception of the functional use of the pronoun also incorporates Givón's principle of iconicity (...). When applied to the choice between full phrases, strong or weak pronouns and zero anaphora for encoding a third-person referent, the principle (...) predicts that the more predictable an information, the less encoding it receives.

She also states that “the choice of the pronoun is strongly correlated with the non-maintenance of the same referent as a subject” (PAREDES SILVA, 2003, p. 104) – cf. also Bentivoglio (1980) and Lira (1982) on the relation of topic continuity and null subjects in Spanish and Portuguese, respectively.

From a generative point of view, Gravina (2008) has already pointed out that the interpretation of null subjects in BP could depend on an antecedent expressed in the discursive context. According to her, “the empty category in the subject position could have a different nature, with its interpretation guaranteed not by inflection, but by an antecedent expressed in the syntactic, discursive or pragmatic context” (GRAVINA, 2008, p. 21), following Huang (1984)'s suggestion for Chinese.

In the next section, we will show how all these four factors interact in analyzing the null subject phenomenon in BP. We will argue that, together, they can provide an explanation that fully accounts for the phenomenon of the null subject in this language.

2. Data Analysis

Here we present our results after careful corpus examination. All the data refers to our analysis of the oral language corpus *LínguaPOA* (a corpus composed of sociolinguistic interviews with informants from the city of Porto Alegre, Brazil, cf. Battisti 2019 for information about the corpus and its informants).

2.1 Total data analysis

After analyzing all the interviews, we found 4,136 occurrences of null and pronominal subjects: 2,884 were pronominals and 1,252 were null subjects, as we show in graph 4.

Graph 4: Null and pronominal subjects in the corpus

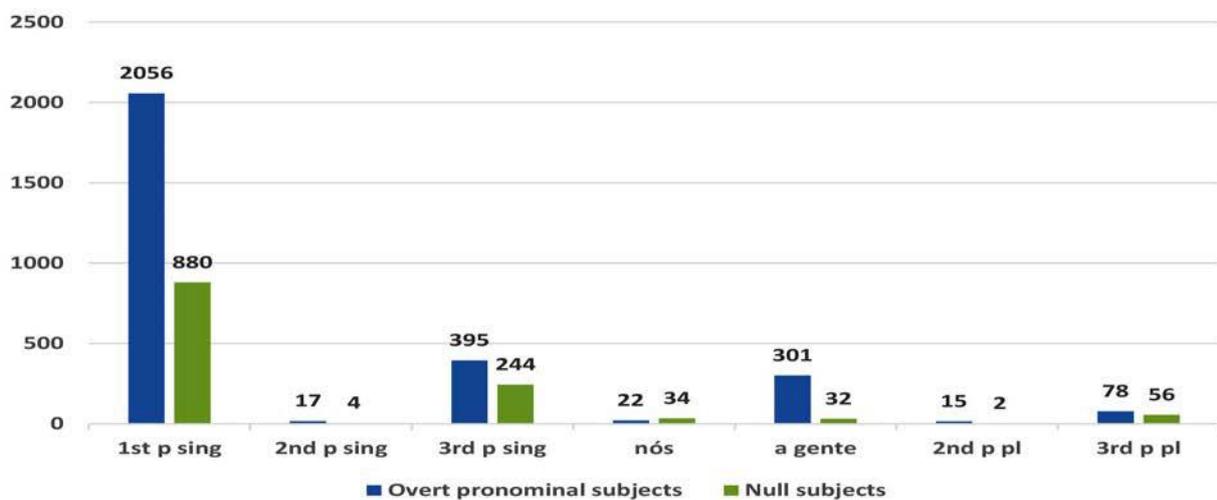


Source: the authors

Null subjects, which, as mentioned above, constituted the majority of subjects in previous periods of BP, are now only residual, 30% of the occurrences. This distribution (~70/~30) has been long found in the literature, at least since the 1990s (cf. BERLINCK; DUARTE; OLIVEIRA, 2015 for a review of the 1990s data). Even though this might lead to think that we are facing a stable phenomenon, as some may suggest (cf. AYRES, 2021, for example), analysis of this same corpus has shown clear indicatives of change in apparent time (cf. OTHERO; LAZZARI, 2022).

With regard to the distribution of null and pronominal subjects and grammatical persons, the vast majority of occurrences were with the 1st person singular, both for null subjects and for pronominal subjects, as shown in the following graph:

Graph 5: Null and pronominal subjects according to grammatical person



Source: the authors

These data are interesting, since they already give a spoiler of the role of morphology: the exponentiation of the subject in verbal morphology seems to be interacting with the null subject phenomenon. For example, notice that BP has two pronouns to express the notion of first person plural: ‘nós’ and ‘a gente’. When we compare these two pronouns, we find more null subjects precisely with ‘nós’, since in this case the verb exhibits exclusive first person plural morphology (60.8%). On the other hand, the pronoun ‘a gente’, which does not have AGR morphology, is almost always realized as a pronoun (90.3%).

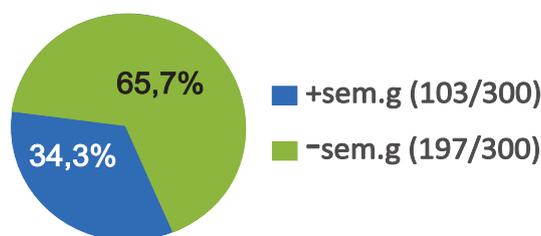
2.2 Analysis of null subject data

As we have seen, the literature has been reporting that overt subjects are the default case in the Brazilian Portuguese. Null subjects are a ‘marked’, less frequent strategy. After investigating the four factors presented in the previous sections, we analyzed all occurrences of null subjects in the corpus to check if they can all be explained – or justified – by one or more of these factors that are known for favoring null subjects.

Our main hypothesis here is that there is no ‘one’ licensing context or factor for null subjects in BP. Instead there are four: **semantic gender of the referent** (referents [-gs] favor null subjects), **topic-chain** or discourse saliency (salient referents in discourse tend to be null subjects), **linear V2 pattern** (null subjects are not favored in V1 clauses) and **verbal morphology** (overt AGR morphology favors null subjects). If this hypothesis is right, these four factors should be necessary and sufficient to ‘explain’ all null subject occurrences in any given corpus of contemporary spoken BP.

Let’s start with the first factor mentioned in the previous paragraph: the semantic gender. That applies only to 3rd personal subjects, since these are the pronouns that have morphological forms for masculine (*ele\eles*, he\they.masc) and feminine (*ela\elas*, she\they.fem). They compete with the null, underspecified form for gender (\emptyset). We found 300 occurrences of 3rd person null subjects, 65.7% of them were coreferent to [-sem.g] referents. That means the other 34.3% occurrences were ‘unexpected’, considering only the semantic gender hypothesis.

Graph 6: Null subjects and their referents [\pm sem.g]



Source: the authors

Examples:

- 5) *A internet mesmo eu uso praticamente vinte e quatro horas por dia... [O meu celular] tá sempre conectado, quando eu tô em casa, \emptyset tá sempre no wifi, quando eu tô na rua, \emptyset tá sempre nos dados.* (LínguaPOA – Interview 06)⁷
- 6) *Eu vejo bastante jovens trabalhando desde/ de cedo, eu tenho [uma prima] que tá com dezesseis anos e \emptyset tá fazendo o jovem aprendiz.* (LínguaPOA – Interview 92)⁸

In (5), we see two null subjects referring to [O meu celular] (‘my cellphone’), as expected, considering only the semantic gender factor. In (6), the null subject refers to [uma prima] (‘a female cousin’), a referent with semantic (apparent) gender. In this case, the null subject is not licensed by the semantic gender of the referent; it is licensed, though, by the discursive context: it appears in a coordinate clause (after *e*, ‘and’), a universal context favoring null subjects.

⁷ Roughly translated as “The internet itself, I use it practically 24h per day... [My cellphone] is always connected when I’m home, it is always with the wi-fi on when I’m on the street, it is always in 4G mode.”

⁸ Roughly translated as “I see lots of young people working sin/since an early age, I have [a cousin] who is 16 and she is in a young trainee program.”

The second factor, referent saliency in discourse (i.e. referents participating in a topic chain), could not cover all the cases of null subjects either: 54.3% of the null subjects were in a context of topic chain, but 45.7% were not. Here we analyzed all the 1,252 occurrences of null subjects in the corpus.

Graph 7: Null subjects and discourse saliency

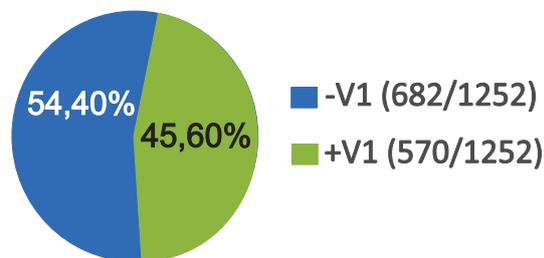


Source: the authors

The third factor, Avoid V1 linear order, could not explain all the occurrences of null subjects we found as well: from the 1,252 cases of null subjects, only 682 (54.4%) did not occur in V1 clauses. That is to say that examples like (7) were relatively frequent:

- (7) *E tu gosta de cozinha(r)?*
 Ø *Gosto muito.* (LínguaPOA – Interview 30)⁹

Graph 8: Null subjects and the *V1 effect



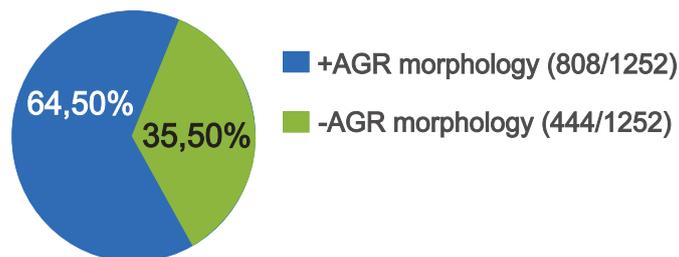
Source: the authors

Finally, when we analyzed the fourth factor, verbal inflectional AGR morphology, we found a slightly higher influence on null subjects: 64.5% of the null subjects found in the corpus occurred with verbs that have distinctive AGR morphology.

- (8) *E tu gosta de cozinha(r)?*
 Ø *Gosto muito.* (LínguaPOA – Interview 30)

The verb ‘gosto’ (I like) presents the ending -o (*gost-o*), marking 1st person singular exclusively.

⁹ Roughly translated as: “And you, do you like cookin’? I like it a lot.”

Graph 9: Null subjects and verbal AGR morphology

Source: the authors

In sum, we can see that none of the factors, in isolation, can be *the* factor licensing null subjects in BP. In the next section, we will show an analysis considering combinations of two factors, in order to check whether these combinations are able to explain all occurrences of null subjects in the corpus.

2.3. Two factors combined

Let's analyze all the possible combinations of the factors, grouping them together, in pairs, namely:

- i) semantic gender and discourse saliency (sem.g & ds, cf. table 2);
- ii) semantic gender and *V1 (sem.g & *V1, see table 3);
- iii) semantic gender and verbal morphology (sem.g & morph, see table 4);
- iv) discourse saliency and *V1 (ds & *V1, cf. table 5);
- v) discourse saliency and verbal morphology (ds & morph, see table 6);
- vi) *V1 linear pattern and morphology (*V1 & morph, cf. table 7).

In order to make it clearer for the reader, we will use the symbols ♡ when the context favors null subject and ☹ when it does not favor null subject. In the table below, we present the combination of semantic gender and discourse saliency:

Table 2: Semantic gender and discourse saliency of the referent

Semantic gender	Discourse saliency	Null subjects
[-sem.g] ♡	[+ds] ♡	93/300 (31%)
[-sem.g] ♡	[-ds] ☹	104/300 (34,7%)
[+sem.g] ☹	[+ds] ♡	72/300 (24%)
One or two favoring factors licensing the null subject		269/300 (89.7%)
[+sem.g] ☹	[-ds] ☹	31/300 (10,3%)

Source: the authors

With the combination that has the two factors that do not favor the null subject (the combination [+sem.g] and [-ds]), we would expect not to find any null subjects. However, as we can see in table 2, there are 31 occurrences (10.3%) of null subjects in this context; the other 269 occurrences (89.7%) appeared in a context with at least one licensing factor (either -sem.g or +ds, or both).

At this point, we can already see that with just these two factors it is not possible to explain the data of null subjects found, nor to make predictions about the data of null subjects that might be found in BP.

The second pair of factors we have analyzed is the semantic gender with linear order. The distribution of null subjects according to these traits is presented in the table below:

Table 3: Semantic gender of the referent and & V1 linear order.

Semantic gender	linear order	Null subjects
[-sem.g] 	[-V1] 	118/300 (39,3%)
[-sem.g] 	[+V1] 	79/300 (26,3%)
[+sem.g] 	[-V1] 	58/300 (19,3%)
One or two favoring factors licensing the null subject		255/300 (85%)
[+sem.g] 	[+V1] 	45/300 (15%)

Source: the authors

Again, we should expect that ([+sem.g]) referents and clauses where the verb is in the initial position (+V1) would not allow null subjects. However, we found 45 occurrences of null subjects in this scenario. Once again, it is not possible to explain the occurrences of null subjects found in the corpus.

The third possible combination of factors is the combination of semantic gender and verbal AGR morphology. The results found are in table 4:

Table 4: Semantic gender of the referent and AGR morphology

Semantic gender	AGR morphology	Null subjects
[-sem.g] 	[+morph] 	0/300 (0%)
[-sem.g] 	[-morph] 	197/300 (65.7%)
[+sem.g] 	[+morph] 	0/300 (0%)
One or two favoring factors licensing the null subject		197/300 (65.7%)
[+sem.g] 	[-morph] 	103/300 (34.3%)

Source: the authors

Here we cannot find any occurrence in a [+morph] context, since we are analyzing only 3rd person subjects and the verbs in 3rd person always present poor morphology (cf. discussion in section 2.1). Most occurrences of null subjects are in a context where they refer to [-gs] referents (as we have seen in 3.2). Nonetheless, we have found more than a third of the null subjects in the ‘unexpected’ context of [+sem.g] and [-morph].

The data presented in the tables above contained only occurrences of 3rd person null subjects, since we were considering the semantic gender feature. In the next tables, we will consider the combination of the other factors, and they are relevant for all three discourse persons.

In table 5, we present the combination of discourse saliency and V1 linear pattern.

Table 5: Discourse saliency of the referent and V1 linear order

Discourse saliency	V1 linear order	Null subjects
[+ds] 	[-V1] 	300/1,252 (23.9%)
[+ds] 	[+V1] 	380/1,252 (30.3%)
[-ds] 	[-V1] 	382/1,252 (30.5%)
One or two favoring factors licensing the null subject		1,062/1,525 (84,9%)
[-ds] 	[+V1] 	190/1,252 (15.1%)

Source: the authors

We still find a great number of null subjects in ‘unexpected’ scenarios, in this case, in [-ds] and [+V1] contexts. That clearly shows that these two factors are far from explaining the contexts where null subjects can occur in BP.

The penultimate combination of two factors is the one that brings together discourse saliency and morphology.

Table 6: Discourse saliency of the referent and verbal AGR morphology

Discourse saliency	AGR morphology	Null subjects
[+ds] 	[+morph] 	426/1,252 (34%)
[+ds] 	[-morph] 	254/1,252 (20.2%)
[-ds] 	[+morph] 	382/1,252 (30.5%)
One or two favoring factors licensing the null subject		1,062/1,525 (84,9%)
[-ds] 	[-morph] 	190/1,252 (15.1%)

Source: the authors

The majority of occurrences were in a context licensed by both factors, i.e., the null subject was in a context where it could be retrieved both by verbal morphology and discourse saliency, as we show in the example below:

- (9) *E tu trabalha? Estuda? Que tu faz?*
 Ø *Trabalho aqui, ahm, como auxiliar administrativo* (LínguaPOA – Interview 60)¹⁰.

Nevertheless, 15.1% of the cases could not be explained by neither of these factors, i.e. null subjects that occurred in [-ds] and [-morph] contexts.

Finally, the last combination of two factors is the one that unites linear V1 order and verbal AGR morphology.

¹⁰ Roughly translated as: “And you, do you work? Study? What do you do? I work here, hum, as an administrative assistant.”

Table 7: V1 and verbal AGR morphology

V1 linear order	AGR morphology	Null subjects
[-V1] 	[+morph] 	415/1,252 (33,1%)
[-V1] 	[-morph] 	267/1,252 (21,3%)
[+V1] 	[+morph] 	393/1,252 (31,3%)
[+V1] 	[-morph] 	177/1,252 (14,1%)

Source: the authors

Again, the majority of null subjects appeared in a context licensed by both factors, i.e. the null subject was in a context where it could be retrieved by verbal morphology and was not in a clause where the verb was the first and leftmost element of the clause. Nonetheless, that was, again, just roughly a third of the cases found. The combination of two factors is clearly not enough to cover all null subjects in the corpus.

Our next step was, then, to check whether three factors could account for all null subjects occurrences, as we will see in the next section.

2.4. Three Null Subject Licensing Factors

In this section we present sets of three factors and see if these sets can explain all occurrences of null subjects in the corpus. The combinations of factors are (i) semantic gender, discourse saliency and V1 linear order (table 8); (ii) semantic gender, V1 linear order and AGR morphology (table 9); (iii) semantic gender, discourse saliency and AGR morphology (table 10); and (iv) AGR morphology, V1 linear order and discourse saliency (table 11).

Table 8: Semantic gender of the referent, discourse saliency and V1 linear order

Semantic Gender	Discourse Saliency	V1 Linear Order	Null Subjects
[+sem.g] 	[+ds] 	[-V1] 	32/300 (10,6%)
[+sem.g] 	[+ds] 	[+V1] 	40/300 (13,3%)
[+sem.g] 	[-ds] 	[-V1] 	26/300 (8,6%)
[+sem.g] 	[-ds] 	[+V1] 	5/300 (1,6%)
[-sem.g] 	[+ds] 	[-V1] 	49/300 (16,3%)
[-sem.g] 	[+ds] 	[+V1] 	44/300 (14,6%)
[-sem.g] 	[-ds] 	[-V1] 	69/300 (23%)
[-sem.g] 	[-ds] 	[+V1] 	35/300 (11,6%)

Source: the authors

Once again, we could not cover all cases of null subjects, considering these three licensing factors. We found null subjects (only 5 occurrences though) in the unexpected scenario of [+sem.g], [-ds] and [+V1]. That means we could understand and ‘explain’ the other 295 occurrences of null subjects, but bear in mind we are only analyzing 3rd person subjects here.

Next we analyze semantic gender, V1 linear order and AGR morphology.

Table 9: Semantic gender of the referent, V1 linear order and AGR morphology

Semantic Gender	V1 linear order	AGR morphology	Null subjects
[+sem.g] 🗑️	[-V1] 👉	[+morph] 👉	0/300 (0%)
[+sem.g] 🗑️	[-V1] 👉	[-morph] 🗑️	58/300 (19,3%)
[+sem.g] 🗑️	[+V1] 🗑️	[+morph] 👉	0/300 (0%)
[+sem.g] 🗑️	[+V1] 🗑️	[-morph] 🗑️	45/300 (15%)
[-sem.g] 👉	[-V1] 👉	[+morph] 👉	0/300 (0%)
[-sem.g] 👉	[-V1] 👉	[-morph] 🗑️	118/300 (39,3%)
[-sem.g] 👉	[+V1] 🗑️	[+morph] 👉	0/300 (0%)
[-sem.g] 👉	[+V1] 🗑️	[-morph] 🗑️	79/300 (26,3)

Source: the authors

Again, we could not explain the phenomenon entirely, since we found 45 occurrences of null subjects in a very disfavoring context: [+sem.g] referents, [+V1] linear order and verbal forms with poor morphology [-morph]. Comparing Tables 9 and 8, it is clear the discourse saliency is a factor to be considered.

The penultimate combination concerns the semantic gender of the referent, the discourse saliency of the referent and AGR morphology. The occurrences of null subjects were distributed as follows:

Table 10: Semantic gender of the referent, discourse saliency and AGR morphology

Semantic gender	Discourse saliency	AGR Morphology	Null subjects
[+sem.g] 🗑️	[+ds] 👉	[+morph] 👉	0/300 (0%)
[+sem.g] 🗑️	[+ds] 👉	[-morph] 🗑️	72/300 (24%)
[+sem.g] 🗑️	[-ds] 🗑️	[+morph] 👉	0/300 (0%)
[+sem.g] 🗑️	[-ds] 🗑️	[-morph] 🗑️	31/300 (10,3%)

[-sem.g] 🇧🇷	[+ds] 🇧🇷	[+morph] 🇧🇷	0/300 (0%)
[-sem.g] 🇧🇷	[+ds] 🇧🇷	[-morph] 🇵🇹	93/300 (31%)
[-sem.g] 🇧🇷	[-ds] 🇵🇹	[+morph] 🇧🇷	0/300 (0%)
[-sem.g] 🇧🇷	[-ds] 🇵🇹	[-morph] 🇵🇹	104/300 (34,6%)

Source: the authors

Here we found 31 occurrences (10,3%) of null subjects appearing in the unexpected context of [+sem.g], [-ds] and [-morph]. So we still could not explain all the occurrences of null subjects in the corpus.

The last combination of three factors involves discourse saliency, V1 linear order and AGR morphology. Here, we analyzed all occurrences of null subjects in the corpus.

Table 11: Discourse saliency, V1 pattern and AGR morphology

Discourse saliency	V1 linear order	AGR Morphology	Null subjects
[+ds] 🇧🇷	[-V1] 🇧🇷	[+morph] 🇧🇷	174/1,252 (13,8%)
[+ds] 🇧🇷	[-V1] 🇧🇷	[-morph] 🇵🇹	126/1,252 (10%)
[+ds] 🇧🇷	[+V1] 🇵🇹	[+morph] 🇧🇷	252/1,252 (20,1%)
[+ds] 🇧🇷	[+V1] 🇵🇹	[-morph] 🇵🇹	128/1,252 (10,2%)
[-ds] 🇵🇹	[-V1] 🇧🇷	[+morph] 🇧🇷	241/1,252 (19,2%)
[-ds] 🇵🇹	[-V1] 🇧🇷	[-morph] 🇵🇹	141/1,252 (11,2%)
[-ds] 🇵🇹	[+V1] 🇵🇹	[+morph] 🇧🇷	141/1,252 (11,2%)
[-ds] 🇵🇹	[+V1] 🇵🇹	[-morph] 🇵🇹	49/1,252 (3,9%)

Source: the authors

We are certainly narrowing down the cases of null subjects. If we assume that three of the four factors commonly assumed by the literature are acting together to license null subjects in BP, we can already explain lots of cases we found in the corpus. That's pretty good, but we are not telling the whole story yet. That is why we advocate for a broader analysis: in the next section we will demonstrate that all four factors are needed to fully understand the null subject phenomenon in BP.

2.4.1. Four Null Subject Licensing Factors

As we have seen so far, taking in consideration only one licensing factor in isolation did not lead to any generalization. As we increased the number of factors to two or three in the analyzed sets, we found better results when analyzing the occurrences of null subjects in the corpus. However, even taking into account three factors, we still found 49 cases of null subjects that were unexpected (see

table 11). Since we aim to reach an explanation of the totality of the null subjects in the corpus (i.e., the 1,252 occurrences found), we will analyze the data using the four factors presented here. The combinations of four null subject licensing factors that occurred were shown in the next table¹¹.

Table 12: Semantic gender of the referent, discourse salience, V1 linear order and AGR morphology

Semantic gender	Discourse salience	V1 linear order	AGR Morphology	Null subjects
[+sem.g] 👉	[+ds] 👉	[-V1] 👉	[+morph] 👉	👉 174/1,252 (13.8%)
[+sem.g] 👉	[+ds] 👉	[-V1] 👉	[-morph] 👉	👉 75/1,252 (5.9%)
[+sem.g] 👉	[+ds] 👉	[+V1] 👉	[+morph] 👉	👉 252/1,252 (20,1%)
[+sem.g] 👉	[+ds] 👉	[+V1] 👉	[-morph] 👉	👉 92/1,252 (6.5%)
[+sem.g] 👉	[-ds] 👉	[-V1] 👉	[+morph] 👉	👉 241/1,252 (19.2%)
[+sem.g] 👉	[-ds] 👉	[-V1] 👉	[-morph] 👉	👉 72/1,252 (5,7%)
[+sem.g] 👉	[-ds] 👉	[+V1] 👉	[+morph] 👉	👉 141/1,252 (11.2%)
[+sem.g] 👉	[-ds] 👉	[+V1] 👉	[-morph] 👉	👉 3/1,252 (1%)
[-sem.g] 👉	[+ds] 👉	[-V1] 👉	[+morph] 👉	👉 0/1,252 (0%)
[-sem.g] 👉	[+ds] 👉	[-V1] 👉	[-morph] 👉	👉 51/1,252 (4%)
[-sem.g] 👉	[+ds] 👉	[+V1] 👉	[+morph] 👉	👉 0/1,252 (0%)
[-sem.g] 👉	[+ds] 👉	[+V1] 👉	[-morph] 👉	👉 46/1,252 (3.6%)
[-sem.g] 👉	[-ds] 👉	[-V1] 👉	[+morph] 👉	👉 0/1,252 (0%)
[-sem.g] 👉	[-ds] 👉	[-V1] 👉	[-morph] 👉	👉 69/1,252 (5.5%)
[-sem.g] 👉	[-ds] 👉	[+V1] 👉	[+morph] 👉	👉 0/1,252 (0%)
[-sem.g] 👉	[-ds] 👉	[+V1] 👉	[-morph] 👉	👉 36/1,252 (2.8%)

Source: the authors

With the combination of these four factors, we achieved a better result: we account for 99.8% (or 1,249 occurrences) of all the 1,252 null subjects found in the corpus. Among all the occurrences of null subjects, only 3 were, at a first sight, unexpected. That is an encouraging result. It means that these four factors play a role for defining the contexts for null subjects in Brazilian Portuguese.

Final remarks

As we said at the beginning, the null subject phenomenon has already been extensively investigated in BP. Much has been discovered, but some questions remain, whose answers are open.

¹¹ We have incorporated all the occurrences in the table, 1st, 2nd and 3rd persons, singular and plural. The semantic gender trait is only valid for 3rd persons.

One of these, which motivated our research, was to narrow down the contexts that allow null subjects in BP. Unlike most of the literature on the topic, we proposed here that four factors play a role in determining the optimal contexts for null subjects. They are both necessary and sufficient to understand and explain all null subject occurrences in any given corpus of contemporary spoken BP. We have presented here a multifactorial analysis, uniting the factors semantic gender, verbal morphology, optimal discursive connection and linear order to account for the null subject phenomenon in BP. We have analyzed the data taking only one factor in consideration, then two, three and finally all the four factors. As we increased the number of factors, the more consistent our analysis proved to be. With one factor alone, we could not point out robust tendencies, but as we started to analyze the factors as a set, we obtained better results, until we have arrived at the occurrences analyzed from the set of four factors, which, as we have seen, explains 99.8% of data, leading us to a better understanding of the null subject phenomenon in BP, at least when it comes to identifying its licensing contexts.

References

- ARIEL, Mira. Referring and accessibility. *Journal of linguistics*, v. 24, n. 1, pp. 65-87, 1988.
- AYRES, M. R. *Aspectos condicionadores do objeto nulo e do pronome pleno em português brasileiro: uma análise da fala infantil*. Dissertação de mestrado, Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS, 2016.
- AYRES, M. R. *Contextos licenciadores de sujeitos nulos em português brasileiro*. 123 f. PhD dissertation. Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2021.
- BATTISTI, E. *O acervo de entrevistas sociolinguísticas LínguaPOA: constituição, possibilidades e desafios*. Comunicação oral proferida no I Fórum Internacional de Sociolinguística, UFRJ, 26 nov. 2019.
- BENTIVOGLIO, P. A. *Why ‘canto’ and not ‘yo canto’? The problem of first person subject pronoun in spoken Venezuelan Spanish*. 72 f. M.A. thesis. University of California, 1980.
- BERLINCK, R. A. A construção V SN no português do Brasil: uma visão diacrônica do fenômeno da ordem. In: TARALLO, F. (org.) *Fotografias Sociolinguísticas*. Campinas: Pontes, 1989. pp. 95-112,
- BERLINCK, R. A. As posposições do sujeito em português. In: 1º Encontro do CelSul (Círculo de Estudos Lingüísticos do Sul), 1997, Florianópolis. *Anais do 1º Encontro do CelSul*, 1995. v. 2. pp. 507-19.
- BERLINCK, R. A.; DUARTE, M. E. L.; OLIVEIRA, M. Predicação. In: KATO, M.; NASCIMENTO, M. (org.) *A Construção da Sentença: Gramática do Português Culto Falado no Brasil*. São Paulo: Contexto, 2015.
- BIBERAUER, T. Semi pro-drop languages, expletives and expletive pro reconsidered. In: BERLINCK, R. A. et al. (org.) *Parametric Variation: null subjects in Minimalist theory*. Cambridge: Cambridge University Press. pp. 153-99, 2010.
- BIBERAUER, T. et al. (org.) *Parametric Variation: Null Subjects in Minimalist Theory*, Cambridge: CUP, 2010.

CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.

COELHO, I. L.; OTHERO, G. A.; VIEIRA-PINTO, C. A. Reanálise de variáveis semânticas no condicionamento do objeto nulo e do pronome pleno na fala de Florianópolis. *Fórum Linguístico*, v. 14, n. 4, pp. 2606-17, 2017.

CREUS, S.; MENUZZI, S. O papel do gênero na alternância entre objeto nulo e pronome pleno em português brasileiro. *Revista da ABRALIN*, Florianópolis, v. 3, n. 1-2, pp. 149-76, 2004.

CYRINO, S. M. L.; DUARTE, M. E.; KATO, M. A. Visible subjects and invisible clitics in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V. (org.) *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt: Vervuert-Iberoamericana, pp. 55-104, 2000.

CYRINO, S. M. *O objeto nulo no português do Brasil: um estudo sintático-diacrônico*. Londrina: Ed. da Universidade Estadual de Londrina, 1997.

CYRINO, S. M. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: ROBERTS, I.; KATO, M. A. (orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp, 1993.

DUARTE, I.; FIGUEIREDO SILVA, M. C. The Null Subject Parameter and the Structure of the Sentence in European and Brazilian Portuguese. In: WETZELS, L.; MENUZZI, S.; COSTA, J. *The Handbook of Portuguese Linguistics*. West Sussex: John Wiley & Sons, 2016. pp. 234-53.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; KATO, M. *Português brasileiro: uma viagem diacrônica: homenagem a Fernando Tarallo*. São Paulo: Contexto, 2018. pp. 83-104.

DUARTE, M. E. L. *A perda do princípio “evite pronome” no português brasileiro*. 151 f. PhD dissertation. Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1995.

DUARTE, M. E. L. A remarcação em curso no valor do parâmetro do sujeito nulo. *Cuadernos de la Alfal*, v. 12, pp. 71-99, 2020.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; & KATO, M. A. (orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp, 1993. pp. 107-28.

DUARTE, M. E. L. *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola Editorial, 2012.

DUARTE, M. E. L. ReVEL na Escola: Sobre pronomes pessoais na fala e na escrita. *ReVEL*, v. 16, pp. 1-12, 2018.

DUARTE, M. E. L.; MARINS, J. Brazilian Portuguese: a ‘partial’ null subject language. *Cadernos de Estudos Linguísticos*, v. 63, 2021.

GRAVINA, A. P. *A Natureza do Sujeito Nulo na Diacronia do PB: estudo de um corpus mineiro (1845 a 1950)*. 146 f. M.A. thesis. Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2008.

- GRAVINA, A. P. Diacronia e sujeito nulo no português brasileiro: um estudo comparativo. *Filol. Linguíst. Port.*, São Paulo, v. 16, n. spe, pp. 199-231, dez. 2014.
- GUILHERME, M. R. C. Português Brasileiro: uma língua V2? *Interfacis*, Belo Horizonte, v. 2, n. 1, pp. 82-97, 2016.
- HOLMBERG, A. Is There a Little Pro? Evidence from Finnish. *Linguistic Inquiry* 36, pp. 533-64, 2005.
- HOLMBERG, A.; NAYUDU, A.; SHEEHAN, M. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish, and Marathi, *Studia Linguistica*, v. 63, n. 1, pp. 59-97, 2009.
- HUANG, C.-T. J. On the distribution and reference of empty pronouns. *Linguistic Inquiry*, v. 15, pp. 531-74, 1984.
- JAEGGLI, O.; SAFIR, K. *The Null Subject Parameter*, Dordrecht: Springer, 1989.
- KATO, M. A. *Determinantes prosódicos em mudança sintática*. Abralin ao vivo, 2020. [<https://www.youtube.com/watch?v=t3BLRPloZJI&t=2925s>]
- KATO, M. A. Strong and weak pronominals in the null subject Parameter. *Probus*, Berlin, v. 11, n. 1, pp. 1-37, 1999.
- KATO, M. A. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. In: KATO, M. A. (org.). *Brazilian Portuguese and the Null Subject Parameter*, Vervuert- Iberoamericana, Frankfurt, Germany, 2000. pp. 223-58.
- KATO, M. A.; DUARTE, M. E. L. Pre-verbal position in BP: a reinterpretation of “avoid pronoun principle”. *Diadorim*, v. 20 n. esp., pp. 610-26, 2018.
- LIRA, S. A. *Nominal, Pronominal and Zero Subject in Brazilian Portuguese*. PhD Dissertation, University of Pennsylvania, 1982.
- MARTINS, M. A. R.; CAVALCANTE, S.; COELHO, I. L. Ordem do sujeito e colocação de clíticos na escrita brasileira dos séculos XIX e XX. *Cadernos de Estudos Linguísticos*, v. 62, pp. 1-22, 2020.
- MEILLET, A. Convergences des développements linguistiques. *Revue Philosophique de la France et de l'Étranger*, v. 85, pp. 97-110, 1918.
- MONTEIRO, J. L. *Pronomes pessoais: subsídios para uma gramática do português do Brasil*. Fortaleza: Edições UFC, 1994.
- OTHERO, G. A.; GOLDNADEL, M. Omissão de sujeito pronominal anafórico e as construções de dupla negação. *Cadernos de Estudos Linguísticos* (Unicamp), v. 62, pp. 1-21, 2020.
- OTHERO, G. A.; LAZZARI, M. Null subjects and null objects in Brazilian Portuguese: correlations and change. *Revista de Estudos da Linguagem*, v. 30, n. 4, 2022.
- OTHERO, G. A.; SPINELLI, A. C. Sujeito expresso e nulo no começo do séc. XXI (e sua relação com o objeto nulo em PB). *Domínios de Linguagem*. Uberlândia, v. 13, n. 1, 2019a.
- OTHERO, G. A.; SPINELLI, A. C. Um tratamento unificado da omissão e da expressão de sujeitos e objetos diretos pronominais de 3ª pessoa em português brasileiro. *Caderno de Estudos Linguísticos*, v. 61, n. 1, 2019b.

PAREDES SILVA, V. L. *Cartas cariocas. A variação do sujeito na escrita informal*. Phd dissertation, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1988.

PAREDES SILVA, V. L. Motivações funcionais no uso do sujeito pronominal: uma análise em tempo real. In: PAIVA, M. C.; DUARTE, M. E. L. *Mudança linguística em tempo real*. Rio de Janeiro: Contra Capa, 2003. pp. 97-114.

PERLMUTTER, D. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston, 1971.

ROBERTS, I. Posfácio. O português brasileiro no contexto das línguas românicas. In: ROBERTS, I.; KATO, M. A. (orgs.) *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp, 1993. pp. 409-25.

RODRIGUES C. N. *Effects of Loss of Morphology in Partial pro-drop Languages*. PhD dissertation, University of Maryland, Maryland, 2004.

SOARES, E. C. *Anaphors in discourse: anaphoric subjects in Brazilian Portuguese*. PhD Thesis. 482 f. University Sorbonne Paris City, Paris, France, 2017.

SOARES, E. C.; MILLER, P.; HEMFORTH, B. The Effect of Semantic and Discourse Features on the Use of Null and Overt Subjects - A Quantitative Study of Third Person Subjects in Brazilian Portuguese. *DELTA*, v. 36, pp. 1-38, 2020.

SOARES, E. C.; MILLER, P.; HEMFORTH, B. The effect of verbal agreement marking on the use of null and overt subjects. *Forum lingüístico*, Florianópolis, v. 16, n. 1, pp. 3579-600, 2019.

SPINELLI, A. C. *Pronomes e sua ausência: por um tratamento unificado da omissão de sujeitos e objetos diretos pronominais de 3ª pessoa*. 78 f. M. A. thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

TARALDSEN, K. T. On the NIC, vacuous application, and the that-trace filter. *Indiana University Linguistics Club*, 1978.

TARALLO, F. *Relativization strategies in Brazilian Portuguese*. Tese de Doutorado. University of Pennsylvania, Pennsylvania, 1983.

WIDERA, C.; KAISER, G. A. Asterix e os pronomes sujeitos: uma análise contrastiva do emprego dos pronomes sujeitos no português europeu e brasileiro. In: *I Congresso Internacional em Variação Linguística nas Línguas Românicas*, pp. 143-65, 2019.

RESTRIÇÕES SEMÂNTICAS DE ANIMACIDADE DO OBJETO NULO DO PORTUGUÊS BRASILEIRO

SEMANTIC ANIMACY CONSTRAINTS OF THE BRAZILIAN PORTUGUESE NULL OBJECT

Gabriel Guimarães Peixoto da Silva¹

RESUMO

Com este trabalho, objetivamos fazer algumas reflexões sobre o objeto nulo do português brasileiro e suas restrições semânticas de animacidade. Muitos trabalhos sincrônicos e diacrônicos, formalistas e variacionistas (GALVES, 1984; DUARTE, 1989; CYRINO, 1994, SOLEDADE, 2011; PEIXOTO, 2021) atestam que o português brasileiro (PB) é uma língua que permite objeto nulo praticamente sem restrições. Isso significa que ele pode ocorrer em diversas estruturas sintáticas – coordenação e subordinação (relativas, adverbiais) – e com diversos tipos de antecedente/referente: [+/- animado, +/- específico], sendo preferencialmente [- animado]. Entretanto, um tipo de configuração enfraquece a possibilidade de objeto nulo no PB, apesar de ainda ser possível: quando seu antecedente é [+ animado, + específico]. Vários autores gerativistas propuseram interpretações para o fato, relacionando-o a diferentes naturezas da categoria vazia. Neste artigo, baseados em um *corpus* de cartas pessoais privadas escritas por brasileiros nascidos durante o século XX, a partir dos resultados de Peixoto (2021), trazemos algumas dessas discussões e resultados de análises quantitativas de dados acerca do objeto nulo no PB e de suas relações com a animacidade do seu referente, evidenciando que a restrição do objeto nulo não é de natureza estritamente sintática, mas semântica, o que Galves (2001) chama de restrição de logoforicidade.

PALAVRAS-CHAVE: Objeto nulo. Restrições de animacidade. Português Brasileiro.

ABSTRACT

With this paper, we aim to make some reflections on the Brazilian Portuguese null object and its semantic animacy constraints. Many synchronic and diachronic, formalist and variationist works (GALVES, 1984; DUARTE, 1989; CYRINO, 1994, SOLEDADE, 2011; PEIXOTO, 2021) attest that Brazilian Portuguese (BP) is a language that allows null object with virtually no restrictions. This means that it can occur in several syntactic structures – coordination and subordination (relative, adverbial) – and with several types of antecedent/referent: [+/- animate, +/- specific], preferably being [- animate]. However, one type of configuration weakens the possibility of null object in BP, although it is still possible: when its antecedent is [+ excited, + specific]. Several generativist authors have proposed interpretations for this fact, relating it to different natures of the null category. In this paper, based on a corpus of private personal letters written by Brazilians born during the twentieth century, based on the results of Peixoto (2021), we bring some of these discussions and results of quantitative data analysis about the null object in BP and its relations with the animacy of its referent, showing that the null object constraint is not strictly syntactic in nature, but semantic, what Galves (2001) calls a logoforicity constraint.

KEYWORDS: Null object. Animacy restrictions. Brazilian Portuguese.

Agradecimento: Agradeço à professora Charlotte Galves pelas valiosas contribuições. Declaro ainda que, havendo problemas, são de minha inteira responsabilidade.

¹ Doutorando em Linguística na Universidade Estadual de Campinas. Bolsista da CNPq, guimaraesgabriel_ps@hotmail.com, <https://orcid.org/0000-0003-0322-1063>.

Introdução

Uma das marcas do português brasileiro (PB) é a praticamente livre ocorrência do objeto nulo (1), em comparação com as outras línguas românicas, inclusive o português europeu (PE). São vários os trabalhos, seja de base sociolinguística, seja de base gerativista (GALVES, 1984; DUARTE, 1989; CYRINO, 1994, SOLEDADE, 2011; PEIXOTO, 2021), com dados sincrônicos e diacrônicos, de escrita e de fala, que atestam que o objeto nulo no PB pode ocorrer em todo tipo de oração, incluindo as que formam as ilhas sintáticas, como orações relativas (2) e adverbiais (3), diferentemente do PE, com todo tipo de grade argumental transitiva (4 - 9), e com basicamente todo tipo de antecedente (10 - 12). As sentenças a seguir, assim como todas as retiradas dos *corpora*, mantêm a grafia como feita pelos missivistas; por isso algumas podem estar em desacordo com a ortografia vigente.

- (1) Vi [o gato]_i na feira e comprei (0)_i.
- (2) Parece que é [uma musica antiga]_i, chama-se “Flying on a jet plane” e quem canta (0)_i me parece que é não-sei-o-que Denver. (SL, 1951-1975)
- (3) Não repare [a carta]_i quando acabar de ler (0)_i rasga. (JM, 1901-1925)
- (4) Eu na noite de terça feira para quarta eu tive [um sonho com voçe]_i, eu depois te comto (0)_i. (JM, 1901-1925)
- (5) A minha distração são [as tuas cartas]_i, eu leio (0)_i toda as noites antes de me deitar. (JM, 1901-1925)
- (6) Já arrumei [a papelada toda que o Brasil pede sobre o diploma]_i. Mandeí (0)_i para o Consulado do Brasil. (SL, 1951-1975)
- (7) Tenho que confessar que sou terrível em [português]_i mas acho (0)_i muito legal. Japonês então nem se fala. (MV, 1976-2000)
- (8) Mas a culpa não foi minha, [eu escrevo-te todos os dias]_i, tu podes ver (0)_i pelas datas. (JM, 1901-1925)
- (9) [As oportunidades]_i aparecem e quem sabe enxergar (0)_i não deixa (0)_i passar. (MV, 1976-2000)²
- (10) Sinto-me envergonhado de te escrever a lápis mas se [te enfastiases de minhas letras]_i, manda-me dizer(0)_i. (JM, 1901-1925)
- (11) Comprei [um vestido]_i para voce eu acho que tu não vaes gostar, mas se tu não quiseses (0)_i serve para a Hilda. (SL, 1951-1975)

² O complemento (nulo) do verbo ‘enxergar’ pode ser interpretado como um complemento genérico, um tipo diferente de objeto nulo. Ver Peixoto (2021). Outra interpretação é considerar o verbo como intransitivo. Outros exemplos ao longo deste trabalho também podem ter essas interpretações. Isso só evidencia o quanto ainda há o que se discutir sobre o tema.

- (12) [Ela]_i é muito bonita e vistosa e esse é o problema. Todo mundo olha (0)_i. (SL, 1951-1975)³

Entender o comportamento do objeto nulo no/do PB é também entender sua relação com as demais formas de acusativo anafórico de terceira pessoa: pronome clítico (13), pronome lexical (14) e pronome demonstrativo neutro ‘isso’ (15).

- (13) Pega [essa carta]_i e a_i rasgue.

- (14) Pega [essa carta]_i e rasga ela_i.

- (14) Pega [essa carta]_i e rasga isso_i.

Diacronicamente, o objeto nulo sempre foi possível na história do português (CYRINO, 1994, 1997), com o antecedente [+ proposicional], e sua generalização aos outros tipos de antecedente ocorreu em paralelo à queda dos clíticos de terceira pessoa, com menor grau de referencialidade, de acordo com a Escala de Referencialidade (CYRINO, DUARTE, KATO, 2000). Essa generalização, por meio de reanálises diacrônicas, através dos séculos, provocada por uma mudança linguística, chegou até os antecedentes [+ animados, + referenciais]. No atual século XXI, esse contexto, porém, ainda se mostra como o mais resistente ao objeto nulo, embora ele ainda seja gramatical de acordo com testes de gramaticalidade, com testes de intuição de falantes nativos, e com dados de *corpora*.

O pronome lexical com retomada anafórica, uma inovação do PB em relação ao PE continental, aparece como alternativa ao objeto nulo e ao clítico de terceira pessoa nos contextos mais informais, na fala ou na escrita, e está ligado tanto aos antecedentes nominais [+ animados] quanto aos [- animados], majoritariamente [+ referenciais] (PEIXOTO, 2021). Analisado por Galves (1989) como um pronome fraco no PB, o início de sua ocorrência em *corpus* escrito de cartas pessoais e peças de teatro data de meados do século XX (SOLEDADE, 2011; MARQUES DE SOUSA, 2017; PEIXOTO, 2021).

De acordo com a relação entre objeto nulo e o pronome demonstrativo neutro, pelo seu percurso diacrônico, esse pronome tem a mesma origem que o finado clítico neutro no português: o demonstrativo latino *illud*. Assim, teríamos no português sua forma clítica – o – e uma forma não clítica – isso/isto. Essa forma clítica foi a primeira a cair (CYRINO, 1994, 1997) na gramática materna do PB, e foi o primeiro contexto concorrente com o objeto nulo. Na maioria dos casos, o pronome demonstrativo neutro tem como antecedente uma oração/proposição, isto é, um evento. As construções cujo antecedente seja um SD parecem um pouco marginais, embora ainda possíveis, como (15), e alguns desses casos adquirem ainda um caráter pejorativo. Também por razão diacrônica, preserva-se, ainda hoje, o estabelecimento da correlação entre o clítico neutro, em contextos mais

³ Em determinadas gramáticas do PB, o complemento do verbo ‘olhar’ nesse contexto pode ser um PP: “todo mundo olha para ela”.

formais, e o demonstrativo neutro (cf. MATOS, 1985), em muitos casos intercambiáveis, na posição de objeto direto (16) e na de predicativo (17). “Isso” também funciona como sujeito da passiva (18), substituindo o clítico acusativo da voz ativa, conforme vemos nos exemplos retirados de Cyrino e Matos (2006, p. 135).

(16) O presidente declarou-o/isso em entrevista à televisão.

(17) Maria é folgada, e Laura também o é/(?)é isso.

(18) O presidente declarou-o em entrevista à televisão/Isso foi declarado pelo presidente em entrevista à televisão.

Visto isso, vemos que o objeto nulo no PB pode ocorrer em basicamente todo contexto sintático, com todo tipo de antecedente. Porém, sentenças como as seguintes desafiam essa livre ocorrência do objeto nulo.

(19) *[Maria]_i chegou à festa, e eu beijei (0)_i.

(20) ?[Marta]_i sempre chora quando eu ponho (0)_i na cama.

Tanto em (19) quanto em (20), o problema de agramaticalidade ou marginalidade das sentenças, segundo alguns autores (CYRINO, 1994, 1997, 2000), não se deve a razões estruturais, mas semânticas e interpretativas, haja vista que o tipo de oração e a classe do verbo não são bloqueadores do objeto nulo, como vimos mais acima. Para que as sentenças fossem bem formadas, um pronome clítico ou lexical poderia ser usado no lugar do objeto nulo. Uma explicação bem aceita para isso se refere ao fato de o antecedente do objeto nulo ser [+ animado].

Nesse sentido, tendo em vista a marginalidade ou agramaticalidade para o antecedente [+ animado] para o objeto nulo do PB, este trabalho, de base gerativista, investigará alguns contextos levantados pela literatura que poderiam licenciar mais facilmente o objeto nulo. Para isso, por meio de uma análise de dados retirados de cartas pessoais, usados em Peixoto (2021), faremos a quantificação das ocorrências de objeto nulo com diversos tipos de antecedente, principalmente em relação ao [+ animado], a fim de entender melhor quais são as restrições mais fortes a esse contexto. Tendo isso em vista, nossas hipóteses são as seguintes:

- I - De acordo com o que prevê Figueiredo Silva (2009), deverá haver uma relação entre o papel temático do antecedente e do objeto nulo;
- II - A partir de Peixoto (2021), a animacidade conjugada à referencialidade evidenciará cronologicamente o aumento dos contextos do objeto nulo no PB;
- III - Tópicos [+ animados] favoreceriam o objeto nulo de antecedente [+ animado], segundo a proposta de Cyrino (2019); e
- IV - A animacidade e referencialidade do sujeito da oração que seleciona o objeto nulo podem favorecer a sua realização, a partir dos resultados de Peixoto (2021) sobre a significância do preenchimento do sujeito.

O presente texto estará dividido da seguinte forma: na primeira seção, serão levantados os pressupostos teóricos relevantes sobre o objeto nulo no PB; na segunda, a metodologia de análise; na terceira, nossos resultados, seguida de nossas conclusões.

1. Definição do objeto (direto)

O objeto nulo não é um fenômeno único, especialmente no PB. Isso significa que há vários “tipos” de objeto nulo sob o mesmo rótulo. Eles são, contudo, essencialmente diferentes. Caracterizar o nosso objeto de estudo é importante para identificar quais são suas próprias características, e como elas interagem com a animacidade do antecedente.

Nosso objeto (nulo) de estudo é “uma categoria vazia, de certa natureza, que é material e foneticamente nula, mas que apresenta conteúdo semântico e sintático, na posição de acusativo, mantendo o mesmo significado com seu referente” (PEIXOTO, 2021, p. 76), além de possibilitar, em alguns casos, a dupla leitura interpretativa. O referente/antecedente do objeto nulo pode ser uma entidade (1), (3-7), (9) e (11-12), ou uma oração, ou uma parte do predicado, (2, 4, 8, 10). Dessa forma, nosso objeto anafórico nulo terá sempre um antecedente linguístico, não ancorado no contexto externo às cartas.

Além disso, outra distinção importante se refere ao material que está fonética e materialmente nulo relacionado ao complemento do verbo, ao que está elidido na sentença. No PB, além do objeto nulo, as principais categorias que podem ser nulas são: anáfora do complemento nulo, elipse proposicional e elipse de VP⁴. Essa distinção é importante, visto que nem toda categoria vazia objeto será aquilo com o que vamos trabalhar neste artigo.

1.1. Acusativo anafórico de terceira pessoa e a animacidade do referente

Desde Duarte (1989) e Cyrino (1994), a animacidade do referente do objeto nulo é vista como um forte fator para o seu favorecimento. Nas suas pesquisas, os resultados indicam que o traço [- animado] do antecedente favorece o objeto nulo. Em contrapartida, quando o antecedente é [+ animado], as opiniões se dividem. O trabalho com dados de fala e entrevistas dos anos 1980 – 1990, de Duarte (1989), atestam a existência de objeto nulo com tal grau de animacidade, embora sua taxa de ocorrência seja menor se comparada com a do traço [- animado] (64,7% vs. 23,7%). Já para Cyrino (1994, 1997), um objeto nulo [+ animado] só poderia ocorrer dentro de uma estrutura de Elipse de VP, ou, quando muito, já próximo ao século XX, com o antecedente [+ animado, – específico]. Em trabalhos posteriores (CYRINO, 2000), a autora admite a tímida existência de objetos nulos com antecedentes [+ animados], fornecendo uma explicação sobre o fato.

Abaixo, apresentamos uma tabela que mostra os índices adaptados de animacidade do objeto nulo em diversos trabalhos sincrônicos, diacrônicos, de fala e de escrita, referentes ao século XX. Em Cyrino (1994) são usadas peças de teatro de autores brasileiros e portugueses. Em Duarte (1989), a

⁴ Para melhor entendimento, ver Peixoto (2021).

autora usa dados de fala e entrevistas de tevê. Cyrino (2018) faz uso de anúncios de jornais. Marques de Sousa (2017) usa peças de teatro de autores brasileiros. E Peixoto (2021) usa cartas pessoais de brasileiros não ilustres.

Tabela 1: Índice de objeto nulo com relação ao traço de animacidade em diversos trabalhos

Traços de animacidade	Cyrino (1994)	Duarte (1989)	Cyrino (2018)	Marques de Sousa (2017)	Peixoto (2021)
[+ proposicional]	90%	79,7%	38,5%	41,2%	78,8%
[- animado]	88,3%	64,7%	45%	51,03%	76,6%
[+ animado]	14,3%	23,7%	13,4%	13,8%	20,4%

Fonte: PEIXOTO, 2021

Observando o traço de animacidade do antecedente, vemos que a tendência dos resultados é semelhante em todos os outros trabalhos. No que se refere ao antecedente [+ proposicional], os resultados de Peixoto (2021) (78,8%) estão entre os maiores índices (90% e 79,7% vs. 41,2% e 38,5%). É interessante notar que estão mais próximos dos resultados de fala (79,7%), de Duarte, do que dos resultados de peças teatrais, com 41,2% (MARQUES DE SOUSA, 2017). Isso confirma a tendência observada quando analisamos a distribuição das formas morfossintáticas em relação ao tempo: o que viria a ser produzido na fala nos anos 80-90 já estava disponível como *input* desde antes, como indicam os dados da escrita produzida em períodos anteriores. Isso significa que esses dados de escrita fornecem um panorama de como estavam organizadas as formas morfossintáticas no PB em períodos cronológicos anteriores aos anos 1980-1990. Esse panorama, então, dá indícios de como seria o *input* em períodos próximos, sendo possível conceber como as próximas gerações o receberam.

A tendência dos contextos [+ animados] é semelhante ao que acontece com os contextos [+ proposicionais]: aqueles aparecem sempre como os mais desfavorecidos, enquanto estes são os mais favorecidos pelo objeto nulo. A tendência percentual sendo parecida, mas com um comportamento oposto entre esses contextos indica que se trata da mesma gramática que gera ambas as estruturas. Os casos de [- animados] variam percentualmente entre os trabalhos, embora estejam, em geral, no “meio”, entre os mais favorecidos e os menos.

Além disso, os resultados de Peixoto (2021), selecionados pelo programa *GoldVarb* (SANKOFF, TAGLIAMONTE, SMITH, 2005), indicam que o fator mais influente para o favorecimento do objeto nulo em detrimento do clítico acusativo, a partir do maior peso relativo (PR), é a animacidade, seguido da realização (plena ou foneticamente nula) do sujeito do predicado que seleciona o objeto nulo, e do período de nascimento do missivista.

Tabela 2: Atuação do contexto morfossintático “animacidade do antecedente” para o objeto nulo em cartas pessoais brasileiras

Traços de animacidade	N/Total	%	PR
[+ sentencial/proposicional]	102/105	97,1%	0.887
[- animado]	281/362	77,6%	0.438
[+ animado, ±humano]	9/42	21,4%	0.047

Fonte: adaptado de PEIXOTO, 2021

O contexto favorecedor é o [+ proposicional] (.p 0,887). O antecedente proposicional se sobrepõe ao inanimado e ao animado, confirmando a hipótese da Escala de Referencialidade (KATO, DUARTE, CYRINO, 2000), segundo a qual os antecedentes [+ proposicionais] são aqueles que licenciam mais facilmente o apagamento do objeto, seguido pelos antecedentes [- animados]. O contexto mais resistente ao objeto nulo, conforme previsto, foi o [+ animado].

Outros trabalhos, de natureza mais teórica do que quantitativa, discutem a relação entre o objeto nulo e sua natureza como categoria vazia, alinhando a discussão à animacidade do antecedente do objeto nulo. Kato e Raposo (2001), ao trabalharem com romances e peças, escritas no século XX, de autores brasileiros e suas “traduções” para o PE, e também com testes de julgamento de gramaticalidade, constataam que não só há uma diferença grande das porcentagens entre as duas variedades do português, como também veem o objeto nulo como um tipo de estrutura resultante das construções de tópico (a saber, topicalização e deslocamento à esquerda com clítico). Além disso, para eles, a natureza desse clítico nulo corresponde à sua gramaticalização como afixo, o que resulta na sua não subida para I (Flexão) e sua afixação junto ao verbo principal.

Ao relacionarem o clítico expresso e o clítico nulo às diferentes construções de tópico, Kato e Raposo (2001) consideram tais construções como sendo o mesmo fenômeno, visto que a diferença entre elas se encontra na presença foneticamente nula do clítico ou foneticamente realizada (a', a'', abaixo, respectivamente).

a'. [Esse livro]_i, eu só _____i encontrei na FNAC. (Topicalização)

a''. [Esse livro]_i, eu só o_i encontrei na FNAC. (Deslocamento à esquerda clítica (CLLD))

Assim, eles discutem a distribuição de objetos nulos em relação ao seu antecedente nas duas variedades do português, verificando se eles podem ocorrer independentemente de seu antecedente estar em posição A', isto é, independentemente de estar em posição de tópico ou não. Dessa forma, procuram ver, na prática, a distribuição do clítico nulo, do clítico expresso e do pronome lexical, levando em consideração a posição do seu antecedente.

Relacionando o objeto nulo com a posição A ou A' e o traço de animacidade do antecedente, os autores chegam a uma conclusão semelhante à de Cyrino (1994): o traço [+ humano] do antecedente ainda apresenta resistência ao objeto nulo quando esse antecedente é sujeito. Isso serviria tanto para o PE quanto para o PB. Uma diferença entre eles se encontra no fato de que é possível no PB a

existência de objeto nulo tendo seu referente no sujeito, desde que esse, como foi dito, não seja humano, i. e., [- humano]. Já em PE, nem mesmo com esse traço o antecedente do objeto nulo pode ser correferencial ao sujeito. Dessa forma, teríamos o seguinte esquema segundo a posição A ou A' e o traço de animacidade do antecedente:

Tabela 3: Esquema segundo a posição A ou A' e o traço de animacidade do antecedente, em (KATO E RAPOSO, 2001)

Posição do antec. do Objeto Nulo	Tópico	Objeto	Sujeito
PB	[+/- humano]	[- humano]	[- humano]
PE	[+/- humano]	[- humano]	*

Fonte: adaptado de PEIXOTO, 2021

Por último, os autores reafirmam que as conclusões da tabela acima não se aplicam quando a realização do objeto anafórico é o clítico ou o pronome lexical, no caso do PB.

Outra proposta é evidenciada em Cyrino (2019), em que a autora refina algumas de suas questões teóricas e defende que o verbo em PB se move para um núcleo aspectual entre vP e VP: este seria *inner aspect*, “aspecto interno”, (AspInn). Esta posição seria mais baixa do que aquela ocupada pelo verbo em PE. Além disso, a autora também propõe que haja uma categoria funcional responsável pela animacidade, o núcleo F[pessoa], localizado logo abaixo de vP e acima de AspInn e V. Segundo ela, essa projeção explicaria a restrição do objeto nulo do PB em relação aos antecedentes [+ animados]. “Essa categoria tem o traço não interpretável [uPessoa] e um traço EPP, traços que atraem os DPs animados. Os DPs inanimados [e os nomes plurais nus] permanecem *in situ*”. Assim, os DPs animados se movem para fora de VP, Spec de F(pessoa), oposto ao que acontece com os DPs inanimados e plurais nus, e, por isso não podem ser elididos, porque, estando fora de VP, não serão licenciados pelo verbo em AspInn, pois estão fora do seu domínio de c-comando. Por essa razão, só os pronomes lexicais podem retomar um DP animado, além de inanimados, pois se movem para “fora de vP para obedecer outros requisitos da interface sintaxe-semântica”⁵.

Além disso, a justificativa de Figueiredo Silva (2009) e Cyrino (2019) para os casos que não seguem a restrição de objeto nulo cujo referente seja o sujeito da matriz é, na verdade, a de que estes são exemplos com verbos inacusativos, e a de que o sujeito, na verdade, é o complemento do verbo. Essa justificativa segue Cyrino (2008, 2013), para quem o objeto nulo do PB, além da restrição de animacidade, também deve ocorrer em uma “estrutura paralela” ao do seu antecedente. Para Cyrino, o objeto nulo, assim como a Elipse de VP, é resultado da Reconstrução, como já dissemos. Baseada em Fiengo e May (1994), a autora considera que “a reconstrução é uma condição necessária, mas não suficiente para a elipse” (CYRINO, 2018), e que outros fatores também devem estar envolvidos para o licenciamento da elipse. Para ela:

⁵ A autora original não põe a preposição ‘a’ como regida pelo verbo ‘obedecer’ nesta sentença.

A natureza da reconstrução, no entanto, deve ser a mesma nas línguas, pois é uma condição sobre identidade sintática. Ela torna explícitas quais ocorrências são as mesmas em um marcador frasal: aquelas que preservam a categoria gramatical e as relações lineares e de dominância dentro da categoria gramatical. Em outras palavras, todas as ocorrências devem ser compostas estruturalmente/configuracionalmente da mesma forma. É o requerimento do “paralelismo”.

O antecedente do objeto nulo do PB também deve ocorrer em uma estrutura paralela a ele, isto é, dados que o objeto nulo é um complemento de um verbo, seu antecedente deve também estar em uma posição de complemento (CYRINO, 2018, pp. 238-9).

Apesar desse requerimento, a autora assume que o traço [+ animado] do antecedente pode suplantar o paralelismo estrutural em determinadas sentenças e tornar a sentença agramatical. Comparemos (i) e (ii) abaixo, retirados de Cyrino (2019, p. 177):

- (i) [O livro]_i afundou t_i antes de o Pedro pegar (0)_i.
- (ii) *[A menina]_i afundou t_i antes de o Pedro pegar (0)_i.

Nos dois exemplos acima (i - ii), temos casos de sujeitos derivados (argumentos internos do verbo ‘afundar’). Dessa forma, se o requerimento do paralelismo fosse extremamente necessário, as duas sentenças deveriam ser agramaticais, o que não é o caso. (ii) só seria gramatical se um pronome (pleno ou clítico), objeto de ‘pegar’, retomasse o antecedente [+ animado].

Em Cyrino (2019, p. 190), a autora mostra que as ocorrências de objeto nulo com antecedente sendo tópico (estrutural ou discursivo) não são exatamente uma propriedade exclusiva do objeto nulo do PB, sendo possível em outras línguas, como russo e hebraico. Ela indica que esse pode ser um caso de objeto nulo diferente do que ela trata, levantando também a questão de se verificar se há uma relação entre a animacidade do antecedente desse tipo de objeto nulo e a posição de tópico. Resolvemos, como será visto, tratar dos casos com antecedente em tópico, porque conseguiremos comparar o comportamento dos constituintes internos à oração com o do externo (o tópico), e suas relações com o objeto nulo, que, segundo alguns estudos, no PB, mantém uma relação particular com a posição de tópico. A autora também indica que a Elipse proposicional, segundo essas considerações, seria um tipo de elipse de vP.

1.2. A polêmica natureza do objeto nulo

A heterogeneidade do objeto nulo, em especial no PB, teve como resultado as inúmeras propostas, de acordo com a teoria gerativa, para determinar o tipo, a natureza, da categoria vazia *in situ*, complemento do verbo. As duas principais propostas para a natureza do objeto nulo do PB, dentre as mais discutidas, como visto em Peixoto (2021), são (i) *pro* (GALVES, 1984, 1989, 2001; FIGUEIREDO SILVA, 2009), e (ii) elipse de DP com Reconstrução em FL (CYRINO, 1994, 1997, 2000, 2019).

No modelo de Regência e Ligação da Teoria de Princípios e Parâmetros, *pro* é uma categoria vazia sujeita ao Princípio B, segundo o qual uma categoria vazia é um pronome se for referencial, a um sintagma que a c-comande fora do domínio de regência em que se encontra. Além disso, *pro* também seria licenciado por T e relacionado aos traços de concordância que têm com o núcleo verbal de TP. Esse é um dos motivos pelo que se considera um *pro* a categoria vazia do sujeito nulo. Entretanto, no caso do PB, o objeto não mantém concordância com o verbo, o que já levou Huang (1984) a desconsiderar a possibilidade de *pro* para o objeto nulo.

Huang chega a essa conclusão analisando a livre ocorrência do objeto nulo nos domínios frásicos de ilhas, por exemplo, sentenças como (21) e os casos de topicalização no PB (22).

Sobre a possibilidade de ter mais de um referente, tal como um pronome foneticamente expresso, Farrell (1987, apud FIGUEIREDO SILVA, 2009) utiliza sentenças semelhantes às seguintes.

- (21) a. [Pedro]_i contratou [Paulo]_j antes que eu tivesse conhecido ele_{i/j/k}.
 b. [Pedro]_i contratou [Paulo]_j antes que eu tivesse conhecido (0)_{i/j/k}.

Além de poder ter mais de um referente dentro da mesma sentença, o objeto nulo, sendo um *pro*, poderia ter um referente discursivo, extralinguístico, como vemos em (21b).

No que se refere à topicalização, o fato de o PB permitir movimento, como as topicalizações em orações encaixadas, mostra que o objeto nulo pode se referenciar a esse tópico. Se isso é possível, deveria ser possível também ao PB o objeto nulo ser correferenciado a um tópico nulo, sendo, então uma variável. Porém, o contraste entre (22) e (23) evidencia que não é isso que ocorre. Em (23), bloqueadas pelo Princípio C, as sentenças em que o antecedente do objeto nulo é o sujeito da oração matriz, que são agramaticais, revelam-nos que, se realmente fosse uma variável, deveria haver um operador nulo na periferia esquerda da sentença encaixada, fazendo essa referência (cf. FARRELL, 1987, p. 16).

- (22) a. O Raimundo_k insiste que [ele]_i ninguém viu (0)_i.
 b. Ouvi falar que [o bolo]_i todo mundo adorou (0)_i.

- (23) * [Maria]_i acha [que OP_i o José viu (0)_i].

O fato, então, de o antecedente do objeto nulo, que está dentro de uma oração completiva, ser sujeito de uma oração matriz tornar a sentença agramatical no PB faz com que Farrell (1987, 1990) desenvolva uma generalização de que esse contexto seria bloqueador do objeto nulo. Segundo Maia (1990) (apud CYRINO, 1994, p. 39), porém, a generalização de Farrell de que as sentenças adjuntas permitem a realização do objeto nulo cujo antecedente seja o sujeito matriz e de que as sentenças completivas são um contexto de restrição não funciona em todos os casos. A depender da semântica verbal, as situações de (a)gramaticalidade podem se inverter. Cyrino (1994), dentre alguns outros autores, vê como agramatical o exemplo em (24a).

- (24) a. [A Julia]_i sempre chora quando ponho (0)_i no berço.
 b. *[A Julia]_i sempre chora quando vejo (0)_i.
- (25) [João]_i acha que Pedro ajudou (0)_i a fazer o exercício.

Além disso, há mais uma dificuldade em se considerar o objeto nulo no PB como um *pro*. A presença de um pronome foneticamente expresso no lugar do objeto nulo, segundo Cyrino (1994), modifica as possibilidades de interpretação para o objeto nulo. Entretanto, para Galves (2001), essa impossibilidade de alternância do objeto nulo com o pronome expresso em alguns contextos se explica pela diferente natureza da categoria vazia e do pronome e pelas diferentes derivações e projeções envolvidas na categoria vazia e no pronome expresso.

- (26) a. [Pedro]_i disse que Marta não conhece (0)_{*i/j}.
 b. [Pedro]_i disse que Marta não conhece ele_{i/j}.

Ademais, como decorrência do referido Princípio B, o antecedente do objeto nulo deve estar fora do seu domínio. Entretanto, se o objeto nulo no PB funciona tal qual um pronome (*pro*), como explicar a agramaticalidade de sentenças com o objeto nulo (27a), mas sua gramaticalidade com um pronome (27b), em que, em ambos os casos, o antecedente do objeto anafórico está fora do seu domínio?

- (27) a. *Como [a menina]_i estava na festa, João entregou (0)_i para a polícia.
 b. Como [a menina]_i estava na festa, João entregou ela_i para a polícia.

Outra proposta para a defesa de *pro* como objeto nulo é descrita em Galves (1989, 2001). A autora tenta resolver o impasse evidenciado em Huang (1984) de que um objeto nulo *pro* não seria possível em nenhuma língua, pois essa categoria nessa posição violaria o Princípio B e o Princípio C, segundo o qual é necessário haver um elemento nominal bem próximo a *pro* que o controle e que seja seu antecedente.

Para isso, Galves (1989, 2001) desconsidera os casos em que o referente do objeto nulo pode ser extrassentencial, considerando apenas o contexto sentencial. Para a autora, que se baseia em Williams (1987, apud GALVES, 1989), é possível que os dois princípios acima – Princípio B e C – sejam satisfeitos, pelo menos no PB, de outra forma, que não a ilustrada por Huang (1984). Sua proposta é de que os domínios de aplicação podem ser diferentes para cada um dos princípios.

De acordo com Williams, há uma dissociação entre vestígios [*aqui entendido como um objeto nulo que pode receber papel temático, mas não pode satisfazê-lo*] e movimento, o que permite dar conta da relação entre a categoria vazia e seu antecedente através da coindexação da CV a seu antecedente realizador do papel temático. Para isso, ele distingue “atribuição” e “satisfação” do papel temático. Nessa perspectiva, Galves (1989) propõe que a coindexação da CV na posição de objeto direto a seu antecedente se dá através do requerimento de

“satisfação” do papel temático e que ela é um vestígio de NP, pois pode receber papel temático, mas não pode satisfazê-lo. (FIGUEIREDO SILVA, 2009, p. 148) [adições nossas].

Dessa forma, a autora segue os pressupostos estabelecidos em Huang (1984) ao entender que “a categoria vazia é livre no que se refere à teoria da ligação, mas dependente dentro da teoria temática” (FIGUEIREDO SILVA, 2009, p. 149).

Como consequência dessa linha argumentativa, Galves (1989, 2001) consegue explicar a agramaticalidade das sentenças em que o antecedente do objeto nulo, que está em uma sentença completiva, é o sujeito da matriz, como (28).

(28) * [Maria]_i acha [que o José viu (0)_i].

Entendemos que a explicação, resumidamente, a partir da proposta Figueiredo Silva (2009), baseada em Galves (1989, 2001), versa em torno dos diferentes papéis temáticos envolvidos na sentença, visto que, em determinados contextos, como o referente do sujeito sendo também o tópico discursivo (que deve ser ativado pelo contexto, o que explica a dificuldade – mas não impossibilidade – de haver essa interpretação), ou construções, como (29), essa sentença pode ser gramatical.

(29) O padre disse que Maria_i acha que o José viu (0)_i na feira.

O que explicaria, então, a agramaticalidade de (28), segundo essa proposta, é que o sujeito da matriz [Maria] pode ser considerado agente, enquanto que o objeto nulo pode ser considerado paciente. Para que a sentença fosse gramatical, o índice do objeto nulo deveria ser outro, diferente do índice do sujeito da matriz. Essa explicação também funciona para os casos de agramaticalidade como o da sentença abaixo (30-31), retiradas de Peixoto (2021):

(30) [Essa encomenda]_{i (tema)} pede que o carteiro entregue (0)_{i (alvo)} em mãos.

(31) *[O supervisor]_{i (experenciador)} pede que os carteiros cumprimentem (0)_{i (alvo)} na hora da saída.

Ademais, motivados, também, por as hipóteses anteriores não explicarem todos os casos de licenciamento e restrição de objeto nulo no PB, segundo a visão da autora, os trabalhos de Cyrino (1994, 1997, 2000, 2016, entre outros) dão conta de casos como a impossibilidade de alternância entre objeto nulo e pronome lexical e as duplas leituras evidenciadas com o objeto nulo, além de explicar o porquê da ocorrência de objeto nulo em ilhas sintáticas. Para isso, suas hipóteses clássicas consideram o objeto nulo no PB como fruto de uma reconstrução em Forma Lógica (FL) e elipse em Forma Fonética, já as mais recentes admitem a possibilidade de objetos nulos com antecedentes [+ animados]; com isso, para a autora, o objeto nulo é preferencialmente [- animado, ± específico], o que mostra que ele sofre restrição de animacidade, não de referencialidade.

Além de *pro*, PRO, anáfora e variável, clássicos tipos de categorias vazias existentes nas diversas línguas, há também um outro tipo de categoria vazia, a elipse, que é motivada por “fatores

discursivos a fim de evitar redundâncias e sua interpretação se dá no contexto discursivo ou no contexto situacional” (FIGEIREDO SILVA, 2009 apud MATOS, 2003; LOBECK, 1995, 1999; FIENGO E MAY, 1994; CYRINO, 1994; 1997; dentre outros). A elipse é uma lacuna fonética e material que corresponde a diferentes categorias, como CP, IP, InfP, SV, SD e SN. Para o objeto nulo do PB, a autora, em seus diversos trabalhos, classifica-o como uma elipse de SD, reconstruída em FL.

Sua hipótese para o PB é a de que há uma Reconstrução do antecedente do objeto nulo em Forma Lógica, e sua elisão em Forma Fonética. A elipse de VP, a Anáfora do Complemento Nulo e a Elipse proposicional, semelhante ao que acontece com objeto nulo, para a autora (cf. CYRINO, 2003, 2004), também seriam casos de Reconstrução, mas eles não teriam restrição de animacidade, o que não acontece com o objeto nulo do PB; inclusive, esse é um dos motivos que pode fazer com que essas construções se confundam. Dessa forma, o processo de Reconstrução é descrito por Cyrino como

A reconstrução é uma relação de identidade sobre marcadores frasais (phrase markers) e todas as ocorrências são tokens, e como tal, podem ser audíveis ou não. (...) os membros de uma reconstrução preservam a categoria gramatical e as relações lineares e dominância dentro da categoria gramatical: todas as ocorrências serão compostas estruturalmente da mesma forma (CYRINO, 1997, p. 107).

A Reconstrução – um processo de identidade estrutural, portanto – explica, então, a possibilidade de duplas leituras (leitura estrita e leitura imprecisa) evidenciado pelo objeto nulo em certas sentenças. Como o seu antecedente é reconstruído para a interpretação, as duas leituras disponíveis também se mostram possíveis em FL.

(32) Lucas_i [beijou a sua amiga]_k, mas Joaquim_j não quis (0)_k.

(33) José utiliza [seu celular]_i nos dias de jogos, enquanto Antônio usa (0)_i todos os dias.

(34) O João cumprimentou [sua mãe]_i, mas o Pedro não cumprimentou (0)_i.

Como dito anteriormente, em (32), (33) e (34), o referente do pronome “seu”, ambíguo no PB, é reconstruído, “repetido”, após a derivação na interpretação em FL, o que faz com que haja dúvida sobre seu referente, possibilitando a leitura imprecisa e a estrita: Joaquim não quis beijar a própria amiga – leitura imprecisa – ou não quis beijar a amiga de Lucas – leitura estrita; o celular que Antônio usa todos os dias é o seu próprio – leitura imprecisa – ou é o do José – leitura estrita; e Pedro não cumprimentou a sua própria mãe – leitura imprecisa – ou a mãe de João – leitura estrita.

Visto isso, apresentaremos agora nossa metodologia de análise.

2. Metodologia

A metodologia utilizada foi a quantificação e análise de dados retirados de *corpus* composto por cartas pessoais escritas por brasileiros nascidos no século XX. A partir do trabalho de Peixoto (2021), nosso *corpus*, pertencente ao acervo do HistLing – UFRJ, é composto por uma amostra representativa de três períodos de nascimento distintos: I- Casal Jayme-Maria (1900-1925); II- Família Salgado Lacerda (1950-1975); e III- Missivistas variados do século XXI (1976-2000).

Baseando-nos nos parâmetros referentes à distância ou proximidade comunicativa (cf. KOCH; OESTERREICHER, 1990), como grau baixo de publicidade das cartas, grau alto de familiaridade entre os missivistas e implicatura emocional entre eles, por exemplo, podemos, portanto, traçar um perfil do nosso *corpus*. Com isso, consideramos que o gênero carta (com seus subtipos: pessoal e amistosa) figura entre dois extremos em um *continuum* de concepção oral e escrita da linguagem: não se encontra nem no extremo da concepção oral, por apresentar elementos próprios da escrita, como o uso dos clíticos, como planejamento textual, mas também não se encontra no extremo da concepção escrita, visto ter ele característica que aproxima o gênero da interação face a face, como a presença de vocativos e de pronomes de segunda pessoa. Por isso, justamente por esse gênero não possuir todas as características da concepção formal, como artigos de jornais, nem da concepção informal, como as entrevistas orais gravadas e transcritas, além de textos que simulam a fala, como as peças de teatro, consideramo-lo ideal para nossas finalidades com esta pesquisa, pois conseguimos por meio dele resgatar traços da oralidade. Ou seja, o gênero carta pessoal amistosa seria um texto em que haveria a emergência da gramática vernácula dos missivistas, o que aponta à gramática do português brasileiro.

Foram selecionados dados de objeto nulo com antecedente nominal e proposicional (elipse proposicional). Com base na discussão apresentada na seção anterior, selecionamos os seguintes fatores para a análise quantitativa: o papel temático do antecedente e do objeto nulo; o cruzamento do tipo de animacidade com o grau de referencialidade; o referente sendo tópico ou não; a animacidade e especificidade do sujeito da oração que seleciona o objeto nulo; e o período de tempo de nascimento do missivista que produziu o dado.

3. Análise de resultados

Quantificamos um total de 399 dados de objeto nulo e elipse proposicional. Eles foram analisados de acordo com o nosso grupo de fatores a fim de testarmos nossas hipóteses sobre a relação entre a animacidade e o objeto nulo no PB.

3.1. Uniformidade de papel temático

Segundo a proposta de Figueiredo Silva (2009), a partir de Bianchi e Figueiredo Silva (1994) e Galves (2001), em geral, uma cadeia formada por papéis temáticos diferentes – entre o antecedente e a categoria vazia – não licencia o objeto nulo, fruto de uma reconstrução/elipse de DP. Nesse sentido, buscamos testar se papéis temáticos diferentes poderiam licenciar o objeto nulo em nosso *corpus*. Para isso, consideramos os seguintes papéis temáticos: agente, experienciador, tema, meta, beneficiário e outro⁶ (cf. FIGUEIREDO SILVA, 2009) – relacionado, muitas vezes, ao tópico (ainda não há consenso na literatura sobre a falta de papel temático dos tópicos).

⁶ Alguns exemplos dos respectivos papéis temáticos:

Agente: **João** abriu a porta.

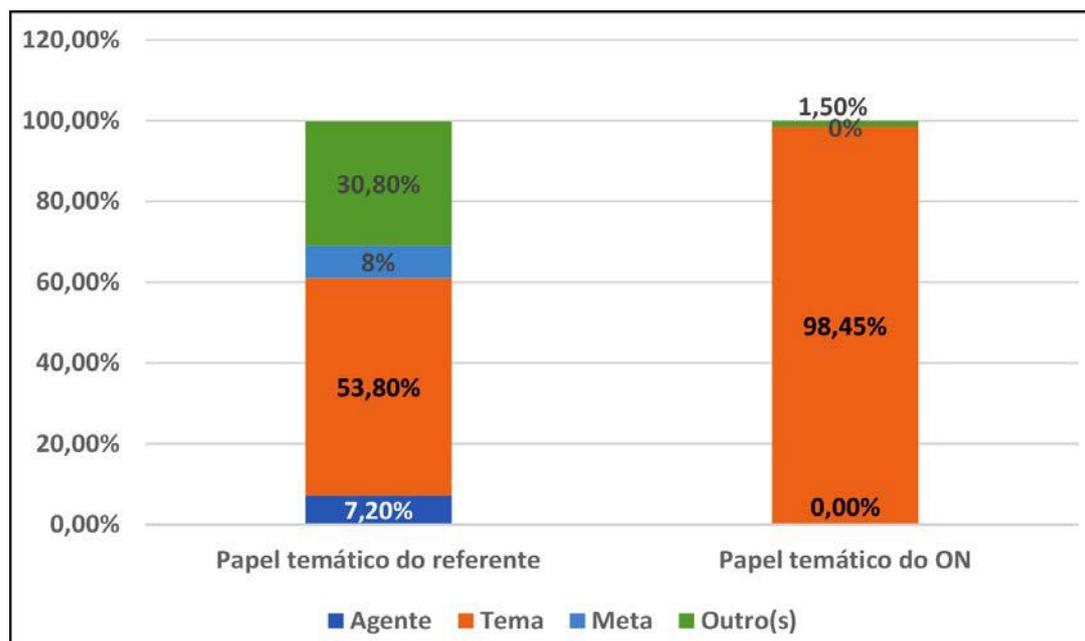
Paciente/tema: João comeu **o bolo**.

Experienciador: **João** gosta de Maria.

Meta: João quer **seus óculos**.

Beneficiário: Isso foi feito para **o João**.

Outro, como os tópicos: **O João**, ele comeu tudo.

Gráfico 1: papéis temáticos do referente e do objeto nulo em corpus de cartas pessoais

Fonte: elaboração do autor.

Por meio do gráfico 1, vemos que o papel temático do objeto nulo parece ser quase categórico: tema, com 98,45% das realizações (35). A taxa de 1,5% corresponde ao papel temático de possuído, típica dos verbos “ter” e “haver”, (36). No nosso *corpus*, em relação ao referente, não encontramos dados com papéis temáticos de beneficiário e de experienciador.

- (35) comprei [um vestido]_i para voce eu acho que tu não vaes gostar, mas se tu não quiseres (0)_i serve para a Hilda. (JM, 1901-1925)
- (36) Não tenho [filhos]_i, mas pretendo ter (0)_i futuramente quando estiver formada, empregada e tiver minha casa/meu canto. (MV, 1976-2000)

Já o papel temático do antecedente/referente é mais heterogêneo. Embora haja a predominância do papel de tema (37), com quase 54% de ocorrências, ele também possui quase 31% de outros papéis temáticos, ou falta de papel temático, como no caso dos tópicos, (38), em que temos um caso da chamada *topicalização*. Os dois menores índices, 8% de meta, (39), e 7,2% de agente, (40 e 41), não aparecem no objeto nulo, como já era esperado.

- (37) eu até agora ainda não recebi [tua carta desta semana]_i, se o Nelzinho não trouxer nenhuma na hora do almoço, de noite vou na tua casa, porque tenho certeza que terá (0)_i. (JM, 1901-1925)
- (38) [as fotografias]_i ja perdi as esperanças de ver (0)_i (JM, 1901-1925)
- (39) Obs: Não esquenta com [o papel e [o sobrenome]_i] você escreveu (0)_i certo. (MV, 1976-2000)

- (40) Se [estas cartas]_i não me deixassem tão felizes, eu iria parar de escrever (0)_i para o pessoal, pois meu dinheiro está indo todo embora. (SL, 1950-1975)
- (41) Ela_i é muito bonita e vistosa e esse é o problema. Todo mundo olha (0)_i. (SL, 1950-1975)

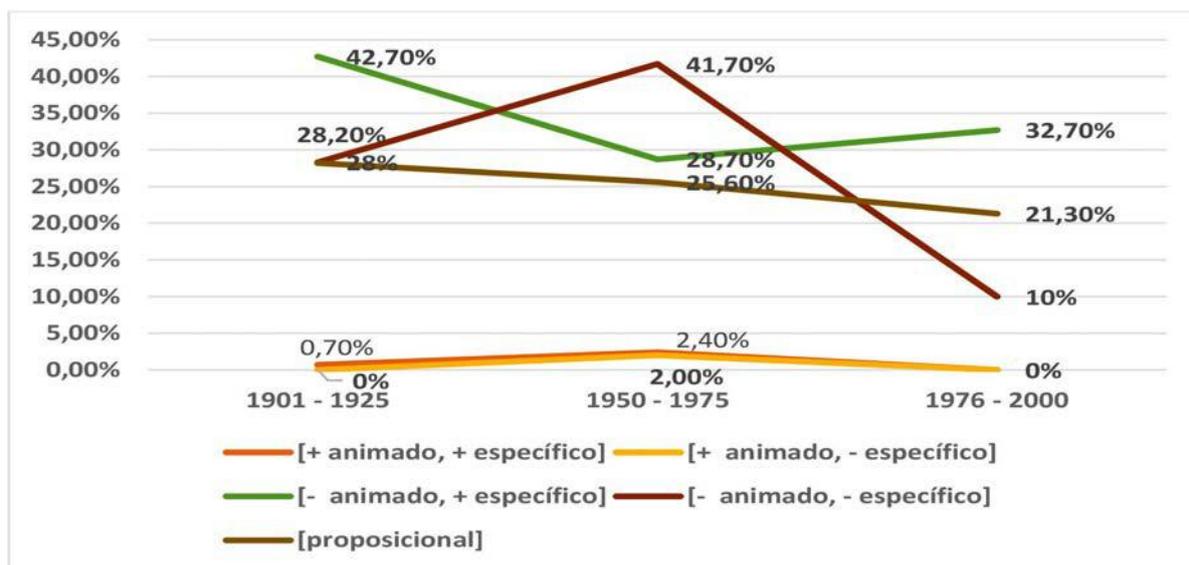
Embora os verbos com objeto nulo (39-41) possam ter uma interpretação intransitiva, decidimos considerá-los casos de objeto nulo pela possibilidade de recuperação do antecedente linguístico, e pelo contexto discursivo das cartas em questão. Tendo em vista os resultados acima, podemos dizer que parece não haver uma ligação tão estreita entre o papel temático do antecedente e o do objeto nulo, embora convirjam na maioria das vezes (mais de 50%). Além disso, a proposta de Figueiredo Silva (2009) também prevê que o antecedente do objeto nulo pode ocorrer se for um argumento do verbo ou um tópico. Os resultados de Peixoto (2021) apontam que adjuntos também podem ocorrer como antecedentes, apesar de estar longe de representar uma grande realização (em torno de 5% dos dados de objeto nulo, somando as ocorrências de antecedentes adjuntos adnominais e adverbiais, complemento nominal e predicativo, em comparação com o antecedente do objeto nulo com as funções sintáticas completivas verbais e tópico).

Nesse sentido, ou seria preciso refinar os papéis temáticos usados, ou então repensarmos o licenciamento pelas cadeias temáticas para o licenciamento do objeto nulo, pensado por Figueiredo Silva (2009).

3.2. Animacidade e referencialidade ao longo do tempo

De acordo com a escala da Hierarquia de Referencialidade (CYRINO, DUARTE, KATO, 2000), os pronomes que tendem a cair nas línguas são os não argumentais [+ proposicionais], assim como os pronomes de terceira pessoa com antecedentes [+ proposicionais] e [- animados], como os clíticos. Isso explicaria a queda acentuada dos clíticos acusativos de terceira pessoa do PB ao longo do tempo, e, por consequência, o expressivo aumento de objeto nulos. Essa escala conjuga a animacidade do antecedente com sua referencialidade. Dessa forma, o cruzamento de animacidade com a referencialidade mostraria se os antecedentes menos robustos semanticamente do objeto nulo, como os não argumentos de terceira pessoa, seguem a tendência observada pelas autoras de se tornarem foneticamente nulos ao longo do tempo. Assim, dividimos os dados da seguinte maneira para verificarmos se houve mudança nos contextos de licenciamento do objeto nulo ao longo dos três períodos de tempo analisados, correspondentes ao século XX: [+ animado, + específico], [+ animado, - específico], [- animado, + específico], [- animado, - específico], e [+ proposicional]⁷.

⁷ Entendemos [+ ou - específico] a partir do grau de conhecimento e/ou especificação gramatical que o referente possui dentro do contexto discursivo: partindo da sentença em que ele é primeiramente mencionado até suas outras menções. Por exemplo: *Vi um professor que me deu aula naquele colégio*. A especificação gramatical dada pela oração relativa especifica [+ específico] o referente. Assim, se podemos conhecer sem dúvidas o referente, dentro de um determinado contexto, ele é [+ específico]; se, por outro lado, o referente é mais genérico, como uma noção de coletividade ou como um conceito, ou ainda como um nome/uma ideia mais geral, por exemplo, ele é [- específico].

Gráfico 2: cruzamento da animacidade e referencialidade do antecedente do objeto nulo ao longo do tempo

Fonte: elaboração do autor.

O gráfico 2 evidencia a trajetória ao longo do século XX dos contextos de animacidade e referencialidade do antecedente do objeto nulo. Assim como afirmam as numerosas pesquisas sobre o tema (CYRINO, 1994, 1997, 2000; DUARTE, 1989; SOLEDADE, 2011; PEIXOTO, 2021), o objeto nulo do PB é preferencialmente [- animado, + específico], (42). Essa tendência se confirma no gráfico, visto que, nos três períodos analisados, esse contexto se mantém como o mais expressivo para a preferência do objeto nulo, mesmo tendo uma queda percentual do primeiro ao último períodos: saindo de 42,7% e chegando a 32,7%. O antecedente proposicional, o primeiro contexto diacrônico a licenciar o objeto nulo (CIRYNO, 1994), que já apresentou índices mais altos em séculos passados, embora ainda se mantenha presente, cai de 28% a 21%, (43). Isso talvez se explique pelo uso do pronome demonstrativo neutro ‘isso’ como retomada proposicional (PEIXOTO, 2021).

(42) eu recebi [a tua cartinha]_i no dia 22 as 2 horas da tarde eu quando recebi fiquei tam comtente mais as saudades au mentarã muito fiquei tam comtente mais eu fiquei lendo (0)_i na estação na quele banco. (JM, 1901-1925)

(43) [é assim que escreve]_i?! Não sei (0)_i, rs (MV, 1976-2000)

Os contextos [- animado, - específico], (44), e [- animado, + específico], (45), parecem ter o comportamento complementar ao longo do tempo, ainda que necessariamente não o sejam.

(44) Quanto a[o esteptocópio]_i eu não vi (0)_i ainda. (SL, 1950-1975)

(45) Ah! Ela adorou [o cordão]_i. Não para de mostrar (0)_i para os outros. (SL, 1950-1975)

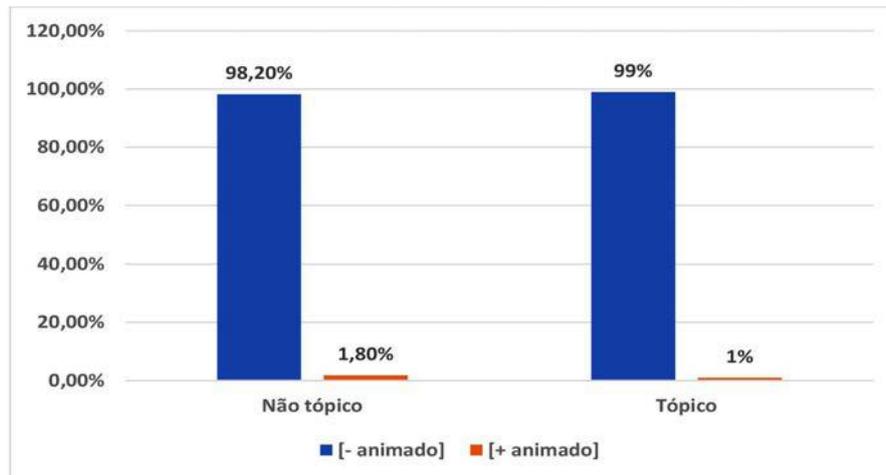
Quanto aos antecedentes [+ animados], independentemente da especificidade, têm baixíssimos índices percentuais, não passando, no segundo período, de 2,4% nos [+ específicos], contexto inovador do objeto nulo (CYRINO, 1994,1997). Dessa forma, os antecedentes [+ animados], embora ocorram, ainda são um contexto de resistência para o objeto nulo no PB, o que confirma os resultados para a língua oral (DUARTE, 1989). Mesmo não ocorrendo no último período, isso não é sinônimo de agramaticalidade, como nos dizem os dados de intuição, apenas reforça a resistência desse contexto à implementação do objeto nulo.

Interessante, ainda, observar que o período de 1950-1975 apresenta a queda percentual da maioria dos contextos de objeto nulo. Só há aumento do contexto [- animado, - específico], e o surgimento dos contextos [+ animados, +/- específicos]. Esse cenário pode ter sido causado tanto pela natureza da escrita da família de missivistas desse período, como também por fatores linguísticos, como o aumento dos pronomes lexicais na posição de objeto direto (que, por mais que sejam mais estigmatizados na escrita, aparecem como uma opção deste período), e outros que fogem ao escopo deste trabalho. Mais estudos são, portanto, necessários.

Esses resultados, então, fornecem uma amostra da relação entre animacidade e referencialidade ao longo do tempo.

3.3. Animacidade e o tópico estrutural

Na proposta de Cyrino (2000, 2019), a autora levanta a questão de que talvez o objeto nulo no PB que tenha um antecedente [+ animado], sobretudo [+ animado, + específico], seja devido ao fato de esse antecedente ser um tópico. Entendemos tópico estrutural neste trabalho como o resultado ou de um movimento de um constituinte interno à oração para a periferia esquerda da sentença, como a topicalização, ou um constituinte gerado nessa mesma posição, como as estruturas de tópico pendente (BRITO, DUARTE, MATOS, 2003). Assim, o resultado da topicalização é uma categoria vazia, um objeto nulo nesse caso, na posição *in situ* dentro da oração. No caso do tópico pendente, o referente tópico serve de antecedente para a categoria vazia na posição de objeto direto dentro da oração. Como dissemos, não há consenso na literatura sobre o tópico – até mesmo o tópico que é resultado de movimento – ter papel temático ou não. O que nos interessa nesta análise é em relação à animacidade do referente, não ao papel temático do constituinte, havendo ou não. São discussões diferentes, que fogem ao escopo deste trabalho. Nesse sentido, controlamos o tipo do antecedente entre tópico vs. não tópico, cruzando essa informação com a animacidade do antecedente.

Gráfico 3: animacidade do antecedente e posição de tópico e de não tópico em cartas pessoais

Fonte: elaboração do autor.

A partir da análise do gráfico 3, vemos que aparentemente não há diferenças percentuais entre os antecedentes tópicos e não tópicos (dentre os quais estão os adjuntos) e a animacidade do antecedente. Com os não tópicos, 98% são de antecedente [- animado], (46), e menos de 2% de antecedente [+ animado], (47). Com os tópicos, a tendência é a mesma: 99% de antecedentes [- animados], (58), contra 1% de antecedente [+ animado], (49). Esses resultados confirmam, basicamente, a natureza preferencialmente [- animada] do objeto nulo do PB, independentemente do estatuto, natureza e posição do seu antecedente.

(46) eu falei [a minha frase]_i e ela entendeu (0)_i. (SL, 1950-1975)

(47) [Eles]_i vão me chamar para ir para o hospital quando o bebê for nascer e Mom falou que se o bebê nascer dentro do carro eu estou lá para ajudar (0)_i. (SL, 1950-1975)

(48) Quanto a[o dinheiro]_i, você manda (0)_i para mim. (JM, 1901-1925) – *Tópico Pendente sem retomada fonética*

(49) [Os brasileiros que moram em Bloomington]_i eu não conheço (0)_i ainda. (SL, 1950-1975) – *Topicalização*

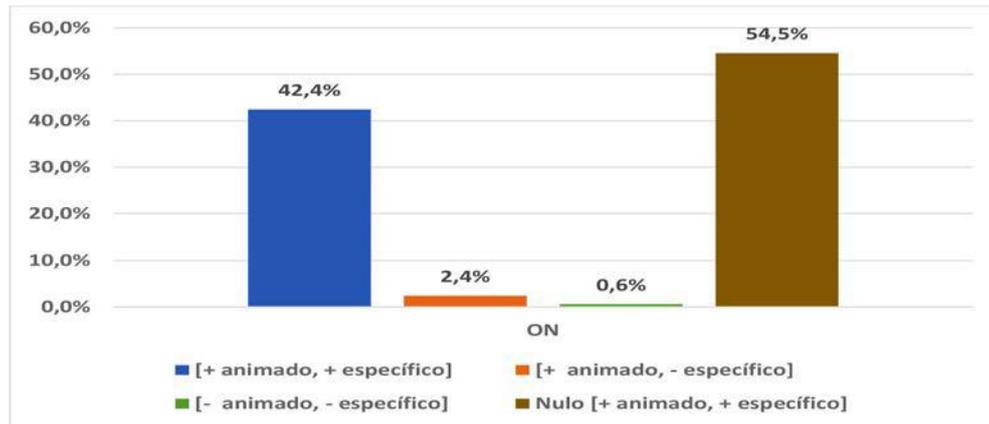
Esses resultados parecem indicar que os objetos nulos com antecedentes animados não são licenciados pelo fato de os antecedentes estarem em uma posição de tópico. Consequentemente, a hipótese de Cyrino (2019) precisar ser reformulada de modo a acomodar melhor o objeto nulo com antecedente animado e específico.

3.4. Animacidade do antecedente do objeto nulo e a do sujeito da sua oração

Em Peixoto (2021), o preenchimento do sujeito é um fator que favorece (do ponto de vista gramatical) o objeto nulo em detrimento do clítico acusativo. Além disso, a partir do trabalho de Tarallo (1993) e de outros trabalhos na mesma esteira, são levantadas questões sobre uma possível

interferência da animacidade do sujeito – preenchido ou nulo – da oração que contém o objeto nulo e da animacidade do seu referente no seu licenciamento. Nesse sentido, analisamos os seguintes contextos de animacidade do sujeito da oração que contém o objeto nulo, que foram os que apareceram nos dados: [+ animado, + específico], [+ animado, – específico], [– animado, – específico], nulo [+ animado, + específico]. Excluimos os dados com sujeito: PRO, PRO/pro, e verbo impessoal, por não possuírem uma contraparte foneticamente realizada.

Gráfico 4: preenchimento ou não do sujeito de acordo com a animacidade nas sentenças com objeto nulo



Fonte: elaboração do autor.

Os resultados do gráfico 4 apontam uma distribuição bem irregular entre os fatores do grupo. Os contextos que mais favoreceram o objeto nulo foram os de sujeito [+ animados, + específicos], independentemente de estarem preenchidos, (50), ou nulos, (51), com 35% e 43%, respectivamente.

(50) Tu me pediste o retrato que te amostrei-te no Domingo. [este retrato]_i eu dei (0)_i a minha irman ja a muito tempo. (JM, 1901-1925)

(51) não repare [a minha carta]_i e (*pro*) rasge (0)_i outraves. (JM, 1901-1925)

Os sujeitos [+ animado, – específico], (52), e [– animado, – específico], (53), não são muitos diferentes percentualmente, mesmo com 2,4% do primeiro contra 0,6% do segundo.

(52) Falei de voce e ela disse que [toda e qualquer menina virgem que vai para la]_i volta totalmente diferente, pois lá o pessoal não dispensa (0)_i. (SL, 1950-1975)

(53) Vê se pode pode arranjar com o Helio Lourença [um atestado dental]_i. A escola está me pedindo (0)_i. (SL, 1950-1975)

Sobre o dado (53), é interessante dizer que consideramos ‘escola’ como [– específico], pois ela seria uma metáfora para alguém que trabalha no recinto; podendo, inclusive também ser [– animado] caso seja uma metáfora para ‘o sistema’, ‘a burocracia da escola’, dentre outras interpretações.

O que nossos resultados, de acordo com nosso *corpus*, sugerem é que os sujeitos das orações que selecionam o objeto nulo são quase categoricamente [+ animados]. É interessante comparar esses resultados com os obtidos por Peixoto (2021). Analisando os fatores de favorecimento do objeto nulo em detrimento do clítico acusativo de terceira no mesmo *corpus*, o programa estatístico *GoldVarb X* (SANKOFF, TAGLIAMONTE, SMITH, 2005) selecionou o preenchimento do sujeito (vs. sujeito nulo) como o segundo fator – logo depois da animacidade [+ proposicional] – de maior favorecimento (Peso Relativo = 0,639 vs. 0,427 de sujeito nulo, não significativo) para o objeto nulo. Os resultados percentuais deste trabalho, por outro lado, mostram que os sujeitos nulos [+ animados] têm uma leve vantagem em relação ao sujeito preenchido; sendo assim, como explicar essa aparente contradição?

Podemos pensar que, em uma relação de categoria vazia com clítico, este aumenta as chances de haver um sujeito nulo, por isso o preenchimento do sujeito está ligado do objeto nulo. Assim, nessa relação, o sujeito nulo diminui as chances de haver objeto nulo, visto que esse sujeito está, aparentemente, ligado ao clítico acusativo. Já na análise levando em conta apenas o objeto nulo, não é necessariamente o preenchimento ou não do sujeito da oração que o contém que o favorece, mas a animacidade do seu referente, embora pareça haver, de fato, uma leve preferência do sujeito nulo [+ animado, + específico] para a realização do objeto nulo.

Conclusões

Neste trabalho, fizemos algumas reflexões sobre um tema bem debatido na literatura gerativista e variacionista sobre os efeitos da animacidade no favorecimento ou restrição do objeto nulo no Português brasileiro.

Por meio dos nossos resultados, testamos algumas hipóteses relacionadas e derivadas do trabalho de Peixoto (2021). O papel temático de tema parece ter uma grande influência na relação entre o antecedente e o objeto nulo. Mais que função sintática, a identidade de papel temático entre o referente e a categoria vazia favorece, em geral, a formação de uma cadeia temática. Em consonância com o esperado pela Escala de Referencialidade, o objeto nulo está ampliando seus contextos de realização ao longo do tempo, ou seja, embora percentualmente pequenas, há possibilidade da existência de objetos nulos [+ animado, + específicos]. A sua não realização no último período não significa impossibilidade, apenas desfavorecimento, como já dissemos. Um tópico [+ animado] não favorece necessariamente o licenciamento do objeto nulo.⁸ Pelo contrário, tanto em posição de tópico estrutural quanto em posição interna à sentença, o referente do objeto nulo é quase categoricamente [- animado]. Uma questão a se perguntar para trabalhos futuros é se a seleção temática do verbo poderia estar associada ao tipo de referente do seu complemento. Um sujeito, realizado ou não, [+

⁸ O Objeto Nulo resultado da Topicalização no PB também foi considerado como *pro*, mas ainda é preciso definir, em relação a esse contexto, os licenciamentos e a natureza desse *pro*: se um N(ome) nulo, um D(eterminante) nulo. Essa descrição e análise são importantes para entendermos, além dos fenômenos da periferia esquerda e sua relação com o interior da sentença-comentário, as diferenças da Topicalização e do seu consequente Objeto Nulo *pro* do PB de outras línguas românicas, como o Português europeu, em que também é possível a Topicalização, mas cujo objeto nulo em outros contextos seja considerado uma *variável*.

animado, + específico] parece estar presente na maioria dos predicados que autorizam o objeto nulo, o que poderia questionar Peixoto (2021), se não fossem análises diferentes.

A partir de Peixoto (2021), os fatores mais significativos para o licenciamento do objeto nulo foram a animacidade do antecedente, o preenchimento da posição de sujeito da oração que seleciona o objeto nulo e o período de nascimento do missivista. Com os nossos resultados, vimos por outro ângulo a relação do preenchimento da posição do sujeito com o objeto nulo. Assim, como, em geral, as orações com sujeito nulo têm um referente [+ animado], quando tiramos esse traço e nos detemos apenas no aspecto estrutural (preenchimento ou não), temos mais sujeitos preenchidos ligados ao objeto nulo, em comparação com o clítico, já que o PB tem se mostrado uma língua de sujeito cada vez mais preenchido (BERLINCK, 1989; COELHO, 2000; ROBERTS, 2016; VERÍSSIMO, 2017) e com cada vez menos clíticos acusativos de terceira pessoa (DUARTE, 1989; TARALLO, 1993; CYRINO, 1994, 1997; KATO, CYRINO, DUARTE, 2000).

Em uma relação muito próxima ao que Tarallo (1993) adianta do contínuo preenchimento do sujeito e do paulatino apagamento fonético do objeto no PB, é interessante que essa tendência pareça similar à animacidade nas duas estruturas. Enquanto o sujeito da oração que contém o objeto nulo tenda a ser [+ animado, + específico], o objeto nulo tende a ser [- animado, +/- específico], o que confirma a aparente assimetria e relação inversamente proporcional entre as duas estruturas argumentais.

Tanto para hipótese de natureza *pro*, quanto para a de natureza elipse, a animacidade parece não ser um requisito sintático, mas semântico, relacionado à interpretação, que satisfaz a interface perceptual-conceptual no licenciamento do objeto nulo, que, conforme Galves (2001), pode ser descrita como uma restrição de logoforicidade. Podemos dizer que não se trata de um requerimento sintático, pois as estruturas argumentais e frásicas que contêm o objeto nulo podem ser as mesmas, mas, com diferentes graus de animacidade, podem ter diferentes resultados de (a)gramaticalidade. Vale lembrar ainda que sob determinados contextos discursivos, como o referente do objeto nulo ser também o tópico discursivo, as sentenças podem variar quanto ao julgamento de (a)gramaticalidade. Isso sugere que essas construções são possíveis na sintaxe, pois, de acordo com Chomsky (1977), o discurso não pode tornar gramatical uma sentença que é agramatical para a sintaxe; logo, se o discurso torna as sentenças mais aceitáveis, significa que a sintaxe permite tais derivações.

Além disso, o fato de as construções com objeto nulo com referente [+ animado] serem possíveis – mesmo com baixíssimas proporções – em alguns períodos do PB sugere que não se trata de uma falta de licenciamento sintático, pelo menos não desde o século XVI (CYRINO, 1994, pp. 172-4). Ainda, alguns trabalhos (MARAFONI, 2004; PEIXOTO, 2021), que consideram gramaticais algumas sentenças com objeto nulo com antecedente animado, que são tratadas por outros autores como agramaticais, defendem um enfraquecimento da restrição de animacidade [+ animado] do objeto nulo. Esses indícios sugerem que, de fato, estamos diante de uma restrição interpretativa do PB, não estritamente sintática. Nossos resultados reforçam, portanto, que a natureza do objeto nulo no PB parece não ser a de uma elipse, mas de um pronome, um *pro*, tal como a proposta de Galves

(1989, 2001), visto que as impossibilidades gramaticais impostas pelas teorias da elipse, ainda que percentualmente baixas, são possíveis, como a existência do objeto nulo [+ animado] e da aparente não relação entre o objeto nulo [+ animado] ter seu antecedente em posição de tópico. Se suas restrições fossem, de fato, estruturais, essas construções não seriam gramaticais em determinados períodos de tempo, como mostram os dados. As restrições ao objeto nulo são, então, mais semânticas que sintáticas/estruturais.

Se a animacidade não é um traço legível à computação sintática, como justificamos o favorecimento, inclusive atestado por peso relativo, dos contextos [proposicionais] e [- animados, +/- específicos] para o objeto nulo? Podemos pensar em dois fatores que, conjugados, podem explicar esse resultado. Primeiro que estamos lidando com dados de Língua-E para chegarmos aos parâmetros da Língua-I; assim, lidamos sempre com dados positivos, ainda mais em se tratando de corpus histórico, não tendo, então, dados negativos de objeto nulo. As análises empíricas, baseadas em intuição e em corpora, mostram apenas que a restrição de antecedente [+ animado] do objeto nulo está/é enfraquecida no PB. Segundo que, sendo uma restrição semântica, como mostramos, do antecedente [+ animado] do objeto nulo, podemos imaginar que os dados desse contexto a que temos acesso sejam mais escassos, mas não impossíveis, em relação aos dados de antecedente [proposicional] e [- animado, +/- específico]. Dessa forma, é possível pensar que a animacidade [proposicional] e [- animada, +/- específica] do antecedente do objeto nulo faça diferença estatística; e por isso também é selecionada pelo programa estatístico.

Com isso, os traços de animacidade e referencialidade não sendo gramaticais, nesse sentido, não são lidos pela computação sintática, que é cega para a semântica, mas se relacionam com a formação de uma boa cadeia tópica, associando-se a questões semânticas da interface com a interpretação, com o que Galves (2001), baseada em Farrell (1987), chama de restrição semântico-discursiva de logoforicidade, como dissemos mais acima.

Portanto, o objeto nulo do PB ainda tem muitos caminhos a serem traçados. Nosso trabalho trouxe alguns indícios de respostas e levantou outras questões. Assim, a relação entre animacidade e objeto nulo não está próxima de ser resolvida.

Referências

CYRINO, Sônia Maria Lazzarini. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: KATO, M.; ROBERTS, I. (orgs.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Ed. Unicamp, 1994, pp. 163-84.

CYRINO, Sônia Maria Lazzarini. *O ON no português do Brasil – um estudo sintático-diacrônico*. Londrina: Editora UEL, 1997.

CYRINO, Sônia Maria Lazzarini. Para a história do português brasileiro: observações sobre a presença de complementos verbais nulos e ausência de clíticos de 3ª. pessoa. In: XV Congresso da ANPOLL, 2000, Niterói. *Resumos dos trabalhos do Grupo de Teoria da Gramática*, 2000.

- CYRINO, Sônia Maria Lazzarini. Complementos nulos em anúncios de jornal do século XIX. Em ALKMIM (org.) *Para a história do português brasileiro*. Vol. III. São Paulo, Humanitas, 2002, pp. 221-45.
- CYRINO, Sônia Maria Lazzarini. Para a história do Português brasileiro: a presença de objetos nulos e a ausência de clíticos. *Letras de Hoje*. Porto Alegre, v. 38, n. I, 2003, pp. 31-47.
- CYRINO, Sônia Maria Lazzarini. O problema da experiência detonadora na mudança sintática do português brasileiro. *Estudos Lingüísticos* (São Paulo), Campinas, v. XXXIII, 2004.
- CYRINO, Sônia Maria Lazzarini. Anáfora do complemento nulo na história do Português brasileiro. In: LOBO, Tania ... [et al.] (orgs). *Para a história do português brasileiro*. Salvador: EDUFBA, 2006. v. 6, t. 1-2.
- CYRINO, Sônia Maria Lazzarini, MATOS, Gabriela. Anáfora do Complemento Nulo: anáfora profunda ou de superfície? Evidência do Português Brasileiro e Europeu. *Letras de Hoje*. Porto Alegre. v. 41, n. 1, março, 2006, pp. 121-41.
- CYRINO, Sônia Maria Lazzarini. Objetos nulos no português brasileiro e espanhol da América do Sul. *Relatório final para a Fapesp*. Processo no. 06/03343-2, 2008.
- CYRINO, Sônia Maria Lazzarini. On richness of tense and verb movement in Brazilian Portuguese. In: CAMACHO-TABOADA, Maria Victoria et al. (org.). *Information structure and agreement*. 1. ed. Amsterdam: John Benjamins, 2013, pp. 297-317
- CYRINO, Sônia Maria Lazzarini. O objeto nulo. In: CYRINO, Sonia; TORRES MORAIS, Maria A. (org.). *Mudança sintática do português brasileiro: perspectiva gerativista*. 1. ed. São Paulo: Contexto, 2018, v. 6, pp. 210-51.
- CYRINO, Sônia Maria Lazzarini. O objeto nulo no português brasileiro: sincronia e diacronia. In: GALVES, C.; KATO, M.; ROBERTS, I. (orgs.). *Português brasileiro: uma segunda viagem diacrônica*. São Paulo: Editora Contexto, 2019, pp. 173-99.
- CYRINO, Sônia Maria Lazzarini; DUARTE, Maria Eugênia Lammoglia; KATO, Mary Aizawa. Visible subjects and invisible clitics in Brazilian Portuguese. In: KATO, Mary A.; NEGRÃO, E. V. (eds.) *Brazilian Portuguese and the Null Subject Parameter*. Madrid: Iberoamericana; Frankfurt am Main: Vervuert, 2000, pp. 55-71.
- DUARTE, Maria Eugênia Lammoglia. Clítico acusativo, pronome lexical e categoria vazia no Português do Brasil. In: TARALLO, F. (org.) *Fotografias Sociolinguísticas*. Campinas: Pontes, 1989, pp. 19-34.
- FARRELL, Patrick. *Empty objects in brazilian Portuguese*. Mimeo, 1987. pp.1-38.
- FARRELL, Patrick. Null Objects in Brazilian Portuguese. *The Linguistic Review* 8, 1990, pp. 325-46.
- FIGUEIREDO SILVA, Maria Cristina de. *Objeto direto anafórico no dialeto rural afro-brasileiro*, 2004. Dissertação de mestrado, UFBA, 2004.
- FIGUEIREDO SILVA. *O objeto nulo no português rural baiano*. Teoria temática e eclipse de DP. Tese de Doutorado. Salvador: UFBA, 2009.
- FREIRE, Gilson Costa. *Os Clíticos de Terceira Pessoa e as Estratégias para sua Substituição na Fala*, 2000. Dissertação (Mestrado em Língua Portuguesa) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2000.

FREIRE, Gilson Costa. *A realização do acusativo e do dativo anafóricos de terceira pessoa na escrita brasileira e lusitana*. Tese de Doutorado em Língua Portuguesa. Faculdade de Letras, UFRJ, Rio de Janeiro, 2005.

GALVES, Charlotte. Pronomes e categorias vazias em português do Brasil. *Cadernos de Linguística*, Unicamp, v. 7, 1984, pp. 107-36.

GALVES, Charlotte. A gramática do português brasileiro. *Línguas e Instrumentos linguísticos*. Campinas: Pontes Editores, v. 1, 1989, pp. 79-96.

GALVES, Charlotte. *Ensaio sobre as gramáticas do português*. Campinas. Editora da Unicamp, 2001.

HUANG, C-T. James. On the determination and reference of empty pronouns. *Linguistic Inquiry*, v. 15: 1984. pp. 531-74.

KATO, Mary Aizawa; RAPOSO, Eduardo. O objeto nulo definido no português europeu e no português brasileiro: convergências e divergências. In: CORREIA, Clara Nunes; GONÇALVES, Anabela (orgs.) *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa, Associação Portuguesa de Linguística, 2001, pp. 673-84.

KOCH, Peter; & OESTERREICHER, Wulf. Oralidad y escrituralidad a la luz de la teoría del lenguaje. In: KOCH, Peter; OESTERREICHER, Wulf; LÓPEZ SERENA, Araceli. *Lengua hablada en la Romania: español, francés, italiano*. Madrid, Gredos, 1990.

MARAFONI, Renata Lopes. *A Realização do objeto direto anafórico: um estudo em tempo real de curta duração*. Dissertação (Mestrado em Letras Vernáculas - Língua Portuguesa). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.

MARQUES DE SOUSA, Antonio. *As realizações do acusativo anafórico no Português europeu e brasileiro: um estudo diacrônico*. Dissertação de Mestrado em Letras Vernáculas (Língua Portuguesa). Universidade Federal do Rio de Janeiro, Faculdade de Letras, 2017.

MATOS, Gabriela. *Clítico Verbal Demonstrativo*. Monografia apresentada em Provas de Aptidão Pedagógica e Capacidade Científica. Lisboa: Faculdade de Letras da Universidade de Lisboa, 1985.

PEIXOTO, Gabriel Guimarães. “*Pega essa carta e rasga*”: a implementação do objeto nulo a partir do acusativo anafórico de 3ª pessoa no Português brasileiro, 2021. Dissertação de Mestrado em Letras Vernáculas (Língua Portuguesa). Universidade Federal do Rio de Janeiro, Faculdade de Letras, 2021.

TARALLO, Fernando. Diagnosticando uma Gramática Brasileira: O Português d’aquém e d’além-mar ao final do século XIX. In: ROBERTS, Ian; KATO, Mary (orgs.). *Português Brasileiro: uma viagem diacrônica*, Campinas: Editora da Unicamp, 1993, p. 69-105.

SANKOFF, David; TAGLIAMONTE, Sali; SMITH, Eric. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

SOLEDADE, Carolina de La Vega. *A realização do objeto direto anafórico em peças de autores brasileiros dos séculos XIX e XX: dados empíricos para a observação da mudança no português brasileiro*, 2011. Dissertação de mestrado, UFRJ, 2011.

PRONOMES PLENOS E CLÍTICOS PRONOMINAIS NO PORTUGUÊS BRASILEIRO E NO FRANCÊS

STRONG PRONOUNS AND CLITIC PRONOUNS IN BRAZILIAN PORTUGUESE AND FRENCH

Luciano de Oliveira¹

RESUMO

Pronomes clíticos correspondem, geralmente, a argumentos verbais e dependem de uma base verbal para se apoiar, sendo esta uma das causas para não poderem ocorrer como elementos isolados em uma frase, ao contrário do que pode ser verificado para os pronomes plenos e os DPs e PPs. Ainda em relação a estes elementos, que podem representar igualmente argumentos verbais, os clíticos pronominais apresentam outros comportamentos sintáticos próprios, como a impossibilidade de movimento para uma posição de tópico. Considerando os pronomes clíticos, o português brasileiro (PB) apresenta divergências em relação à gramática tradicional, a qual toma como norma a variedade europeia da língua; uma dessas diferenças é a perda dos pronomes clíticos acusativos de terceira pessoa, os quais não pertencem mais à língua-I da(do) falante brasileira(o). No francês, entretanto, os clíticos pronominais de terceira pessoa são muito produtivos, independentemente do caso sintático considerado. Também diferentemente do PB, o francês possui clíticos pronominais nominativos, locativos e partitivo/genitivo. Além disso, até mesmo na língua falada, combinações de pronomes clíticos são usadas no francês, o que não se verifica no PB. Todas essas diferenças podem representar dificuldades para a aquisição de francês por brasileiras(os).

PALAVRAS-CHAVE: Pronomes clíticos. Português Brasileiro. Francês.

ABSTRACT

Clitic pronouns generally correspond to verbal arguments and depend on a verbal base to support them, which is one of the reasons why they cannot occur as isolated elements in a sentence, contrary to what can be verified for strong pronouns and DPs and PPs. Still in relation to these elements, which can also represent verbal arguments, pronominal clitics present other syntactic behaviors of their own, such as the impossibility of moving to a topic position. Considering the clitic pronouns, Brazilian Portuguese (BP) presents divergences from the traditional grammar, which takes the European variety of the language as a norm; one of these differences is the loss of third-person accusative clitic pronouns, which no longer belong to the Brazilian speaker's I-language. In French, however, third-person pronominal clitics are very productive, regardless of the syntactic case considered. Also unlike BP, French has nominative, locative and partitive/genitive pronominal clitics. Furthermore, even in spoken language, clusters of clitic pronouns are used in French, which is not the case in BP. All these differences may represent difficulties for Brazilians to acquire French.

KEYWORDS: Clitic pronouns. Brazilian Portuguese. French.

Introdução

De acordo com Dubois *et al.* (2006 [1978], p. 489), pronomes são “palavras que se empregam para reenviar a ou substituir uma outra já utilizada no discurso (emprego anafórico), ou para representar uma(um) participante na comunicação, um ser ou um objeto presentes no momento do enunciado (emprego dêitico)”. No seu uso anafórico, os pronomes costumam servir à substituição de DPs (ou de PPs), mas também podem substituir adjetivos ou sentenças, como nos exemplos dados a seguir em (1) e (2).

¹ Doutor pelo Programa de Pós-graduação em Linguística pela Universidade Federal de Santa Catarina, lucioliv1975@gmail.com, <https://orcid.org/0000-0001-6628-6409>.

- (1) a. Tu és [corajoso]_i?
 b. Sim, eu [o]_i sou.
- (2) a. Vais [escrever à tua mãe]_j?
 b. Pretendo fazê-_j[lo].

(exemplos (1) e (2): adaptados de DUBOIS *et al.* (2006 [1978], p. 489))

Neste trabalho, são considerados os pronomes pessoais, pois é nessa subclasse que se encontram os pronomes plenos e os clíticos pronominais que serão aqui analisados nas línguas envolvidas: o português brasileiro (doravante, PB) e o francês. Para tanto, inicia-se com as propriedades dos pronomes clíticos, passando-se então para o inventário de pronomes pessoais nas duas línguas em questão. Na sequência, é fornecido um quadro com a equivalência desses clíticos nas línguas. Ao final do artigo, tem-se algumas considerações finais relacionando o tema do artigo à aquisição/aprendizagem de francês por brasileiras(os), assim como as referências bibliográficas utilizadas.

1. Os clíticos pronominais

Os clíticos pronominais constituem uma (sub)classe de palavras dentro do que a gramática tradicional (GT) chama de pronomes pessoais átonos (DUBOIS *et al.*, 2006 [1978], pp. 112-3). Considerando-se a Gramática Gerativa, clíticos pronominais podem ser os “pronomes” (que correspondem, na GT, aos pronomes átonos não reflexivos) ou as “anáforas” (pronomes átonos reflexivos/recíprocos).

Martins (2013, pp. 2231-2) define clítico como um elemento lexical sem acento prosódico (como os afixos e contrariamente às palavras), mas com alguma liberdade de posição (como as palavras e contrariamente aos afixos). A ausência de acento faz com que o clítico dependa de uma palavra adjacente acentuada, denominada “palavra hospedeira”, ou “hospedeiro” do clítico. “Cliticização” é o processo de ligação do clítico ao seu hospedeiro.² Seguindo essa definição, são também clíticos do português “os artigos definidos, os pronomes interrogativos *que* e *porque*, o pronome relativo

² Autores como Anderson (2005, pp. 9-14) e Caink (2005, pp. 492-3) apresentam critérios que permitem diferenciar palavras, clíticos e afixos. Apesar da grande variação interlinguística, a ordem dos afixos em uma palavra é bastante rígida, já as palavras possuem uma maior liberdade de posicionamento em uma frase, nem sempre relacionada a mudanças de significado; afixos são formas presas, devendo ocorrer associadas a outras, enquanto que palavras são formas livres, podendo mesmo aparecer sozinhas em uma frase ou sintagma; afixos não podem receber acento prosódico, o que é possível em relação às palavras; afixos estão sujeitos a processos fonológicos, o que não tende a acontecer com palavras adjacentes. Assim, algumas características dos clíticos pronominais, como a ausência de acento prosódico e a relativa liberdade de posicionamento em relação à sua base (como a próclise e a ênclise), permite situar essa categoria entre os afixos e as palavras.

De acordo com Câmara Jr. (1999 [1970]), formas livres são aquelas que podem ocorrer isoladas (como “releia” em “O que faço com esse livro? – Releia.”). Já formas presas só funcionam ligadas a outras e sem possibilidade de variação de posição em relação a estas (como o prefixo “re-” em “releia”). Entre as duas, o autor define as formas dependentes, que são aquelas que, apesar de necessitarem de uma base, podem encontrar-se separadas desta por outra(s) forma(s) livre(s) (como o artigo “a” em “a menina” e “a bela menina”) ou possuem liberdade de posição em relação à sua base (como o pronome clítico “te” em “te quero” e “quero-te”).

que, as conjunções *que*, *se*, *mas*, *e*, *ou*, as preposições *de*, *para*, *por*, *com*, *em* e o quantificador *cada*” (MARTINS, 2013, pp. 2231-2, itálicos da autora). Entretanto, nas línguas românicas, a maioria dos clíticos são pronominais (TSEDRYK, 2014, p. 7). Por esta razão, neste trabalho, são considerados somente os clíticos pronominais, que, por selecionarem como hospedeiro um verbo, são chamados também de “clíticos verbais”.

Cardinaletti (1999) indica propriedades que alguns pronomes devem apresentar para serem classificados como clíticos. Com base nessas propriedades, Kanthack (2002, pp. 6-21) apresenta uma série de testes em PB para comprovar que o comportamento dos clíticos pronominais é diferente do dos argumentos verbais internos expressos por DPs ou PPs plenos, ou ainda por pronomes tônicos. Esses testes, realizados a seguir com exemplos do PB (mas válidos igualmente para o francês), permitem confirmar as três principais conclusões indicadas por Kanthack (2002) em relação às propriedades dos clíticos pronominais, as quais são indicadas a seguir.

i. Os clíticos pronominais não se comportam como DPs (ou PPs) lexicais ou pronomes não clíticos.

i.a. Os clíticos pronominais não podem ocorrer em posições argumentais: por exemplo, na posição de argumento interno do verbo, o pronome clítico provoca a agramaticalidade da sentença, como exemplificado em (3);

- (3) a. O Carlos ajuda sempre **a Joana**.³
 b. O Carlos ajuda sempre **ela**.
 c. *O Carlos ajuda sempre **a**.

i.b. Os clíticos pronominais não podem ocorrer em posições não argumentais: um exemplo de posição não argumental é a posição de tópico, indicada em (4), onde um pronome clítico não é aceitável;

- (4) a. **A Joana**, o Carlos ajuda sempre.
 b. **Ela**, o Carlos ajuda sempre.
 c. ***A**, o Carlos ajuda sempre.

i.c. Os clíticos pronominais não podem ocorrer como itens isolados: não se pode responder a uma pergunta, por exemplo, usando somente um pronome clítico, como indicado em (5c);

- (5) a. O Carlos ajuda sempre quem? **A Joana**.
 b. O Carlos ajuda sempre quem? **Ela**. / **Você**.
 c. O Carlos ajuda sempre quem? ***A**. / ***Te**.

³ Nos conjuntos de sentenças (3) a (11), os elementos de interesse encontram-se em negrito, e advérbios foram inseridos em (3) para deixar as posições argumentais mais evidentes – sem o advérbio “sempre”, a sentença (3c) poderia ser considerada gramatical, mas o clítico não ocuparia a posição argumental.

i.d. Os clíticos pronominais não podem ser coordenados: a coordenação não pode ocorrer nem entre pronomes clíticos, nem entre um pronome clítico e outro elemento, o que é indicado pelas sentenças agramaticais em (6);

- (6) a. O Carlos deu o presente para o **Antônio** e **a Joana**.
 b. O Carlos deu o presente para **mim** e **a Joana**.
 c. O Carlos deu o presente para **mim** e **ela**.
 d. *O Carlos deu-**me** e **para a Joana** o presente.
 e. *O Carlos deu-**me** e **para ela** o presente.
 f. *O Carlos **me** e **lhe** deu o presente.

i.e. Os clíticos pronominais não podem receber acento contrastivo: como mostra (7c), eles não podem veicular foco contrastivo, recebendo, assim, acento;

- (7) a. O Carlos deu o presente **PARA O ANTÔNIO**, e não **PARA A JOANA**.
 b. O Carlos deu o presente **PARA ELE**, e não **PARA ELA**.
 c. *O Carlos **ME** deu o presente, e não **LHE** deu.⁴

i.f. Os clíticos pronominais não podem ser modificados: como se vê em (8c), eles não podem ser precedidos por advérbios associados ao foco, como “somente” e “também”.

- (8) a. O Carlos ajuda somente **a Joana**.
 b. O Carlos ajuda somente **ELA**.
 c. *O Carlos ajuda somente **A**.

Após demonstrar que as propriedades dos argumentos lexicais e dos pronomes não clíticos não se aplicam aos clíticos pronominais, falta fazer o oposto: verificar que os pronomes clíticos não compartilham as suas propriedades com aqueles elementos. Isso é realizado a seguir.

ii. Os DPs (e os PPs) lexicais e os pronomes não clíticos não se comportam como os pronomes clíticos.

ii.a. A posição do clítico pronominal objeto antes de um verbo finito é imprópria para um DP (ou PP) lexical ou para um pronome não clítico: somente um clítico pronominal pode ocupar a posição pré-verbal, como mostrado em (9);⁵

⁴ Para algumas(alguns) falantes nativas(os) de PB, sentenças como (7c) podem ser aceitáveis, mesmo com a acentuação dos clíticos (KANTHACK, 2002, p. 11).

Hvidsten em Helland (2018, p. 63) mencionam que, em francês, a única situação em que um clítico pode ser acentuado (contrastivamente ou não) é no imperativo afirmativo (com ênclise), como em *Prends-le*, [prã'lə], “Pegue-o”.

⁵ Como “posição pré-verbal” está sendo considerada aqui a posição proclítica ocupada pelo clítico pronominal não nominativo, adjacente ao verbo; não se confunda essa posição com a posição de sujeito, que aparece geralmente antes do verbo. Isso porque, por exemplo, entre o sujeito e o verbo, pode haver outros elementos, como em “Ele sempre compra os mesmos produtos”.

- (9) a. O Carlos **a** ajuda sempre.
 b. *O Carlos **a Joana** ajuda sempre.
 c. *O Carlos **ela** ajuda sempre.⁶

ii.b. A posição intermediária de um grupo verbal ocupada pelo clítico pronominal também é imprópria para um DP (ou PP) lexical ou para um pronome não clítico: somente um pronome clítico pode ocupar a posição entre dois verbos que compõem um grupo verbal, o que é exemplificado em (10);

- (10) a. O Carlos quer **me** ajudar sempre.
 b. *O Carlos quer **o Antônio** ajudar sempre.
 c. *O Carlos quer **ele** ajudar sempre.

ii.c. A posição pós-verbal ocupada pelo clítico pronominal é diferente daquela ocupada por um DP (ou PP) lexical ou por um pronome não clítico: os DPs (ou PPs) lexicais têm maior liberdade de posição do que os pronomes clíticos, pois estes devem permanecer adjacentes à sua base verbal, como é possível perceber em (11) a seguir.

- (11) a. Ajuda **o Antônio** sempre!
 b. Ajuda sempre **o Antônio**!
 c. Ajuda **ELE** sempre!
 d. Ajuda sempre **ELE**!
 e. Ajuda-**o** sempre!
 f. *Ajuda sempre **o**!

Até este ponto, foi possível comprovar que argumentos lexicais e pronomes não clíticos não compartilham propriedades com clíticos pronominais. Para concluir os testes, na sequência, será demonstrado qual é a “palavra hospedeira” desses clíticos, o que já foi sugerido pelo teste anterior.

iii. Os pronomes clíticos necessitam de uma base verbal para se incorporar.

Dada a diferença de comportamento em relação aos argumentos internos plenos (representados por DPs, PPs ou pronomes não clíticos), uma característica dos pronomes clíticos é o fato de eles necessitarem de uma base (verbal) para se incorporar, uma vez que são elementos átonos.

Nas sentenças em (12) e (13), são indicados entre colchetes o pronome clítico e a palavra considerada como base para o mesmo, utilizando-se um verbo ou um grupo verbal.

- (12) a. O meu namorado [me] [dava] presentes.
 b. *O meu [namorado]-[me] dava presentes.
 c. ?O meu namorado [dava]-[me] presentes.
 d. *O meu namorado dava [me] [presentes].
 e. *O meu namorado dava [presentes]-[me].

⁶ Nas sentenças em (9) e (10), está sendo considerado que o sujeito é sempre [O Carlos].

- (13) a. O meu namorado queria [me] [dar] presentes.
 b. *O meu namorado [queria]-[me] dar presentes.
 c. ??O meu namorado [me] [queria] dar presentes.⁷
 d. *O meu [namorado]-[me] queria dar presentes.
 e. ?O meu namorado queria [dar]-[me] presentes.
 f. *O meu namorado queria dar [me] [presentes].
 g. *O meu namorado queria dar [presentes]-[me].

Nas sentenças em (12) e (13) é possível verificar que, no PB, os pronomes clíticos somente podem ocorrer adjacentes a um verbo, como constatado para o verbo “dar” em (12a), em que se tem a próclise, e em (12c), em que há ênclise para este mesmo verbo (o sinal “?” justifica-se pelo fato de a ênclise não ser tão natural em PB quanto a próclise). Não pode haver adjacência do pronome clítico a um substantivo, como verificado pela agramaticalidade das sentenças (12b), (12d) e (12e) em relação aos nomes “namorado” e “presentes”. Quando há dois verbos (um grupo verbal), como em (13), o clítico pronominal pode colocar-se em próclise ou em ênclise em relação ao segundo verbo (sentenças (13a) e (13e) respectivamente) ou em próclise em relação ao primeiro (sentença (13c)), ainda que haja dúvidas quanto à aceitabilidade de (13c) e de (13e) no PB (sendo (13c) agramatical para muitos falantes). Qualquer outra possibilidade é considerada agramatical no PB, como em (13b), (13d), (13f) e (13g).

Além de incorporar-se somente a um verbo, o pronome clítico (ou o grupo de pronomes clíticos) não pode se encontrar separado deste, de forma que entre eles não pode haver outro elemento, como um advérbio, por exemplo. As sentenças em (14) ilustram esse fato, nas quais um advérbio posto entre o clítico pronominal e o verbo (ambos indicados entre colchetes) causa agramaticalidade.⁸

- (14) a. O meu pai sempre [me] [mandava] dinheiro.
 b. *O meu pai [me] sempre [mandava] dinheiro.
 c. O meu pai podia sempre [me] [mandar] dinheiro.
 d. *O meu pai podia [me] sempre [mandar] dinheiro.

A única possibilidade para que um pronome clítico se encontre separado da sua base verbal é se, entre eles, houver outro clítico, o que não tende a ocorrer em PB. No entanto, em francês, tal situação pode ser observada, como em (15).

- (15) [*Ce plaisir*]_i, *je ne t'en*_i *priverai pas* !
 “Deste prazer, eu não te privarei!”

⁷ Conforme menciona Martins (2012, pp. 171-3) em sua pesquisa sobre a escrita do PB, a próclise ao verbo finito nos grupos verbais diminuiu drasticamente do século XIX ao XX.

⁸ Luís e Kaiser (2016, pp. 218-31), porém, mencionam que no português europeu (doravante, PE) é possível haver interpolação, causada principalmente pela negação, como em “Se me não engano, ele faz anos a 21 de janeiro”. Os autores também mencionam que este fato se constitui em um resíduo de estágios anteriores do PE, em que sintagmas inteiros podiam colocar-se entre o clítico e o verbo.

Na sentença em (15), *je, ne, te (t')* e *en* são todos clíticos, não havendo então, entre qualquer um deles e o verbo, nenhum elemento que não seja também clítico.⁹

Tendo sido apresentadas, nesta Introdução, as propriedades dos clíticos pronominais, que os diferenciam dos argumentos verbais representados por pronomes plenos, DPs ou PPs, nas próximas seções serão apresentados os sistemas de pronomes pessoais em PB e em francês, englobando pronomes clíticos e não clíticos.

3. Sistema de pronomes pessoais do PB

No PB, o inventário dos pronomes pessoais tem sofrido alterações devido a mudanças sintáticas observadas principalmente a partir do século XIX (CYRINO, 2018 [1993]; TARALLO, 2018 [1993]), de modo que aquilo que a GT preconiza, e que é apresentado nos bancos escolares, encontra-se cada vez mais distante do que se verifica realmente na língua em uso no Brasil. O quadro 1, no qual é apresentado o sistema de pronomes pessoais no PB, indica tais mudanças.

⁹ A única exceção a essa regra ocorre em francês na negação do infinitivo, como em *J'ai bu du café pour ne pas dormir tôt*, "Eu bebi café para não dormir cedo", em que *pas* se interpõe entre o clítico *ne* e o verbo (JONES, 1996, p. 255). Entretanto, *ne* tem sido cada vez mais omitido no francês coloquial (BALL, 2000, p. 13). *Je* (clítico pronominal nominativo) e *ne* (partícula de negação) são classificados como clíticos pois satisfazem todos os testes apresentados anteriormente, dentre eles a impossibilidade de ocorrerem como itens isolados; nesta situação, as formas tônicas correspondentes *moi* e *non* devem ser empregadas, como em *Qui a parlé ? – *Je / Moi*, "Quem falou? – Eu" e *Tu viens ? – *Ne / Non*, "Você está vindo? – Não").

Quadro 1: Pronomes pessoais no PB – fala e escrita.¹⁰

		FORMAS NOMINATIVAS (TÔNICAS)	FORMAS ÁTONAS NA FALA E NA ESCRITA (PRONOMES CLÍTICOS)			FORMAS TÔNICAS OBLÍQUAS NA FALA E NA ESCRITA	
			Acusativo (acc.)	Dativo (dat.)	Reflexivo (refl.)	Indefinido	Formas nominativas em função acusativa e oblíqua ¹¹
SINGULAR	P1	eu	me	me	me		mim, comigo
	P2	tu, você	te, <u>o</u> , <u>a</u> , <i>lhe</i>	te, <u>lhe</u>	te, se		ti, contigo, você ^o , <u>si</u> , <u>consigo</u> , você mesmo
	P3	ele, ela	<u>o</u> , <u>a</u> , <i>lhe</i>	<u>lhe</u>	se	<u>se</u> , ¹² <i>tu</i> , <i>você</i> , <i>a gente</i> ¹³	<u>si</u> , <u>consigo</u> , <i>ele</i> ^o , <i>ela</i> ^o , <i>ele(a) mesmo(a)</i>
PLURAL	P4	<u>nós</u> , <i>a gente</i>	<u>nós</u>	<u>nós</u>	<u>nós</u> , <i>se</i>		<u>nós</u> , <u>conosco</u> , <i>a gente</i> ^o
	P5	<u>vós</u> , vocês	<u>vós</u> , <u>os</u> , <u>as</u> , <i>lhes</i>	<u>vós</u> , <u>lhes</u>	<u>vós</u> , se		<u>vós</u> , <u>convosco</u> , vocês ^o , vocês mesmos
	P6	eles, elas	<u>os</u> , <u>as</u> , <i>lhes</i>	<u>lhes</u>	se		<u>si</u> , <u>consigo</u> , <i>eles</i> ^o , <i>elas</i> ^o , <i>eles(as) mesmos(as)</i>

Fonte: adaptado de Duarte (2013, p. 120).¹⁴

O quadro 1 não se limita a apresentar os pronomes pessoais admitidos pela GT; também não se limita aos pronomes utilizados no PB escrito, trazendo também aqueles utilizados na língua falada, sendo esta última menos conservadora que a escrita não somente no PB mas, de modo geral, em todas as línguas naturais.

¹⁰ A decisão de referir-se às pessoas do discurso como de P1 a P6 (e não como 1ª pessoa do singular, 1ª do plural, etc.) nos quadros 1, 2 e 3 se deve à interpretação de que, por exemplo, a 1ª pessoa do plural (P4) não se refere a “eu+eu” (não é o plural de “eu”), mas sim a “eu+tu/você” ou “eu+uma (ou várias) pessoas” (LOPES, 2018 [2007], p. 117).

¹¹ Na última coluna do quadro 1, as formas que coincidem com as nominativas e são acompanhadas do sobrescrito o podem ser utilizadas tendo caso tanto acusativo quanto oblíquo, como exemplificado em (i) a seguir.

- (i) a. Eu vi [ele] na feira ontem.
b. Eu fui [com ele] na feira ontem.

¹² Na coluna dos indefinidos, “se”, usado como partícula de indeterminação do sujeito, é também um pronome clítico.

¹³ A forma “a gente”, apesar de compor um DP, pode ser analisada como um pronome (LOPES, 2018 [2007]).

¹⁴ Seguindo a simbologia de Duarte (2013, p. 119-120), no quadro 1 são apresentados:

tachados os pronomes extintos da fala e da escrita (o que acontece com o paradigma de “vós” – embora esses pronomes ainda estejam presentes em textos bíblicos e em outros contextos religiosos, e ainda vivos em Portugal (BECHARA, 2015, p. 173));

sublinhados os pronomes de uso mais restrito a certos grupos sociais (em geral de maior prestígio) e até mesmo em extinção na fala espontânea;

em **itálico e negrito** as formas inovadoras, que já começam a ser implementadas na escrita;

sem grifos e em fonte normal os pronomes que fazem parte tanto da língua oral quanto da escrita.

Para uma explanação mais detalhada do que é apresentada aqui acerca dos pronomes mencionados no quadro 1 (inclusive sobre a variação diatópica dos mesmos, tendo como referência pesquisas de outras(os) autoras(es)), consulte-se Duarte (2013, p. 120ss.).

Como Duarte (2013, pp. 119-20) menciona, no PB falado todos os pronomes nominativos podem exercer as funções acusativa, dativa e oblíqua, estas últimas com o uso de preposições (DUARTE, 2013, p. 122). Dessa forma, seus traços de caso estão subespecificados no léxico, uma vez que uma única forma (um único pronome) pode receber vários casos, a depender da sentença em que se encontra. Haveria, entretanto, uma maior estigmatização para o uso dos pronomes “eu” e “tu” nessas funções sem serem seguidos de infinitivo pessoal ou compondo uma *small clause* (tanto que estes pronomes não foram incluídos na última coluna do quadro 1, pois estes usos não são admitidos na escrita, como o título da respectiva coluna indica). Exemplos desses usos são indicados a seguir.

- (16) a. Comprei esse presente pra tu!
 b. Cruzei com ela na rua mas ela não viu eu.
 c. Eu comprei essas frutinhas pra tu comer.
 d. Por que você tá me olhando? Tá achando eu bonita?

As sentenças (16a) e (16b) ilustram os empregos mais estigmatizados dos pronomes “eu” e “tu”; em (16c), ocorre o pronome “tu” seguido de um infinitivo impessoal e, em (16d), “eu” compondo uma *small clause* com o adjetivo “bonita” (usos também estigmatizados).

A observação do quadro 1 revela ainda que P1 é a única pessoa em que os pronomes pessoais permanecem inalterados em PB, apresentando as mesmas formas para as línguas falada e escrita.

Com relação à segunda pessoa, a substituição de “tu” – “vós” por “você” – “vocês” na posição de sujeito em algumas variedades do PB fez com que as formas clíticas de terceira pessoa (“o” / “a”, “os” / “as”, “lhe”, “lhes”, “se”) passassem a ser possíveis para referir-se também à segunda, o que tornou o inventário de pronomes clíticos no PB mais complexo (FARACO, 2017, pp. 116; 25). O uso acusativo de “lhe” em P2 (relacionado a “você”) e em P3 (LUÍS; KAISER, 2016, p. 211) pode ser visto como uma substituição aos pronomes “o” e “a” (sentença (17a) a seguir), estando também em variação livre com “te” (sentença (17b)) (GALVES, 2001, p. 155). O uso dativo de “lhe” vem sendo substituído por um PP, com o uso da preposição “a” na escrita e “pra” na fala (sentença (17c)).

- (17)
- a. Não se esqueça que estamos preparados para deixar-lhe em plena forma a qualquer momento. (Jornal – texto de propaganda)
- (DUARTE, 2013, p. 132)
- b. Te/Lhe vi ontem, mas você nem me cumprimentou.
- c. Eu dei um presente pra você.

Como aponta Galves (2001, p. 143; 155), o PB perdeu a oposição “tu/você”, ainda presente no PE, sendo que, nesta variedade, “tu” é empregado como forma íntima, e “você”, como forma polida. No Brasil, segundo a autora, “tu” só subsiste regionalmente, principalmente no Rio Grande do Sul. Entretanto, “tu” e “você” são pronomes que também vêm sendo usados em contextos de

referência indeterminada, como em (18a), que tem sua equivalência em (18b) (nesse exemplo, em uma construção existencial).

- (18) a. Você/Tu tem uma loja lá na rua que só vive em liquidação.¹⁵
 b. Existe/Há uma loja lá na rua que só vive em liquidação.

(adaptado de LOPES (2018 [2007], p. 104))

Segundo Lopes (2018 [2007], p. 104) e Quarezemin e Fuchsberger (2020, p. 100), o uso de “a gente” substituindo “nós” apresenta um caráter indeterminador, de modo que a(o) falante pode evitar um maior comprometimento com o discurso, tornando-o mais vago e genérico, em oposição a uma nuance mais específica de “nós”. Isso porque “a gente” pode ser interpretado como “eu+você(s)+ela(ele)(s)+todo mundo ou qualquer um” (ainda que se combine com verbos em P3). Um fato que pode corroborar essa afirmação é o uso desse termo em contextos de indeterminação, como em (19a), com seu equivalente em (19b).¹⁶

- (19) a. A gente não pode mais ir tranquilamente à praia aqui.
 b. Não se pode mais ir tranquilamente à praia aqui.

Ainda considerando os pronomes nominativos, percebe-se uma propriedade no português que se repete no italiano e no francês, mas não no espanhol: há marcas de gênero somente em P3 e P6, pois aí os pronomes assumem formas diferentes segundo o gênero dos seus referentes (o que acontece também com os pronomes clíticos acusativos);¹⁷ nas demais pessoas, o gênero formal dos pronomes é subespecificado, de modo que a interpretação de gênero da(do) falante ou da(do) ouvinte só pode ser realizada a partir do contexto, como em (20), onde só se identifica o gênero da(do) falante a partir do uso dos adjetivos no feminino ou no masculino.

- (20) a. Eu estou velha.
 b. A gente ficou arrasada.
 c. A gente ficou arrasado.

(LOPES, 2018 [2007], pp. 111-2)

No PB, muitos dos clíticos pronominais não fazem mais parte da língua-I da(do) aprendiz (ou gramática da(do) aprendiz), sendo essa língua-I o conhecimento linguístico de uma pessoa, presente na sua mente e que lhe permite usar uma língua-E para compreender mensagens e guiá-la

¹⁵ A combinação de “tu” com verbos em P3 tem se mostrado cada vez mais frequente no PB falado.

¹⁶ Outras estratégias de indeterminação do sujeito no PB podem ser consultadas em Oliveira (2018) e em Quarezemin e Fuchsberger (2020).

¹⁷ No espanhol, P4 e P5 também possuem marcas de gênero: em P4 há *nosotros* e *nosotras*, e em P5, *vosotros* e *vosotras* (embora essas formas de P5 praticamente sejam utilizadas somente na Espanha, sendo substituídas por *ustedes* nos demais países hispânicos) (FANJUL, 2005, p. 20).

Como é possível observar no quadro 1, os clíticos acusativos e dativos de P3 e P6 podem ocorrer em P2 e P5 também, respectivamente, seguindo-se um uso mais restrito no PB.

na estruturação das suas próprias sentenças (CHOMSKY, 1986, pp. 21-2). A língua-E constitui o código linguístico compartilhado pelos indivíduos que integram uma mesma sociedade, um construto compreendido independentemente das propriedades da mente/cérebro (CHOMSKY, 1986, p. 20).¹⁸ Com relação aos pronomes clíticos acusativos “o(s)” e “a(s)”, nota-se que os mesmos encontram-se praticamente extintos da fala da(do) brasileira(o), não fazendo mais parte do vernáculo do PB.¹⁹ Esses clíticos são associados ao aprendizado escolar, identificados com língua escrita e estilo formal, geralmente revelando grau de instrução elevado da(do) interlocutora(interlocutor) (KATO; CYRINO; CORRÊA, 2009; NUNES, 2018 [1993], p. 161). Geralmente, o emprego dos clíticos pronominais de terceira pessoa ocorre em momentos de grande monitoração da fala, como em ocasiões solenes ou interlocuções com superiores hierárquicos. No PB oral, tais clíticos são geralmente substituídos pelas formas nominativas em função acusativa. Esse fato é ilustrado pelos dados de (21) a (24), dos quais as sentenças (a) poderiam corresponder à língua escrita e as (b) e (c), à língua falada.

- (21) a. Papai_i, sempre o_i vejo trabalhando tanto!
 b. Papai_i, sempre vejo você_i trabalhando tanto!
- (22) a. [Essa sua namorada]_j, eu não a_j suporto!
 b. [Essa sua namorada]_j, eu não suporto ela_j!
- (23) a. [Queridos sobrinhos]_k, gostaria de convidá-los_k para um Natal em família.
 b. [Queridos sobrinhos]_k, gostaria de convidar vocês_k para um Natal em família.
- (24) a. Quando você quiser [os livros]_i, basta vir buscá-los_i.
 b. Quando você quiser [os livros]_i, basta vir buscar eles_i.
 c. Quando você quiser [os livros]_i, basta vir buscar _____i.

Em (22b) e (24b), ocorre o emprego do pronome “ele” (podendo variar em gênero e/ou número) em forma acusativa, especialmente estigmatizado pela norma e banido da língua escrita culta (GALVES, 2001, p. 162). Esse uso distingue o PB não somente do PE, mas também das outras línguas românicas, que preferem o emprego do pronome clítico de terceira pessoa nessa situação.²⁰ Ainda, como exemplificado por (24c), no PB falado existe uma preferência pelo objeto nulo (DUARTE, 2013, p. 123), ou seja, não realizado foneticamente (CYRINO, 1994, 2018 [1993], 2019, 2020; CYRINO; MATOS, 2016).

¹⁸ Entretanto, no francês, todos os pronomes clíticos ainda compõem a gramática da(do) falante, no sentido de que todos eles são amplamente utilizados no dia-a-dia, como é possível verificar em trabalhos como os de Ball (2000, p. 78) e de Morin (1979).

¹⁹ Também praticamente extintos da língua falada se encontram os dativos “lhe” e “lhes”, sendo substituídos por sintagmas preposicionadas (PPs), como já mencionado (ver sentença (17c)).

²⁰ Exemplos de sentenças equivalentes a (22b) em outras línguas seriam *Essa sua namorada, não a suporto!* (PE), *Ta petite amie, je ne la supporte pas !* (francês), *La tua ragazza, non la sopporto!* (italiano) e *Tu novia, no la soporto!* (espanhol), sendo as formas *a/la* pronomes clíticos acusativos.

Ainda sobre os clíticos pronominais em P3 e P6, sejam dativos ou acusativos, Duarte (2013, p. 129), assim como Galves (2001, p. 129), menciona que, uma vez que esses clíticos não existem na Gramática Nuclear (GN) do indivíduo (aquela que se forma na criança durante o processo de aquisição da linguagem, resultando da seleção dos valores dos parâmetros da GU), tais elementos são incorporados por este indivíduo através do letramento ou de um contato maior com a escrita, pois estão em uma parte periférica à L1 ouvida durante a infância.²¹ Esse seria um dos pontos que forma a Língua-I (Individual) do Letrado (LIL), a qual seria então a união entre a GN e o conjunto de regras aprendidas pela(o) adulta(o) letrada(o), conjunto este que pode ser visto como uma L2 (AGUIAR, 2020; KATO, 2013, p. 150).

Considerando a perda dos clíticos acusativos de terceira pessoa, o paradigma dos clíticos no PB reduz-se praticamente a “me”, “te”, “lhe” e “se”, formas que apresentam sincretismo (LUÍS; KAISER, 2016, p. 229), sendo ambíguas quanto ao caso e ao gênero (considerando o uso acusativo de “lhe”, como mencionado anteriormente, além do dativo). A ambiguidade de “lhe” em relação ao caso estende-se também para a terceira pessoa, em algumas variantes do PB (GALVES, 2001, p. 139). Como menciona Galves (2001, p. 159), a categoria lexical pronominal objeto de terceira pessoa no PB é o pronome fraco “ele” (ou “eles”, “ela(s)”), não admitido no PE.²²

Kato (2013, p. 149) menciona que “em todas as línguas se observam diferenças entre a fala e a escrita, mas no Português do Brasil (PB) existe um fosso de tal ordem que para a criança aprender a ler e a escrever, ela enfrenta a tarefa da aprendizagem de uma língua estrangeira”. Isso se justifica pelo fato de a norma padrão do PB ainda seguir a GT do PE em muitos pontos, dentre eles o emprego dos pronomes clíticos de terceira pessoa.

Finalmente, no PE é possível a combinação de dois pronomes clíticos, sendo um correspondente ao objeto indireto e o outro, ao objeto direto em uma mesma sentença, como em “Este livro pertence a José, tenho de devolver-lho” (lho = lhe + o). No PB, essas construções encontram-se banidas, sendo uma das causas justamente o desaparecimento do pronome clítico de terceira pessoa (GALVES, 2001, p. 139).²³ Em PB, os verbos não podem ter mais de um argumento expresso por um pronome clítico ao mesmo tempo, ainda que o verbo tenha mais de um argumento interno (LUÍS; KAISER, 2016, pp. 211-31).

Após terem sido apresentadas as particularidades do sistema pronominal do PB, na próxima seção, será feito o mesmo para o francês.

²¹ Os clíticos acusativos e dativos em P3 e P6, no entanto, não se encontram ausentes nas falas de crianças em fase pré-escolar que têm o italiano, o espanhol ou o francês como L1 (GALVES, 2001, p. 13).

Chomsky (1988 [1981]), assim como Kato (2005), dá o nome de “periferia marcada” àquilo que Duarte (2013) chama de “parte periférica”.

²² Galves (2001, pp. 162-4) classifica o pronome não clítico (e não tônico) “ele” (ou “eles”, “ela(s)”) como “pronome fraco”, e o pronome pleno “ele”, como “pronome forte”. Isso porque, apesar de ambos serem representados pelo mesmo item lexical, o pronome forte não pode se referir a um antecedente inanimado, como em “O que você bebeu? – *Ele”, referindo-se ao suco, por exemplo. Em Galves (2001, pp. 172-6) podem ser obtidas também mais informações acerca dos pronomes fracos e nulos.

²³ Para detalhes sobre a ocorrência dos grupos clíticos no PE, ver Martins (2013, pp. 2235-7).

4. Sistema de Pronomes Pessoais do Francês

Uma grande diferença do sistema de pronomes pessoais do francês em relação ao do PB é que, na língua francesa, existe um maior número de formas pronominais que apresentam o fenômeno do sincretismo, ou seja, que acumulam funções, seja em relação ao caso sintático, seja em relação ao número das pessoas do discurso (CARVALHO, 2017, p. 30). No quadro 2 são apresentados os pronomes pessoais existentes no francês.

Quadro 2: Pronomes pessoais no francês.

		FORMAS ÁTONAS (PRONOMES CLÍTICOS)							FORMAS TÔNICAS E OBLÍQUAS
		Nominativo (nom.) ²⁴	Acusativo (acc.)	Dativo (dat.)	Reflexivo (refl.)	Indefinido	Locativo (loc.)	Partitivo (part.)/ Genitivo (gen.)	
SINGULAR	P1	<i>je</i>	<i>me (m')</i>	<i>me (m')</i>	<i>me (m')</i>				<i>moi</i>
	P2	<i>tu, vous</i>	<i>te (t'), vous</i>	<i>te (t'), vous</i>	<i>te (t'), vous</i>				<i>toi, vous</i>
	P3	<i>il, elle, ce (c')</i>	<i>le (l'), la (l')</i>	<i>lui</i>	<i>se (s')</i>	<i>il, se (s'), on</i>	<i>y, en</i>	<i>en</i>	<i>lui, elle, soi, ça / cela</i>
PLURAL	P4	<i>nous, on</i>	<i>nous</i>	<i>nous</i>	<i>nous, se (s')</i>				<i>nous</i>
	P5	<i>vous</i>	<i>vous</i>	<i>vous</i>	<i>vous</i>				<i>vous</i>
	P6	<i>ils, elles, ce (c')</i>	<i>les</i>	<i>leur</i>	<i>se (s')</i>		<i>y, en</i>	<i>en</i>	<i>eux, elles, soi</i>

Fonte: adaptado de Oliveira (2016, p. 33).

No francês, existe um só pronome de cortesia: *vous*, que sempre se combina com verbos em P5 e não apresenta marcas morfológicas de gênero, número ou caso. Como *vous* é usado também no plural em situações informais, será o contexto a identificar se se trata de uma situação de formalidade (como em (25a)) ou de informalidade ((25b)).

- (25) a. *Monsieur, vous pourriez attendre dix minutes ?*
“Senhor, o senhor poderia esperar dez minutos?”
- b. *Chères amies, vous êtes très importantes dans ma vie.*

“Minhas caras amigas, vocês são muito importantes na minha vida.”

O francês é uma língua de sujeito não nulo (KATO; DUARTE, 2014, p. 2; dentre outros), assim como o inglês: o sujeito, ou o pronome nominativo que o representa, sempre deve ser foneticamente

²⁴ Como é possível verificar em Kaiser (1992, p. 31), clíticos pronominais nominativos somente podem ser separados da sua base verbal pela negação e/ou por um outro clítico. É o que ocorre em *je ne le connais pas*, “eu não o conheço”; outro elemento não pode causar essa separação, como em **je seulement aime LUI*, “eu somente amo ELE”. Não havendo clíticos nominativos no PB, o pronome “eu”, por exemplo, não possui essas restrições, como se pode ver na tradução da última sentença aqui, em que um advérbio o separa do verbo.

realizado. Isso se atribui principalmente ao fato de o paradigma verbal para a maioria dos tempos e modos ser fonologicamente igual para várias pessoas gramaticais; por exemplo, para o verbo *travailler*, “trabalhar” no presente do indicativo, tem-se *je travaille* ([tra’vaj]), *tu travailles* ([tra’vaj]), *il/elle/on travaille* ([tra’vaj]), *nous travaillons* ([travaj’õ]), *vous travaillez* ([travaj’e]), *ils/elles travaillent* ([tra’vaj]). Para os verbos ditos impessoais, deve-se usar o pronome expletivo *il*, como ilustrado em (26).

- (26) *Il pleuvra beaucoup cette année.*
 “Vai chover muito este ano.”

Ça (*Ce/C’*) é um pronome que pode corresponder a *isso*, mas que também pode ser usado em contextos coloquiais para substituir o expletivo *il* em algumas situações, como a que está indicada em (27).

- (27) *Ça ne vaut pas le coup de s’en faire pour si peu.*
 “Não vale a pena preocupar-se por tão pouco.”

On é um pronome pessoal que serve à indeterminação e que pode ser equivalente a “alguém” ou a “as pessoas”. No entanto, pode também ser um substituto para “nós”, sendo equivalente a “a gente” em PB. É sempre combinado com verbos em P3, embora possa ser visível a concordância de adjetivos ou participios quando o sentido é de plural ou de feminino. Exemplos desses usos são dados em (28).

- (28) a. *On frappe à la porte.*
 “Alguém está batendo na porta.”
 b. *Au Brésil on mange de la feijoada.*
 “No Brasil as pessoas comem feijoada.”
 c. *Hier on est allées au cinéma, Jeanne et moi.*
 “Ontem nós fomos ao cinema, a Jeanne e eu.”

Em (28c) há a concordância do participio passado *allé* no feminino plural (acrescentando-se “*e + s*”), única indicação de que o sujeito é constituído somente de pessoas do gênero feminino.

No francês, é obrigatória a elisão diante de verbos (ou de outros clíticos) começando por som vocálico para os pronomes que a admitem (cujas formas elididas se encontram entre parênteses no quadro 2), como se verifica para o grupo clítico *se + en = s’en* na sentença em (27).²⁵

²⁵ Há ainda, no francês, o que a GT chama erroneamente de “h aspirado” ou h aspiré (pois a letra “h” nunca é pronunciada no francês, contribuindo, no máximo, para o som [ʃ] do dígrafo “ch”), diante do qual não pode haver elisão, em oposição ao “h mudo” ou h muet, diante do qual é obrigatória a elisão. Um exemplo do h aspiré é dado em (i) a seguir, em que seria incorreto dizer ou escrever *Je l’hais !

(i) Je le hais !
 “Eu o odeio.”

Não há regra para a identificação do h aspiré; é necessário saber de cor as palavras que o apresentam ou consultar um dicionário. Léon (2007, p. 226, tradução livre) menciona que o h aspiré é “uma lembrança do tempo em que ele era pronunciado”.

Uma grande diferença do francês em relação ao PB é a presença de clíticos pronominais com função locativa e partitiva/genitiva: *y* e *en* (PRÉVOST, 2009, p.116; SABIO, 2005). Esses clíticos são conhecidos tradicionalmente como “clíticos adverbiais” (*clitiques adverbiaux*), ainda que, em muitas situações, eles correspondam a sintagmas não adverbiais (MILLER; MONACHESI, 2003). Em (29) há exemplos do uso desses pronomes.

- (29) a. *Je voudrais connaître Paris_i et y_i habiter.*
 “Eu gostaria de conhecer Paris e morar lá.”
- b. *Elle aime [le chocolat]_j, mais elle ne peut plus en_j manger.*
 “Ela gosta de chocolate, mas não pode mais comer.”
- c. *Je connais [ce médicament]_k et je peux en_k garantir l’efficacité.²⁶*
 “Eu conheço esse remédio e posso garantir a sua eficácia.”

Ainda em relação à função locativa, tanto *y* quanto *en* podem corresponder a um lugar no contexto comunicativo. Se o verbo na frase em que o pronome se encontrar exigir a preposição *à*, o pronome a utilizar será *y*; se a preposição exigida for *de*, o pronome empregado será *en*. Exemplos são fornecidos nas sentenças em (30) e (31).

- (30) *Je passe les vacances toujours [en Argentine]_p, mais cette année je ne pourrai pas y_i aller.*
 “Eu passo as férias sempre na Argentina, mas esse ano eu não vou poder ir lá.”
- (31) a. *Quand est-ce qu’ils sont revenus [de la campagne]_j ?*
 “Quando eles voltaram do campo?”
- b. *Ils en_j sont revenus hier.*
 “Eles voltaram (de lá) ontem.”

Na sentença em (30), ainda que *y* substitua [*en Argentine*], e que a construção seja *aller en Argentine*, o que deve ser considerado é a forma “geral” do verbo, exigindo a preposição *à* (*aller à un endroit*, “ir a um lugar”), o que caracteriza o uso desse pronome locativo. Já na resposta em (31), o verbo usado é *revenir*, que exige a preposição *de* (*revenir d’un endroit*, “voltar de um lugar”), e o pronome locativo a utilizar é então *en*.

Ainda, os clíticos *y* e *en* podem substituir complementos introduzidos pelas preposições *à* e *de*, respectivamente, como nas sentenças em (32) (verbos: *croire à*, *s’occuper de*).

- (32) a. *C’est [ce que tu dis]_p, mais je n’y_i crois pas.*
 “É o que você está dizendo, mas eu não acredito nisso.”
- b. *J’adore [les statistiques]_j, alors je m’en_j occupe.*
 “Eu adoro estatísticas, então deixo(m) que eu faço.”

No entanto, nessas situações, esses clíticos não podem se referir a antecedentes com traço [+animado], como em **Amène tes enfants, je m’en occupe !*, “Traz os teus filhos, eu tomo conta

²⁶ Com relação ao caso genitivo, apesar de *en* representar o complemento de um sintagma nominal, como se verifica na sentença (29c), ele se incorpora à base verbal, como qualquer outro pronome clítico.

deles!"; deve-se usar então a preposição seguida do pronome tônico correspondente, *Je m'occupe d'eux*. Entretanto, na língua coloquial constata-se o uso de *en* nessas situações, referindo-se a pessoas (GRÉGOIRE; KOSTUCKI, 2012, p. 98).²⁷

A marcação de gênero nos pronomes pessoais do francês ocorre somente em P3 e em P6 no nominativo, no acusativo e com os pronomes tônicos, como ilustra o quadro 3 (que é, na verdade, um subconjunto do quadro 2).

Quadro 3: Pronomes nominativos, acusativos e tônicos em P3 e P6 no francês.

	NOMINATIVOS		ACUSATIVOS		TÔNICOS	
	MASC.	FEM.	MASC.	FEM.	MASC.	FEM.
SINGULAR (P3)	<i>il</i>	<i>elle</i>	<i>le (l')</i>	<i>la (l')</i>	<i>lui</i>	<i>elle</i>
					<i>soi</i>	
PLURAL (P6)	<i>ils</i>	<i>elles</i>	<i>les</i>		<i>eux</i>	<i>elles</i>
				<i>soi</i>		

Fonte: Oliveira (2016, p. 36).

Para a identificação de gênero dos pronomes nas demais pessoas (assim como de *les* e *soi* no quadro 3), é necessário recorrer ao contexto, similarmente ao verificado em (28c).

A combinação de pronomes clíticos no francês é um pouco mais complexa do que no português. Para o uso com verbos que não estejam no imperativo afirmativo, as combinações possíveis são mostradas no quadro 4, que seguem a ordem *me, te, se, nous, vous* < *le, la, les* < *lui, leur* < *y* < *en*, segundo Anderson (2005, p. 125).²⁸ Nesse quadro, as células hachuradas indicam combinações não possíveis (CALLAMAND, 1987, p. 46).²⁹

²⁷ Y e en podem ainda referir-se a complementos de adjetivos introduzidos pela preposições à e de, respectivamente (RIEGEL; PELLAT; RIOUL, 2004, p. 201), como indicado em (i) a seguir.

- (i) a. Ne mets pas [de film d'horreur]i, j'yi suis très sensible.
"Não coloca filme de terror, eu sou muito sensível a isso."
b. Julie [a réussi son examen]k et elle enk est très fière.
"A Julie obteve uma boa nota na prova e ela está muito orgulhosa disso."

²⁸ É ainda possível, no francês, a ocorrência de grupos clíticos com três elementos, em estruturas como em (i) a seguir (MANZINI; SAVOIA, 2018, p. 26).

- (i) *Jean me le lui fera porter.*
"O Jean vai me fazer levar isso para ela(ele)."

²⁹ Para a ocorrência de pronomes duplos com verbos no imperativo afirmativo em francês, ver obras como Boularès e Frérot (1997, p. 128) e Callamand (1987, p. 46).

Quadro 4: Combinações de pronomes clíticos no francês para verbos que não estejam no imperativo afirmativo.³⁰

		ACUSATIVOS			LOCATIVO	PARTITIVO / GENITIVO
		<i>le</i>	<i>la</i>	<i>les</i>	<i>y</i>	<i>en</i>
DATIVOS OU REFLEXIVOS	<i>me</i>	<i>me le (me l')</i>	<i>me la (me l')</i>	<i>me les</i>	<i>m'y</i>	<i>m'en</i>
	<i>te</i>	<i>te le (te l')</i>	<i>te la (te l')</i>	<i>te les</i>	<i>t'y</i>	<i>t'en</i>
	<i>lui</i>	<i>le lui</i>	<i>la lui</i>	<i>les lui</i>		<i>lui en</i>
	<i>se</i>	<i>se le (se l')</i>	<i>se la (se l')</i>	<i>se les</i>	<i>s'y</i>	<i>s'en</i>
	<i>nous</i>	<i>nous le (nous l')</i>	<i>nous la (nous l')</i>	<i>nous les</i>	<i>nous y</i>	<i>nous en</i>
	<i>vous</i>	<i>vous le (vous l')</i>	<i>vous la (vous l')</i>	<i>vous les</i>	<i>vous y</i>	<i>vous en</i>
	<i>leur</i>	<i>le leur</i>	<i>la leur</i>	<i>les leur</i>	<i>leur y</i>	<i>leur en</i>
LOCATIVO	<i>y</i>	<i>l'y</i>	<i>l'y</i>	<i>les y</i>		<i>y en</i>
PARTITIVO / GENITIVO	<i>en</i>	<i>l'en</i>	<i>l'en</i>	<i>les en</i>	<i>y en</i>	

Fonte: adaptado de Oliveira (2016, p. 36).

Para construções verbais que não envolvam imperativo afirmativo no francês, na combinação de pronomes clíticos o acusativo se coloca à direita, exceto quando um dos clíticos for *lui*, *leur*, *y* ou *en* (combinações em negrito no quadro 4). Nesses casos, a ordem crescente de preferência para ocupar a posição da direita no grupo clítico é *lui/leur* → *y* → *en*.³¹ Exemplos destes usos são dados nas sentenças (33) a (35).

- (33) a. *Tu peux me_i prêter [ton livre de grammaire]_j ?*
 “Tu podes me emprestar o teu livro de gramática?”
- b. *Bien sûr, je te_i le_j prête.*
 “Claro, eu te empresto.”
- (34) a. *Ils ont déjà envoyé [la facture]_k [au client]_l ?*
 “Ele já enviaram a fatura ao cliente?”
- b. *Oui, ils la_k lui_l ont déjà envoyée.*
 “Sim, eles já enviaram.”
- (35) a. *Ta sœur t'_ma donné [des bonbons]_n ?*
 “Tua irmã te deu balas?”
- b. *Non, elle n'a pas voulu m'_men_n donner.*
 “Não, ela não quis me dar.”

³⁰ Segundo Jones (1996, p. 254), a incompatibilidade entre *lui* e *y* é tradicionalmente explicada em termos fonéticos: ou haveria um hiato entre os dois clíticos, ou *y* seria “absorvido” pela vogal de *lui*. *Leur* e *y* não apresentam essa restrição, como em *Pierre leur y a parlé*, “Pierre falou com elas(eles) ali (/naquele lugar)”.

³¹ Assim, para os grupos de pronomes clíticos no francês, a ordem nem sempre é “dativo + acusativo”, ordem esta que se verifica no português, e também no italiano e no espanhol.

Concluindo esta seção, são apresentadas, no quadro 5, as equivalências entre os vários clíticos pronominais e os pronomes tônicos em francês e em PB.

Quadro 5: Formas pronominais no francês e no PB.³²

PESSOA	NÚMERO	CASO / TIPO	PRONOMES CLÍTICOS		PRONOMES TÔNICOS (PLENOS)	
			FRANCÊS	PB	FRANCÊS	PB
1 ^a	Singular (P1)	Nom	<i>je (j')</i>		–	
		Acc	<i>me (m')</i>		me	
		Dat				
		Refl				
	Plural (P4)	Nom		(<i>on</i>)	–	
		Acc	<i>nous</i>		<u>nos</u>	
		Dat				
		Refl				
2 ^a	Singular (P2)	Nom	<i>tu</i>	(v <i>ous</i>)	–	
		Acc	<i>te (t')</i>			(<u>o/a</u>) (<i>lhe</i>)
		Dat			te	(<u>lhe</u>)
		Refl				(<i>se</i>)
	Plural (P5)	Nom		<i>vous</i> <i>vos</i>		–
		Acc	(<u>os/as</u>) (<u>lhes</u>)			
		Dat	(<u>lhes</u>)			
		Refl	(<i>se</i>)			
3 ^a	Singular (P3)	Nom	<i>il / elle / ce (c')</i>		–	
		Acc	<i>le (l') / la (l')</i>		<u>o / a</u> (<i>lhe</i>)	
		Dat	<i>lui</i>		<u>lhe</u>	
		Refl	<i>se (s')</i>		se	
	Plural (P6)	Nom	<i>ils / elles / ce (c')</i>		–	
		Acc	<i>les</i>		<u>os / as</u> (<i>lhes</i>)	
		Dat	<i>leur</i>		<u>lhes</u>	
		Refl	<i>se (s')</i>		se	
Locativo		<i>y (en)</i>		–		
Partitivo/Genitivo		<i>en</i>		–		

Fonte: elaboração do autor.

³² O quadro 5 foi construído com base nas informações obtidas em Kaiser (1992, pp. 29-47). São seguidas as mesmas indicações utilizadas para o quadro 1 em termos de formas inovadoras, em desuso e de uso restrito. As formas entre parênteses representam as variações possíveis dos pronomes; as barras separam as formas masculinas e femininas.

Finalmente, a seguir são realizadas as considerações finais acerca dos sistemas de pronomes do PB e do francês.³³

Considerações finais

Como mencionado, no PB os pronomes clíticos de terceira pessoa não pertencem mais à língua-I da(do) falante, ao contrário do que acontece no francês, em que esses pronomes são muito produtivos. Além disso, diferentemente do PB, o francês possui clíticos pronominais locativos e partitivo/genitivo, fato este que poderia representar mais uma dificuldade na aquisição da L2 por uma(um) aprendiz brasileira(o).³⁴ Além disso, até mesmo na língua falada, combinações de pronomes clíticos são usadas no francês, o que não se verifica no PB, que preferirá a omissão de um dos clíticos e/ou o emprego de preposição + pronome para substituir o dativo, como em *Ce livre-là, tu peux me le passer?*, “Aquele livro, você pode me passar ___?” e *Jean a acheté ce livre pour Julie e va le lui donner ce soir*, “Jean comprou este livro para a Julie e vai dar ___ pra ela esta noite”.

As propriedades da L1 da(do) aprendiz constituem um fator que influencia a aquisição de clíticos pronominais, assim como outros aspectos na aquisição da L2 (GASS; SELINKER, 2008, pp. 89-120; VANPATTEN; BENATI, 2010, pp. 14-5). Devido a isso, nos estágios iniciais do desenvolvimento da interlíngua, a(o) aprendiz brasileira(o) tende a substituir os clíticos pronominais por objetos nulos (como em *Il va chanter ___ à Juliette* em vez de *Il va la chanter à Juliette*, “Ele vai cantá-la (uma canção) para a Juliette”) ou por pronomes tônicos (*Cette musique empêche lui de lire* em vez de *Cette musique l’empêche de lire*, “Esta música o impede de ler (por causa do barulho)” ou ainda *Elle offre des cadeaux à eux* em vez de *Elle leur offre des cadeaux*, “Ela lhes oferece presentes (a eles)”), como ocorre no PB (OLIVEIRA, 2022). Além disso, Oliveira (2022) constatou na sua pesquisa com informantes brasileiras(os) a influência da perda dos clíticos acusativos e dativos de terceira pessoa no PB para a aquisição desses clíticos em francês: mesmo as(os) informantes mais proficientes não demonstraram, na sua produção nas entrevistas, taxas de utilização desses elementos próximas às verificadas para as(os) nativas(os) (os clíticos acusativos foram constatados em 3,57% das sentenças construídas com

³³ Para mais informações sobre os pronomes do francês, clíticos ou não clíticos, é possível consultar o excelente Capítulo 6 de Jones (1996).

³⁴ Na área de Linguística Aplicada a Ensino e Aprendizagem de Segunda Língua, assim como na Psicolinguística, costuma-se fazer a diferenciação entre “aquisição” e “aprendizagem”. A aquisição aconteceria natural e inconscientemente, emergindo espontaneamente quando da interação normal na língua em questão, tendo o significado como foco, ou seja, sem preocupação direta com a forma. A aprendizagem, ao contrário, envolveria conhecimento explícito sobre a língua, como suas regras e padrões, com foco na forma. Esse processo aconteceria tomando a L2 como objeto da instrução, mas não necessariamente como o meio para tal. (MACARO; VANDERPLANK; MURPHY, 2010, pp. 77-8; VANPATTEN; WILLIAMS, 2015, p. 25). Entretanto, dado que nos estudos gerativistas os autores têm usado a expressão “Aquisição de Segunda Língua” de modo generalizado (PAIVA, 2014, p. 9), referindo-se a situações em que aconteça seja aprendizagem, seja aquisição, optou-se por fazer o mesmo no presente trabalho.

Igualmente, muitos autores fazem a distinção entre “segunda língua” (L2) e “língua estrangeira” (LE). Segundo eles, L2 seria considerada a língua oficial ou dominante na comunidade local, necessária para educação, trabalho, ou outros propósitos básicos. Seria considerada LE a língua estudada que não é falada na comunidade local, podendo ser usada para propósitos de viagens futuras ou como disciplina escolar, por exemplo (BOT; LOWIE; VERSPOOR, 2005, p. 7; LEFFA, 1988, pp. 212-3; SAVILLE-TROIKE, 2006, p. 4). Sendo que, nos estudos gerativistas, admite-se que os processos de aquisição de L2 ou de aprendizagem de LE são os mesmos, aqui também não é feita a distinção entre LE e L2.

verbos contendo dois ou mais argumentos na produção de nativas(os), contra 1,39% na das(dos) brasileiras(os) mais proficientes; finalmente, em relação aos clíticos dativos, as taxas foram de 1,04% e 0,00%, respectivamente; as(os) brasileiras(os) omitiram três vezes mais clíticos acusativos do que as(os) nativas(os)). Tal fato não foi constatado em relação aos pronomes clíticos das demais pessoas, cuja aquisição foi identificada como mais avançada do que a dos clíticos de terceira pessoa, estando as suas taxas de produção próximas às verificadas para as(os) nativas(os). Entretanto, a aquisição dos clíticos reflexivos de terceira pessoa não se mostrou problemática, dadas as semelhanças entre as suas formas na L1 e na L2, assim como a grande produtividade desses clíticos no PB; exemplos de dados de produção verificados são *On se connaît depuis longtemps*, “A gente se conhece há tempo” e *C’est ce qui s’est passé*, “Foi isso que aconteceu (que se passou)”. Os grupos clíticos e os pronomes *y* e *en*, por não encontrarem elementos correspondentes em português, também oferecem grande dificuldade para a aquisição por brasileiras(os), como também constatou Oliveira (2022): exemplos de produções que envolveriam grupos clíticos foram dados no início dessas Considerações; *y* e *en* tendem a ser omitidos, como em *En 2016 je suis restée ___ pendant 16 jours* em vez de [...] *j’y suis restée [...]*, “Em 2016 eu fiquei 16 dias (na França)” e *Je les aime, mais je n’ ___ ai pas* em vez de [...] *je n’en ai pas*, “Eu gosto deles, mas não tenho (nenhum animal de estimação)”.

A atenção à morfologia funcional da L2, assim como o *input* explícito, em que pode ser dada ênfase a pontos da língua que não costumam ser identificados implicitamente, pode servir a facilitar a aquisição da L2. Conhecer a língua materna da(do) aprendiz também é essencial, para que se possam concentrar os esforços de ensino nos pontos da L2 que diferem dos da L1, como ocorre em relação à aquisição dos clíticos pronominais do francês.

Referências

- AGUIAR, Ana Carolina Nunes de. *Clíticos Acusativos de Terceira Pessoa e o Input em Livros Didáticos no Ensino da Escrita do PB*. 2020. 240 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade de Brasília, Brasília, DF, 2020. Disponível em: https://repositorio.unb.br/bitstream/10482/39480/1/2020_AnaCarolinaNunesdeAguiar.pdf. Acesso em: 7 maio 2022.
- ANDERSON, Stephen R. *Aspects of the Theory of Clitics*. New York: Oxford University Press, 2005.
- BALL, Rodney. *Colloquial French Grammar: A Practical Guide*. Oxford: Blackwell Publishers, 2000.
- BECHARA, Evanildo. *Moderna Gramática Portuguesa*. Rio de Janeiro: Editora Nova Fronteira, 2015.
- BOT, Kees de; LOWIE, Wander; VERSPOOR, Marjolijn. *Second Language Acquisition: An advanced resource book*. London/New York: Routledge, 2005.
- BOULARÈS, Michèle; FRÉROT, Jean-Louis. *Grammaire Progressive du Français: Niveau avancé*. Paris: CLE International, 1997.
- CAINK, A. D. Clitics. In: BROWN, Keith. *Encyclopedia of Language and Linguistics*. Elsevier, 2005, pp. 491-5.
- CALLAMAND, Monique. *Grammaire Vivante du Français*. Paris: Larousse, 1987.

- CÂMARA Jr., Joaquim Matoso. *Estrutura da Língua Portuguesa*. 30ª ed. Petrópolis/RJ: Vozes, 1999 [1970].
- CARDINALETTI, Anna. Pronouns in Germanic and Romance Languages: An overview. In: RIEMSDIJK, Henk van. (org.) *Clitics in the Languages of Europe*. Berlin/New York: Mouton de Gruyter, 1999, pp. 33-82.
- CARVALHO, Danniell da Silva. Da natureza dos traços-phi. In: CARVALHO, Danniell da Silva. (org.) *Traços-phi: contribuições para a compreensão da gramática do português*. Salvador: EDUFBA, 2017, pp. 15-36.
- CHOMSKY, Noam. *Knowledge of Language: Its Nature, Origin, and Use*. Series: Convergence. New York: Praeger Publishers, 1986.
- CHOMSKY, Noam. *Lectures on Government and Binding: The Pisa Lectures*. 5th ed. Dordrecht (Holland) / Providence RI (USA): Foris Publications, 1988 [1981].
- CUNHA, Celso Ferreira da; CINTRA, Luís Filipe Lindley. *Nova Gramática do Português Contemporâneo*. 5. ed. Rio de Janeiro: Lexikon, 2011.
- CYRINO, Sonia Maria Lazzarini. *O Objeto Nulo no Português do Brasil: um estudo sintático-diacrônico*. 1994. 229 f. Tese (Doutorado em Ciências) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, SP, 1994. Disponível em: http://repositorio.unicamp.br/bitstream/REPOSIP/270355/1/Cyrino_SoniaMariaLazzarini_D.pdf. Acesso em: 22 nov. 2020.
- CYRINO, Sonia Maria Lazzarini. Observações sobre a Mudança Diacrônica do Português do Brasil: Objeto Nulo e Clíticos. In: ROBERTS, Ian; KATO, Mary A. (org.) *Português Brasileiro: Uma Viagem Diacrônica*. 3. ed. São Paulo; Contexto, 2018 [Campinas: Editora da Unicamp, 1993], pp. 129-42.
- CYRINO, Sonia Maria Lazzarini. Objetos nulos/pronomes plenos e topicalidade no português brasileiro. *Forum Linguístico*, Florianópolis, SC, v. 16, n. 1, p. 3482-3498, 2019. Disponível em: <https://periodicos.ufsc.br/index.php/forum/article/view/1984-8412.2019v16n1p3482/39205>. Acesso em: 7 maio 2022.
- CYRINO, Sonia Maria Lazzarini. Objetos Nulos em Português Brasileiro. *Cuadernos de la ALFAL*, n. 12 (2), pp. 387-410, noviembre 2020. Disponível em: https://www.mundoalfal.org/sites/default/files/revista/12_2_cuaderno_017.pdf. Acesso em: 7 maio 2020.
- CYRINO, Sonia Maria Lazzarini; MATOS, Gabriela. Null Objects and VP Ellipsis in European and Brazilian Portuguese. In: WETZELS, W. Leo; COSTA, João; MENUZZI, Sergio. (org.) *The Handbook of Portuguese Linguistics*. Oxford: Wiley Blackwell, 2016, pp. 294-317.
- DUARTE, Maria Eugénia Lammoglia. O Papel da Sociolinguística na Descrição da Gramática da Escrita Contemporânea. In: MARTINS, Marco Antônio; TAVARES, Maria Alice. (org.) *Contribuição da Sociolinguística e da Linguística Histórica para o Ensino de Língua Portuguesa*. Coleção Ciências da Linguagem Aplicadas ao Ensino, v. 5. Natal, RN: EDUFRN, 2013, pp. 115-42.
- DUBOIS, Jean; GIACOMO, Mathée; GUESPIN, Louis; MARCELLESI, Jean-Baptiste; MEVEL, Jean-Pierre. *Dicionário de Linguística*. São Paulo: Cultrix, 2006 [1978].
- FANJUL, Adrián. (org.) *Gramática y Práctica de Español para brasileños*. São Paulo: Moderna, 2005.

- FARACO, Carlos Alberto. O tratamento “você” em português: uma abordagem histórica. *LaborHistórico*, Rio de Janeiro, RJ, v. 3, n. 2, pp. 114-32, jul./dez. 2017. Disponível em: <https://revistas.ufrj.br/index.php/lh/article/view/17150/10437>. Acesso em: 7 maio 2022.
- GALVES, Charlotte. *Ensaio sobre as Gramáticas do Português*. Campinas, SP: Editora da Unicamp, 2001.
- GASS, Susan M.; SELINKER, Larry. *Second Language Acquisition: An Introductory Course*. Third Edition. New York: Routledge, 2008.
- GRÉGOIRE, Maïa; KOSTUCKI, Alina. *Grammaire Progressive du Français: Niveau perfectionnement avec 600 exercices*. France: CLE International, 2012.
- HVIDSTEN, Eirik; HELLAND, Hans Petter. L’acquisition du pronom « en » en français langue étrangère. *Synergies Pays Scandinaves*, n. 13, pp. 51-63, 2018. Disponível em: http://gerflint.fr/Base/Paysscandinaves13/hvidsten_helland.pdf. Acesso em: 7 maio 2020.
- JONES, Michael Allan. *Foundations of French syntax*. Melbourne, Australia: Cambridge University Press, 1996.
- KAISER, Georg A. *Die klitischen Personalpronomina im Französischen und Portugiesischen: Eine synchronische und diachronische Analyse*. Frankfurt am Main: Vervuert Verlag, 1992.
- KANTHACK, Gessilene Silveira. *Clíticos no Português Brasileiro*. 2002. 174f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, SC, 2002. Disponível em: <http://repositorio.ufsc.br/xmlui/handle/123456789/83641>. Acesso em: 7 maio 2022.
- KATO, Mary Aizawa. A Gramática do Letrado: Questões para a Teoria Gramatical. In: MARQUES, M. A.; KOLLER, E.; TEIXEIRA, J.; LEMOS, A. S. (org.) *Ciências da Linguagem: trinta anos de investigação e ensino*. Braga (Portugal): CEHUM (Universidade do Minho), 2005, pp. 131-45.
- KATO, Mary Aizawa. A Gramática Nuclear e a Língua-I do Brasileiro. In: MARTINS, Marco Antônio. (org.) *Gramática e Ensino*. Coleção Ciências da Linguagem Aplicadas ao Ensino, v. 1. Natal, RN: EDUFRN, 2013, pp. 147-64.
- KATO, Mary Aizawa; CYRINO, Sonia L.; CORRÊA, Vilma Reche. Brazilian Portuguese and the recovery of lost clitics through schooling. In: PIRES, Acrísio; ROTHMAN, Jason. *Minimalist Inquiries Into Child and Adult Language Acquisition: Case Studies Across Portuguese*. Berlin/New York: Walter de Gruyter, 2009, pp. 245-71.
- KATO, Mary Aizawa; DUARTE, Maria Eugênia Lammoglia. Restrições na distribuição de sujeitos nulos no Português Brasileiro. *Veredas: Sintaxe das Línguas Brasileiras*, Juiz de Fora, MG, v. 18/1, pp. 1-22, 2014. Disponível em: https://www.ufjf.br/revistaveredas/files/2014/07/01-Kato_Duarte1.pdf. Acesso em: 7 maio 2022.
- LEFFA, Vilson J. Metodologia do Ensino de Línguas. In: BOHN, Hilário Inácio; VANDRESEN, Paulino. (org.) *Tópicos de Linguística Aplicada: O Ensino de Línguas Estrangeiras*. Florianópolis: Editora da UFSC, 1988, pp. 211-36.
- LÉON, Pierre R. *Phonétisme et prononciations du français*. 5. éd. Paris: Armand Colin, 2007.
- LOPES, Célia Regina. Pronomes Pessoais. In: VIEIRA, Silvia Rodrigues; BRANDÃO, Silvia Figueiredo. (org.) *Ensino de Gramática: descrição e uso*. 2. ed. São Paulo: Contexto, 2018 [2007], pp. 103-19.

- LUÍS, Ana R.; KAISER, Georg A. Clitic Pronouns: Phonology, Morphology, and Syntax. In: WETZELS, W. Leo; COSTA, João; MENUZZI, Sergio. (org.) *The Handbook of Portuguese Linguistics*. West Sussex: Wiley Blackwell, 2016, pp. 210-33.
- MACARO, Ernesto; VANDERPLANK, Robert; MURPHY, Victoria A. A Compendium of Key Concepts in Second Language Acquisition. In: MACARO, Ernesto. (org.) *The Continuum Companion to Second Language Acquisition*. New York/London: Continuum, 2010, pp. 29-106.
- MANZINI, M. Rita; SAVOIA, Leonardo M. What orders the romance clitic string? Against precompiled hierarchies. In: CARVALHO, Danniell da Silva; BRITO, Dorothy Bezerra Silva de. (org.) *Pronomes: Morfossintaxe e Semântica*. Salvador: EDUFBA, 2018, pp. 15-37.
- MARTINS, Marco Antônio. *A colocação de pronomes clíticos na escrita brasileira: Para o estudo das gramáticas do português*. Natal/RN: Editora da UFRN, 2012.
- MARTINS, Ana Maria. Posição dos Pronomes Pessoais Clíticos. In: PAIVA RAPOSO, Eduardo Buzaglo; BACELAR DO NASCIMENTO, Maria Fernanda; COELHO DA MOTA, Maria Antônia; SEGURA, Luísa; MENDES, Amália. (org.) *Gramática do Português: Volume II*. Lisboa: Fundação Calouste Gulbenkian, 2013, pp. 2231-302.
- MILLER, Philip; MONACHESI, Paola. Les pronoms clitiques dans les langues romanes. In: GODARD, Danièle. *Les langues romanes: Problèmes de la phrase simple*. Paris: Editions du CNRS, 2003. Disponível em: <http://docplayer.fr/22076414-Les-pronoms-clitiques-dans-les-langues-romanes.html>. Acesso em: 7 maio 2022.
- MORIN, Yves-Charles. La morphophonologie des pronoms clitiques en français populaire. *Cahier de linguistique*, v. 9, pp. 1-36, 1979. Disponível em: <<https://doi.org/10.7202/800076ar>>. Acesso em: 7 maio 2022.
- NUNES, Jairo. Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em português brasileiro. In: ROBERTS, Ian; KATO, Mary A. (org.) *Português Brasileiro: Uma Viagem Diacrônica*. 3. ed. São Paulo; Contexto, 2018 [Campinas: Editora da Unicamp, 1993], pp. 161-73.
- OLIVEIRA, Luciano de. *Estudos dos Clíticos e das Construções com Clitic Dislocation em Línguas Neolatinas*. 2016. 169 f. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal e Santa Catarina, Florianópolis, SC, 2016. Disponível em: <http://tede.ufsc.br/teses/PLLG0639-D.pdf>. Acesso em: 7 maio 2022.
- OLIVEIRA, Luciano de. Estratégias de Indeterminação do Sujeito em Línguas Românicas. *Revista do GELNE*, Natal, RN, v. 20, n. 2, pp. 30-45, 2018. Disponível em: <https://periodicos.ufrn.br/gelne/article/view/14356/10907>. Acesso em: 7 maio 2022.
- OLIVEIRA, Luciano de. *Análise de Aquisição dos Clíticos Pronominais e de Posições Sintáticas Não Nulas do Francês por Aprendiz(es) Brasileiras(os) Adultas(os)*. 2021. 407 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, SC, 2022. Disponível em: <https://repositorio.ufsc.br/handle/123456789/231250>. Acesso em: 7 maio 2022.
- PAIVA, Vera Lúcia Menezes de Oliveira e. *Aquisição de Segunda Língua*. São Paulo: Parábola, 2014.
- PRÉVOST, Philippe. *The Acquisition of French: The development of inflectional morphology and syntax in L1 acquisition, bilingualism, and L2 acquisition*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2009.

QUAREZEMIN, Sandra; FUCHSBERGER, Gabriel. Indeterminação do sujeito no português brasileiro: o pessoal tá inventando moda. In: CASTAGNA, Vanessa; QUAREZEMIN, Sandra. (org.) *Travessias em Língua Portuguesa: Pesquisa linguística, ensino e tradução*. Sail 16. Venezia (Itália): Edizioni Ca' Foscari, 2020, pp. 93-106. Disponível em: <https://edizionicafoscari.unive.it/media/pdf/books/978-88-6969-461-5/978-88-6969-461-5.pdf>. Acesso em: 7 maio 2022.

RIEGEL, Martin; PELLAT, Jean-Christophe; RIOUL, René. *Grammaire méthodique du français*. 3^{ème} ed. PUF: 2004.

SAVILLE-TROIKE, Muriel. *Introducing Second Language Acquisition*. New York: Cambridge University Press, 2006.

SABIO, Frédéric. Les pronoms clitiques et l'expression du lieu: l'usage de y et en dans le français contemporain. *Le français parlé au XXI^e siècle*, Oxford, UK, pp. 87-105, jun. 2005. Disponível em: <https://hal.archives-ouvertes.fr/hal-00576847/>. Acesso em: 7 maio 2022.

TARALLO, Emilio Gozze. Clíticos, Mudança e Seleção Natural. In: ROBERTS, Ian; KATO, Mary A. (org.) *Português Brasileiro: Uma Viagem Diacrônica*. 3. ed. São Paulo; Contexto, 2018 [Campinas: Editora da Unicamp, 1993], pp. 143-60.

TSEDRYK, Kanstantsin. *Les interactions morphosyntaxiques dans l'interlangue des apprenants: le cas de l'acquisition des pronoms clitiques en français L2*. 2014. 271 f. Tese (Doctor of Philosophy) – Graduate Program in French Studies, The University of Western Ontario, Ontario, 2014. Disponível em: <https://ir.lib.uwo.ca/etd/2062>. Acesso em: 7 maio 2022.

VANPATTEN, Bill; BENATI, Alessandro G. *Key Terms in Second Language Acquisition*. London: Continuum, 2010.

VANPATTEN, Bill; WILLIAMS, Jessica. (org.) *Theories in Second Language Acquisition: An Introduction*. New York/London: Routledge, 2015.

SE-PASSIVO, SE-IMPESSOAL E SUJEITOS NULOS INDEFINIDOS: UMA ABORDAGEM MINIMALISTA PARA A PERDA E O SURGIMENTO DE CONSTRUÇÕES NA GRAMÁTICA DA LÍNGUA¹

PASSIVE-SE, IMPERSONAL-SE, AND INDEFINITE NULL SUBJECTS: A MINIMALIST APPROACH TO THE LOSS AND THE EMERGENCE OF CONSTRUCTIONS IN A LANGUAGE'S GRAMMAR

Humberto Borges²

RESUMO

Contrariamente ao que tem sido proposto por parte da literatura gerativa, que vê a perda das construções com o chamado *se*-passivo no português brasileiro como resultado do enfraquecimento da concordância dos verbos transitivos, este artigo argumenta com base em dados empíricos de que a perda das construções com *se*-passivo na história do português brasileiro está crucialmente associada à perda de uma projeção funcional mais alta que T, nomeadamente: FP. Essa é a posição em que o verbo dessas construções estabelece concordância com o argumento interno no português europeu (RAPOSO; URIAGEREKA, 1996). Arrolamos esse achado às mudanças na gramática do português brasileiro desencadeadas pela perda de FP previstas por Pires (2006). Paralelamente, argumentamos que o surgimento das construções com sujeitos nulos indefinidos de 3ª pessoa singular, como em “Vende casas”, também não está associada à perda da concordância nos verbos transitivos e ao desuso do clítico *se* (passivo ou impessoal), como se tem tentativamente argumentado, mas, sim, à perda do traço-D de T, o que permitiu o português brasileiro – assim como outras línguas de sujeito nulo parcial – licenciar sujeitos nulos indefinidos de 3ª pessoa singular.

PALAVRAS-CHAVE: *Se*-passivo. *Se*-impessoal. Sujeitos nulos. Português brasileiro. Português europeu.

ABSTRACT

Contrary to what has been proposed in most generative approaches which see the loss of passive-*se* constructions in Brazilian Portuguese as an outcome of the impoverishment of the verbal morphology, in this article, we show empirical evidence that the loss of constructions with the so-called passive-*se* in the history of Brazilian Portuguese is crucially associated with the loss of a higher functional projection than T, namely: FP. This was the position in which the transitive verb of these constructions established agreement with the internal argument in European Portuguese (RAPOSO; URIAGEREKA, 1996). We link this finding to the changes in Brazilian Portuguese grammar triggered by the loss of FP predicted by Pires (2006). At the same time, we argue that the emergence of constructions with indefinite null subjects in the 3rd person singular, as in “*Vende casas*” (“(One) sells houses”), is also not associated with the loss of agreement in transitive verbs and the disuse of the clitic *se* (passive or impersonal), as it has been tentatively argued in the literature, but to the loss of the D-feature of T, which allowed Brazilian Portuguese – as well as other partial null-subject languages – to license indefinite null-subjects in the 3rd person singular.

KEYWORDS: Passive-*se*. Impersonal-*se*. Null subjects. Brazilian portuguese. European portuguese.

¹ Trabalho realizado com financiamento da Capes (código de financiamento 001/Capes-Print e processo n. 007549/2015-06) e da Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF).

² Professor Adjunto de Linguística e Formação de Professores de Língua Portuguesa da Universidade Federal de Jataí, humbertoborges@ufj.edu.br, <http://orcid.org/0000-0001-5492-7264>.

Introdução

Argumenta-se que o empobrecimento da concordância dos verbos transitivos, especialmente da 3ª pessoa do plural, é responsável pela perda das construções com *se*-passivo no português brasileiro (PB) coloquial (GALVES, 1985; NUNES, 1990; dentre outros). Nessa perspectiva, Martins e Nunes (2016) afirmam que as raras e inconsistentes construções com *se*-passivo no PB são exclusivas da escrita e ao estilo guiados pela tradição gramatical prescritivista, que condena o uso do clítico *se* com verbo transitivo no singular e DP argumento interno plural. Similarmente, Cavalcante (2018) argumenta que a aceitabilidade das construções com *se*-passivo por falantes do PB é resultante de saberes linguísticos construídos especialmente por falantes letrados que se espelham em modelos de uma outra gramática na construção de dados linguísticos. Silva (2017) mostra, por meio de estudo em aquisição de língua, que as construções com *se*-passivo e *se*-impessoal não fazem parte da gramática adquirida na infância pelo falante do PB e que elas são, portanto, estruturas adquiridas principalmente via escolarização.

Neste artigo, com base em evidências de um *corpus* diacrônico e dialetal, propomos que a perda das chamadas construções com *se*-passivo no PB pode ser interpretada, conforme previu Pires (2006), como uma consequência do enfraquecimento e da perda de FP (do inglês, *Functional Phrase*; cf. URIAGEREKA, 1995) na gramática do PB. Argumentamos, ainda, que a perda do clítico *se*-impessoal e o surgimento das construções com sujeitos nulos indefinidos são consequências da perda do traço-D de T na gramática do PB, i.e., é resultado de uma mudança paramétrica: de língua de sujeito nulo (LSN) consistente para LSN parcial.

O artigo está estruturado como segue. A seção 1 apresenta estudos descritivos e formais sobre as construções com *se*-passivo e *se*-impessoal no português europeu (PE) e no PB. Tomamos a análise de Raposo e Uriagereka (1996) para argumentar que nas construções com *se*-passivo o clítico ocupa a posição de Spec-TP, e o verbo é projetado para FP, domínio em que é possível, no PE, ser amalgamada a concordância entre o DP argumento interno e o verbo transitivo. Consoante Pires (2006), defendemos que a projeção FP não está mais ativa no PB. Na seção 2, assumimos, conforme Holmberg (2010), que o PB é uma LSN sem um traço de definitude não interpretável [*uD*] em T. Por meio de dados históricos de Goiás, mostramos que a perda das construções com *se*-passivo e *se*-impessoal no PB podem ser interpretadas, respectivamente, como resultado da perda: da projeção FP; e do traço-D de T. Por fim, temos as considerações finais.

1. Aspectos descritivos e formais das construções com *se*-passivo e *se*-impessoal

Segundo Duarte (2013), as construções com *se*-passivo no PE têm o mesmo alinhamento entre papéis temáticos que as passivas verbais, pois seus argumentos internos se alinham à função sintática de sujeito, e o argumento externo está implícito, como, respectivamente, em (1) e (2). Com efeito, a autora afirma que as construções com *se*-passivo no PE se restringem aos verbos transitivos ou bitransitivos, pois nessas construções o argumento interno deve controlar os traços de número do verbo.

- (1) a. O raciocínio algébrico ensina-se propositadamente na escola. (PE)
 b. As obras mais recentes publicaram-se numa editora espanhola para garantir maiores tiragens. (PE) (DUARTE, 2013, p. 444)
- (2) a. O raciocínio algébrico é ensinado propositadamente na escola.
 b. As obras mais recentes foram publicadas numa editora espanhola para garantir maiores tiragens. (PE) (DUARTE, 2013, p. 444)

Para Duarte (2013), nas construções com *se*-passivo, o DP argumento interno controla os traços de número do verbo no singular e no plural, como em (1a) e (1b). Ela explica que a interpretação do agente da passiva nas construções com *se*-passivo pode ser extraída pela possibilidade de sua ocorrência: com um advérbio orientado para o agente (1a); e com orações subordinadas finais formadas a partir de um adjunto cujo sujeito só pode ser controlado pelo argumento externo implícito da passiva (1b). A autora propõe que, nessas construções, o clítico denota uma entidade arbitrária relacionada com o papel temático de agente da passiva, de modo que o argumento externo está implícito. Isto é, ela argumenta que o *se*-passivo atua como um morfema apassivador, expressando a diátese passiva da construção sem realizar qualquer alteração na morfologia do verbo e bloqueando a atribuição de papel temático à posição do argumento externo e de Caso acusativo ao argumento interno. A autora insere o *se*-passivo dentro do escopo dos clíticos quase argumentais.

Duarte (2013) afirma que, quando o verbo de uma construção com *se* é conjugado na 3ª pessoa do singular, há ambiguidade de interpretação entre a leitura ativa ou passiva (3).

- (3) Descobriu-se um gene responsável pela obesidade. (PE) (DUARTE, 2013, p. 446)
 (Interpretação 1, passiva pronominal: ‘Foi descoberto um gene responsável pela obesidade.’)
 (Interpretação 2, se-impessoal: ‘Alguém descobriu um gene responsável pela obesidade.’)

Quando analisado, na leitura ativa, como impessoal ou nominativo, o clítico *se* é interpretado como o sujeito da oração (BRITO; DUARTE; MATOS, 2003; DUARTE, 2013). Brito, Duarte e Matos (2003) apresentam dois argumentos em favor da análise de que o *se*-impessoal possui conteúdo argumental e funcional: o primeiro é o seu caráter referencial, que permite a extração simultânea do clítico em orações coordenadas (4a); o segundo é o de que esse caráter referencial não permite que ele ocorra associado a uma posição de expletivo (4b).

- (4) a. Informa-se que se aluga apartamentos e __ vende moradias. (PE)
 b. *Há-se muitos livros nesta biblioteca. (PE)

(BRITO; DUARTE; MATOS, 2003, p. 837)

Brito, Duarte e Matos (2003) indicam, ainda, que as construções com *se*-passivo no PE contemporâneo não admitem o redobro do clítico, de modo que um argumento interpretado como agente não pode ser expresso por meio da preposição *por* (5).

- (5) *Venderam-se hoje muitos livros por alguém na feira do livro. (PE)

(BRITO; DUARTE; MATOS, 2003, p. 839)

Duarte (2013) pontua que o *se*-impessoal no PE é licenciado: (i) com verbos transitivos na 3ª pessoa do singular que expressam um traço de indeterminação do sujeito, parafraseável por “alguém”, “as pessoas (em geral)” ou por “há pessoas que” (6a); (ii) com verbos intransitivos que não selecionam argumento interno (6b); e (iii) com verbos inacusativos (6c).

- (6) a. Diz-se que esse ministro vai ser demitido. (PE) (DUARTE, 2013, p. 445)
 b. Incorre-se frequentemente nesse erro. (PE)
 c. Morria-se de sarampo no princípio do século passado. (PE)

Raposo (2013) fornece outra descrição para a estrutura argumental das orações com o *se*-passivo e o *se*-impessoal no PE. Ele reconhece haver um argumento externo implícito nas construções com *se*-passivo (ou passivas pronominais/reflexas), o que as leva a se assemelhar com a estrutura temática das passivas verbais, possibilitando, por exemplo, uma interpretação agentiva quando elas ocorrem com orações subordinadas, como em (7a).

- (7) a. Declamaram-se vários poemas de Fernando Pessoa (para comemorar o aniversário do poeta). [frase passiva pronominal] (PE) (RAPOSO, 2013, p. 391)
 b. Vários poemas de Fernando Pessoa foram declamados (pelo autor) (para comemorar o aniversário do poeta). [frase passiva verbal] (PE)

Raposo (2013) pontua que nas construções com *se*-passivo, assim como nas passivas verbais, o DP paciente possui a função gramatical de sujeito, isto é, controla os traços de concordância do verbo; por outro lado, diferentemente das passivas verbais, o argumento externo implícito nunca pode ser manifesto, como em (5). Para o autor, apesar de as construções com *se*-impessoal também compartilharem dessa propriedade das construções com *se*-passivo, nelas o DP argumento interno não controla os traços de concordância do verbo, como evidenciado mais especificamente em (8c).

- (8) a. Dançou-se muito na festa de ontem. (PE) (RAPOSO, 2013, p. 392)
 b. No começo do jogo, assobiou-se ao tocar o hino nacional. (PE)
 c. Comeu-se morangos à sobremesa. (PE)

Raposo (2013) defende que o *se*-passivo se restringe às construções em que o DP argumento interno plural controla os traços de concordância do verbo. Ele esquematiza as seguintes estruturas argumentais para as construções com *se*-passivo e *se*-impessoal:

Martins e Nunes (2016) salientam que o argumento externo implícito das passivas com *se* no PE contemporâneo, que não mais admitem a *by-phrase*, ainda só é compatível com uma leitura [+humana], ao passo em que o argumento externo implícito das passivas verbais é compatível com uma interpretação [+humana] ou [-humana], como exemplificado a seguir.

- (11) a. As casas foram destruídas. (MARTINS; NUNES, 2016, p. 326)
 b. Destruíram-se as casas.

Em (11a), pode-se interpretar que as casas podem ter sido destruídas tanto por seus proprietários, que têm traço [+humano], quanto pela chuva, que tem traço [-humano]: “as casas foram destruídas pelo dono” ou “as casas foram destruídas pela chuva”. Em contraste, em (11b), somente é permitida uma interpretação [+humana] para o argumento externo implícito – ou seja, (11b) não pode ser empregada para descrever a destruição de casas pela chuva. Os autores afirmam que as construções com *se*-impessoal herdaram das construções com *se*-passivo a restrição de uma interpretação [+humana] para o argumento externo implícito, como evidenciado em (12).

- (12) Destruiu-se as casas. (MARTINS; NUNES, 2016, p. 326)

Consoante Naro (1976), a reanálise na voz verbal das construções com *se*-passivo teria possibilitado não somente o surgimento das construções com *se*-impessoal com verbos transitivos, mas também com verbos intransitivos. Com variações na gramaticalidade de acordo com a variedade e o dialeto do falante, Martins e Nunes (2016) exibem exemplos da expansão do uso do *se*-impessoal com outros verbos não transitivos no PE e no PB:

- (13) a. Precisa-se de funcionários. (Verbo transitivo com complemento preposicional).
 b. Trabalha-se muito nesta cidade. (Verbo inergativo).
 c. Chegava-se cedo ao trabalho. (Verbo inacusativo).
 d. Quando se é promovido, as coisas ficam mais fáceis. (Verbo passivo).
 e. Não se ficou contente com a nova situação. (Verbo de ligação).
 f. Parecia-se ir ganhar o jogo. (Verbo de alicamento: PE^{OK}; PB*).

(MARTINS; NUNES, 2016, pp. 328-9)

Martins e Nunes (2016) chamam atenção para o fato de o DP argumento interno das construções com *se*-passivo no PE poder se mover para uma posição pré-verbal (14a)-(15a), ocupando a posição de sujeito tal como ocorre nas passivas verbais (15b). As construções com *se*-impessoal já não permitem o deslocamento do objeto (14b).

- (14) a. Os bolos comeram-se ontem. (PE)(MARTINS; NUNES, 2016, p. 323)
 b. *Os bolos comeu-se ontem. (PE)

- (15) a. Estas maçãs todas apanharam-se sem estarem maduras. (PE)
 b. Estas maçãs todas foram apanhadas sem estarem maduras. (PE)

(MARTINS; NUNES, 2016, p. 325)

Os autores salientam que a construção impessoal em (14b) pode ser julgada como aceitável no PE com uma entonação marcada, graficamente representada por uma vírgula: “Os bolos, comeu-se ontem”. Por sua vez, as construções (14a) e (15a) não requerem qualquer entonação especial para serem licenciadas. Os autores tratam o contraste entre (14a) e (14b) como uma evidência de que o clítico *se* nas construções em que o DP plural não concorda com o verbo deve ser analisado como impessoal, e não como passivo. Aqui se nota que Martins e Nunes (2016), assim como Raposo (2013), divergem da análise de Duarte (2013) e restringem o *se*-passivo às construções em que o verbo concorda com o DP plural. Nessa direção, Martins e Nunes (2016) pontuam, por meio dos exemplos em (16)-(18), que as construções com *se*-passivo no PE não atribuem Caso acusativo ao argumento interno:

- (16) a. *Compraram-se-os ontem. (PE) (MARTINS; NUNES, 2016, p. 324)
 b. Comprou-se-os ontem. (PE)
- (17) a. *Podem-se comprá-los amanhã. (PE) (MARTINS; NUNES, 2016, p. 324)
 b. Pode-se comprá-los amanhã. (PE)
- (18) a. *Histórias de lobisomens, ouviam-se-lhas vezes sem conta. (PE)
 b. Histórias de lobisomens, ouvia-se-lhas vezes sem conta. (PE)

(MARTINS; NUNES, 2016, p. 324)

A respeito dos dados em que o clítico *se* vem seguido por um clítico acusativo, Martins e Nunes (2016) afirmam que: (16a), (17a) e (18a) mostram que as construções com *se*-passivo não atribuem Caso acusativo ao argumento interno de seus verbos; ao passo em que as construções com *se*-impessoal atribuem Caso acusativo ao argumento interno em diferentes contextos: quando os clíticos estão encadeados em um mesmo hospedeiro (16b); quando ambos são colocados em hospedeiros diferentes (17b); e quando um clítico dativo intervém entre o *se* e o clítico acusativo (18b). Para Martins e Nunes (2016), essas evidências corroboram a análise de que, apesar de apresentarem propriedades semânticas semelhantes às construções com *se*-passivo, as construções com *se*-impessoal não são sintaticamente caracterizadas como passivas, isto é, nas construções com *se*-impessoal, o argumento interno (plural ou singular) não controla os traços de flexão do verbo, enquanto nas passivas com *se* o argumento interno deve estar no plural e controlar os traços de flexão do verbo. Para os autores, a agramaticalidade de (16a), (17a) e (18a) confirma o argumento de que o *se*-passivo se comporta como um participio de morfologia passiva na medida em que desativa a propriedade de o verbo atribuir Caso acusativo.

Martins e Nunes (2016) destacam, ainda, o fato de que no PE construções com *se*-impessoal licenciam predicado secundário orientado para o sujeito (19).

- (19) a. *Criam-se avestruzes despreocupado. (MARTINS; NUNES, 2016, p. 327)
 b. Cria-se avestruzes despreocupado.

A restrição de um predicado secundário orientado para o sujeito ocorrer somente em construções com *se*-impessoal leva os autores a argumentar que o *se*-impessoal checa o Caso nominativo disponível na derivação e é, portanto, o sujeito da oração. Raposo e Uriagereka (1996) também propuseram que o clítico checa o Caso nominativo nas construções com *se*-impessoal, entretanto, para esses autores, o *se*-passivo no PE não atua como um item de morfologia passiva, como sugerem Duarte (2013) e Martins e Nunes (2016).

Raposo e Uriagereka (1996) propõem que o *se*-passivo ocupa a mesma posição que o *se*-impessoal na estrutural derivacional, Spec-TP, com a diferença de que nas construções com *se*-passivo o clítico recebe Caso nulo em Spec-TP e o DP argumento interno recebe Caso nominativo em uma projeção mais alta que TP. O trabalho de Raposo e Uriagereka (1996) lida com as seguintes questões sobre as construções com *se*-passivo no PE: (i) a posição do DP argumento interno e do clítico; (ii) a atribuição de Caso ao DP argumento interno e ao clítico; e (iii) o estabelecimento da concordância entre o DP argumento interno e o verbo.

Testando o comportamento do DP argumento interno das construções com *se*-passivo no PE, os autores comparam-no com o DP sujeito de sentenças infinitivas complementos de predicados adjetivos, que pode ocupar uma posição pré-verbal nas formas ativa e passiva (20a-b). Na comparação, verifica-se que o DP argumento interno das passivas com *se* não é aceito em posição pré-verbal (20d), o que sugere que esse DP não ocupa Spec-TP.

- (20) a. Vai ser difícil os tribunais aceitarem os documentos. (PE)
 b. Vai ser difícil os documentos serem aceites. (PE)
 c. Vai ser difícil aceitarem-se os documentos. (PE)
 d. *Vai ser difícil os documentos aceitarem-se. (PE)

(RAPOSO; URIAGEREKA, 1996, p. 754)

Em outro dos testes empregados, os autores analisam o comportamento do DP das passivas com *se* em construções cuja semântica do verbo permite extrair tanto uma leitura reflexiva quanto uma leitura passiva do clítico, a depender da posição que o DP ocupa. Assim, em (21a), há leitura ambígua (passiva ou recíproca), e, em (21b), obrigatoriamente uma leitura recíproca.

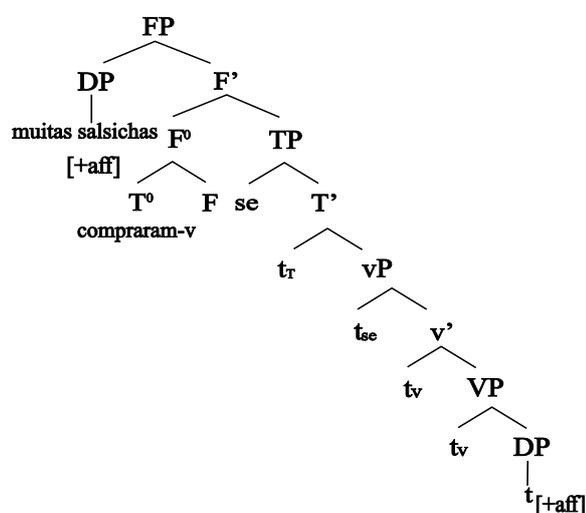
- (21) a. Os especialistas consultaram-se durante a operação. (PE)
 b. Em que momento da operação os especialistas se consultaram? (PE)

(RAPOSO; URIAGEREKA, 1996, pp. 765-6)

Raposo e Uriagereka (1996) afirmam que a leitura da construção em (20a) é ambígua entre uma interpretação passiva ou recíproca devido ao traço [+humano] do DP e à possibilidade de ele ocupar duas posições: na leitura recíproca, Spec-TP; na leitura passiva, uma posição à esquerda de T, nomeadamente, FP. Já a construção em (21b) possui apenas leitura recíproca, pois o adjunto é deslocado para uma posição de tópico, e, assim, o DP assume a posição de sujeito. É justamente por o DP ocupar a posição de sujeito que (21b) tem apenas leitura recíproca. Assim, a leitura passiva de (21a) ocorre com o DP em Spec-FP, enquanto a leitura recíproca de (21b) ocorre com o DP em Spec-TP. (22) e (23) ilustram, respectivamente, o DP em posição pré e pós-verbal nas construções com *se*-passivo no PE.³

(22) a. Muitas salsichas compraram-se. (RAPOSO; URIAGEREKA, 1996, p. 786)

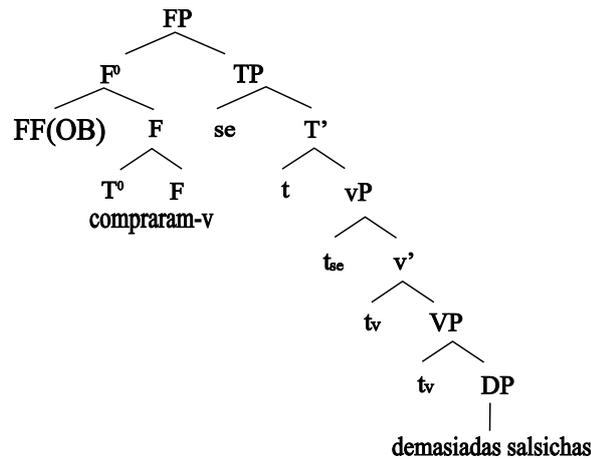
b.



Raposo e Uriagereka (1996) argumentam que em (22) o DP argumento interno atua como um sintagma afetivo (do inglês, *affective phrase*), ou seja: como um elemento sintático que ocupa Spec-FP por ter um traço de afetividade forte [+Aff], estabelecer a concordância de número e pessoa com o verbo e checar o Caso nominativo da derivação. Os autores destacam que, em (22), o clítico *se* não bloqueia o movimento do DP, pois não possui os traços que poderiam ser atraídos por F, como [+Aff] e [+número]. As operações de movimento em (22) são, portanto: a subida do clítico *se* para Spec-TP, a subida do complexo V-v para T e o subsequente movimento de T contendo o complexo V-v para F. Para os autores, o movimento do DP para Spec-FP em (22) é uma operação de último recurso.

³ Raposo e Uriagereka (1996) assumem que um sintagma verbal transitivo ou inergativo consiste em uma configuração [v-VP] responsável pela atribuição dos papéis temáticos, onde o sujeito é atribuído ao argumento externo em Spec-v e o objeto (OB) ao argumento interno em V.

- (23) a. Compraram-se demasiadas salsichas. (RAPOSO; URIAGEREKA, 1996, p. 787)⁴
 b.



Em (23), Raposo e Uriagereka (1996) argumentam que o DP não ocupa Spec-FP – é, portanto, um argumento interno *in situ* – e não atua como um sintagma afetivo, de modo que nessas construções há a checagem dos traços formais do DP – com exceção de [+Aff] – por F sem movimento manifesto. Isto é, em (23), o que é atraído para FP não é a categoria completa, mas traços formais particulares, mais precisamente os traços de número, pessoa e Caso nominativo. Seguindo Chomsky (1995), os autores argumentam que os traços de uma categoria completa são checados em movimento manifesto apenas por razões de convergência na PF, o que implica dizer que o movimento implícito na LF é apenas para a atração de traços formais, mas não da própria categoria. Assim, quando um traço formal específico de uma determinada categoria é atraído na LF, o movimento carrega automaticamente o conjunto de traços formais da categoria.

Em relação à natureza do clítico *se* nos dados em tela, Raposo e Uriagereka (1996) afirmam que o *se* é um DP mínimo que possui traços semânticos reduzidos, como PRO; assim, nas construções com *se*-passivo, o clítico checa o traço-D de T, recebe Caso nulo (*minimal Case*) em Spec-TP e não impede a atribuição de Caso nominativo por FP, que verifica os traços de número e pessoa do DP argumento interno. Isso sugere que nas construções com *se*-impessoal, o clítico checa o traço-D de Spec-TP e recebe Caso nominativo, tendo em vista ser o único elemento disponível para essa operação na derivação.

O tratamento de Raposo e Uriagereka (1996) para dar conta da concordância do argumento interno quando movido e *in situ* nas construções com *se*-passivo levanta a questão a respeito de o que é o sujeito da oração no PE: se o elemento que ocupa Spec-TP ou se o elemento que carrega o Caso nominativo e desencadeia a concordância verbal. No primeiro caso, o DP argumento interno das construções com *se*-passivo não pode ser o sujeito da oração, mas, sim, o clítico *se*. No segundo caso, é necessário expandir a noção de sujeito para além do domínio de T e admitir que, em línguas com F

⁴ Em (23b), FF se refere aos traços formais (do inglês, *formal features*) do objeto.

sintaticamente ativo, o sujeito também pode ser checado em F. Destaca-se que Raposo e Uriagereka (1996) denominam o *se*-passivo de *se*-indefinido, pois, para eles, os fatos de o DP estabelecer a concordância com o verbo por meio da projeção F e de o clítico receber Caso nulo em Spec-TP evidenciam que essas construções são ativas. Já o chamado *se*-impessoal é denominado pelos autores de *se*-genérico.

Raposo e Uriagereka (1996) fornecem uma análise elegante para explicar o comportamento das construções com *se*-passivo/indefinido no PE. Diferentemente do que é proposto por Duarte (2013) e Martins e Nunes (2016), Raposo e Uriagereka (1996) argumentam que o chamado *se*-passivo não atua como um particípio de morfologia passiva: o que acontece é que ele recebe Caso nulo em Spec-TP, ao passo em que o argumento interno recebe Caso nominativo por meio da checagem de seus traços formais por F. A necessidade de atribuir Caso nominativo ao argumento interno nas construções com *se*-passivo explicaria a agramaticalidade de sentenças como aquelas apresentadas por Martins e Nunes (2016) em (16)-(18) sem a necessidade de postular que o clítico atuaria como um particípio de morfologia passiva. A explicação de Raposo e Uriagereka (1996), todavia, não dá conta do PB, uma vez que a projeção F não está mais sintaticamente ativa nessa língua (cf. PIRES, 2006), impossibilitando, por consequência, a ocorrência de construções com *se*-passivo/indefinido.

Pires (2006) salienta que a projeção FP deve ter estado ativa na gramática do PB até pelo menos o século XIX, o que torna plausível sugerir que a perda dessa projeção pode estar relacionada com várias mudanças desencadeadas concomitantemente na gramática do PB. Pires (2006) observa que, contanto que o movimento de um núcleo ocorra localmente, no movimento de V-para-C no português, V teria que se mover primeiramente para F, e só então para C. Dessa forma, a perda de V-para-C pode ter sido um gatilho crucial para a perda de V-para-F no PB. O autor cita como exemplos de mudanças vinculadas à perda da projeção FP: mudança na colocação de clíticos, perda do movimento de V-para-F e perda da inversão nos infinitivos.

Pires (2006) mostra que a perda do núcleo funcional F no PB eliminou o movimento aberto do verbo para a posição à esquerda de um clítico, provocando a preferência da próclise no lugar da ênclise (24) – sua análise incide sobre os clíticos de 1ª e 2ª pessoa, dado que as propriedades de colocação dos clíticos no PB não se estendem aos clíticos acusativos de 3ª pessoa, cujo comportamento tem outros aspectos.

- (24) a. Ela me/te viu ontem. (*viu-me/*te) (PB) (PIRES, 2006, p. 158)
 b. (Ela) viu-me/te ontem. (*me/*te viu) (PE/século XVIII)

Pires (2006) incrementa seu argumento mencionando os resultados do trabalho de Cyrino (1993, pp. 167-8) sobre a posição do clítico em peças de teatro, os quais indicam uma perda sistemática da ênclise com verbos no infinitivo impessoal e no gerúndio no PB. Ele sistematiza os resultados de Cyrino (1993) em relação à posição enclítica dos clíticos com verbos no infinitivo impessoal e no gerúndio conforme a tabela 1.

Tabela 1: Clíticos em posição enclítica em Cyrino (1993).

Período	Infinitivos impessoais	Gerúndios
1700	100%	100%
1800	100%	100%
1850	86%	100%
1900	56%	25%
1950b	25%	66%
1950b	0%	0%

Fonte: Pires (2006, p. 149).

Pires (2006) afirma que esses resultados fornecem evidência consistente de que o movimento do verbo para F foi em geral perdido na segunda metade do século XX. Ademais, o pesquisador pontua que: “Se a hipótese aqui considerada estiver no caminho certo, deve-se esperar padrões semelhantes de mudança para outras propriedades possivelmente associadas à ativação da projeção FP” (PIRES, 2006, p. 160, tradução nossa).⁵ Nessa direção, sugerimos que a perda das construções com *se*-passivo na gramática do PB pode ser arrolada nesse rol de mudanças provocadas pela perda do núcleo funcional F na gramática do PB. Voltaremos a essa questão na seção seguinte.

Martins e Nunes (2016) observam que, apesar de a gramática do PB ser proclítica, o *se*-impessoal é sempre enclítico, salvo quando há uma partícula atratora de próclise.

- (25) a. Contratou-se/*Se contratou um novo professor. (PB)
 b. Não se contratou/*contratou-se um novo professor. (PB)
 c. *Deve, sem sombra de dúvida, se contratar um novo professor. (PB)
 d. Deve-se, sem sombra de dúvida, contratar um novo professor. (PB)
 e. *Não deve, sem haver justa causa, se despedir um professor. (PB)
 f. *Não deve-se, sem haver justa causa, despedir um professor. (PB)
 g. Não se deve, sem haver justa causa, despedir um professor. (PB)

(MARTINS; NUNES, 2016, p. 331)

- (26) a. *Pode-se se sentar em qualquer lugar. (PB) (MARTINS; NUNES, 2016, p. 333)
 b. Não se pode se divertir com um barulho desses. (PB)
 c. Não se deve se levantar tarde. (PB)

A restrição da gramaticalidade do *se*-impessoal à próclise em (25), exceto nos casos com partículas atradoras, leva Martins e Nunes (2016) a concluir que o *se*-impessoal no PB possui especificações lexicais que o diferenciam dos outros clíticos da língua, incluindo o clítico reflexivo de 3ª pessoa, que é homófono do *se*-impessoal. Os autores também argumentam que o que resulta a

⁵ No original: “If the hypothesis entertained here is on the right track, one should expect similar patterns of change for other properties arguably associated with the activation of the FP projection”.

agramaticalidade de (26a) é a adjacência. Isto é, as sentenças em (26b-c) são bem formadas porque os pronomes não estão adjacentes um ao outro, enquanto em (26a) o *se*-impessoal e o *se*-reflexivo não são permitidos num mesmo domínio. Assim, os autores presumem que sentenças agramaticais como (26a) se tornam gramaticais se um material parentético romper a adjacência entre os dois clíticos, como em (27):

(27) Pode-se, salvo engano, se sentar em qualquer lugar. (PB)

(MARTINS; NUNES, 2016, p. 333)

Martins e Nunes (2016) argumentam que como o PB tornou-se uma língua de sujeito nulo parcial sua gramática também começou a favorecer sujeitos abertamente expressos em contextos sintáticos onde uma LSN do tipo do PE escolhe um pronominal nulo como a opção não marcada. Uma consequência disso é que o *se*-impessoal passou a ser mais frequentemente expresso em orações com verbos infinitivos no PB quando comparado com o PE. Nessa perspectiva, o uso do *se*-impessoal seria uma estratégia para apoiar uma interpretação genérica para o sujeito de infinitivos. Esse contraste entre PB e PE é especialmente evidente nas orações infinitivas em (28b), nas quais vemos que o PE visivelmente desfavorece ou exclui a presença de *se*-impessoal – da mesma forma que exclui outros sujeitos, como *a gente* (28e), que tem leitura genérica no PB.

(28) a. O mar está perigoso para nadar. (PB^{OK}; PE^{OK}) (MARTINS; NUNES, 2016, p. 330)

b. O mar está perigoso para se nadar. (PB^{OK}; PE^{??})

c. Nova Iorque e Barcelona são cidades fáceis de gostar. (PB^{??}; PE^{OK})

d. Nova Iorque e Barcelona são cidades fáceis de se gostar. (PB^{OK}; PE^{*})

e. Nova Iorque e Barcelona são cidades fáceis da gente gostar. (PB^{OK}; PE^{*})

Cavalcante (2006) já havia mostrado que há uma grande diferença entre a porcentagem de uso do clítico *se* sujeito de infinitivo no PB e no PE. Com dados dos séculos XIX e XX, os resultados da tese da autora trazem: 8% de *se* junto a infinitivos na fala e na escrita no PE; e 20% na fala e 50% na escrita de *se* junto a infinitivo no PB, além de 20% de *a gente* e *você* (ambos genéricos) como sujeitos de infinitivo na fala do PB. A partir desses resultados, a autora admite que o PB é uma língua *pro-drop* parcial e associa essa propriedade ao sistema de Agr.

Martins e Nunes (2016) também debatem o fato de o PB licenciar, conforme observou Galves (1985), construções sem o clítico *se*-impessoal, como em (29).

(29) a. Não usa mais esse estilo de redação. (PB/*PE) (MARTINS; NUNES, 2016, p. 329)

b. Casava muito cedo no século passado. (PB/*PE)

c. No futuro vai descobrir remédio para tudo quanto é doença. (PB/*PE)

Os autores afirmam que alguém pode se perguntar se as sentenças em (29) resultariam da queda no uso do *se*-impessoal ou do enfraquecimento da morfologia de 3ª pessoa do plural, que também licencia sujeitos nulos de leitura arbitrária, como em (30).

- (30) a. Telefonaram para você. (PB/PE) (MARTINS; NUNES, 2016, p. 329)
 b. Estão batendo na porta. (PB/PE)

Martins e Nunes (2016) ressaltam que uma diferença importante entre as sentenças em (29) e (30) é a de que o sujeito indefinido nulo necessariamente exclui o falante em (30), mas não o faz em (29). Para reforçar o argumento de que os sujeitos nulos indefinidos de 3ª pessoa singular no PB incluem o falante (ao passo que os indefinidos de 3ª pessoa plural o excluem), Martins e Nunes (2016) mostram que, em determinados dialetos do PE, construções com *se*-impessoal podem até ser desambiguizado via duplicação/redobro do clítico – conforme ilustrado em (31), que, segundo os autores, também é uma evidência para propor que o *se*-impessoal recebe Caso nominativo e ocupa a posição de sujeito da oração.

- (31) a. Chama-se-lhe a gente espigas. (%PE) (MARTINS; NUNES, 2016, p. 327)
 b. Há várias qualidades que até ainda nos não se conhecemos. (%PE)
 c. Sei é de real certeza que isto era com o que se eles batiam o centeio. (%PE)

Não concordamos que os dados de duplicação/redobro do clítico do PE atestam e explicam a leitura das construções com e sem *se* para o PB. Primeiramente porque, conforme descreveram Lunguinho e Medeiros Júnior (2009), as construções com sujeitos nulos indefinidos de 3ª pessoa singular no PB também possuem leitura arbitrária (ou seja, excluem o falante), como visto em (32).

- (32) a. Matou um rapaz no show do Zezé di Camargo e Luciano ontem.
 b. Montou o armário lá em casa semana passada.
 c. Telefonou aí da CEB para você.
 d. Lava sofá. (LUNGUINHO; MEDEIROS JÚNIOR, 2009, p. 10)

Assim, é necessário admitir que sujeitos nulos indefinidos de 3ª pessoa singular no PB denotam pessoas em geral e podem incluir o falante e o ouvinte (leitura inclusiva), bem como denotam pessoas em geral, mas excluem o falante e o ouvinte (leitura exclusiva, arbitrária).

Consoante Borges (2019), atribuímos a posição enclítica do *se*-impessoal no PB ao fato de ele – assim como o sujeito nulo indefinido de 3ª pessoa singular – ocupar a posição de Spec-*v*P. Assumimos, portanto, que o pronome indefinido de 3ª pessoa singular e o *se*-impessoal no PB ocupam Spec-*v*P e são sujeitos pós-verbais. A seguir, apresentamos os dados de Borges (2019) e propomos, a partir deles, uma derivação para o sujeito nulo indefinido de 3ª pessoa singular e outra para o *se*-impessoal no PB.

2. *Se-indefinido (passivo), se-impessoal e sujeitos nulos indefinidos no português goiano*

Neste trabalho, adotamos a explicação de Holmberg (2010) para as LSN parciais. O autor defende que as LSN parciais não possuem um traço-D não interpretável (rotulado [μ D]) como parte dos traços formais de T e, por isso, essas línguas apresentam certas restrições no licenciamento de sujeitos nulos definidos em contraste com as LSN consistentes. Para o autor, quando um sujeito nulo de 3ª pessoa singular entra em uma relação de *Agree* com T em uma LSN sem um traço-D em T, esse sujeito deve ser interpretado indefinidamente (33b) – diferentemente do que ocorre nas LSN consistentes, que permitem apenas o licenciamento de sujeitos indefinidos marcados abertamente (33a). Assim, a relação sonda-alvo entre T e um ϕ P de 3ª pessoa singular não supre o valor de definitude nas LSN parciais, e o resultado dessa operação é um pronome nulo indefinido de 3ª pessoa singular.⁶

- (33) a. É assim que se faz o doce. (PE/PB) (HOLMBERG, 2010, p. 92)
 b. É assim que __ faz o doce. (*PE/PB)

Fornecemos, a seguir, dados da pesquisa de Borges (2019), os quais mostram resultados quantitativos da distribuição de propriedades gramaticais relacionadas ao parâmetro do sujeito nulo codificadas nas amostras do *corpus*. O *corpus* é constituído de dois documentos escritos em Goiás – um do século XVIII (quando Goiás foi inserido no processo colonial português na América e se tornou capitania), outro do século XIX. 1250 orações de cada um dos manuscritos no formato diário/memórias foram codificadas – o *corpus* é, portanto, composto por 2500 sentenças no total. Ressalta-se que o primeiro documento – *Coleção de notícias da Capitania de Goyaz: diários de viagem do Barão de Mossâmedes* (manuscrito 1596, Biblioteca Geral da Universidade de Coimbra, Portugal) – foi escrito majoritariamente pelo escriba português Tomás de Souza Vila Real e por outro escriba de origem não identificada durante viagem a Goiás nos anos de 1772 e 1773. O segundo manuscrito – *Memorial de lembrança de Anna Joaquina da Silva Marques* (arquivado no Instituto de Pesquisas e Estudos Históricos do Brasil Central (IPEHBC), da Pontifícia Universidade Católica de Goiás, em Goiânia) – é de autoria de Anna Joaquina da Silva Marques, nascida em 1885 na cidade de Goiás. Seus diários começaram a ser escritos em 1881 e foram finalizados em 1930. Foram codificados dados referentes aos anos de 1882-1884.⁷ Na tabela 2, fornecemos os resultados das amostras codificadas dos documentos setecentista e oitocentista. Ressaltamos que os dados – orações finitas e não finitas – se referem à 1ª ou à 3ª pessoa, tendo em vista a ausência de dados de 2ª pessoa nos documentos.

⁶ De acordo com Holmberg (2010), pronomes são DPs, com a estrutura [_{DP} D [_{ϕ P} ϕ [_{NP} N]]], ou ϕ Ps, e pronomes nulos são ϕ Ps.

⁷ Consideramos aqui a hipótese de Borges (2014, 2019) de que as memórias de Anna Joaquina da Silva Marques registram aspectos gramaticais de uma variedade do PB que emergiu em Goiás no século XIX. O argumento para essa hipótese é o de que esses diários possuem propriedades gramaticais *sui generis* e constituem a evidência empírica mais antiga encontrada até então com propriedades similares à sintaxe do sujeito do PB atual, o que permite inferir que eles refletem, de modo minimamente consistente, uma gramática particular que emergiu naquela comunidade linguística. Foram encontrados no documento dados cruciais para a hipótese em tela, como, por exemplo: o aumento de sujeitos manifestos definidos em detrimento de sujeitos nulos definidos e o licenciamento sujeitos nulos indefinidos de 3ª pessoa singular no pretérito perfeito, que só recentemente foram identificados como gramaticais no PB (LUNGUINHO; MEDEIROS JR., 2009).

Tabela 2: Distribuição geral dos dados.

	Construções com sujeitos definidos		Construções com <i>se</i> (passivo e/ou impessoal)	Construções impessoais
	Sujeitos nulos	Sujeitos manifestos		
Século XVIII:	59,36% 742(1250)	20,88% 261(1250)	16,88% 211(1250)	2,88% 36(1250)
Século XIX:	22,24% 278(1250)	64,32% 804(1250)	0,4% 5(1250)	13,04% 163(1250)

Fonte: o autor.

A tabela 2 ilustra dados significativos em relação aos sujeitos definidos (nulos ou manifestos). Entretanto, apesar de esse fenômeno ser revelador de mudança na sintaxe do sujeito no PB em Goiás, nossa análise aqui não recairá sobre a sintaxe dos sujeitos definidos. Analisaremos o conjunto de construções com *se*-passivo e *se*-impessoal.⁸ Nosso objetivo é mostrar que essas construções podem, respectivamente, ser explicadas pela perda da projeção funcional FP (presente na gramática do PE, mas não na do PB) e pela perda do traço-D de T na gramática do PB. Na tabela 2, vemos que a amostra do manuscrito setecentista apresenta um uso produtivo de construções com *se*-passivo ou *se*-impessoal (16,88% do total da amostra), principalmente com *se*-impessoal, enquanto o manuscrito oitocentista traz pouquíssimos usos de construções com *se*-impessoal (apenas 5 dos 1250 dados) e carece de construções com *se*-passivo.⁹

Nossos dados evidenciam que a perda das construções com *se*-passivo no PB não está ligada a um empobrecimento da morfologia de 3ª pessoa plural dos verbos transitivos, pois não há ausência significativa de concordância de flexão entre verbo transitivo e sujeito nos dados. No manuscrito setecentista, foram registrados somente quatro casos de ausência de concordância entre um sujeito de 3ª pessoa e o verbo. Três desses casos envolvem DPs coordenados em estrutura VS, tendo ocorrido apenas um caso de DP coordenado em estrutura SV. Assim, nas estruturas VS a relação de concordância é parcial, ou seja, ocorre apenas entre o verbo e o primeiro elemento da coordenação (34a) – o que, nesse caso, está de acordo com o sistema gramatical do PE, no qual a concordância parcial é licenciada em estruturas com sujeitos compostos pós-verbais (cf. RAPOSO, 2020).

- (34) a. Sua Excelência intentou ir ao regimento do Duro; porém, como **veio** ao Arraial de Natividade o vigário e o capitão Roque, regente dos índios, que ali ainda se conservam aldeados, desistiu desta ideia.
- b. O seu esmoler, ânimo, e a sua liberalidade, [o] **fez** despender nesta digressão mais de 4 mil cruzados.

⁸ Mais adiante, analisaremos também os resultados da tabela 2 referentes às construções impessoais, isto é, com sujeitos nulos não argumentais registradas nas amostras do *corpus*, a fim de corroborar a proposta de que a inversão locativa é um fenômeno emergente na gramática oitocentista do PB em Goiás.

⁹ Embora tenhamos utilizado uma amostra de 1.250 sentenças do manuscrito oitocentista, destacamos que não encontramos registro de construções com *se*-indefinido (passivo) no documento

Raposo (2020) denomina o fenômeno em (34a) de concordância com o DP mais próximo, pois, nesses casos, obrigatoriamente, o primeiro DP da estrutura coordenada, com o qual T estabelece a concordância, deve estar no singular (e, evidentemente, numa posição abaixo de T), ao passo que o segundo DP pode ser tanto singular quanto plural.¹⁰

O único caso de ausência de concordância em estrutura SV encontrado na amostra setecentista (34b), por sua vez, pode ser explicado pela proposta de Raposo (2020) de que, no português, a aparente ausência de concordância entre sujeitos pré-verbais compostos e o verbo derivam, de modo excepcional, “de uma coordenação de orações, aplicando-se na primeira oração elipse regressiva do material repetido, o que explicaria, assim, a concordância no singular” (p. 2438). O autor fornece os seguintes exemplos, em que o tachado corresponde ao material elidido na componente fonológica.

- (35) a. [A grandeza e a significação das coisas] resulta do grau de transcendência que encerram.
 b. A grandeza ~~das coisas resulta do grau de transcendência que encerram~~ e a significação das coisas resulta do grau de transcendência que encerram.
 c. Morro, se [a graça e a misericórdia de Deus] não me acode.
 d. Morro, se a graça ~~de Deus me não acode~~ e a misericórdia de Deus me não acode.

Seguindo essa proposta, ilustramos a seguir a elipse regressiva do dado inicialmente apresentado em (34b).

- (36) O seu esmoler ~~fez despende~~ nesta digressão mais de 4 mil cruzados, [o seu] ânimo **fez despende** nesta digressão mais de 4 mil cruzados, e a sua liberalidade [o] **fez despende** nesta digressão mais de 4 mil cruzados.

Dessa forma, a partir de (36), verifica-se que o dado singular em (34b) é, na verdade, uma estrutura oracional com apagamento na componente fonológica, na qual apenas (os núcleos d)os sujeitos das orações elididas são realizados e ligados à última oração por meio de uma conjunção aditiva.

Em relação à amostra oitocentista, menos de 12% das orações com sujeito pluralizado (manifesto ou nulo), ou seja, 54 dados de um total de 464, não apresentaram concordância explícita entre o verbo e o sujeito, conforme ilustrado na tabela 3.

¹⁰ Uma evidência dada por Raposo (2020) para a hipótese de que há concordância parcial nos casos com sujeitos compostos pré-verbais, isto é, concordância com o DP mais próximo de uma estrutura coordenada aditiva, é o fato de que, no PE, a concordância no plural é obrigatória se o primeiro DP coordenado estiver no plural:

- (i) a. **Apareceram** no jardim [*um gato* e três cães].
 b. **Apareceu** no jardim [*um gato* e três cães].
 c. **Apareceram** no jardim [*três cães* e um gato].
 d. ***Apareceu** no jardim [*três cães* e um gato].

Tabela 3: Concordância com verbo no plural no século XIX.

	Concordância	Ausência de concordância
<i>Século XIX</i>	88.36% 410/464	11.64% 54/464

Fonte: o autor.

Na codificação das 464 sentenças da tabela 3 foram considerados todos os tipos de verbos não impessoais (transitivos, intransitivos, inergativos etc.), com sujeito manifesto ou nulo. Todos os casos de sujeitos referenciais nulos no plural apresentaram concordância manifesta com o verbo, independentemente da natureza semântica do verbo. Os 11.64% de casos de incompatibilidade de concordância na amostra oitocentista envolvem, crucialmente, a 3ª pessoa do plural de verbos de natureza existencial (37a) ou inacusativa (37b-d).

(37) Ausência de concordância com verbo existencial e inacusativo no PB oitocentista

- a. Dia 14 **esteve** aqui o Luiz de Castro e Lulu Pitaluga de noite.
- b. Dia 22 **chegou** aqui o Ilustríssimo Excelentíssimo Senhor Dr. Antônio Gomes Pereira Júnior, Presidente para Goyaz com sua digna consorte; e mais alguns seus colegas chefe de polícia.
- c. De noite **veio** aqui as filhas do Carneiro.
- d. Dia 17 Lulu Pitaluga e o Cavalcante **veio** aqui.

Em relação à ausência de concordância verbal em dados como (37a) e (37b), pode-se, inicialmente, pensar que a concordância é parcial, ou seja, é estabelecida entre o verbo (existencial ou inacusativo) e o elemento nominal singular mais próximo dentro de um constituinte com dois nominais coordenados por uma conjunção aditiva, como ocorre no PE. Haveria, assim, nos dados da amostra oitocentista, a coexistência da concordância parcial e da concordância plena em estruturas com verbos existenciais e inacusativos e sujeitos compostos pós-verbais constituídos de dois DPs coordenados. Considerando, no entanto, aspectos recentes da gramática do PB, em que, por hipótese, fatores discursivos estão presentes nas estruturas VS da língua devido à observação de que um elemento dêitico ou um dêitico nulo recuperado anaforicamente é exigido nessas estruturas (cf. PILATI, 2006; TEIXEIRA; SALLES, 2016; TEIXEIRA; BORGES, 2022), vislumbramos outra possibilidade de análise: as estruturas em (37a-b) são casos de inversão locativa. Avaliamos essa proposta como mais adequada porque ela é capaz de dar conta tanto dos dados em que um DP (nu) de natureza temporal (ou locativa) ocupa a posição pré-verbal, como em (37a-b), quanto dos dados em que um advérbio ou um PP adverbial ocupam a posição pré-verbal, como exemplificado, respectivamente, em (38a) e (38b).

- (38) a. *Lá estava* o Dr. Pitaluga T. Cav. e Jm Mel. Correa.
- b. *Nesse dia faleceu* a viúva do Te. Col. Custodio.

Remetemos o leitor à análise de Teixeira e Borges (2022) sobre as especificidades de (i) DPs, (ii) advérbios e (iii) PPs – todos de natureza locativa ou temporal – em Spec-TP em estruturas de inversão locativa no PB. Os autores propõem que DPs locativos/temporais têm traços de pessoa e número, valorando plenamente os traços- ϕ e assinalando Caso nominativo em Spec-TP, enquanto advérbios e PPs adverbiais locativos ou temporais manifestam o Caso oblíquo (inerente), o que sugere uma situação de *Quirky Case*. Assim, o Caso oblíquo é determinado no nível lexical no caso do advérbio e no nível sintático na presença de preposição. O Caso oblíquo não se choca com a operação *Agree* em TP, pois, no caso do advérbio, é verificado apenas o traço de pessoa em T, valorando o traço de número como singular, a opção *default*; na presença de PPs, há apenas o traço-EPP licenciado, uma vez que os traços de número e pessoa não são acessíveis, sendo, portanto, valorados na opção *default*.

Outra diferença crucial na ausência de concordância nos dados oitocentistas com verbos existenciais e inacusativos é verificada em ocorrências como em (37c), na qual há a ausência de concordância entre o DP plural (não coordenado) e o verbo em uma oração com estrutura VS. Ao todo, há cinco construções desse tipo na amostra do manuscrito oitocentista. Vejamos mais duas delas.

- (39) a. De tarde **esteve** aqui os meninos do Luiz Jardim.
 b. Dia 20 **faleceu** duas crianças com coqueluche; a fa. do João Athanazio e o fo. Do Te. Macedo.

Para esses dados, também seguimos a proposta de que são casos de inversão locativa, de modo que os adjuntos adverbiais em posição pré-verbal nessas estruturas – um PP e um DP nu nos exemplos (39a) e (39b) – ocupam Spec-TP. Nessa perspectiva, há duas opções para o estabelecimento da concordância com verbos existenciais e inacusativos na gramática oitocentista do PB em Goiás: com o DP argumento interno pluralizado, quer seja em posição pré-verbal ou pós-verbal, tal como já ocorria na gramática setecentista; ou com um elemento de natureza locativa ou temporal à esquerda do verbo, quando se considera a hipótese do surgimento da inversão locativa na gramática oitocentista. Em resumo, isso direciona nossa explicação para a ausência de concordância nas estruturas VS com DPs plurais com a estrutura DET + SN ou com DPs coordenados à emergência de estruturas com inversão locativa na gramática oitocentista – enquanto ainda é substancialmente mais produtiva a concordância entre o verbo existencial ou inacusativo e o DP.¹¹

Para incrementar o argumento de que a inversão locativa havia emergido na gramática oitocentista do PB em Goiás, destacamos que as construções impessoais do *corpus*, representando 13,04% dos dados oitocentistas contra apenas 2,88% dos dados setecentistas (cf. tabela 2), também podem ser analisadas nessa perspectiva. As construções impessoais da amostra oitocentista, além de substancialmente mais recorrentes, estão sempre acompanhadas por um elemento dêitico em posição

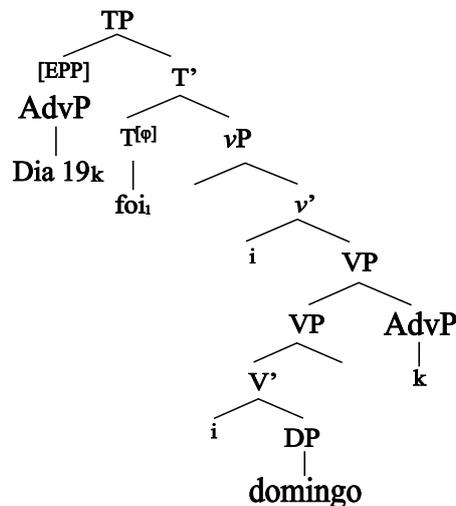
¹¹ Borges (2019, 2022) defende que falantes de línguas nigero-congolesas foram agentes da mudança e, conseqüentemente, da emergência da gramática do PB em Goiás e que essa influência consistiu na massiva ampliação da ordem SVO na gramática oitocentista do PB goiano, o que ocasionou a perda da inversão livre nessa gramática e, curiosamente, o surgimento da inversão locativa, presente nas línguas nigero-congolesas.

pré-verbal ou facilmente recuperado anaforicamente, como exemplificado em (40).¹²

- (40) a. Dia 2 de tarde choveu a 1ª chuva e muito grande.
 b. Nessa noite choveu muito.
 c. Dia 19 foi domingo.
 d. Dia 15 houve espetáculo no Teatro: A Sogra e o entremez Lucrecia Borges.

Nesse sentido, a derivação de construções impessoais na gramática oitocentista pode ser representada como em (41) – ressalta-se que constituintes não argumentais movidos para Spec-TP para satisfazer o traço-EPP em construções impessoais não se restringem à variedade goiana do PB.

- (41) Derivação de uma construção impessoal no século XIX (cf. (40c))



Curiosamente, dados com ausência de concordância em estruturas SV, como em (37d), que representa estatisticamente a maioria dos casos de ausência de concordância no manuscrito oitocentista (57,4%), não fazem parte, de modo sistemático, da gramática do português (cf. COLAÇO, 2016; RAPOSO, 2020), evidenciando outro comportamento inovador na sintaxe do sujeito na gramática oitocentista em estruturas com verbos existenciais e inacusativos. Contudo, é razoável supor que dados como em (37d) são casos de apagamento de oração na componente fonológica, tal como propusemos para um caso específico registrado na amostra setecentista. Vejamos.

- (42) a. Dia 17 Lulu Pitaluga e o Cavalcante **veio** aqui.
 b. Dia 17 Lulu Pitaluga **veio** aqui e o Cavalcante **veio** aqui.

Por hipótese, a análise em (42) infere a inexistência de pluralidade semântica nos DPs coordenados por conjunção aditiva, pois rejeita a hipótese da concordância parcial e argumenta em favor da hipótese de que, nesses casos em específico, os constituintes coordenados pela conjunção aditiva são de natureza

¹² As construções impessoais nas amostras do nosso *corpus* (cf. tabela 2) incluem construções com verbos impessoais (*haver/ter* etc.) e com verbos meteorológicos (*chover/amanhecer* etc.), que no português possuem um sujeito expletivo nulo e não argumental.

oracional. Uma evidência a favor da argumentação de que há elipse nessas estruturas coordenadas da gramática oitocentista está em dados como em (43a), em que todo o material da primeira oração é pronunciado, mas o material da coordenada é fragmentado, como ilustrado em (43b).

- (43) a. Dia 7 detarde Nicota **veio** aqui e denoite Lulu Pitaluga.
 b. Dia 7 detarde Nicota **veio** aqui e denoite Lulu Pitaluga ~~**veio**~~ aqui.

De certo modo, propomos que as inovações na sintaxe de concordância entre sujeito e verbos existenciais e inacusativos na amostra oitocentista estão ligadas mais ao surgimento da inversão locativa na sintaxe do PB do que a algum fenômeno mais específico de concordância, pois: (i) a concordância parcial com sujeitos pré-verbais é fenômeno comum ao português; (ii) e, apesar de não ser atestado sistematicamente no português setecentista em Goiás como o foi na amostra do português goiano oitocentista, o apagamento na componente fonológica de material de estruturas oracionais nos casos de ausência de concordância com sujeitos compostos pré-verbais pode ser explicado da mesma forma em ambas as variedades. Diante disso, nossos dados sugerem que a morfologia verbal não teria tido um papel significativo na mudança na gramática do PB no que concerne à sintaxe do sujeito. Assim, considerando especialmente estruturas transitivas, que não tiveram qualquer alteração significativa na sua morfologia de concordância na amostra oitocentista, propomos que a perda da projeção funcional FP e de um traço categorial de definitude [um traço-D] em T teria tido um papel mais determinante no desenvolvimento da sintaxe do sujeito do PB em Goiás. Baseando-nos em Holmberg (2010), propomos que a perda do traço-D em T na gramática do PB oitocentista em Goiás está envolvida na diminuição de construções com *se*-impessoal e no surgimento de um tipo inovador de sujeito nulo em comparação com o PE, qual seja, o sujeito nulo indefinido de 3ª pessoa singular, como em (44).¹³

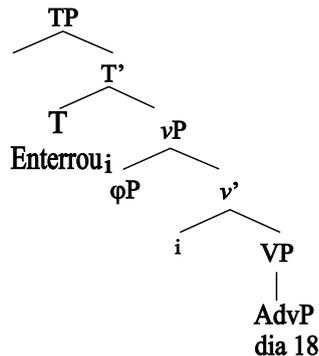
- (44) Pronome nulo indefinido de 3ª pessoa singular no século XIX em Goiás
- a. Dia 18 o senhor Joaquim Fernandez esteve aqui de noite, quando ele saiu entrou a Anna Joaquina daqui de frente que veio despedir para ir para Barra. Dia 17 faleceu o cadete Cândido Gonzaga, e __ **enterrou** dia 18.
- b. Dia 14 de novembro de 1883 3ª-feira às 5 horas da manhã faleceu o cônego Pio à tarde foi sepultado nessa mesma tarde faleceu Pedro cunhado do Antônio Manoel __ **sepultou** no dia seguinte.

Dados como em (44) constituem, portanto, evidência empírica da perda do traço-D em T na gramática do PB oitocentista em Goiás, a qual consiste na possibilidade de licenciar sujeitos

¹³ Ressaltamos que, na amostra de Borges (2019) do manuscrito oitocentista, são codificados e analisados 5 sujeitos nulos indefinidos de 3ª pessoa singular. Borges (2014), por sua vez, identificou e analisou 38 sentenças dessa natureza no mesmo manuscrito, considerando o período de 1881 a 1899. Ademais, a proposta de Borges (2019) a respeito da perda do traço-D de T também é o alicerce para explicar a perda de sujeitos nulos definidos na amostra oitocentista em comparação com a amostra setecentista do *corpus* da pesquisa (cf. tabela 2).

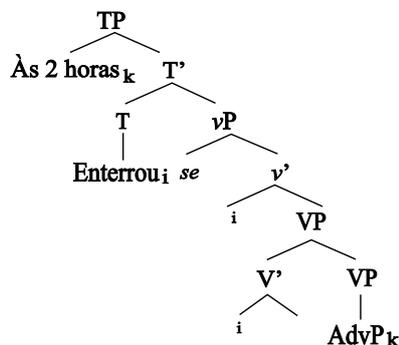
indefinidos de 3ª pessoa singular sem um sujeito manifesto – quer dizer, sem o clítico *se*-impessoal. Assumimos, então, consoante Holmberg (2010), que o sujeito indefinido nulo de 3ª pessoa singular no PB, como o gerado em (44), é um ϕ P. Também assumimos que o pronome nulo indefinido no PB ocupa uma posição pós-verbal: Spec- v P, posição na qual o sujeito é gerado, seguindo Chomsky (1995 e trabalhos posteriores). Igualmente, sua contraparte manifesta também ocupa Spec- v P e ali permanece, restringindo-o à ênclise. Em (45) e (46), propomos, respectivamente, uma derivação para esses sujeitos.

(45) Pronome nulo indefinido de 3ª p. singular na gramática oitocentista (cf. (44a))



(46) *Se*-impessoal na gramática oitocentista

- a. Dia 23 às 3 horas da manhã Manoel daqui morreu, às 2 horas enterrou-se.
b.



É importante destacar que não postulamos ser obrigatório um elemento ocupar Spec-TP nas sentenças com verbos transitivos e sujeitos nulos indefinidos na gramática oitocentista do PB em Goiás. Similarmente ao que argumentam Pilati, Naves e Salles (2017), advogamos que, obrigatoriamente, deve haver nesse tipo de sentença uma âncora adverbial discursiva para que elas sejam licenciadas: assim, em (45), essa âncora está manifesta em “dia 18”; em (46), ela está manifesta em “às 2 horas”, que ocupa Spec-TP.

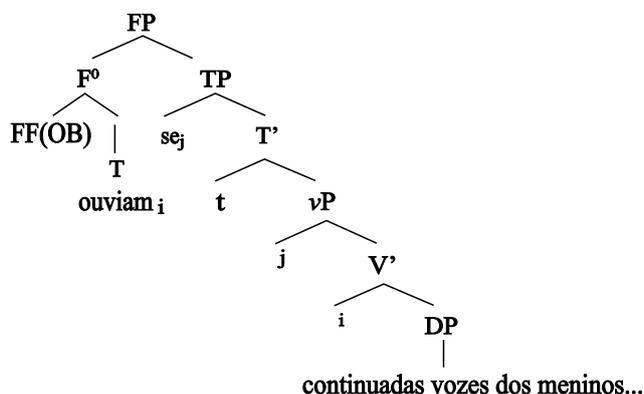
No caso das construções com *se*-passivo/indefinido registradas somente na amostra setecentista, seguimos Raposo e Uriagereka (1996) e propomos que a concordância é estabelecida por meio da atração dos traços formais do DP argumento interno por FP – que, como resultado, permite que, em alguns casos, o DP argumento interno ocupe uma posição à esquerda do verbo. Essa concordância,

como anteriormente visto, se estabelece por conta de FP ser um núcleo funcional periférico especificado para traços- ϕ e Caso nominativo.

(47) Construção com *se*-passivo na amostra setecentista

a. Pelas árvores da vizinhança do povo, se ouviam continuadas vozes dos meninos: “*Viva sua Excelência*”.¹⁴

b.



Seguindo a proposta de Raposo e Uriagereka (1996), temos que o *se*-passivo/indefinido checa o traço-D de T na gramática setecentista, como ilustrado em (47b). Como mencionamos, atribuir a perda das construções com *se*-passivo/indefinido na gramática oitocentista à impossibilidade de o DP argumento interno estabelecer concordância com o verbo por meio da projeção FP vincula-se à predição de Pires (2006) a respeito de a perda do movimento do verbo de T para F no PB ter desencadeado uma série de alterações na gramática da língua. Assim, a perda das construções com *se*-passivo/indefinido na gramática oitocentista está vinculada ao fato de o núcleo FP não estar mais ativo, impossibilitando, portanto, que essa gramática gere construções com *se* em que o DP argumento interno seja atraído pelos traços- ϕ e o Caso nominativo de F.¹⁵

Observa-se que, em (47b), o *se*-passivo/indefinido satisfaz o traço-D de T na gramática setecentista por meio do seu movimento para Spec-TP. No entanto, na gramática oitocentista, em que há a perda do traço-D de T, o clítico *se* passa a atuar apenas com interpretação impessoal e na posição em que é gerado (Spec-vP). Temos, então, que a gramática oitocentista do PB em Goiás registra a perda do núcleo funcional FP e do traço-D não interpretável de T de forma concomitante. Isso explica a perda do *se*-passivo/indefinido, a diminuição radical do uso do *se*-impessoal/genérico

¹⁴ Em (47), associamos o *se* proclítico em vez de enclítico à presença de um sintagma frontado proclisador, o que determina a próclise no português em orações principais. Não é nosso objetivo representar esse fenômeno na derivação. Para um estudo específico sobre a sintaxe sincrônica e diacrônica dos clíticos no português, sugerimos, inicialmente, a leitura de Martins (2016).

¹⁵ Cavalcante e Marcotulio (2012) chegam à conclusão semelhante a respeito da perda das construções com *se*-passivo/indefinido e da projeção funcional FP no PB. Não mencionam, contudo, que Pires (2006) já havia descrito a perda dessa projeção no PB e predito que outras mudanças, além das que ele analisa, poderiam ter sido causadas por ela.

e a emergência de sujeitos nulos indefinidos de 3ª pessoa singular ocorrerem simultaneamente no manuscrito oitocentista.¹⁶

Em (48), exemplificamos construções – registradas no manuscrito oitocentista, mas não pertencentes à amostra da tabela 2 – que poderiam ser parafraseadas com *se*-passivo/indefinido, mas são, respectivamente, construção com *se*-impessoal/genérico (48a) e com sujeito nulo indefinido de 3ª pessoa singular (48b-c).

(48) Sujeitos nulos indefinidos de leitura arbitrária no português oitocentista em Goiás

- a. Dia 10 depois do cazam.to **dansou-se** 2 quadrilha. (BORGES, 2014, p. 87-91)
- b. Dia 4 Faleceu a Dona Lin[...], mulher do capitão João Berquó para amanhecer. Faleceu a senhora Ninica de repente, nesse dia **enterrou** ambas 1 de manhã a outra de tarde.
- c. Dia 20 **Arrancou** as partileira da venda, de tarde Silvia veio.

Interessantemente, uma vez que FP não mais é projetado e T não mais carrega um traço-D não interpretável, a única forma de um argumento interno controlar os traços- ϕ de um verbo transitivo no PB contemporâneo passa a ser quando um DP argumento interno se move para Spec-TP – para a posição canônica de sujeito –, valorando os traços- ϕ do verbo e fornecendo à construção um índice de definitude, como em (49), cujas frases podem facilmente ser parafraseadas na escrita formal e monitorada como em (50).

- (49) a. Apartamento vende bem em Goiânia.
- b. A cada minuto quatro coisas vendem.
- (50) a. Apartamento, vende-se bem em Goiânia. (*se*-impessoal/genérico)
- b. Vende-se bem apartamento em Goiânia. (*se*-impessoal/genérico)
- c. A cada minuto quatro coisas vendem-se. (*se*-passivo/indefinido)
- d. A cada minuto vendem-se quatro coisas. (*se*-passivo/indefinido)

O deslocamento do argumento interno para Spec-TP em estruturas como em (49) é uma evidência favorável para nossa análise de que o traço de definitude do sujeito no PB só é valorado pelo movimento do sujeito lexical para essa posição (cf. BORGES, 2019), razão pela qual o sujeito nulo indefinido somente é licenciado na posição em que é gerado – Spec-vP.¹⁷ Para finalizar, apresentamos

¹⁶ Com base em dados empíricos de cartas escritas por brasileiros, Cyrino (2007) argumenta que todo clítico *se* no PB pode ser interpretado como impessoal desde o século XIX. Uma questão curiosa quando postos em contrastes os dados de Cyrino (2007) e os dados deste trabalho é que o clítico no *corpus* de Cyrino (2007) é sempre proclítico, ao passo em que em nosso *corpus* ele é sempre enclítico, tal como no PB contemporâneo.

¹⁷ Barbosa (2019) propõe que sujeitos nulos nas LSN são uma instanciação do núcleo *n* – um núcleo categorizador sem um radical – e argumenta que os sujeitos nulos indefinidos de línguas como o finlandês e o PB não se elevam para fora do núcleo verbal, ao passo que os sujeitos nulos definidos ocupam uma posição externa ao núcleo verbal. Assim, ela propõe duas configurações para o componente semântico: a) quando [_n Ø] permanece no interior do núcleo verbal: interpretação

no quadro 1 as principais propriedades das construções com *se*-passivo/indefinido e *se*-impessoal/genérico no PE e no PB, assim como das construções com sujeitos nulos indefinidos no PB, que discutimos neste trabalho.

Quadro 1: Panorama das construções com *se*-passivo, *se*-impessoal e sujeitos nulos indefinidos no português.

CONSTRUÇÃO	CARACTERÍSTICAS
<p>Construção com <i>se</i>-passivo prototípica (português clássico):</p> <p>- <i>Aqui se escreverão novas histórias, por gentes estrangeiras que virão.</i> (NARO, 1976, p. 781)</p>	<p>Licenciava um sintagma preposicional introduzindo o agente da passiva até o século XVII; nesse período, inicia-se a reanálise da leitura do <i>se</i>-passivo para uma leitura impessoal devido à perda do agente da passiva e da concordância entre o verbo transitivo e o argumento interno. Surgem as construções com <i>se</i>-impessoal e verbos intransitivos.</p>
<p>Construção com <i>se</i>-passivo (ou <i>se</i>-indefinido) no PE:</p> <p>- <i>As obras mais recentes publicaram-se numa editora espanhola para garantir maiores tiragens.</i> (DUARTE, 2013, p. 444).</p> <p>- <i>Os bolos comeram-se ontem.</i> (MARTINS; NUNES, 2016, p. 323).</p>	<p>Para Duarte (2013) e Martins e Nunes (2016), o <i>se</i> atua como um morfema passivo nessa construção, sem realizar alteração na morfologia do verbo, e bloqueia a atribuição de papel temático à posição do argumento externo e de Caso acusativo ao argumento interno. Assim, o <i>se</i>-passivo se insere dentro do escopo dos clíticos quase argumentais e o argumento interno, que controla os traços-ϕ do verbo, está ligado a uma posição à esquerda de T, pois pode ser alçado a uma posição pré-verbal, onde recebe Caso nominativo. Extrai-se a interpretação do argumento interno implícito da construção com <i>se</i>-passivo por meio de testes com advérbio orientado para o agente e com subordinadas finais formadas a partir de um adjunto cujo sujeito só pode ser controlado pelo argumento externo implícito da passiva. Não se admite o redobro do clítico, mas sua extração simultânea. Para Raposo e Uriagereka (1996), o <i>se</i>-passivo é indefinido (está numa construção ativa), e o argumento interno concorda com o verbo e recebe Caso nominativo por meio da projeção FP, ao passo que o <i>se</i>-indefinido recebe Caso nulo em Spec-TP. Assumimos que, no PB, essa construção não é licenciada como parte da gramática nuclear dos falantes nativos devido à perda da projeção FP.</p>
<p>Construções com <i>se</i>-impessoal (PE/PB):</p> <p>- <i>Precisa-se de funcionários.</i></p> <p>- <i>Trabalha-se muito nesta cidade.</i></p> <p>- <i>Chegava-se cedo ao trabalho.</i> (MARTINS; NUNES, 2016, p. 328-329).</p>	<p>Para a maioria dos autores, o <i>se</i>-impessoal tem valor referencial, é o sujeito frásico e recebe Caso nominativo. Permite-se a extração simultânea e, em dialetos do PE, o redobro do clítico. Além dos transitivos e intransitivos, o <i>se</i>-impessoal também é licenciado com verbos inergativos, inacusativos, passivos, de ligação e, no PE, de alçamento. Na ausência de uma partícula atratora de próclise, o <i>se</i>-impessoal é enclítico no PB, ao passo que, inversamente, os demais clíticos no PB são proclíticos. No PE, o <i>se</i>-passivo e o <i>se</i>-impessoal são enclíticos. Na nossa análise, atribuímos a posição enclítica do <i>se</i>-impessoal no PB à posição em que ele é gerado: Spec-vP.</p>

indefinida; b) quando [_n Ø] se eleva para posição pré-verbal: interpretação definida. A pesquisadora pontua, ainda, que, no PB, em construções como “Aqui vende fruta”, o *nP* é pós-verbal e possui leitura existencial, isto é, se restringe pragmaticamente a um traço [+humano].

<p>Construções com sujeitos nulos indefinidos de 3ª pessoa singular (restritas ao PB): - <i>Casava muito cedo no século passado.</i> (MARTINS; NUNES, 2016, p. 329).</p>	<p>Na nossa análise, a partir de dados empíricos de sujeitos nulos indefinidos de 3ª pessoa singular e leitura arbitrária produzidos em Goiás na segunda metade do século XIX, propusemos que eles surgiram como consequência da perda do traço-D não interpretável de T na gramática do PB e ocupam a mesma posição de sua contraparte manifesta (o <i>se</i>-impessoal): Spec-vP.</p>
<p><i>Se</i>-sujeito de infinitivo (mais produtivas no PB): - <i>Nova Iorque e Barcelona são cidades fáceis de se gostar.</i> (MARTINS; NUNES, 2016, p. 330).</p>	<p>Cavalcante (2006) e Martins e Nunes (2016) argumentam que como o PB é uma LSN parcial favorece sujeitos manifestos em contextos em que uma LSN como o PE escolhe um sujeito nulo; assim, o <i>se</i>-impessoal como sujeito de infinitivos pode ser interpretado como uma estratégia para o preenchimento da posição de sujeito e para apoiar uma leitura genérica para o sujeito do infinitivo.</p>

O quadro 1 não só sistematiza as propostas teóricas discutidas neste trabalho, mas busca evidenciar o percurso pelo qual essas construções passaram na história gramatical do PE e do PB, mais especialmente do PB, que não mais apresenta *se*-passivo e *se*-impessoal no processo inicial de aquisição de língua materna (cf. SILVA, 2017) e cuja configuração paramétrica para o licenciamento de sujeitos nulos e manifestos se difere crucialmente do PE.

Considerações finais

Neste trabalho, trouxemos dados originais de manuscritos goianos escritos no formato diário/memórias para propor uma explicação para a perda de construções com *se*-passivo e *se*-impessoal no PB e o surgimento de construções com sujeitos nulos indefinidos no PB, sem associar esses fenômenos ao empobrecimento da morfologia verbal. Nessa direção, argumentamos que a gramática oitocentista do PB em Goiás evidencia: (i) a perda da projeção FP, responsável pelo estabelecimento da concordância entre verbo transitivo e argumento interno nas chamadas passivas com *se* – as quais estavam presentes no manuscrito setecentista cujo autor principal era de origem portuguesa, mas não foram registradas no manuscrito oitocentista; e (ii) a perda do traço-D não interpretável em T, passando a licenciar, quando um pronome nulo de 3º pessoa singular em Spec-vP entra numa relação de *Agree* com T, uma construção com sujeito nulo indefinido.

Referências

BARBOSA, P. *pro* as a Minimal nP: towards a unified approach to *pro-drop*. *Linguistic Inquiry*, v. 50, n. 3, pp. 487-526, 2019. https://doi.org/10.1162/ling_a_00312.

BORGES, H. *A sintaxe do sujeito na história do português em Goiás: evidências oitocentistas de uma língua de sujeito nulo parcial*. 2019. 273 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade de Brasília, Brasília, 2019.

BORGES, H. Mudança sintática induzida por contato linguístico no português brasileiro: um estudo diacrônico e dialetal baseado em corpus. In: FIGUEIREDO, C.; HAGEMEIJER, T.; OLIVEIRA, M. de. *Novas dinâmicas do português: a África atlântica e o Brasil*. Lisboa: Lisbon Press, 2022, no prelo.

BORGES, H. *Sujeitos nulos indeterminados no português brasileiro: uma investigação diacrônica em Goiás*. 2014. 99 f. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade de Brasília, Brasília, 2014.

BORGES, H. PIRES, A. The emergence of Brazilian Portuguese: earlier evidence for the development of a partial null subject Grammar. *Revista Linguística*, v. 13, pp. 174-90, 2017. <http://dx.doi.org/10.31513/linguistica.2017.v13n2a14033>.

BRITO, A. M.; DUARTE, I.; MATOS, G. Tipologia e distribuição das expressões nominais. In: MIRA MATEUS, M. H. *et al.* (org.). *Gramática da Língua Portuguesa*. Lisboa: Caminho, 2003. pp. 826-48.

CAVALCANTE, S. R. O. Revisitando as construções com *se* na história do português brasileiro. In: CYRINO, S.; MORAIS, M. A. T. de. (org.). *Mudança sintática do português brasileiro: perspectiva gerativista*. São Paulo: Contexto, 2018. pp. 382-419.

CAVALCANTE, S. R. O. 2006. 227f. *O uso de se com infinitivo na história do português: do português clássico ao português europeu e brasileiro moderno*. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2006.

CAVALCANTE, S. R.; MARCOTULIO, L. L. Novo olhar sobre as construções com *se*: para além da questão da concordância. In: Duarte, M. E. L. (org.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola, 2012. pp. 143-60.

CHOMSKY, N. *The Minimalist Program*. Cambridge, Mass.: MIT Press, 1995.

COLAÇO, M. Especificidades das estruturas de coordenação: padrões de concordância. In: MARTINS, A. M.; CARRILHO, E. *Manual de linguística portuguesa*. Berlin, Boston: De Gruyter, 2016, pp. 502-22. <https://doi.org/10.1515/9783110368840-021>.

CYRINO, S. Construções com SE e promoção de argumento no português brasileiro: uma investigação diacrônica. *Revista da Abralin*, v. 6, n. 2, pp. 85-116, 2007.

CYRINO, S. Observações sobre a mudança diacrônica no português brasileiro. In: ROBERTS, I.; KATO, M. (orgs.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993. pp. 163-84.

DUARTE, I. Construções ativas, passivas, incoativas e médias. In: RAPOSO, E. B. P. *et al.* (org.). *Gramática do Português*, v. 1. Lisboa: Fundação Calouste Gulbenkian, 2013. pp. 429-58.

GALVES, C. A sintaxe do português brasileiro. *Cadernos de Linguística e Teoria da Literatura*, n. 13, pp. 33-52, 1985. <http://dx.doi.org/10.17851/0101-3548.7.13.33-52>.

HOLMBERG, A. Null Subject Parameters. In: BIBERAUER, T. *et al.* (org.). *Parametric Variation: null subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 2010. pp. 88-124. <https://doi.org/10.1017/CBO9780511770784.003>.

LUNGUINHO, M. V. S.; MEDEIROS JÚNIOR, P. Inventou um novo sujeito: características sintáticas e semânticas de uma estratégia de indeterminação do sujeito no português brasileiro. *A Interdisciplinar*, v. 9, ano IV, pp. 7-21, 2009.

- MARTINS, A. M. A colocação dos pronomes clíticos em sincronia e diacronia. In: MARTINS, A. M.; CARRILHO, E. *Manual de linguística portuguesa*. Berlin, Boston: De Gruyter, 2016. pp. 401-30. <https://doi.org/10.1515/9783110368840-017>.
- MARTINS, A. M.; NUNES, J. Passives and *se* constructions. In: WETZELS, L.; MENUZZI, S.; COSTA, J. (org.). *The Handbook of Portuguese Linguistics*. Hoboken: Wiley-Blackwell, 2016. pp. 318-37. <https://doi.org/10.1002/9781118791844.ch17>.
- NARO, A. The genesis of the reflexive impersonal in Portuguese: a study in syntactic change as a surface phenomenon. *Language*, v. 52, n. 4, pp. 779-810, 1976. <https://doi.org/10.2307/413295>.
- NAVES, R. R.; BORGES, H. A emergência da gramática do português brasileiro em Goiás. *Confluência*, n. 47, v. 2, pp. 68-102, 2015. <https://doi.org/10.18364/rc.v1i47.62>.
- NUNES, J. *O famigerado se: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. 1990. 172 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1990.
- PILATI, E. *Aspectos sintáticos e semânticos de orações com ordem verbo-sujeito no português do Brasil*. 2006. 242f. Tese (Doutorado em Linguística) – Instituto de Letras, Universidade de Brasília, 2006.
- PILATI, E.; NAVES, R.; SALLES, H. Locative DPs and deictic adverbs/pronouns in subject position in Brazilian Portuguese. In: FERNÁNDEZ-SORIANO, O.; et al. (org.). *Boundaries, phases and interfaces: Case studies in honor of Violeta Demonte*. Amsterdam: John Benjamins, 2017, pp. 63-83. <https://doi.org/10.1075/la.239.04pil>.
- PIRES, A. *The minimalist syntax of defective domains: gerunds and infinitives*. Amsterdam: John Benjamins, 2006. <https://doi.org/10.1075/la.98>.
- RAPOSO, E. Concordância verbal. In: RAPOSO, E. B. P. et al. (org.). *Gramática do Português*. Vol. III. Lisboa: Fundação Calouste Gulbenkian, 2020. pp. 2436-9.
- RAPOSO, E.; URIAGEREKA, J. Indefinite SE. *Natural language and linguistic theory*, v. 14, n. 2, pp. 749-810, 1996. <https://doi.org/10.1007/BF00133363>.
- SILVA, L. R. da. *As diferenças entre o que se fala e o que se escreve no português do Brasil: a aquisição do clítico se indeterminador e apassivador*. 2017. 118 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2017.
- TEIXEIRA, Z. D.; SALLES, H. M. M. L. A. Advérbios locativos na posição de sujeito no PB. In: MARTINS, M. A. et al. (orgs.). *Estudos linguísticos: textos selecionados*. João Pessoa: Ideia, 2016. pp. 164-73.
- TEIXEIRA, Z. D.; BORGES, H. Locative adverbs in the subject position in Brazilian Portuguese. *Revista de Estudos da Linguagem*, 2022, no prelo.
- URIAGEREKA, J. An F position in Western Romance. In: KISS, K. (org.). *Discourse Configurational Languages*. New York: Oxford University Press, 1995. pp. 153-75.

RECONSTRUINDO A HISTÓRIA DO PORTUGUÊS DO BRASIL PELO CORPUS TYCHO BRAHE BRASIL: NOVOS DADOS, NOVOS OLHARES¹

REBUILDING THE HISTORY OF BRAZILIAN PORTUGUESE THROUGH TYCHO BRAHE BRAZIL CORPUS: NEW DATA, NEW PERSPECTIVES

Paulo Ângelo Araújo-Adriano²

Williane Silva Corôa³

RESUMO

Este artigo apresenta novos dados para a investigação histórica do português brasileiro, a serem introduzidos no *Corpus* Histórico Anotado do Português Tycho Brahe Brasil (CTBB), que passa a conter cartas e atas de Homens Bons da Câmara Municipal de Salvador, peças de teatro de comédia, grande parte do Rio de Janeiro, e cantiga. Como exemplificação do CTBB, o presente texto apresenta uma investigação sobre a colocação clítica em contexto V1 e sobre a perífrase progressiva no português brasileiro, pautando-se não apenas na data de nascimento dos autores, mas também na data de publicação dos textos. Discutimos a questão, grandemente debatida, da emergência do português brasileiro como já existente em meados do século XVIII, conforme defendem Ribeiro (1998) e Corôa (2021). Além disso, argumentamos em favor da importância de se fazer linguística histórica contrastando não só a data de nascimento e a de publicação, mas, também, tipos textuais diferentes.

PALAVRAS-CHAVE: Linguística de *corpus*. Metodologia em linguística diacrônica. Emergência do português brasileiro. Clíticos. Estrutura progressiva.

ABSTRACT

This paper presents new data for the historical investigation in Brazilian Portuguese, to be available on the Brazilian Tycho Brahe Parsed *Corpus* of Historical Portuguese (CTBB), which will contain letters and minutes from Good Men from Salvador City Council and comedy plays, essentially from Rio de Janeiro. As an exemplification of this new set of data, this paper presents an investigation on clitic placement in V1 context and on the morphology of progressive periphrasis in Brazilian Portuguese, based not only on the authors' date of birth but also on the publication date of the texts. We discuss the issue of the emergence of Brazilian Portuguese, widely debated, as already existing in the mid-18th century, as defended by Ribeiro (1998) and Corôa (2021). Furthermore, we argue in favour of the importance of research in historical linguistics by comparing not only birth and publication date but also different textual types.

KEYWORDS: *Corpus* linguistics. Methodology in diachronic linguistics. The emergence of Brazilian Portuguese. Clitics. Progressive structure.

¹ Agradecemos aos dois pareceristas anônimos pelas observações e sugestões a este artigo. Ainda que nem todas puderam ser contempladas neste texto, elas serão consideradas em trabalhos futuros.

² O presente trabalho foi realizado com apoio da Fapesp, Processo 2019/17443-9. Doutorando da Unicamp, pauloangelooa@gmail.com, <https://orcid.org/0000-0002-9884-0723>.

³ O presente trabalho foi realizado com apoio da Fapesp, Processo 2017/16581-3. Doutoranda da Unicamp, williscorea@gmail.com, <https://orcid.org/0000-0002-5887-7281>.

Introdução

Neste artigo apresentamos novos elementos para reconstruir a história do português do Brasil (PB) pelo *Corpus* Histórico Anotado do Português Tycho Brahe-Brasil (CTB-Brasil). Partimos de dois *corpora* que integram esse novo *corpus*: o primeiro constituído por *Cartas da Câmara Municipal de Salvador*, produzidas por brasileiros que ocupavam o cargo de escrivão ao longo dos séculos XVII e XVIII; e o segundo, constituído de texto teatrais e cantigas, também produzidos por brasileiros a partir do século XVII. Analisamos os padrões de colocação de clíticos e a perífrase progressiva.

Apesar de estarem disponíveis diversos *corpora* do português brasileiro sincrônico, como NILC – Núcleo Interinstitucional de Linguística Computacional – (USP, UFSCAR, UNESP) e NURC – Norma Urbana Culta –, para citar alguns, há poucos *corpora* que concentram textos históricos brasileiros. Dos mais notórios há apenas dois, o *corpus* do PHPB – Para a história do português brasileiro – (UFPE, UFPBA, UFMG, UFRJ, UFSC, UFPB, USP) e o *Corpus* Tycho Brahe – CTB – (Unicamp). Muito do que se sabe da história do PB atualmente certamente é devido ao extenso trabalho que pesquisadores do Brasil desenvolveram tendo em mãos o PHPB. No entanto, as fontes históricas de tal projeto não estão anotadas nem morfológica tampouco sintaticamente, além de não estarem disponíveis ao público, o que impede que o linguista histórico trabalhe de forma mais rápida, prática, dada a ausência de anotação, e, o mais importante, que o linguista replique a pesquisa desenvolvida – padrões intrínsecos das ciências. Para além dessas questões, quando não se trabalha com *corpora* anotados, os dados acabam sendo colhidos à mão, o que pode dar margem para falhas e lapsos, resultados indesejados em qualquer campo da ciência. Por esse motivo, defendemos a ideia de que é necessário trabalhar com *corpora* históricos anotados morfológica e sintaticamente, tal como o *Corpus* Tycho Brahe.

Adicionalmente, faz-se relevante estar atento às idiosincrasias ligadas às tipologias textuais que integram tais *corpora*, pois cada tipo textual pode favorecer uma ou outra estrutura dependendo dos seus objetivos e do seu formato. Apenas com *corpora* de diferentes tipos textuais tais nuances aparecem, daí a necessidade de se ter em mãos *corpora* grandes, não só em termos de número de palavras, mas também em relação à diversidade de tipos textuais, de modo que a língua representada por aquele recorte não seja enviesada por um ou outro tipo textual.

Ainda, faz parte do processo metodológico do linguista histórico decidir se o recorte temporal a ser analisado irá se basear na data de nascimento do autor ou na data de publicação do texto. Entretanto, uma escolha não deveria excluir a outra, na medida em que, a depender do tipo textual investigado, a data de publicação pode representar melhor a gramática de um período, e o contrário sendo igualmente verdade. Assim, também defendemos ser necessário considerar tanto a data de nascimento dos autores quanto a data de publicação dos textos, atentando às especificidades de cada tipo textual.

De modo a explicitar tal posicionamento, este texto está organizado da maneira como se segue. Em §2, discutimos alguns desafios do labor do linguista histórico, destacando alguns pontos metodológicos, como escolha do gênero e da datação das obras, se pelo nascimento do autor, se

pela publicação da obra. Em §3, apresentamos os novos dados que passam a compor o CTB, que compreendem cartas e atas de *Homens Bons*⁴ da Câmara municipal de Salvador e representações artísticas. Já em §4, analisamos dois fenômenos para ilustrar como esses novos dados podem trazer novos olhares para compreendermos a história do português do Brasil. Por fim, em §5 fazemos algumas considerações finais.

1. Desafios de se trabalhar com linguística histórica/diacrônica

O labor do historiador da língua assenta-se, essencialmente, em fontes escritas que resistiram às intempéries do tempo. Nesse sentido, as informações sobre *quando, como, onde e por quem* o texto foi escrito são indispensáveis para melhor delinear como a mudança linguística acontece (MATTOS E SILVA, 1996). Porém, nos textos históricos, nem sempre é possível acessar tais informações, o que torna a vida do pesquisador em linguística histórica bastante complicada.

Com vistas a sanar esse problema, nos últimos anos, o número de projetos de investigação que visavam à constituição de *corpora* históricos, tanto no Brasil quanto no exterior, deu um salto (JENSET; MCGILLIVRAY, 2017). Isso se justifica, pois os *corpora* são elementos essenciais da linguística histórica, cuja utilização com fins de investigação permite a coleta e a exploração de um conjunto de dados úteis para descrição e análise linguística de forma mais rápida, prática, confiável e, o mais importante, replicável.

Para um pesquisador de linguística histórica, o estudo de qualquer fenômeno linguístico deve fundamentalmente basear-se em um *corpus*. Um *corpus* histórico pode ser definido como um conjunto de exemplos de linguagem natural – desde frases até um conjunto de textos escritos em estágios pretéritos de uma língua (HUNSTON, 2002). Deve, portanto, ser tão extenso quanto possível, visto que o historiador da língua conta apenas com a documentação histórica remanescente. Além disso, deve, quando possível – cientes das dificuldades inerentes – ser representativo considerando as diferentes variedades da língua e os diferentes tipos textuais, a fim de que se possa fazer generalizações. Uma vez que se é impossível investigar todo material linguístico existente de uma língua, o que seria a única maneira de ter um *corpus* realmente representativo, o trabalho do linguista histórico é, de certa sorte, limitado (cf. GARCÍA GARCÍA, 2002, p. 121). Dada essa limitação, o único método para que o objeto de investigação – o *corpus* – seja o mais representativo o possível, é pela variação de tipos textuais.

No que diz respeito à variabilidade do tipo textual, a eleição do *corpus* está intimamente atrelada à escolha do tipo textual e ao recorte histórico a ser investigado. No bojo dessa discussão, “a inexistência de um mesmo tipo de texto em todas as fases da história do Português” (CAMBRAIA, 1994, p. 11) faz com que o historiador da língua manuseie tipologias textuais diferentes, o que pode impactar nos resultados alcançados, visto que tais resultados se mostram sensíveis à seleção do tipo textual. Um exemplo disso aparece nas monografias organizadas por Cyrino e Torres Morais (2018).

⁴ O termo *homens bons* era utilizado para designar os membros da nobreza local, elegíveis às Câmaras municipais. Os requisitos para ser considerado Homem bom eram: ser maior de 25 anos, casado ou emancipado, católico, e sem nenhuma “impureza de sangue”.

Galves (2018, p. 442), no *posfácio* do livro, aponta para a influência do tipo textual sobre os resultados encontrados, pois tipologias textuais diversas apresentaram pesos sócio-históricos distintos. Desse modo, as especificidades de cada tipo devem ser levadas em consideração pelo linguista histórico ao constituir seu *corpus*⁵.

Outro ponto a ser levado em consideração refere-se a qual momento histórico adotar, se a data de nascimento do autor, se a data de publicação do texto. Dentro do arcabouço da Gramática Gerativa, assume-se que a gramática de um falante é estabelecida no período de aquisição da linguagem, por volta dos 6 anos de idade; adicionalmente, assume-se que a aquisição e a mudança estão intrinsecamente atreladas (cf. LIGHTFOOT, 1979; entre outros). Sob esse ponto de vista, a mudança linguística seria resultado da interação entre a capacidade inata do indivíduo de adquirir uma língua e a experiência linguística vivenciada pelas sucessivas gerações de falantes. Se aquisição e mudança se entrelaçam, é patente adotar a data de nascimento do autor como critério de datação dos textos, isto é, estaríamos diante da *geração biológica* – um termo adotado de Paixão de Sousa (2004).

Por outro lado, não é tão claro assim que esse deva ser o caso para textos históricos que foram escritos com o objetivo de interpretar uma narrativa para um público específico, como é o caso de peças teatrais. Uma vez que autores de teatro tentam representar na fala de seus personagens a linguagem corriqueira do seu público-alvo, parece-nos mais adequada adotar a data de publicação do texto, não a data de nascimento do seu escrevente, a fim de investigarmos uma *geração histórica* – outro termo adotado de Paixão de Sousa (2004).

Assim, embora a gramática do autor da peça tenha sido fixada na sua infância, no período de aquisição, a gramática que ele utiliza, idealmente, nas suas obras teatrais é aquela representativa do contexto de sua criação, que depende do seu público-alvo. Isso pode ser evidenciado no comentário/advertência de Artur Azevedo (1900), autor de peças brasileiras de teatro do século XIX, sobre a tentativa de suas peças agradarem e serem destinadas a todo tipo de público-alvo, mesmo com certa resistência da Companhia Lucinda Simões, responsável à época por promover espetáculos destinados apenas à elite econômica (cf. NEVES, 2006, p. 21):

Esses espetáculos podem ser freqüentados, necessariamente, por todo aquele ou aquela que comprar o seu bilhete e esteja trajado, ou trajada, com certa decência; mas a empresa destina-os especialmente “às mais distintas famílias da elite da nossa sociedade”, e conta que o seu teatro seja, às quartas-feiras, um ponto de reunião para as damas e os cavalheiros do monde, como dizem os franceses, ou do high life, como dizem os ingleses. A tentativa é inteligente e simpática, mesmo porque talvez consiga fazer as pazes entre a boa sociedade e o teatro, que há muito se desavieram.

⁵ Cambraia (1994) chama a atenção para essa questão tanto no que tange à oposição (i) literário versus não literário, quanto no que tange à oposição (ii) prosa versus verso. Com relação à (i), Cambraia aponta a dificuldade em diferenciar a prosa epistolar literária e as correspondências comuns, dado que não há uma clara fronteira entre o que seria ou não literário. Com relação à (ii), Cambraia (1994) exemplifica a questão com os textos de teatro e pontua que se um pesquisador desejar utilizá-los como fonte de investigação deve considerar que, durante um tempo, tais textos eram escritos em verso. Assim, uma mesma tipologia teria o inconveniente de misturar textos de diferente natureza.

Fica patente, portanto, que o autor, a despeito da pressão econômica, representava nas suas peças todo tipo de público, provavelmente adequando e aproximando a linguagem do seu teatro à linguagem da sociedade.

Nota-se que autores teatrais não ignoram solenemente a fala do seu público; ao contrário, precisam estar atentos à linguagem da sua sincronia para que suas obras não soem artificiais. Um exemplo dessa preocupação é o Primeiro Congresso Brasileiro de Língua Falada no Teatro na Universidade da Bahia em 1956, para discutir exatamente questões relativas à língua na dramaturgia teatral. É claro que, muito embora a escrita teatral não seja totalmente verossímil à fala da época representada, o trabalho do linguista histórico “garimpeiro” – para usar um termo de Barbosa (1998) – é sempre desnudar a língua oral do passado, através da escrita, “como uma *tentativa* de aproximação da realidade” (LOBO, 1998, p. 179, grifo nosso). Dessa maneira, parece razoável afirmar que manusear a data de publicação de um texto também é exitoso, quando se quer investigar o conhecimento linguístico de falantes antepassados.

Olhar para a *geração biológica* ou para a *geração histórica* impacta, portanto, a escolha do *corpus* a ser adotado. A impossibilidade de identificar *por quem* o texto foi escrito tem consequências óbvias sobre a eleição da data de nascimento como critério de datação dos textos. Optar por um ou por outro critério também repercute sobre a periodização, pois olhar para a *geração biológica* implica observar a mudança precocemente se comparado com estudos que consideram apenas a data de produção dos textos, isto é, a *geração histórica*. Não é nosso objetivo com essa discussão bater o martelo e definir qual a melhor escolha, muito menos condenar outra; ao contrário o objetivo é problematizar e salientar que tal escolha não deve ser ingênua, mas consciente, dadas as implicações.

Além dessas questões, é indispensável que quaisquer *corpora* sejam armazenados eletronicamente e possam ser pesquisados por programas de computador. Também urge que os *corpora* sejam anotados, dado que a análise feita manualmente abre grande margem para a possibilidade de erros e até de vieses tendenciosos. Assim, os *corpora* anotados são vantajosos, pois o processo de coleta e classificação dos dados, além de contar com uma fonte de evidências muito mais ampla do que se o levantamento for feito manualmente, é mais preciso, mais transparente do ponto de vista metodológico e facilita a reprodutibilidade dos dados investigados (JENSET; MCGILLIVRAY, 2017).

Dessa maneira, o processo de anotação promove a generalização dos dados para um conjunto de dados maior, possibilitando a replicabilidade – princípio intrínseco ao fazer científico –, o que é impossível caso o levantamento dos dados seja feito manualmente. Portanto, a anotação de *corpus* é uma etapa essencial no processo de investigação linguística e que contribui para testar empiricamente hipóteses a partir dos dados.

No Brasil, apesar da grande tradição de constituição de *corpora* históricos, contamos com poucos *corpora* anotados. Um deles é *Corpus Tycho Brahe*, primeiro *corpus* eletrônico anotado em língua portuguesa. Na próxima seção, apresentamos o *Corpus Tycho Brahe*, com dados majoritariamente do português europeu (PE), e a nova vertente na qual se está repaginando: o *Corpus Tycho Brahe Brasil*.

2. O *Corpus* Tycho Brahe: velhos e novos dados

O *Corpus* Histórico do Português Tycho Brahe é um *corpus* eletrônico anotado (cf. GALVES; FARIA, 2010; GALVES *et al.*, 2017), composto de textos majoritariamente em português europeu, escritos por autores nascidos entre o fim do século XIV e o fim do século XIX. Atualmente, 88 textos (3.544.628 palavras) estão disponíveis para pesquisa, com um sistema de anotação morfológica (aplicada em 58 textos, com um total de 2.280.819 palavras) e sintática (aplicada em 27 textos, com um total de 1.234.323 palavras).

A comunidade científica muito ganhou com esse *corpus* considerado médio-grande⁶, que serviu de base para diversos estudos sobre a ordem do sujeito, sistema V2, interpolação, ordem relativa clítico-verbo, o fenômeno do sujeito nulo⁷ etc., dando novos subsídios para o entendimento da emergência do português europeu moderno.

Como já mencionado na introdução, *corpora* anotados seja morfológicamente seja sintaticamente possibilitam o manuseio de uma quantidade grande de dados em pouco tempo. Até o momento, no *Corpus* Tycho Brahe há apenas 11 textos do português brasileiro. Entretanto, é imprescindível ampliar o conjunto de dados para se investigar rigorosamente o passado de uma língua: apenas com uma diversidade de gêneros e de dados é que é possível delinear acuradamente a emergência do português falado no Brasil, e até mesmo conjugando aspectos linguísticos com aspectos sócio-históricos. Galves (2018, p. 456), discorrendo sobre o que vai nos permitir fazer descrições mais precisas sobre o português brasileiro, aponta que “serão necessários, portanto, dados, novos dados, muito mais dados”. Na próxima seção, apresentamos esses novos dados que passam a fazer parte da nova roupagem do *Corpus* Tycho Brahe: o *Corpus* Tycho Brahe Brasil.

2.1. O *Corpus* Tycho Brahe (Brasil): novos dados.

Mais recentemente o projeto, liderado por Charlotte Galves (Unicamp), propõe-se a ampliar a quantidade de textos brasileiros com a criação do *Corpus* Tycho Brahe-Brasil (CTB-Brasil). O CTB-Brasil será constituído de textos escritos no Brasil e por brasileiros entre o século XVI e o século XX. O objetivo é retratar a dinâmica da emergência do português brasileiro e melhor compreender a separação entre a variedade clássica, europeia e brasileira do português. Dentre os textos que constituirão o CTB-Brasil estão o *corpus* de textos de teatro e cantiga e o *corpus* de Cartas e Atas da Câmara Municipal de Salvador. A seguir, ambos são detalhados.

⁶ Berber Sardinha (2004, p. 26) considera a seguinte classificação em relação ao tamanho de um *corpus*, relativo ao número de palavras: pequeno (menos de 80 mil), pequeno-médio (80 a 250 mil), médio (250 mil a 1 milhão), médio-grande (1 milhão a 10 milhões) e grande (10 milhões ou mais).

⁷ Para ter acesso à lista de trabalhos realizados a partir do *Corpus* Tycho Brahe acesse o link: <http://www.tycho.iel.unicamp.br/~tycho/pesquisa/>.

2.2. *Corpus* de textos de teatro

A fim de possibilitar novos olhares para o português brasileiro, 11 peças de teatro publicadas entre o século XVIII e o século XXI e 1 cantiga do século XVIII foram incluídas no *corpus*. Os textos anotados morfológicamente (233.183 palavras), fruto do projeto de tese FAPESP (Processo 2019/17443-9), estão todos disponíveis no site do *Corpus*⁸, com as informações básicas expostas no quadro 1, a seguir.

Quadro 1: Informações das peças que compuseram o *corpus*

Nascimento	Autor	Nome da Peça	Publicação do texto	Número de palavras	Tipo Textual
1705/séc. 18	Antônio José da Silva	Guerra do Alecrim e da Manjerona	1737/séc. 18	27.224	Comédia teatral
1740/séc. 18	Domingos Caldas Barbosa	Viola de Lereno	1798/séc. 18	18.445	Cantiga
1815/séc. 19	Martins Pena	O juiz de paz na roça	1833/séc. 19	6.897	Comédia teatral
1815/séc. 19	Martins Pena	O noviço	1845/séc. 19	17.563	Comédia teatral
1829/séc. 19	José de Alencar	O demônio familiar	1857/séc. 19	25.319	Comédia teatral
1838/séc. 19	França Junior	Caiu o ministério	1883/séc. 19	14.629	Comédia teatral
1855/séc. 19	Artur de Azevedo	O tribofe	1891/séc. 19	22.434	Comédia teatral
1880/séc. 19	Gastão Tojeiro	Onde canta o sabiá	1920/séc. 20	27.495	Comédia teatral
1934/séc. 20	Gianfrancesco Guarnieri	Eles não usam black-tie	1957/séc. 20	22.571	Comédia teatral
1956/séc. 20	Miguel Falabella	A partilha	1990/séc. 20	17.531	Comédia teatral
1978/séc. 20	Paulo Gustavo	Minha mãe é uma peça	2006/séc. 21	33.075	Comédia teatral
1966/séc. 20	Paulo Sacaldassy	Fulana, Sicrana e Beltrana	2007/séc. 21	8.548	Comédia teatral

A escolha pela tipologia peça de teatro não foi ingênua. Dada a abordagem teórica assumida, a da Gramática Gerativa, a pergunta de pesquisa norteadora leva em consideração o conhecimento interno do falante. Entretanto, quando se faz linguística histórica, na ausência de uma gramática interna viva, é impossível acessar o conhecimento de falantes de outras épocas, cuja gramática já não pode fornecer julgamento de gramaticalidade/aceitabilidade.

Dentre os textos de teatro, o gênero comédia teatral foi o escolhido visto que a comédia é um gênero sobre o cotidiano de pessoas e, em geral, são retratados indivíduos de distintas classes sociais, de modo que possibilita que vejamos retratados nesses textos diferentes variedades da língua da época, inclusive aquelas que não estariam sob pressões normativas. Dessa maneira, concordamos com a afirmação de que “as peças de teatro constituem importante material de pesquisa quando se quer tentar uma aproximação com a fala de sincronia passadas” (DUARTE, 2012, p. 19) e, portanto, defendemos que peças de teatro também são fontes confiáveis e plausíveis de se acessar essa gramática interna de outras sincronias.

⁸ <http://www.tycho.iel.unicamp.br/~tycho/corpus/en/catalogo.html>

No que se refere às obras ineditamente inseridas no *Corpus* Tycho Brahe, a quantidade de textos para cada século foi determinada levando em consideração o início, o meio e o fim do século em questão. Por exemplo, para representar o século XX, uma peça de 1920, uma em torno do meio do século, de 1957, e uma de 1990, na virada do século. Para o século XIX, foram incluídas mais peças, pois é o século cujo número de peças disponíveis é maior. Todos os textos foram coletados levando em consideração a nacionalidade do autor. Na medida do possível, foram selecionadas peças de autores de vários locais do Brasil, porém a maioria é oriunda do Rio de Janeiro, dada a concentração artística e populacional do país nesse estado (cf. SILVA NETO, 1977).

Na próxima seção, vai ser apresentado o conjunto de cartas e atas da Câmara Municipal, escritas nos séculos XVII e XVIII por brasileiros, que também passam a compor o *Corpus* Tycho Brahe Brasil.

2.3. Cartas e atas de Homens Bons da Câmara Municipal de Salvador

As cartas e atas que compõem o *corpus* elaborado ao longo do projeto de tese (Processo Fapesp: 2017/16581-3) fazem parte do Fundo da Câmara Municipal de Salvador, composto por documentos provenientes da administração municipal. A vasta documentação da Câmara de Salvador encontra-se sob a guarda do Arquivo Histórico Municipal, inaugurado em 10 de abril de 1932, administrado atualmente pela Fundação Gregório de Matos.

As cartas e atas, ao lado de alvarás, posturas régias, requerimentos da população, eram os principais meios de comunicação entre a colônia e a metrópole. A responsabilidade de escrever os documentos era do escrivão. Na Câmara da Bahia, no século XVI, as principais funções eram as de escrivão, almotacé de execuções e tesoureiro, mas apenas o cargo de escrivão era exercido por “homens das letras”: para exercê-lo, os indivíduos deveriam ser nomeados por provisão real, ter domínio da escrita e conhecimento da legislação portuguesa.

Quadro 2: Informações das cartas e atas que compuseram o *corpus*

Nascimento	Autor	Publicação do texto	Tipo Textual
1602/séc. 17	Rui de Carvalho Pinheiro	1650/séc. 17	Carta
1630/séc. 18	Domingos Garcia de Aragão	1653/séc. 17	Carta
1650/séc. 17	Pedro Dias Pereira	1700/séc. 18	Carta
1670/séc. 17	João de Couros Carneiro Filho	1695/séc. 17	Carta
1670/séc. 17	Manuel Silveira de Magalhães	1715/séc. 18	Carta
1672/séc. 17	Manuel Pessoa de Vasconcelos	1699/séc. 17	Carta
1672/séc. 17	Manuel Pessoa de Vasconcelos	1715/séc. 18	Carta
1700/séc. 18	João de Couros Carneiro Neto	1725/séc. 18	Carta
1710/séc. 18	Jerônimo Sodré Pereira	1765/séc. 18	Ata
1720/séc. 18	Joaquim Rodrigues da Silveira	1770/séc. 18	Ata
1720/séc. 18	Manuel José de Azevedo	1768/séc. 18	Ata
1721/séc. 18	José Álvaro Pereira Sodré	1768/séc. 18	Ata
1725/séc. 18	João Duarte Silva	1775/séc. 18	Ata

O conjunto de cartas e atas faz parte da Série *Documentos Históricos do Arquivo Histórico Municipal* que foi publicada em 1949, na ocasião da comemoração dos 400 anos da Cidade de Salvador. As *Cartas do Senado a Sua Magestade* reúnem, em seis volumes, as cartas produzidas pela câmara municipal entre 1638 e 1730. Já o *Livro de Atas da Câmara* reúne, em dez volumes, as atas produzidas entre 1625 e 1765. Para a tese em elaboração foram anotados todos os volumes de cartas e os volumes 10 e 11 de atas, o que totaliza 269.777 palavras.

Na próxima seção, exemplificamos a praticidade de se fazer linguística histórica com o morfológica e sintaticamente anotado *Corpus Tycho Brahe Brasil*, a partir dos novos dados supramencionados.

3. O *Corpus Tycho Brahe*: novos olhares.

A partir do conjunto de dados organizados e elaborados, **já é possível delinear algumas questões** relativas à emergência do PB e à dinâmica da mudança. Para demonstrá-las, escolhemos dois fenômenos característicos do PB a fim de investigar quando esses fenômenos aparecem nos dados históricos: (i) a estrutura progressiva e a (ii) colocação de clíticos.

Uma característica bastante peculiar na história do PB é a colocação de clíticos, que, apesar de ser um fenômeno bastante estudado (cf. PAGOTTO, 1992; CARNEIRO, 2005; MARTINS, 2009), ainda guarda algumas questões sem resposta. Uma delas é saber se a colocação de clíticos no PB mudou em relação ao PE ou em relação a fases mais pretéritas do português, visto que o PB e o PE são claramente distintos atualmente: enquanto o primeiro apresenta próclise generalizada, o segundo apresenta variação quer em frases finitas quer em frases infinitas.

(1) Colocação clítica do PB

- a. João **te** olhou
- b. **Me** abrace

(2) Colocação clítica do PE

- a. João olhou-**te**
- b. Abrace-**me**

Neste artigo, para o fenômeno da colocação clítica, discutimos os casos em (1b) e (2b), em cujo contexto o PB difere muito fortemente do PE, visto que o clítico não aparece em primeira posição na sentença, isto é, em um contexto V1⁹.

Dentre as estruturas típicas que diferem o PE do PB aparece a estrutura progressiva. Enquanto na primeira variedade a perífrase progressiva é veiculada pelo auxiliar *estar* junto de um domínio não finito infinitivo encabeçado pela preposição *a* (*estar + a + infinitivo*), na segunda, a leitura progressiva também é estruturada pelo verbo *estar*, mas esse auxiliar é seguido por um domínio gerundivo (*estar*

⁹ São consideradas sentenças V1 aquelas em que o verbo é o primeiro elemento da sentença e também as que são iniciadas por interjeições, conjunções coordenativas, vocativos e clíticos. Neste trabalho analisamos apenas as sentenças iniciadas por verbo e por um clítico, deixando os demais contextos de lado.

+ gerúndio), conforme os contrastes a seguir ilustram¹⁰.

- (3) a. Olha! A Maria **está a correr**. (PE, *PB)
 b. Olha! A Maria **está correndo**. (*PE, PB)

Nota-se que, muito embora algumas regiões do sul de Portugal, como Alentejo e Algarve, também façam uso de (3b) com marcas de gerúndio flexionado (cf. *estão me chamandem* retirado de LOBO, 2003, p. 381), o uso de (3a) é bastante disseminado no PE (cf. HRICSINA, 2014; LOBO, 2003; MIRA MATEUS *et al.*, 2003). Observar tal variação se faz, portanto, relevante pois pode fornecer evidências para o período em que a gramática do PB emergiu.

3.1. A colocação de clítico em contexto V1 e a estrutura progressiva no *corpus* de representações artísticas

Quando o verbo aparece em primeira posição na sentença, tanto no Português Clássico (PCI) (século XVI-XVII) quanto no PE a ênclise é obrigatória. Dessa maneira, a gramática do PB mostra-se inovadora permitindo a ocorrência de próclise em sentenças com verbo em primeira posição absoluta. Inclusive, na sincronia, a próclise é categórica. No sistema de busca do *Corpus* Tycho Brahe, a anotação eletrônica do *corpus* possibilita rápida e prontamente a busca pela colocação clítica. A fim de observar sua distribuição ao longo do tempo, a busca pelos dados foi construída considerando a posição em que o clítico aparece. Nos dados de representações artísticas analisados, encontramos 37 casos (7%) de próclise e 496 casos (93%) de ênclise em sentenças com o verbo em posição inicial.

Para os casos de próclise, buscamos¹¹ pela etiqueta CL*|SE* (para clítico e clítico reflexivo *se*) considerando que o clítico deveria aparecer em primeira posição, por isso a etiqueta *isFirst*. Também foram desconsideradas as vírgulas em relações de precedência imediata. Alguns resultados obtidos para cada século estão apresentados abaixo.

- (4) Algoritmo de busca fornecido para encontrar próclise em contexto V1
 CL|SE|CL+*|SE+* isFirst
- (5) Próclise em contexto V1
- Me** dá coceira ver tanto bagunça e tanta coisa fora do lugar! (1705)
 - Te** arrependeste a tempo. (1829)
 - Te** peguei, seu capitalista! (1934)
 - Me** explica melhor o que você está sabendo ... (1978)

¹⁰ O asterisco usado no exemplo representa a agramaticalidade da sentença seja na variedade brasileira seja na europeia.

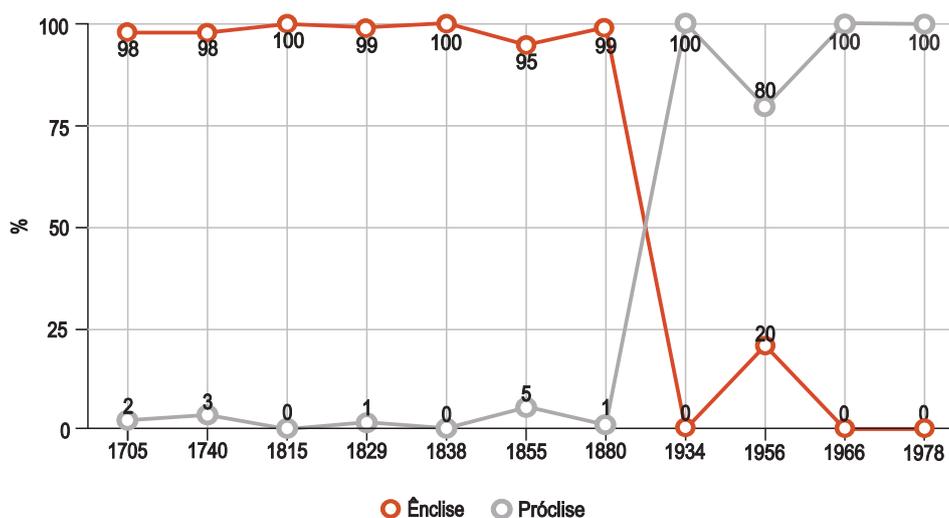
¹¹ É possível realizar buscas usando o *Corpus* Tycho Brahe de dois modos: (i) a partir da página “Consulta aos textos etiquetados” (<http://www.tycho.iel.unicamp.br/~tycho/corpus/texts/csquery/csquery.html>), em que se pode construir buscas graficamente – com o auxílio de uma interface – ou manualmente – seguindo a sintaxe de busca usada na ferramenta *CorpusSearch* (disponível em <http://corpussearch.sourceforge.net/CS-manual/Contents.html>) ou (ii) a partir da instalação da ferramenta *CorpusSearch*. A grande diferença entre os dois modos é que a busca a nível sintático apenas pode ser realizada com a ferramenta *CorpusSearch* instalada na máquina.

Para os casos de ênclise, buscamos pela etiqueta VB*+CL em primeira posição (*isFirst*) considerando que o clítico deveria aparecer após qualquer tipo de flexão verbal (representado pelo asterisco na etiqueta VB para verbo). Também foram desconsideradas as vírgulas em relações de precedência imediata. Os exemplos abaixo ilustram alguns dos resultados de ênclise para os séculos analisados.

- (6) Algoritmo de busca fornecido para encontrar ênclise em contexto V1
VB*+CL *isFirst*
- (7) Ênclise em contexto V1
- Favoreça-**me** vossa mercê, Senhora Fagundes, com o seu voto, que eu terei bom despacho no tribunal de Cupido. (1705)
 - Vi-**as** em Paris. (1829)
 - Orgulhem-**se** de mim. (1956)

O gráfico 1 mostra em porcentagem o contraste entre próclise e ênclise no contexto de verbo em primeira posição ao longo do tempo. Nela é possível perceber que, muito embora a próclise já permeava o sistema da língua desde 1705, a ênclise era a colocação categórica até o fim do século XIX. A partir de 1880, o cenário inverte e a próclise passa a ser generalizada.

Gráfico 1: Contraste próclise vs. ênclise em contexto V1 no *corpus* de representações artísticas



Fonte: elaboração dos autores

No que se refere à busca da estrutura progressiva, para a estrutura prototípica do PE, *estar + a + infinitivo*, buscamos pela etiqueta ET*, que inclui nos resultados apenas o verbo *estar*; sendo o asterisco uma instrução à busca para se obter resultados com qualquer tipo de flexão no verbo. Para a busca da preposição *a*, foi especificado que se desejava uma preposição, pela etiqueta P, mas não qualquer uma, apenas *a*, com a subespecificação *a*. Por fim, para verbos no infinitivo, a etiqueta que compôs a busca foi VB, de verbo no infinitivo. Alguns dos resultados gerados ao longo dos séculos estão exemplificados abaixo.

(8) *estar + a + infinitivo*

- Vá, senhora, não **esteja a choramingar**. (1815)
- O que **está** o senhor **a dizer**? (1815)
- Não, abaixei-a para evitar um cascudo que o patrão pretendia dar-me em um belo dia em que **estava a olhar** para a rua, em vez de servir as freguesas, e não voltei mais à loja. (1838)
- Estava a olhar** para a luz da lua. (1880)

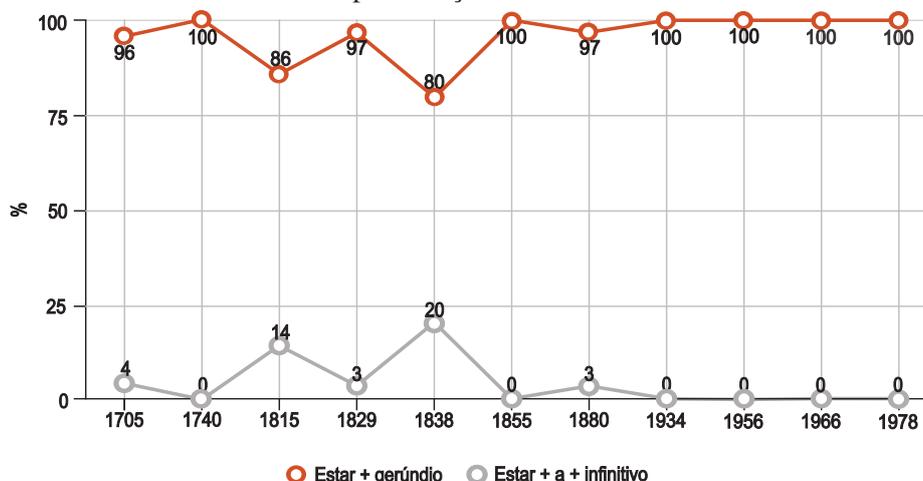
Para a estrutura típica brasileira, a busca foi praticamente a mesma, com a diferença de que não havia nela a preposição, e, no lugar de um verbo no infinitivo, havia um verbo no gerúndio, representado por VB-G (não um infinitivo), compondo a estrutura *estar + gerúndio*. Exemplos dos resultados aparecem a seguir.

(9) *estar + gerúndio*

- Homem, acabemos com isso, venha Dona Clóris, por quem **estou suspirando**. (1705)
- Venha, mamãe **está** nos **esperando**. (1829)
- Ei, Tião, **está** me **ouvindo**. (1934)
- Carlos Alberto, **estou** te **ligando** pelo seguinte. (1978)

A variação ao longo do tempo está ilustrada no gráfico 2, em que é possível observar que em todos os textos a frequência da estrutura típica brasileira *estar + gerúndio* é maior que a variante europeia, padrão semelhante ao do PCI (cf. discussão em §4). Há certa disparidade nos autores nascidos em 1815 e 1838, que usam a variante europeia; porém ainda assim o progressivo é veiculado majoritariamente pela perífrase inconfundível do Brasil, que é a estrutura usada categoricamente a partir do século XIX. Nota-se que a estrutura progressiva está presente desde o início do século XVIII, o que é diferente do comportamento encontrado para a próclise, que só emerge a partir do fim do século XIX.

Gráfico 2: Contraste *estar + gerúndio* vs. *estar + a + infinitivo* ao longo do tempo no *corpus* de representações artísticas



Fonte: elaboração dos autores

3.2. A colocação de clítico em contexto V1 e a estrutura progressiva no *corpus* de cartas e atas da Câmara Municipal de Salvador

Por mérito da anotação morfossintática, foi possível recuperar facilmente os dados sobre a estrutura progressiva e a colocação de clíticos. Diferente do *corpus* de representações artísticas, nossos dados contam também com anotação sintática. Por isso, a busca por dados é feita com o auxílio da ferramenta *CorpusSearch*¹², que deve ser instalada no computador. A sintaxe das buscas, portanto, difere um pouco do que foi apresentado anteriormente.

Para os dados de ênclise, seguimos as etapas de busca elaboradas por Andrade e Namiuti (no prelo), com a seguinte sintaxe:

- (10) Algoritmo fornecido para encontrar ênclise em contexto V1
 node: IP-MAT
 query: (IP-MAT iDomsMod NP* CL)
 AND (IP-MAT iDomsFirst flex_vb)

A sintaxe da busca informa o seguinte: busque todas as sentenças matrizes que dominam imediatamente um clítico (CL|SE), e os únicos nós que intervêm no caminho do IP matriz a um clítico são membros de um sintagma nominal (NP*) e que o IP matriz domina imediatamente como primeiro filho um verbo flexionado (flex_vb).

Para os dados de próclise, a sintaxe foi a seguinte:

- (11) Algoritmo fornecido para encontrar próclise em contexto V1
 node: IP-MAT
 query: (IP-MAT iDomsFirst NP*)
 AND (NP* idoms CL|SE)
 AND (NP* iprecedes flex_vb)

Busque todas as sentenças matrizes que dominam imediatamente como primeiro filho um sintagma nominal (NP*) e o sintagma nominal domine imediatamente um clítico (CL|SE) e preceda imediatamente um verbo flexionado (flex_vb).

O resultado das buscas soma 57 ocorrências, 49% de ênclise (28 casos) e 51% de próclise (29 casos). Os percentuais e contexto de algumas ocorrências são apresentados a seguir:

- (12) Ênclise em contexto V1

Concedemo-**lo** enquanto a avaria feita a requerimento das partes aos nossos donativos (va_PINHEIRO, 1602).

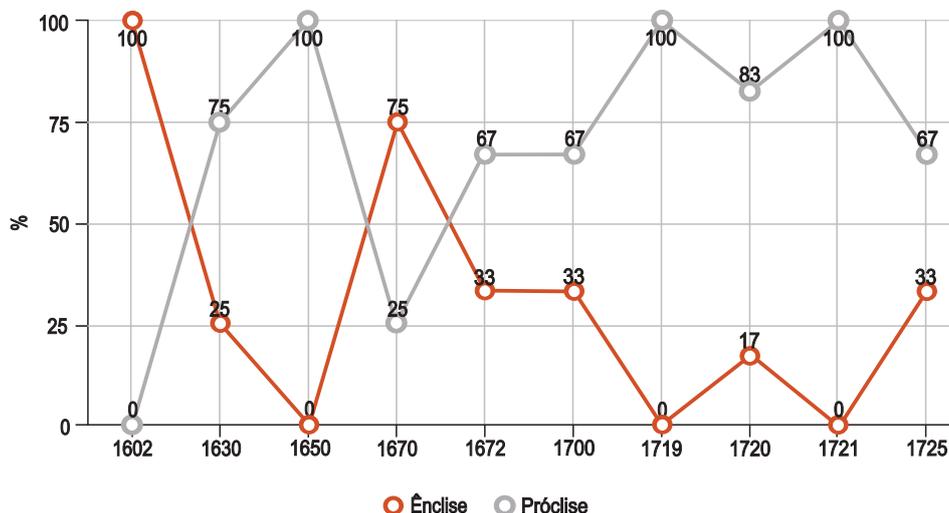
¹² <http://corpussearch.sourceforge.net/>

(13) Próclise em contexto V1

A nossa obrigação ao Serviço de Vossa Alteza, eade procurar obem Comum | desta Republica da B[ahi]a [*s:syn-clause*] **nos** leva por esta aos Reaes Pés de V[ossa] A[lteza] a Repre- | zentar o prejuizo q[ue] se seguem de schirem fundando, e fazendo pella | terra dentro muitos Eng[enh]os deaSuc[are]S junto hus dos outros sem fun- | damento deterra bastante aoq[ue] demandão delenhas para oseu gasto que he muito grande (va_Aragão, 1630).

O gráfico 3 mostra em porcentagem o contraste entre próclise e ênclise em sentenças V1 entre nascidos entre os séculos XVI e XVII. Observa-se que já na primeira metade do século XVII, o padrão de colocação de clítico encontrado em sentenças V1 é proclítico, diferindo do PCI (GALVES, BRITO; PAIXÃO DE SOUSA, 2005), cujo padrão é enclítico. O padrão se mantém entre os nascidos ao longo do século XVIII.

Gráfico 3: Contraste ênclise vs. próclise em contexto V1 ao longo do tempo no *corpus* de cartas e atas da Câmara Municipal de Salvador



Fonte: elaboração dos autores

Nos dados, já há evidências de que o padrão de colocação de clíticos começa a se afastar do padrão encontrado na gramática clássica, apontando para a emergência de uma gramática brasileira.

Os dados relacionados à estrutura progressiva também foram coletados com o auxílio da ferramenta *CorpusSearch*. A sintaxe da busca foi a seguinte para a estrutura *estar + a + infinitivo*:

(14) Algoritmo fornecido para encontrar *estar + a + infinitivo*

node: IP*

query: (flex_estar precedes PP)

AND (PP iDoms P)

AND (P idoms a)

AND (PP iDoms IP-INF*)

Busque todas as sentenças em um nó IP qualquer em que o verbo *estar* flexionado (*flex_estar*) preceda um PP (sintagma preposicionado) e o PP domine imediatamente uma P(reposição) *a*, que domine imediatamente uma sentença infinitiva. Já a sintaxe para a estrutura *estar* + gerúndio informa o seguinte:

- (15) Algoritmo fornecido para encontrar *estar* + *gerúndio*
 node: IP*
 query: (*flex_estar* HasSister IP-GER)

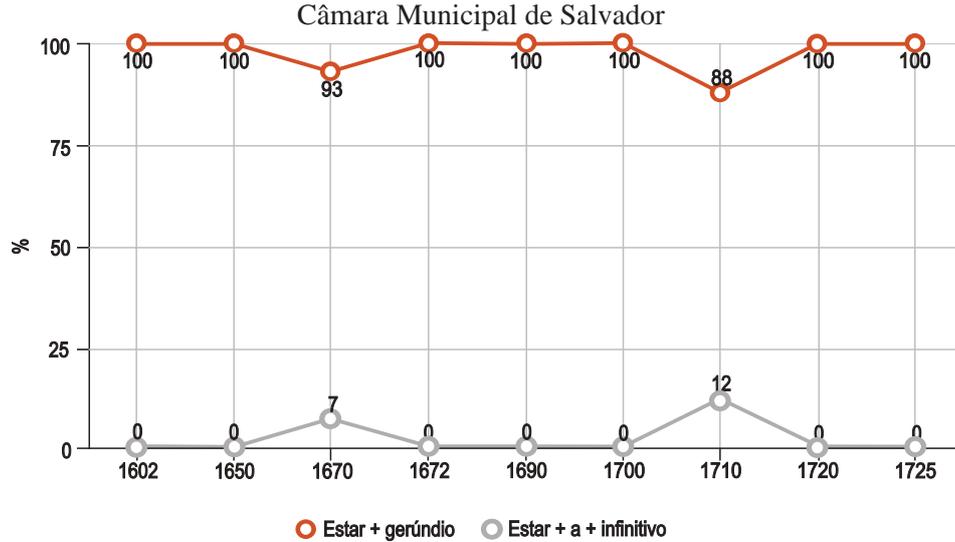
Busque em um nó IP qualquer em que o verbo *estar* flexionado (*flex_estar*) tenha como irmão uma sentença no gerúndio.

Os dados de progressivo somam 49 ocorrências, sendo a estrutura *estar* + *a* + infinitivo a menos encontrada (4%), diferente da estrutura *estar* + gerúndio com 96% dos casos. A seguir reproduzimos algumas ocorrências no *corpus* analisado:

- (16) *estar* + *a* + *infinitivo*
- a. E porque Sua Excelência várias vezes tem recomendado o seu concerto por ser aquela ladeira e por onde se há de conduzir o Marquês e Arcebispo que **estão a chegar** da Índia requereu se fizesse na dita ruína inspeção ocular para à vista dela se lhe dar a providência que pede a necessidade o que visto pela vereação assim o mandaram de que mandaram fazer este termo que assinaram (va_Sodré, 1710).
 - b. A causa que defendo contra os irmãos da Casa da Misericórdia desta cidade com que nos receberam os embargos **está a ressoar** sobre estes (va_Magalhães, 1680).
- (17) *estar* + *gerúndio*
- a. E na dita vereação requereu o procurador atual da Câmara que como capitão André Marques não tinha feito a conta do que era devedor ao donativo real **estava devendo** a este, (va_Sodré, 1710).
 - b. Além da ruína que **está ameaçando** a terra dos quintais[,] por lhe faltar o disparo dos muros por causa do inverno[,] caíram (va_Silveira, ~1725).

De um modo geral, os resultados encontrados apontam para a manutenção da expressão da estrutura progressiva tal como ocorre no PCI. Há prevalência pela estrutura *estar* + *gerúndio* nos dados brasileiros entre os séculos XVII e XVIII, com alguns poucos casos de *estar* + *a* + *infinitivo*, atestado em indivíduos cultos, que viveram por um período de tempo em Lisboa. Os dados, considerando a data de nascimento, são apresentados no gráfico 4:

Gráfico 4: Contraste *estar + gerúndio vs. estar + infinitivo* ao longo do tempo no *corpus* de cartas e atas da Câmara Municipal de Salvador



Fonte: elaboração dos autores

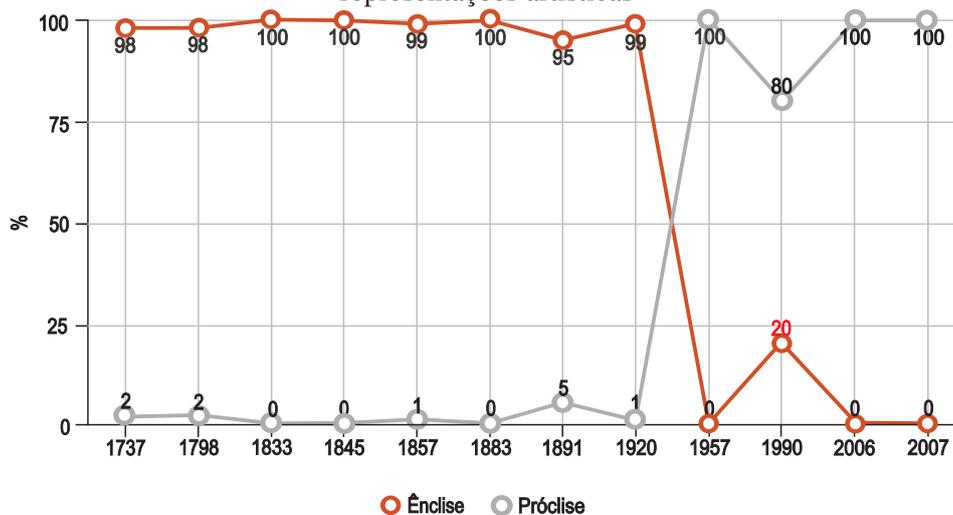
A seguir, discutimos os resultados.

3.3. Discussão dos resultados

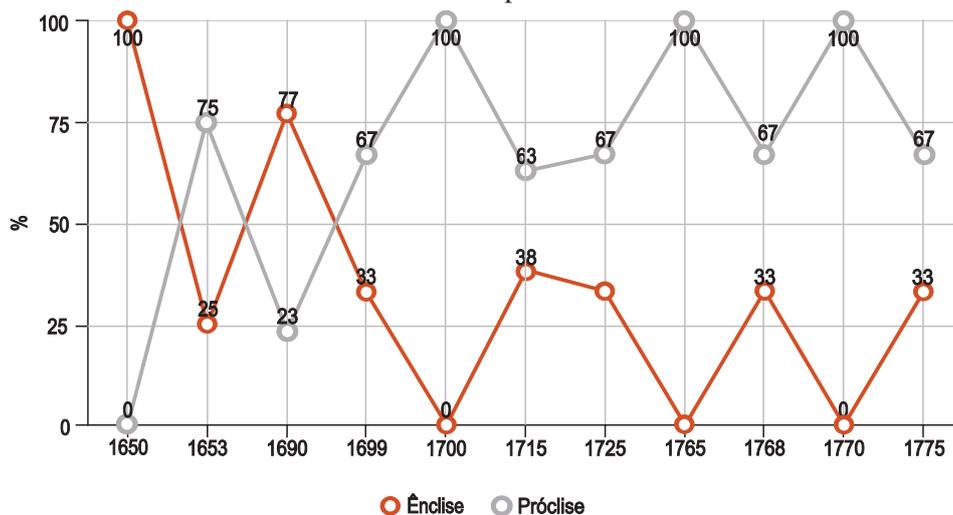
Visto que os dados relativos à perífrase progressiva são similares em ambos os *corpora* e o que de fato difere são os dados de colocação de clítico, passamos a discutir o aspecto ligado à tipologia textual e ao contraste *geração biológica versus geração histórica*.

Do ponto de vista da data de publicação dos textos, os dados diacrônicos em ambos os *corpora* mostram que a mudança é detectada mais tardiamente, comparados com os dados a partir da data de nascimento dos autores. Na seção anterior, mostramos que a mudança do padrão ênclise para próclise ocorreu nos textos de representações artísticas no fim do século XIX, com um crescimento abrupto de 1% para 100% de próclise em contexto V1 (cf. gráfico 1). Já para o *corpus* de Cartas e Atas da Câmara Municipal de Salvador, a mudança aparece ainda na segunda metade do século XVII (cf. gráfico 3), de modo não tão abrupto como nos *corpora* de representações artísticas.

Quando o parâmetro de análise passa a ser a data de publicação dos textos, a mudança ênclise-próclise ocorre não mais no fim do século XIX, mas no começo do século XX nos textos de representação artística, quando a colocação clítica passa a se comportar de maneira oposta ao padrão até então vigente, como fica patente no gráfico 5. O mesmo acontece com relação à percepção da mudança nos dados do *Corpus* de cartas e atas da Câmara Municipal de Salvador, cuja mudança torna-se perceptível no início do século XVIII, como aparece no gráfico 6.

Gráfico 5: Contraste próclise vs. ênclise em contexto V1 por data de publicação no *corpus* de representações artísticas

Fonte: elaboração dos autores

Gráfico 6: Contraste próclise vs. ênclise em contexto V1 por data de publicação no *corpus* de Cartas e Atas da Câmara Municipal de Salvador

Fonte: elaboração dos autores

Olhando para o *corpus* de representação artística, uma das vantagens de se adotar a *geração histórica* é que a data de publicação possibilita observar o comportamento do português brasileiro no século XXI, visto que não é possível olhar para a *geração biológica* no mesmo século: autores que nasceram no advento do século XXI ainda não publicaram peças de teatro. Futuramente, com peças escritas por autores que nasceram no século XXI, vamos ter dados o suficiente para contrastar e compreender mais acuradamente se de fato a data de publicação de uma peça reflete o conhecimento linguístico do período publicado ou se é a data de nascimento do autor. Como ainda não temos o queijo e a faca em mãos, acaba sendo importante e fulcral levar em consideração os dois parâmetros temporais, quando se usa peças de teatro na diacronia.

Alguns estudos sincrônicos mostram que o português contemporâneo é categoricamente proclítico (cf. CARNEIRO, 2016; VIEIRA, 2016), mas desde quando essa característica aparece? Com os dados do *corpus* de Cartas e Atas da Câmara Municipal de Salvador podemos aventar que essa característica aparece já no final do século XVII, apesar de parecer “desaparecer” nos dados de textos teatrais. Aqui fica evidente que a tipologia textual é um fator importante a ser considerado quando olhamos para os dados diacrônicos.

Ribeiro (1998, p. 114), ao avaliar o clítico em início de sentença, lista uma série de dados assistemáticos sobre esse fenômeno e sugere que “só com uma busca em documentos de diferentes tipos e autores se poderá ter uma ideia do valor dessas construções”. É isso que fazemos nos gráficos 7 e 8.

Mas antes, insistimos na ideia de que, para se evitar que façamos linguística histórica de apenas uma tipologia textual, é fundamental “misturar” diferentes tipologias a fim de buscar um *corpus* “equilibrado” de textos, “representativo” o máximo possível da língua de uma época. Desse modo, é essencial fazer

uma história da língua que estuda as diferentes tradições sem limitar-se a uma só, mantendo a diferenciação – uma história da língua menos monolítica que permitirá saber em quais T[rações] D[iscursivas] uma inovação é criada, como se difunde ao longo das TD, e também onde há TD resistentes às inovações, TD que preservam elementos que em outras TD não se usam mais (KABATEK, 2006, p. 516).

Superficialmente, poderíamos afirmar que os textos teatrais por estarem mais próximos da oralidade refletiriam de modo “mais genuíno” a gramática de falantes de determinado período. Porém, observando essa tradição discursiva, vê-se que os textos teatrais parecem ser mais resistentes às inovações do que as cartas e atas. Portanto, é necessário delinear a função desse tipo textual desde o momento do chamado “descobrimento” do Brasil para compreender o porquê de as inovações não serem tão evidentes.

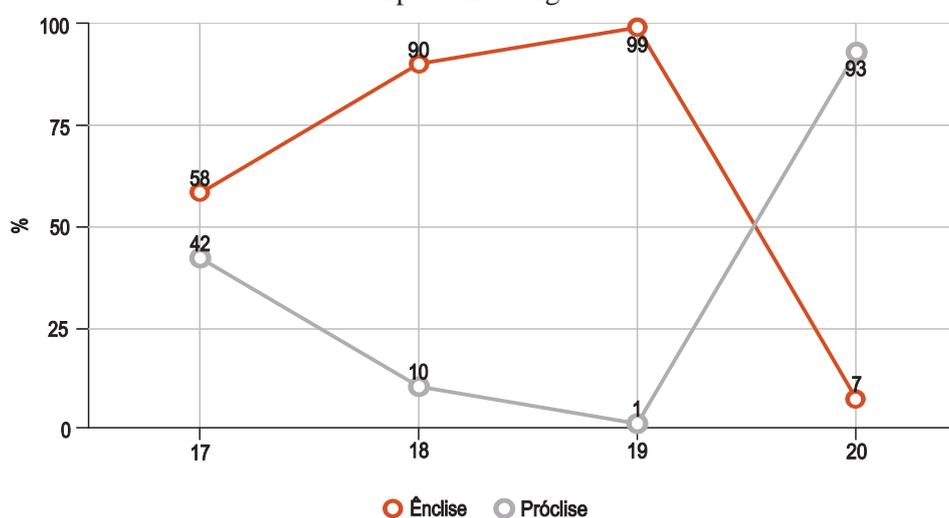
Inicialmente, as manifestações teatrais no Brasil Colônia tinham um caráter pedagógico, pois pretendiam educar religiosamente os indígenas (MAGALDI, 1997). Posteriormente, a partir do século XVIII, segundo Mayor (2015, p. 105), “[...] o teatro na colônia adquiriu mais uma função além de seus significados políticos, pedagógicos e religiosos, pois começou a se tornar também uma mercadoria”. Mayor (2015, p. 106) também considera “as produções culturais da colônia não como autônomas, mas sempre relacionadas com as portuguesas” e chama a atenção para o fato de que, nesse período,

muitas das peças apresentadas, por exemplo, eram de autoria do dramaturgo Antônio José da Silva, o Judeu, que passou toda sua curta vida em Lisboa, apresentando-se no Teatro do Bairro Alto. Por isso, o desafio do estudo dos temas coloniais é justamente investigar quais seriam as especificidades da Colônia, considerando os modelos europeus (MAYOR, 2015, pp. 106-7).

Fica evidente, desse modo, que, apesar de ser uma tipologia mais próxima da oralidade, os textos teatrais cumpriam várias funções na colônia – políticos, pedagógicos e religiosos – de sorte que seguiam muito mais de perto o modelo português do que textos de tradições discursivas mais fortemente fixadas, como as cartas e atas.

Como defende Kabatek (2006), a visão de conjunto da história da língua só poderá ficar mais tangível se olharmos para um *corpus* diacrônico multidimensional. Nesse sentido, o *Corpus Tycho Brahe* cumpre essa “missão”, visto que é resultado da reunião de diferentes tipologias textuais. Dispor de tipologias textuais variadas é fundamental para poder tirar conclusões quantitativas e qualitativas sobre o que se quer analisar. Por isso, unimos os dados de ambos os *corpora* para ter uma visão mais global dos fenômenos investigados. Começamos com os dados apresentados no gráfico 7 de clítico em posição inicial.

Gráfico 7: Contraste próclise vs. ênclise em contexto V1 por século de nascimento em ambos os corpora investigados



Fonte: elaboração dos autores

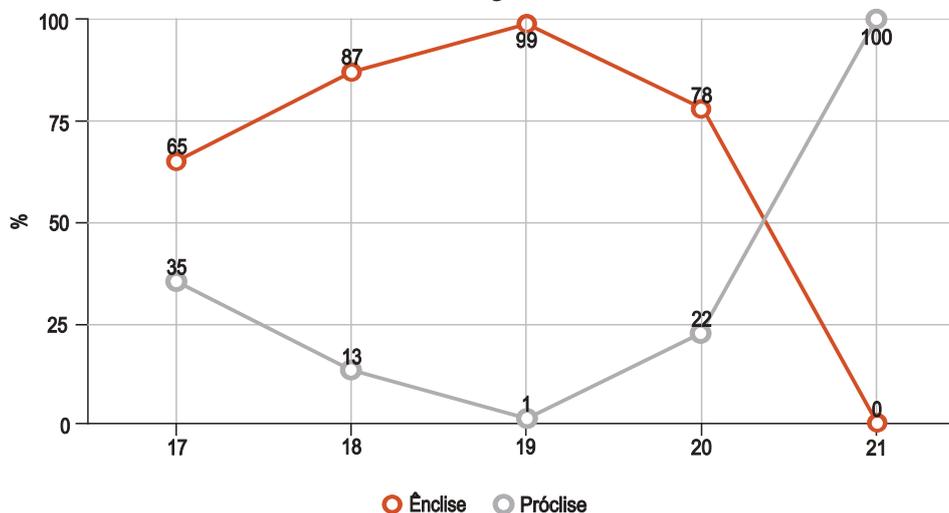
Com base nos resultados apresentados no gráfico 7, as ocorrências de próclise em orações com o verbo em primeira posição absoluta aparecem já no século XVII e somam 10% no século XVIII. Chama a atenção que tanto nos textos teatrais quanto nas cartas e atas da câmara o percentual de próclise é maior do que registrado em trabalhos anteriores (cf. PAGOTTO, 1992; LOBO, 2001; CARNEIRO, 2005; MARTINS, 2010; CARDOSO, 2020).

Entre os nascidos no século XIX, atesta-se um aumento no emprego da ênclise; e a próclise em contextos V1 é bastante marginal, ocorrendo em apenas 1% dos dados. Já entre os nascidos no século XX, o aumento nas taxas de próclise em V1 é abrupto: os valores passam de 1% para 90% dos casos.

Os padrões encontrados, quando olhamos para o século de publicação (cf. gráfico 8), diferem-se. Nos textos publicados no século XX, ainda há um alto percentual de ênclise (78%), padrão que reflete

a gramática do PE. O padrão encontrado nos textos publicados nos séculos 19 e 20 reflete uma pressão da norma culta padrão (PAGOTTO, 1998; CARNEIRO, 2005; MARTINS, 2010), o que não deixa claro quando a gramática brasileira é implementada.

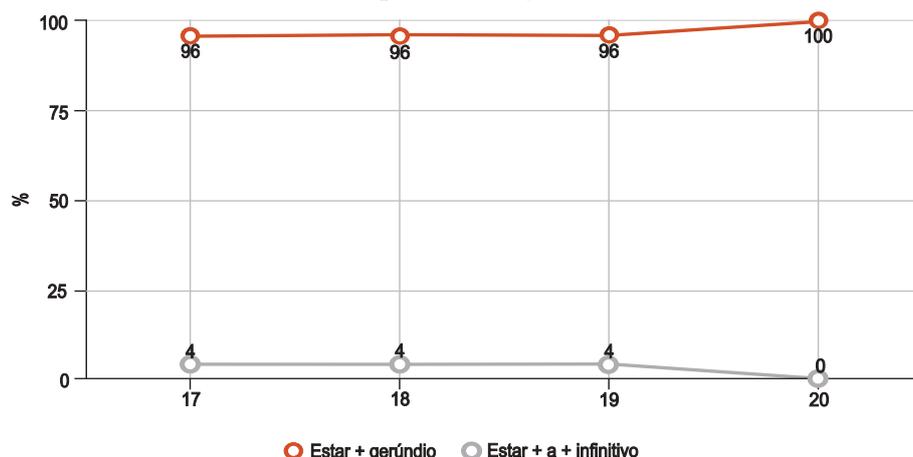
Gráfico 8: Contraste próclise vs. ênclise em contexto V1 por século de publicação em ambos os corpora investigados



Fonte: elaboração dos autores

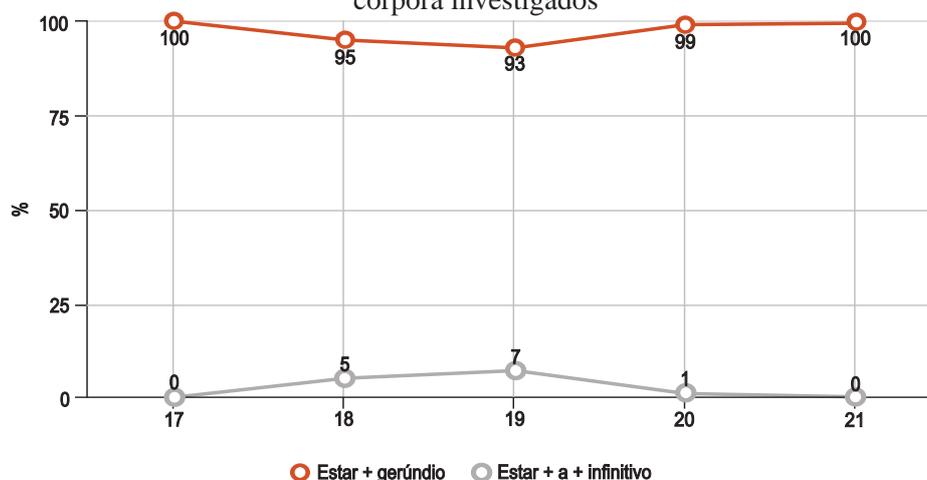
De modo a termos uma visão ampla do fenômeno da variação da estrutura progressiva, unimos os dados de ambos os corpora no gráfico 9. A partir dela, temos evidência de que, independente do tipo textual, a estrutura *estar + gerúndio* para veicular progressividade estaria presente na gramática brasileira desde o século XVII, seguindo o padrão do PCI, sendo a estrutura *estar + a + infinitivo* a inovação do PE: até o século XIX, o PE fazia uso extensivo de complemento infinitivo encabeçado pela preposição *a* com verbos como *começar*, *tornar* e *dever*, mas não com o verbo *estar*: seu complemento típico também era o gerúndio (cf. HRICSINA, 2014 para dados do PE do século XIV ao XX). É no século XIX ou ligeiramente antes que o complemento infinitivo toma lugar generalizado das perífrases verbais¹³ (HRICSINA, 2014). Nota-se pelo gráfico que alguns autores talvez tenham sido influenciados pela metrópole, porém nem mesmo a pressão da norma culta padrão lusitana foi capaz de inverter o cenário em terras tupiniquins. Diante disso, é razoável pensar que se houve alguma tentativa de implementar a estrutura *estar + a + infinitivo* no PB, essa tentativa foi fracassada.

¹³ Hricsina (2014, p. 393-395) nota que o infinitivo substituiu plenamente o complemento das perífrases, sendo usado inclusive com *ir* (e também com gerúndio), *vir* (e também com gerúndio), *continuar*, *andar*, *ficar*. Ainda nota que, muito embora em orações adverbiais reduzidas se use o gerúndio (*Chovendo, não vamos à praia*), em orações adjetivais reduzidas e orações reduzidas também a preferência é pelo infinitivo: ... *peessoas a fazer palhaçadas* (equivalente ao ... *peessoas fazendo palhaçada* do PB) e *O Pedro ficou sentado a falar com o seu pai* (equivalente ao *O Pedro ficou sentando falando com seu pai* do PB).

Gráfico 9: Contraste estar + gerúndio vs. estar + a + infinitivo por século de nascimento em ambos os corpora investigados

Fonte: elaboração dos autores

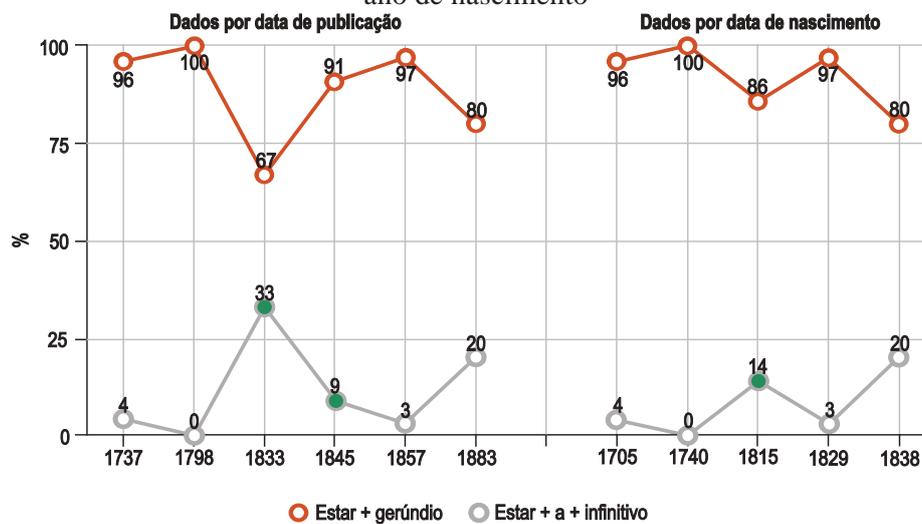
Quando a análise é baseada no século de publicação, os dados do fenômeno em questão também se diferem substancialmente. O gráfico 10 mostra esses resultados, sugerindo que talvez os textos publicados no século XIX tenham sido influenciados pela pressão normativa de Portugal, já que é nesse século que mais aparece a variante europeia nos textos publicados. Considerando os resultados de Hricsina (2014, p. 400), parece que o processo de mudança que “deve estar situado no século XIX ou ligeiramente antes (século anterior)” no PE influenciou a produção de *estar + infinitivo* nos textos publicados aqui, justamente por observamos ausência dessa estrutura nos textos publicados no século XVII, caracterizando textos puramente do PCI, mas um aumento tímido dessa estrutura exatamente quando ela emergiu em Portugal. Dessa maneira, sob o ponto de vista da geração histórica, a hipótese de que a eventual implementação de *estar + a + infinitivo* no PB teria sido uma tentativa fracassada, levantada anteriormente, parece se confirmar.

Gráfico 10: Contraste estar + gerúndio vs. estar + a + infinitivo por século de publicação em ambos os corpora investigados

Fonte: elaboração dos autores

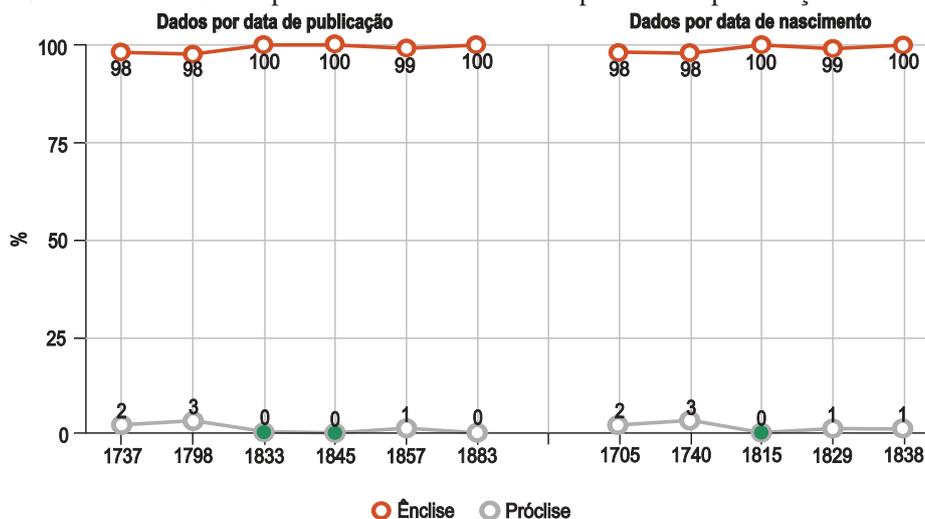
Para dar cabo à questão da *geração biológica* contrastivamente à *geração histórica*, chama atenção os resultados de Martins Pena – autor de duas peças publicadas em épocas diferentes, 1833 e 1845. Caso nos baseemos apenas na data de nascimento, por exemplo, não capturamos o fato de que, muito embora se trate do mesmo autor, há variação na gramática empregada nas peças. Isso é muito claro na comparação feita entre o uso de ênclise e próclise com a estrutura progressiva: muito embora o autor tenha a mesma gramática em relação à colocação clítica em ambas as peças (cf. gráfico 12), houve variação no uso da estrutura progressiva, com diferença de 24 pontos percentuais entre um texto e outro (cf. gráfico 11). Esse fato curioso, por exemplo, só é capturado quando a data de publicação é trazida à análise.

Gráfico 11: Contraste estar + gerúndio vs. estar + a + infinitivo em Martins Pena por ano de publicação e ano de nascimento



Fonte: elaboração dos autores

Gráfico 12: Contraste ênclise vs. próclise em Martins Pena por ano de publicação e ano de nascimento



Fonte: elaboração dos autores

Com relação à tipologia textual, os dados de progressivo mostram-nos um comportamento bastante similar entre os *corpora*. O mais interessante é que os dados de progressivo também nos informam sobre as origens do PB se recolocamos a pergunta de Ribeiro (1998, p. 115): “Pode-se falar em mudança linguística do PB tendo como parâmetro unicamente os dados do PE?”.

A resposta à pergunta acima, olhando para os dados de estrutura progressiva, é negativa. Enquanto o padrão prototípico da gramática do PE *estar + a + infinitivo* aparece apenas a partir do século XIX, segundo Hricsina (2014, p. 399), o padrão encontrado no PB é similar ao padrão do PCI. Nesse sentido, a gramática do PE é mais inovadora do que a gramática do PB e difere do padrão encontrado em outras línguas românicas, com exceção do francês; e diferente do que acontece com a colocação clítica, os dados de progressivo não parecem sofrer pressão da norma culta. Sob esse ponto de vista, ainda, podemos afirmar, consoante Corôa (2021, p. 6757), que “a mudança ocorrida em Portugal a partir do século XVIII não afeta a gramática desenvolvida no Brasil”, pelo menos não inteiramente.

Conclusões

O objetivo central deste artigo foi apresentar os novos *corpora* que passam a fazer parte do *Corpus Tycho Brahe*, quais sejam, cantiga, comédias teatrais cariocas, cartas e atas de Homens Bons da Câmara Municipal de Salvador. Apontamos que com esses novos dados podemos lançar novos olhares para a diacronia do português brasileiro.

Considerando que o objeto da linguística histórica é o texto escrito, mostramos que o linguista diacrônico, mais especificamente o que trabalha sob a perspectiva gerativista, para acessar indiretamente o conhecimento linguístico de falantes de tempos remotos, precisa tomar certas decisões antes de iniciar seu labor. Uma delas é qual tipo/gênero textual vai representar a língua do seu recorte temporal e qual parâmetro temporal vai ser adotado, a data de publicação desses tipos/gêneros textuais ou de nascimento dos seus autores. Neste artigo, demonstramos a significância de tais escolhas a partir de dois fenômenos que distinguem claramente o PE do PB: a colocação do clítico em contexto V1 e a perífrase progressiva.

Quando contrastamos dois tipos textuais, percebemos que certos fenômenos são mais sensíveis a essa escolha que outros. Por exemplo, a leitura de progressividade era categoricamente veiculada pela estrutura *estar + gerúndio* tanto nas representações artísticas quanto nas cartas e atas, o que sugere que tanto em tipologias textuais mais orais, como nas peças, quanto em tipologias textuais mais escritas, a estrutura era a mesma. No entanto, analisando a colocação clítica, observa-se que as cartas e as atas apresentam um comportamento majoritariamente proclítico desde o século XVII, enquanto as peças, um comportamento mais conservador, passando a ser majoritariamente proclítico a partir do século XIX. Sem a comparação desses novos dados, duas análises distintas poderiam emergir: a próclise não/é uma inovação do século XIX.

Adicionalmente, demonstramos que, a depender da periodização, considerar a geração biológica ocasiona uma análise da mudança mais precoce. Caso a análise seja feita a partir da geração histórica, a mudança será delineada mais tardiamente. Por exemplo, a clara mudança do padrão enclítico para o proclítico em contexto V1, tendo como base as peças de teatro, ocorreu de fato no fim do século XIX (com a geração biológica) ou no meio do século XX (com a geração histórica)?

Em suma, neste artigo tornamos públicos textos históricos ineditamente anotados (morfologicamente e morfossintaticamente) que passam a compor o *Corpus* Tycho Brahe, agora com 2.783.779 palavras, contribuindo com novos dados a fim de termos novos olhares para a história do português brasileiro. Ficou notório que apenas com novos dados é possível expandir nossas análises e, sobretudo, compreender o que está em jogo quando se escolhe por certo tipo textual ou por certa periodização temporal. Esperamos ter mostrado que tais escolhas não são isentas de consequências e que, no melhor dos mundos, o linguista diacrônico deve ter consciência das suas implicações.

Referências

ANDRADE, Aroldo; NAMIUTI, Cristina. Gone without the verb: clitic interpolation and clitic climbing in the history of European Portuguese. *Caderno de Estudos Linguísticos*, Campinas, v. 58, n. 2, pp. 201-19, 2016.

BARBOSA, Afranio. O português escrito no século XVIII: fontes reunidas na Biblioteca Nacional do Rio de Janeiro. In: CASTILHO, Ataliba (ed.). *Para a história do português brasileiro*. São Paulo: Humanitas Publicações, 1998, v. Primeiras Idéias, pp. 229-38.

CAMBRAIA, César Nardelli. A pesquisa diacrônica e o problema do *corpus*. *Anais da Semana de Estudos de Língua Portuguesa*, v. 2, n. 2, pp. 11-9, 1994.

CARDOSO, Lara. *A gramática dos pronomes clíticos no Brasil Colônia: o português clássico na história do português brasileiro*. 2020. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana, 2020.

CARNEIRO, Zenaide de Oliveira Novais. *Cartas brasileiras (1809-1904): um estudo lingüístico-filológico*. 2005. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Unicamp. Campinas, SP, 2005.

CARNEIRO, Zenaide de Oliveira Novais. Colocação de clíticos em orações finitas em duas vertentes do português oral feirense: um contexto não variável. In: ALMEIDA, Norma Lucia Fernandes de; ARAUJO, Silvana Silva de Farias; TEIXEIRA, Eliana Sandra Pitombo; CARNEIRO, Zenaide de Oliveira Novais. (org.). *Varição Linguística em Feira de Santana*. Feira de Santana: UEFS Editora, 2016, v. 1, pp. 141-74.

CORÔA, Williane. Novos elementos para a periodização do português no Brasil. *Fórum Linguístico*, v. 18, n. 3, pp. 6748-59, 2021. <https://doi.org/10.5007/1984-8412.2021.e78982>.

CYRINO, Sonia; TORRES MORAIS, Maria Aparecida. *Mudança sintática do português brasileiro: perspectiva gerativista*. História do Português Brasileiro (Coordenador geral: Ataliba de Castilho), São Paulo: Contexto, 2018. v. 6.

DUARTE, Maria Eugênia. Apresentação. In: DUARTE, Maria Eugênia (ed.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola, 2012. pp. 11-9.

GALVES, Charlotte. Posfácio: o retrato da emergência de uma nova gramática. *Mudança sintática do português brasileiro: perspectiva gerativista*. História do Português Brasileiro (Coordenador geral: Ataliba de Castilho), São Paulo: Contexto, 2018. v. 6, pp. 441-56.

GALVES, Charlotte; FARIA, Pablo. *Tycho Brahe Parsed Corpus of Historical Portuguese*. 2010. URL: <http://www.tycho.iel.unicamp.br/~tycho/corpus/en/index.html>.

GALVES, Charlotte; ANDRADE, Aroldo Leal de; FARIA, Pablo. *Tycho Brahe Parsed Corpus of Historical Portuguese*. 2017.

GARCÍA GARCÍA, L. A case study in historical linguistic research. In: BRUTON, Anthony; GARCÍA, Luisa García; DOMÍNGUEZ, Joaquín José Fernández (orgs). *Perspectives on the genitive in english: Synchronic, Diachronic, Contrastive and research*. Universidad de Sevilla: 2000, pp. 118-29.

HRICSINA, Jan. Substituição do gerúndio pela construção a + infinitivo no Português Europeu (estudo diacrônico). *Studia Iberystyczne*, v. 13, pp. 383-401, 2014.

HUNSTON, S. *Corpora in applied linguistics*. Cambridge: Cambridge University Press, 2002.

JENSET, Gard B.; MCGILLIVRAY, Barbara. *Quantitative historical linguistics: A corpus framework*. Oxford University Press, 2017.

KABATEK, Johannes. Tradições discursivas e mudança linguística. In: LOBO, Tânia; *et al.* (orgs.). *Para a história do português brasileiro*. Salvador: EDUFBA, 2006, v. 6, t. 1-2. pp. 505-27.

LIGHTFOOT, David. *Principles of diachronic syntax*. Cambridge, UK: Cambridge University Press, 1979.

LOBO, Maria. *Aspectos da Sintaxe das Orações Subordinadas Adverbiais do Português*. 2003. Dissertação de Doutorado em Linguística/Sintaxe – Universidade Nova de Lisboa, Lisboa, 2003.

LOBO, Tânia. Depoimento sobre a constituição de um *corpus* diacrônico do português brasileiro – Bahia. In: CASTILHO, Ataliba (ed.). *Para a história do português brasileiro*. São Paulo: Humanitas Publicações, 1998. vol. Primeiras Idéias, pp. 171-95.

LOBO, Tânia. *Para uma sociolinguística histórica do português no Brasil*: edição filológica e análise linguística de cartas particulares do Recôncavo da Bahia, século XIX. 2001. 808f. Tese (Doutorado em Filologia e Língua Portuguesa) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo. São Paulo, 2001.

MAGALDI, S. *Panorama do teatro brasileiro*. São Paulo: Global, 1997.

MARTINS, Marco Antônio. *Competição de gramáticas do português na escrita catarinense dos séculos 19 e 20*. 2009. 326 f. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2009.

MATTOS E SILVA, Rosa Virgínia. Notícia sobre o Programa para a História da Língua Portuguesa. *Estudos Linguísticos e Literários*, Salvador, número especial, pp. 231-7, 1996.

MAYOR, Mariana Soutto. O teatro do século XVIII no Brasil: das festas públicas às casas de ópera. *Revista Aspas*, v. 5, n. 2, pp. 103-10, 2015.

- MIRA MATEUS, Maria Helena et al. *Gramática da Língua Portuguesa*, revista e aumentada 5. ed. Lisboa: Caminho, 2003.
- NEVES, Larissa de Oliveira. *As Comédias de Artur Azevedo – Em Busca da História*. 2006. 212 f. Tese (Doutorado em Letras) – Instituto de Estudos da Linguagem, Unicamp, Campinas, 2006.
- PAGOTTO, Emílio. *A posição dos clíticos em Português: um estudo diacrônico*. 1992. 168 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Unicamp. Campinas, SP, 1992.
- PAGOTTO, Emílio. Norma e Condescendência: Ciência e Pureza. *Língua e Instrumentos Linguísticos*, Campinas, v. 2, n. 1, pp.49-68, jul./dez. 1998.
- PAIXÃO DE SOUSA, Maria Clara. *Língua barroca: Sintaxe e história do português nos 1600*. 2004. 455 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Unicamp. Campinas, SP, 2004.
- RIBEIRO, Ilza. A mudança sintática do Português Brasileiro é uma mudança em relação a que gramática? In: CASTILHO, Ataliba T. de. (org.). *Para a história do português brasileiro: primeiras ideias*. V.1. São Paulo: Humanitas, 1998. pp.101-9.
- SARDINHA, Tony Berber. *Linguística de corpus*. Barueri: Manole, 2004.
- SILVA NETO, Serafim da. *Introdução ao estudo da língua portuguesa no Brasil*. Rio de Janeiro: Presença, 1977.
- VIEIRA, Maria de Fátima. *A ordem dos clíticos pronominais nas variedades urbanas europeia, brasileira e são-tomense: uma análise sociolinguística do português no início do século XXI*. 2016. 238 f. Tese (Doutorado) - Universidade Federal do Rio de Janeiro, Faculdade de Letras, 2016.

UMA ANÁLISE DIACRÔNICA DAS FORMAS PRONOMINAIS *CONTIGO* E *COM VOCÊ* EM CARTAS PESSOAIS

A DIACHRONIC ANALYSIS OF PRONOMINAL FORMS “*CONTIGO*” AND “*COM VOCÊ*” IN PERSONAL LETTERS

Thaissa Frota Teixeira de Araujo Silva¹

RESUMO

O presente trabalho apresenta uma análise das variantes oblíquas de 2SG com destaque para a alternância *contigo* ~ *com você* em cartas pessoais fluminenses datadas entre fins do século XIX e decorrer do século XX. Levamos em conta os princípios da sociolinguística histórica (HERNÁNDEZ-CAMPOY; CONDE SILVESTRE, 2012) para análise de um *corpus* diacrônico. Os resultados evidenciam que a alternância *contigo* ~ *com você* se coloca de maneira mais equilibrada que a oposição *prep. + ti* ~ *prep. + você*. Além disso, a forma *contigo* é predominante entre missivas trocadas entre familiares e há indícios de que seja, no terceiro período de tempo analisado, mais resistente a entradas das formas de *você* do que as demais estratégias oblíquas do paradigma de *tu*.

PALAVRAS-CHAVE: Pronomes de segunda pessoa. Tratamento. Oblíquo.

ABSTRACT

The paper aims to analyze the oblique alternating forms of 2nd person singular with special focus to the variation *contigo* ~ *com você* in private letters from Rio de Janeiro written from the end of 19th century and during the 20th century. As theoretical-methodological apparatus, we consider the assumptions of the historical sociolinguistics (HERNÁNDEZ-CAMPOY & CONDE SILVESTRE, 2012) to the analyses of a diachronic corpus. According to the results, the alternation *contigo* ~ *com você* is more balanced than the opposition *prep. + ti* ~ *prep. + você*. Furthermore, the form *contigo* is predominant among missives exchanged between family members and there are indications that the form is, in the third time period analyzed, more resistant to entries from the oblique constructions of *você* than the other oblique strategies of the *tu* paradigm.

KEYWORDS: 2nd person's pronouns. Address. Oblique.

Introdução

No presente trabalho, realizamos uma análise diacrônica com relação às formas oblíquas de segunda pessoa do singular (doravante 2SG) destacando a alternância *contigo* ~ *com você* das demais estratégias oblíquas preposicionadas. Nosso objetivo é verificar diferenças no comportamento entre a variante *contigo* e as construções preposicionadas com *ti* tônico e, para isso, utilizamos como base a *Sociolinguística Histórica* (HERNÁNDEZ-CAMPOY; CONDE SILVESTRE, 2012).

Nosso *corpus* é constituído por cartas pessoais produzidas entre fins do século XIX e durante o século XX no Rio de Janeiro. Essa amostra já possibilitou diversos estudos a respeito da variação *tu* ~ *você* em diferentes contextos morfosintáticos, como posição de sujeito (cf. RUMEU, 2008; SOUZA, 2012); formas acusativas (cf. SOUZA, 2014); dativas (cf. OLIVEIRA, 2014), genitivas (possessivas) (cf. PEREIRA, 2016) e oblíquas (cf. SILVA, 2020).

¹ Doutoranda do Programa de Pós-graduação em Letras Vernáculas da Universidade Federal do Rio de Janeiro, thaissafrota@letras.ufrj.br, <https://orcid.org/0000-0001-7295-8507>.

Na presente análise, trataremos diretamente dos dados levantados por Silva (2020) apresentando um novo olhar para a distribuição das ocorrências ao separar a alternância *contigo* ~ *com você* da variação entre as demais formas oblíquas de 2SG – *preposição + ti* ~ *preposição + você*. Dessa forma, pretendemos observar: (i) possíveis diferenças no comportamento entre *contigo* e *prep.+ti*; (ii) como variáveis independentes linguísticas e extralinguísticas atuam na alternância *contigo* ~ *com você* no período analisado; e (iii) uma possível resistência da forma *contigo* à implementação das estratégias oblíquas de 2SG do paradigma de *você*.

Formulamos como hipótese que (i) há diferença entre *contigo* e *prep.+ti*, uma vez que se observa um maior equilíbrio entre as frequências na alternância *contigo* e *com você* do que é observado para a oposição *prep.+ti* ~ *prep.+você*; (ii) que as variáveis independentes linguísticas e extralinguísticas atuam de maneira equilibrada na alternância *contigo* ~ *com você*; (iii) e que a forma *contigo* apresente uma certa resistência à implementação de estratégias oblíquas com *prep.+você*.

O presente artigo encontra-se estruturado da seguinte maneira: na introdução, apresentamos o objetivo principal da análise e as questões norteadoras do estudo. Em seguida, descrevemos o objeto de estudo: as formas oblíquas de 2SG, com destaque para alternância *contigo*~*com você*. Na sequência, definimos os pressupostos teórico-metodológicos que fundamentaram a análise e o corpus. Discutimos os principais resultados obtidos, comparando os percentuais referentes à variação *contigo* ~ *com você* e à variação *prep.+ti* ~ *prep.+você* na seção subsequente. Por fim, reunimos como considerações finais as principais contribuições do presente trabalho.

1. Descrevendo o objeto de estudo: *contigo* x *com você*, formas oblíquas de 2SG

Uma das primeiras questões se coloca para o estudo das estratégias oblíquas de 2SG é a definição do objeto de estudo. Isso se deve ao fato de a classificação/terminologia dessas formas não se apresentar de maneira consensual nem entre gramáticos tradicionais e nem entre os linguistas que descrevem e/ou discutem o fenômeno (cf. MATEUS *et al.*, 2003; RAPOSO *et al.*, 2013; CASTILHO, 2014)

Dentro de uma perspectiva tradicional, Cunha e Cintra (2001 [1985]) não fazem distinção entre complementos preposicionados dativos e oblíquos, com esses sintagmas sendo analisados como *objetos indiretos*. Rocha Lima (2006 [1972]), por outro lado, diferencia o *objeto indireto* do *complemento relativo* e do *complemento circunstancial*. Mateus *et al* (2003) apresenta a *relação gramatical oblíqua* como uma relação não-central que poderia funcionar como argumento ou adjunto do predicador verbal, enquanto Raposo *et al* (2013) e Castilho (2014) referem-se apenas a *complementos oblíquos*. Além disso, essas descrições e distinções só surgem no que tange aos complementos e adjuntos de predicadores verbais. Os autores não fazem referência a predicadores nominais (Estou com o pensamento *em ti*) ou adverbiais (*estarei sempre junto a ti*).

No presente estudo, consideramos, seguindo a proposta de Silva (2020), que as formas oblíquas de 2SG sejam constituintes obrigatoriamente preposicionados que não podem ser substituídos pelo clítico *lhe* (Eu sonho *com você/contigo*; *Eu *lhe* sonho). Esses constituintes podem estar ligados tanto a predicadores verbais quanto não-verbais, funcionando como complementos ou adjuntos. Além disso, podem ser introduzidos por diversas preposições. Listamos a seguir alguns exemplos:

- (1) Confio *em ti/em você*.
- (2) A saudade *de ti/ de você* só cresce.
- (3) Vou guardar os livros *para ti/ para você*.
- (4) Adoro sonhar *contigo/ com você*.

Como é possível observar, as formas oblíquas de 2SG podem ser divididas em dois paradigmas distintos: o paradigma de *tu* e o de *você*. O primeiro é constituído por *ti* em estruturas preposicionadas e pela forma morfologizada *contigo*. Tais construções têm origem no quadro pronominal latino e formam, portanto, o paradigma mais antigo e original. O segundo é formado por sintagmas preposicionados que apresentam como núcleo a forma *você*, resultado da gramaticalização de *vossa mercê*.

2. A análise de formas de 2SG na escrita epistolar fluminense: uma breve retomada de estudos diacrônicos

A análise de cartas pessoais do Rio de Janeiro possibilitou estudos sobre formas de 2SG em diversos contextos. Lopes e Cavalcante (2011) analisaram missivas do fim do século XIX e do início do século XX e apresentaram resultados para a função de sujeito e de complemento (acusativo, dativo e oblíquo).

Com relação ao acusativo, as autoras observaram uma grande produtividade do clítico *te*, com 89% de frequência. Além disso, as frequências para o clítico se mantiveram altas independentemente das formas pronominais de 2SG utilizadas em posição de sujeito. Com relação ao dativo, as autoras observaram maior diversidade entre as variantes, com maior produtividade no uso do clítico *te*, com 58,9% de frequência. O dativo nulo, foi a segunda variante mais produtiva com 20,7% de frequência. Traçando uma correlação com as formas utilizadas na posição de sujeito, o clítico *te* foi mais produtivo entre as cartas com uso de *tu exclusivo* ou *misto* (alternância *tu* e *você*). Entre as missivas com uso de *você exclusivo* foi percebido um maior equilíbrio entre as variantes, com predomínio do dativo nulo (30%).

No que diz respeito aos complementos oblíquos, Lopes e Cavalcante (2011) observaram que não eram tão frequentes quanto os demais, mas se diferenciavam destes por apresentarem uma maior correlação com a forma pronominal utilizada na posição de sujeito. Em missivas com uso de *você exclusivo*, as autoras apenas registraram estratégias preposicionadas com *você*. Por outro lado, em missivas com uso de *tu exclusivo*, houve predomínio de estratégias oblíquas do paradigma de *tu* (78% de *prep.+ti* e 11% de *contigo*). Em missivas de uso *misto*, as autoras encontraram maior frequência de *prep.+você* (57%).

A partir desse trabalho de Lopes e Cavalcante (2011), novos estudos surgiram com foco em diferentes contextos morfossintáticos para as formas de segunda pessoa. No que se refere ao estudo da variação *tu* e *você* na posição de sujeito, podemos citar o trabalho de Souza (2012), que, a partir

de um *corpus* de 354 cartas cariocas e fluminenses entre 1870 e 1970, realizou um mapeamento da entrada do pronome *você* no quadro pronominal brasileiro. Com base na análise dos 1381 dados obtidos, a autora encontrou a seguinte distribuição das formas ao longo do tempo: o pronome *tu* era mais utilizado nas primeiras décadas analisadas (1870-1890); no início do século XX, as frequências das formas *tu* e *você* são equilibradas; a partir da década de 1930, a nova estratégia suplanta a anterior.

Um estudo para as formas acusativas foi realizado por Souza (2014). A autora constituiu sua análise também com base em um *corpus* de cartas pessoais, produzidas entre 1884 e 1985 por indivíduos oriundos do Rio de Janeiro. Foram selecionados como variantes os clíticos *te*, *lhe* e *o/a*, o pronome lexical *você* e o objeto nulo. A autora observou que o clítico *te* era a estratégia mais frequente (77,8%); em segundo lugar, o clítico *o/a* (9,2%), seguido pelo pronome lexical *você* (6,7%) e pelo clítico *lhe* (4%). O objeto nulo foi a variante menos frequente (2,3%). Sendo assim, foi possível afirmar que a forma acusativa é predominantemente realizada por clíticos.

Souza (2014) constatou que o acusativo seria um contexto de resistência à inserção da forma inovadora de 2P, devido à alta produtividade do clítico *te*, independentemente das formas utilizadas na posição de sujeito (*tu* e/ou *você*). Além disso, também foi a estratégia mais utilizada em todos os tempos e modos verbais, nas missivas familiares e amorosas, nas diversas partes da carta, por ambos os gêneros (masculino e feminino), em quase todas as famílias e ao longo de todo o período analisado. Esses resultados apontaram para uma generalização de *te* para referência da 2P no acusativo.

Uma análise para as formas dativas foi realizada por Oliveira (2014), baseada em um *corpus* formado por cartas pessoais fluminenses e cariocas do período de 1880 a 1980. As formas variantes dativas analisadas foram os clíticos *te* e *lhe*, os sintagmas preposicionados *a ti*, *para ti*, *a você*, *para você* e objeto nulo. Os resultados do autor apontaram que mais da metade dos dados correspondia ao clítico *te* (57,2%), enquanto a segunda estratégia mais utilizada foi o objeto nulo (22,3%) e a terceira, o clítico *lhe* (11,3%). Quanto às variantes preposicionadas, nenhuma atingiu 5% do total da amostra.

Traçando uma correlação com as formas utilizadas na posição de sujeito, observou-se que tanto com sujeito exclusivo *tu* e como em casos de alternância *tu ~ você* houve uma alta frequência do clítico *te* (78,5% e 70,2%, respectivamente). Quando havia *você* exclusivo, constatou-se um maior equilíbrio entre as formas, *te* (30,9%), *zero* (29,1%) e *lhe* (27,3%). Oliveira (2014) observou, também, que o dativo, ao contrário do acusativo, favorece o uso do objeto nulo. Além disso, estas formas “sofrem uma elevação de frequência notável após os anos 1930, ou seja, no período indicado pelos estudiosos em que a forma *você* começa a ser utilizada em textos escritos com estatuto de pronome pessoal” (OLIVEIRA, 2014, p. 158).

Silva (2020) analisou as formas oblíquas de 2SG em cartas pessoais produzidas aproximadamente entre 1870 e 1980 no Rio de Janeiro e constatou que – diferentemente do que havia sido observado com relação ao acusativo e ao dativo, as formas oblíquas de 2SG apresentavam uma forte correlação com a estratégia pronominal de 2SG utilizada na posição de sujeito. Além disso, a autora destacou que as estratégias do paradigma de *tu* foram amplamente substituídas pelas estratégias de *preposição + você* a partir da década de 1940, seguindo um padrão de implementação similar ao que foi postulado por Souza (2012) para a posição de sujeito.

Dessa forma, Silva (2020) delimitou as seguintes fases para a implementação das formas oblíquas de 2SG:

- Fase 1 (1870 – 1899): predomínio do paradigma de *tu*
- Fase 2 (1900 – 1939): maior equilíbrio entre os paradigmas
- Fase 3 (1940 – 1989): predomínio do paradigma de *você*

No último período de tempo analisado, a autora verificou que a maior parte das ocorrências do paradigma de *tu* não eram referentes às estratégias com *ti* preposicionado, mas sim da forma morfologizada *contigo*. Sendo assim, trazemos para o centro da presente análise a alternância entre as formas pronominais *contigo* e *com você* – exemplos de formas oblíquas de 2SG introduzidas pela preposição *com*.

É relevante destacar que a forma *contigo* é a única estratégia pronominal oblíqua de 2SG não expressa por um sintagma preposicionado. Com relação a sua formação, Câmara Jr. (1979, p.97) aponta que a forma tem origem no ablativo latino *te* aglutinado de forma posposta à preposição *cum* (*tecum*). Devido a desgastes fonéticos, *tecum* teria resultado em *tigo* e a noção da preposição foi sendo perdida. Ocorreu, portanto, a aglutinação da preposição “com” ao radical “tigo”, formando “contigo”.

3. Pressupostos teórico-metodológicos

Para nortear a análise realizada no presente artigo, utilizamos como base os fundamentos teóricos Sociolinguística Histórica (HERNÁNDEZ-CAMPOY; CONDE SILVESTRE, 2012), que se apresenta como subcampo híbrido e interdisciplinar relacionado em diferentes graus com diversos campos do conhecimento, como a filologia, paleografia, história, sociopragmática, linguística de corpus, entre outros (LOPES; RUMEU, 2018, p. 164). Seus objetivos se dedicam à compreensão de fenômenos variáveis ao longo do tempo e, dessa maneira, compartilha os ideais postulados por Weinreich, Labov & Herzog (1968) ao tratar a língua como um sistema heterogeneamente ordenado.

Assim, concebemos a análise as formas oblíquas de 2SG como *fenômeno variável*. A alternância entre as formas do paradigma de *tu* e de *você* acontece nos mesmos contextos sem alterações semânticas, mas condicionada por fatores de natureza linguística e extralinguística. Para observar a variação ao longo do tempo se faz necessário recorrer a registros linguísticos de períodos anteriores, o que traz alguns problemas para o trabalho do pesquisador.

Hernández Campoy & Schilling (2012) expõem alguns desses problemas, dentre os quais destacamos a questão da *representatividade* e a da *validade empírica*. Com relação à primeira, os autores pontuam que a preservação dos dados históricos ocorre de maneira aleatória, não sendo possível remodelar os conjuntos de dados disponíveis. Além disso, observamos uma grande dificuldade em manter uma amostra equilibrada durante os períodos de tempo analisados.

Com relação à *validade empírica*, os autores destacam que a *Sociolinguística Histórica* trabalha com coleções de textos cujas dimensões são inevitavelmente limitadas, o que restringe os resultados da análise quantitativa e as medidas estatísticas associadas. Essa limitação atinge quantidade, tipo e

informações socioculturais. Portanto, a validade empírica de um corpus histórico não pode ser avaliada pelos mesmos parâmetros da sociolinguística sincrônica que lida com amostras contemporâneas.

Para constituir a presente análise sob uma perspectiva histórica, optamos pelo uso de cartas pessoais para a composição do *corpus*, uma vez que formas pronominais de 2SG pressupõem a necessidade de um interlocutor. Sendo assim, consideramos que tal gênero seja propício para o estudo desse fenômeno em sincronias passadas. Na tabela a seguir, apresentamos os acervos que foram utilizados para levantamento de dados:

Tabela 1: A amostra analisada

Acervo	Número de documentos
Acervo C.A. (1857 – 1859)	11
Acervo Cup. (1873 – 1895)	23
Acervo P.F.M. (1876 – 1947)	170
Acervo Ot. (1879 – 1889)	37
Acervo R.B. (1886 – 1907)	29
Acervo O.C. (1889 – 1915)	29
Acervo A.P. (1896 – 1926)	30
Acervo L.A. (1907 – 1917)	40
Acervo R. S. (1908)	12
Acervo Br. (1933 – 1972)	47
Acervo J. e M. (1936 – 1937)	93
Acervo F.B. (1956 – 1994)	115
Acervo Lac. (1978 – 1979)	80
Total	716

Fonte: elaboração da autora

O termo acervo corresponde ao conjunto de cartas de uma determinada família, aproveitando assim a designação que consta nos arquivos públicos e/ou nos arquivos digitais de onde o material foi retirado. Além disso, optamos por manter o nome dos acervos abreviados, mantendo em sigilo a identidade dos missivistas.

Trata-se de uma amostra bastante diversificada quanto a origem dos remetentes e relações interpessoais observadas entre os missivistas. Os acervos C.A., Cup., P.F.M., Ot., O.C., A.P. e L.A. apresentam remetentes ilustres e os missivistas, em sua maioria, possuem formação acadêmica, religiosa e até mesmo participação na política e na literatura. Os acervos R.S. e J. e M., por outro lado, não constituem amostras ilustres. O primeiro é formado por bilhetes amorosos extraídos de um processo judicial que investigou o assassinato de um amante cometido pelo companheiro da remetente. O segundo é composto pela troca de missivistas entre noivos nos anos 30. Todas as informações sobre os remetentes foram extraídas das próprias cartas. Todos esses acervos fazem parte de subprojetos vinculados ao Projeto Nacional Para uma História do Português Brasileiro – PHPB e estão disponíveis atualmente no site <http://www.histling.lettras.ufrj.br>.

Dentre os acervos não disponíveis no site, foram analisados os acervos ilustres R.B. – cuja edição está disponível em Callou e Barbosa (2011) – e Br. – editado por Janaina Pedreira de Souza (cf. SOUZA, 2012). Foram analisados também os acervos não-ilustres F.B. e Lac. – editados por bolsistas de Iniciação Científica e que se encontram em fase de preparação para serem disponibilizados no site do Laboratório *Histling*. Esses acervos não-ilustres são constituídos por missivas de cunho familiar. Com relação ao sexo dos missivistas, foram levantadas 349 missivas masculinas e 367 missivas femininas.

A distribuição de documentos por acervos já nos indica as dificuldades mencionadas anteriormente com relação a *representatividade* do *corpus*. No entanto, acreditamos que a análise de processos variáveis em amostras como essas podem fornecer, apesar dos problemas apontados, relevantes considerações e informações a respeito de como fenômenos linguísticos e variáveis se deram no decorrer do tempo.

4. Análise de resultados

No *corpus* analisado, foram levantadas 361 ocorrências de formas oblíquas de 2SG, sendo 220 referentes ao paradigma de *tu* (61%) e 141 ao paradigma de *você* (39%). Como objetivo do presente trabalho é direcionar o olhar para a forma *contigo*, optamos por analisar esses resultados gerais destacando as formas *contigo* e *com você* dos demais sintagmas preposicionados.

Foram levantadas 104 ocorrências de estratégias oblíquas de 2SG introduzidas pela preposição *com*, sendo 57 de *contigo* (54,8%) e 47 de *com você* (45,2%). Com relação às demais preposições, foram registradas 257 ocorrências, sendo 163 de *prep.+ti* (63,4%) e 94 de *prep.+você* (36,6%). O gráfico a seguir apresenta essa distribuição:

Gráfico 1: Distribuição das formas oblíquas de 2SG em cartas pessoais



Fonte: elaboração da autora

Com base na análise das frequências absolutas apresentadas no gráfico anterior, podemos considerar que a variação se mostra de maneira mais equilibrada entre as formas *contigo* ~ *com você* do que o observado em sintagmas preposicionados não iniciados pela preposição *com*. Tendo em vista o objetivo de compreender um comportamento diferenciado de *contigo* em relação às demais estratégias do paradigma de *tu* (*prep. + ti*), observamos como diferentes variáveis atuam na variação *contigo* ~ *com você* em oposição à alternância *prep. + ti* ~ *prep. + você*.

Os dados levantados foram codificados de acordo com seis variáveis independentes de natureza linguística e extralinguística e submetidos ao programa Goldvarb-X². Com relação às primeiras, foram controladas (i) a forma pronominal de 2SG utilizada na posição de sujeito (exclusivo *tu*, exclusivo *você*, alternância *tu* ~ *você*, sem referência de 2SG); (ii) o tipo de núcleo predicador ao qual se conecta a forma oblíqua (verbal; nominal; verbo suporte; adverbial); e (iii) o tipo de relação entre a forma oblíqua o seu núcleo (complementação, adjunção, estrutura predicativa). Quanto aos fatores extralinguísticos, foram controlados (iv) o sexo do remetente das missivas (homem, mulher); (v) relação interpessoal entre os missivistas (familiar, amorosa, pessoal); e (vi) a distribuição das missivas de acordo com as fases postuladas por Silva (2020). Nas subseções a seguir, detalhamos as considerações referentes a cada variável independente seguindo a ordem em que foram apresentadas.

4.1. A forma pronominal utilizada na posição de sujeito

Estudos anteriores, como Lopes e Cavalcante (2011), Lopes *et al.* (2018) e Silva (2020), apontam para uma forte correlação entre a forma pronominal de 2SG utilizada na posição de sujeito e a forma oblíqua de 2SG. Dessa forma, quando, em uma carta, havia na posição de sujeito uso exclusivo de *tu* era observado também uma maior predominância das estratégias oblíquas de paradigma de *tu*. Por outro lado, com uso exclusivo de *você* na posição de sujeito, havia maior predominância de estratégias oblíquas de *prep. + você*. Em casos de alternância na posição de sujeito de 2SG numa mesma carta, a variação entre as formas oblíquas de 2SG se equilibrava. Seguindo os fatores condicionantes controlados em Silva (2020), também foram analisados casos de missivas que não apresentavam nenhuma forma pronominal de 2SG na posição de sujeito sob o código *sem referência*.

Com base nos resultados de Silva (2020), formulamos como hipótese que a forma pronominal utilizada na posição de sujeito seria bastante atuante na alternância *prep. + ti* ~ *prep. + você*. Por outro lado, na presente análise, consideramos que para a oposição *contigo* ~ *com você* haveria frequências mais equilibradas. Apresentamos, na tabela a seguir, os resultados obtidos destacando a alternância *contigo* ~ *com você* da alternância entre demais sintagmas preposicionados oblíquos de 2SG:

² Devido à natureza histórica dos dados, havia muitas células vazias (*knock-outs*) para o estudo das variantes *contigo* ~ *com você*, o que impossibilitou a análise de pesos relativos. Dessa maneira, optamos por realizar o estudo com base em percentuais.

Tabela 2: Forma pronominal de 2SG utilizada na posição de sujeito (valor de aplicação *contigo* / *prep.+ti*)

Posição do sujeito	<i>Contigo</i>		<i>Prep.+ti</i>	
	N/T	%	N/T	%
Exclusivo <i>tu</i>	30/31	93,8%	115/128	89,8%
Exclusivo <i>você</i>	5/24	20,8%	14/66	21,2%
<i>Tu ~ você</i>	17/40	42,5%	31/57	54,4%
Sem referência	5/8	62,5%	3/6	50,0%

Fonte: elaboração da autora

Com base nos resultados obtidos, podemos considerar que, tanto na alternância *contigo ~ com você* quanto na alternância *prep.+ti ~ prep.+você*, o efeito da variável independente *forma pronominal utilizada na posição de sujeito* é similar, contrariando nossa hipótese. O uso de *contigo* e de *prep.+ti* é predominante em cartas com uso exclusivo *tu* na posição de sujeito com 93,8% e 89,8% respectivamente. Em casos de uso exclusivo *você*, as frequências caem drasticamente para 20,8% e 21,2%. Quando há alternância *tu ~ você* na posição de sujeito e em cartas sem referência pronominal de 2SG, as frequências se apresentam de forma mais equilibrada. Dessa forma, com relação ao grupo de fatores analisados, não se observa diferença relevante entre a forma *contigo* e as demais estratégias do paradigma de *tu*.

4.2. O tipo de núcleo

Como foi mencionado na seção 1 do presente artigo, as formas oblíquas de 2SG podem estar relacionadas tanto a predicadores verbais quanto não verbais. Dessa forma, os dados foram codificados de acordo com o tipo de núcleo ao qual a estratégia oblíqua se conecta. Listamos a seguir exemplos das formas alternantes ligadas a predicadores verbais (5) a (8), a predicadores nominais (9) a (12), a verbos suporte (13) a (16) e a predicadores adverbiais (17) e (18).

- (5) Minha querida quando sonho **contigo** não quisera mais acordar (1936 – Acervo J. e M.)
- (6) ela disse que eu podia falar **com você** (1937 – Acervo J. e M.)
- (7) Meu irmão, quanto penso **em ti** e desejo ver-te padre! (1919 – Acervo P.F.M)
- (8) por que você é meu só quem manda **em você** sou eu (1908 – Acervo R. S.)
- (9) Sinharinha está muito queixosa **contigo** (1891 – Acervo O.C.)
- (10) que os teus pais estejam mais calmos **com você** (1937 – Acervo J. e M.)
- (11) eu morro de ciúmes **de ti** (1908 – Acervo R. S.)
- (12) **Para** Yolanda e **você** um grande abraco de Maria (1972 – Acervo Br.)
- (13) Minha flor esta noite tive um sonho tão lindo **contigo** (1936 – Acervo J. e M.)

- (14) Este anno espero ter uma cor-respondencia mais regular **com Você** (1933 – Acervo P.F.M)
- (15) tenho tanta confiança em ti (1937 – Acervo J. e M.)
- (16) eu tive um sonho com você eu depois te conto (1936 – Acervo J. e M.)
- (17) **longe de ti** minha flor, sou um pobre sonhador que vive a sonhar a história do nosso amor!
(1936 – Acervo J. e M.)
- (18) lembrando-me do domingo a noite que eu estava junto **de você** (193 – Acervo J. e M.)

Segundo nossa hipótese, tendo em vista os resultados de Silva (2020), as formas de *preposição + ti* seriam mais frequentes com predicados verbais, enquanto *contigo* apresentaria frequências equilibradas independentemente do tipo de núcleo analisado. Na tabela a seguir, apresentamos os resultados obtidos destacando a alternância *contigo ~ com você* da alternância entre demais sintagmas preposicionados oblíquos de 2SG:

Tabela 3: Tipo de núcleo predicador (valor de aplicação *contigo / prep.+ti*)

Tipo de núcleo	<i>Contigo</i>		<i>Prep.+ti</i>	
	N/T	%	N/T	%
Verbal	46/83	55,4%	100/155	64,5%
Nominal	6/13	46,2%	31/68	45,6%
Verbo suporte	5/8	62,5%	1/2	50%
Adverbial	-	-	31/32	96,9%

Fonte: elaboração da autora

Com relação ao tipo de núcleo, destacamos que os predicadores adverbiais não apresentaram uso de *contigo* ou *com você*. Além disso, a respeito da alta frequência de uso de *prep.+ti* com esses predicadores é relevante destacar que se trata de uma estrutura pouco diversificada – há a repetição de apenas três advérbios (*longe, perto, junto*).

Acerca dos predicadores nominais, o comportamento de *contigo* e de *prep.+ti* se apresenta de forma similar, com percentual de 46,2% para *contigo* e 45,6% para *prep.+ti*. Por outro lado, na análise de predicadores *verbais* e de estruturas com *verbos suporte*, há uma leve diferença no comportamento das frequências de *contigo* e *prep.+ti*. Enquanto *contigo* apresenta maior percentual em estruturas com verbo suporte (62,5%), *prep.+ti* alcança maior percentual com predicadores verbais (64,5%). É importante destacar que há poucos dados de formas oblíquas ligadas a verbos suporte e que, na maior parte das ocorrências, a preposição *com* é necessária. No entanto, podemos considerar que entre essas estruturas a forma *contigo* é predominante.

Dessa forma, consideramos que, a análise dos percentuais referentes a essa variável independente, não nos permite traçar diferenças significativas entre a variante *contigo* e as demais estruturas do paradigma de *tu*.

4.3. A natureza da relação sintática

Conforme foi apontado na seção 1 do presente artigo, as formas oblíquas de 2SG podem estar relacionadas aos predicadores funcionando tanto como complementos quanto como adjuntos. Em consonância com Mateus *et al.* (2003) e Raposo *et al.* (2013), as estruturas foram classificadas como casos de *complementação* quando a forma oblíqua de 2SG era selecionada obrigatoriamente pelo predicador verbal ou não-verbal e como casos *adjunção* quando a seleção não fosse obrigatória, ou seja, não essencial para a concepção do sentido do predicador. Além disso, optamos por analisar construções com verbos copulativos como casos de estrutura predicativa, na qual a forma oblíqua desempenha a função de predicativo do sujeito.

Listamos a seguir exemplos das variantes em estruturas de complementação (19) a (22), de adjunção de (23) a (26) e em estruturas predicativas (27) a (30).

- (19) porem o desejo de escrever-te e fingir que estou conversando **contigo** são incentivos mais poderosos (1887 – Acervo P. F. M.)
- (20) Eu tive um sonho **com você** e tua mãe (1936 – Acervo J. e M.)
- (21) esquecer-me **de ti**, oh M., nunca, nunca, por piedade não me digas mais isto (1891, Acervo O.C.)
- (22) Por exemplo - eu proprio, que gosto muito **de Você** (1895 – Acervo P.F.M)
- (23) aproveito destes momentos livres para passal-os **contigo** (1924 – Acervo P. F. M.)
- (24) que você estava também e que eu estava danzando **com você** (1937 – Acervo J. e M.)
- (25) morro **por ti**. tu é e a minha vida (1908, Acervo R.S.)
- (26) Também estou super animada **por você** (1979, Acervo Lac.)
- (27) Estarei **contigo** no pensamento e no coração todo o dia (1907, Acervo A. P.)
- (28) Os retratos já devem estar **com você** (1936 – Acervo J. e M.)
- (29) Meu pensamento é só **em ti** (1936 – Acervo J. e M.)
- (30) Sinceramente não posso explicar nem a você, nem a mim mesmo, porque quero tão insistentemente que você venha. Não sei sinceramente se é **por você** ou se é por mim mesmo (1970 – Acervo F. B.)

Formulamos com hipótese, com base nos resultados de Silva (2020), que as estratégias de *prep.+ti* seriam mais frequentes que *prep.+você* em todos os tipos de relação. A forma *contigo*, por outro lado, apresentaria apenas maior frequência que *com você* em estruturas predicativas. Apresentamos, na tabela a seguir, os resultados referentes à natureza da relação sintática trazendo a comparação entre a alternância *contigo* ~ *com você* e os demais sintagmas preposicionados oblíquos:

Tabela 4: Natureza da relação sintática (valor de aplicação *contigo* / *prep.+ti*)

Natureza da relação	<i>Contigo</i>		<i>Prep.+ti</i>	
	N/T	%	N/T	%
Complementação	35/55	45,5%	67/101	66,3%
Adjunção	22/36	61,1%	93/152	61,2%
Estrutura predicativa	10/13	76,9%	3/4	75,0%

Fonte: elaboração da autora

A partir dos resultados expostos acima, podemos considerar que há poucas diferenças entre a forma *contigo* e as demais estratégias do paradigma de *tu* com relação à natureza da relação sintática entre núcleo predicador e forma oblíqua de 2SG, contrariando nossa hipótese. Em casos de adjunção e de estrutura predicativa, os percentuais das estratégias do paradigma de *tu* são praticamente equivalentes tanto na oposição *contigo* ~ *com você* quanto na oposição *prep.+ti* ~ *prep.+você*. Observamos, no entanto, que, entre as estruturas de complementação, *contigo* apresentou uma menor frequência em relação a *com você* em comparação às estratégias de *prep.+ti* frente a *prep.+você*.

Mais uma vez, a análise de percentuais não nos direciona para diferenças relevantes entre a variação *contigo* ~ *com você* e a variação *prep.+ti* ~ *prep.+você*.

4.4. O sexo dos missivistas

A variável sexo/gênero tem sido amplamente explorada em estudos sociolinguísticos. Labov (1990) estabelece o *Paradoxo de Gênero*: mulheres tendem a optar pela variante de maior prestígio social em casos de variação, mas impulsionam formas inovadoras em contextos de mudança. Contudo, para analisar essa variável dentro do âmbito da variação *tu* e *você* é preciso ressaltar que as noções de forma *conservadora* x *inovadora* não estão estritamente relacionadas a uma questão de *padrão* x *não padrão* ou *menor prestígio* x *maior prestígio*. As formas do paradigma de *tu* são consideradas *conservadoras* por serem mais antigas e as formas do paradigma de *você* – *inovadoras* – apresentam um uso considerado de prestígio desde o século XIX. (cf. SOTO, 2001, MARCOTULIO, 2008; RUMEU, 2008)

Estudos que se ocupam da variação *tu* ~ *você* têm observado que a forma inovadora *você* possa ter herdado um “valor de cortesia, formalidade, distanciamento ou indiretividade própria da expressão primitiva” (LOPES, 2009, p. 52). Carvalho (2019), ao analisar a avaliação do falante com relação às formas *tu* e *você* na posição de sujeito, também observou esse caráter de maior distanciamento para *você*. Dessa forma, entre os estudos que se ocupam dessa variação sob uma perspectiva diacrônica observam um maior favorecimento das formas de *você* na escrita feminina (cf. LOPES; MACHADO, 2005; RUMEU, 2008; 2014, SOUZA, 2012, SILVA, 2020)

No presente trabalho, formulamos com hipótese que as estratégias de *prep.+você* seriam mais frequentes na escrita feminina. No entanto, para a alternância *contigo* ~ *com você*, haveria percentuais mais equilibrados. Na tabela a seguir, apresentamos os resultados referentes à variável sexo do remetente comparando os resultados da alternância *contigo* ~ *com você* e os resultados para os demais sintagmas preposicionados oblíquos:

Tabela 5: Sexo do remetente (valor de aplicação *contigo* / *prep.+ti*)

Sexo do remetente	<i>Contigo</i>		<i>Prep.+ti</i>	
	N/T	%	N/T	%
Homem	37/54	68,5%	137/184	74,4%
Mulher	21/51	41,2%	26/73	35,6%

Fonte: elaboração da autora

No *corpus* analisado, tanto as formas *contigo* quanto as estratégias de *prep.+ti* foram mais frequentes na escrita masculina, com percentuais de 68,5% e 74,4% respectivamente. No entanto, podemos observar que na oposição *contigo* ~ *com você* os percentuais se encontram levemente mais aproximados. Dessa forma, é possível considerar que, apesar de não haver distinções relevantes entre *contigo* e as demais estratégias do paradigma de *tu* com relação ao sexo do remetente da carta, a oposição *contigo* ~ *com você* é menos acentuada que a oposição *prep.+ti* ~ *prep.+você*.

4.5. A relação interpessoal entre os missivistas

Uma vez que o *corpus* analisado constituía uma amostra muito diversificada e que as relações entre os missivistas eram muito variadas, optamos por elaborar um grupo de fatores que reunisse de certa forma relações mais similares. Dessa maneira, dividimos as cartas em três categorias: as cartas *familiares* agrupam missivas trocadas entre pais e filhos, avôs e netos, primos etc.; as cartas *amorosas*, trocadas entre casais (namorados, noivos e cônjuges) e as cartas entre amigos, intituladas como *pessoais*.

Formulamos como hipótese que, as formas de *prep.+ti* seriam mais frequentes em cartas amorosas e que *contigo* apresentaria frequências mais equilibradas em todas as relações analisadas. Na tabela a seguir, apresentamos os resultados referentes a essa variável independente, trazendo a comparação entre os percentuais observados para variação *contigo* ~ *você* e para variação *prep.+ti* ~ *prep.+você*.

Tabela 6: Teor das cartas (valor de aplicação *contigo/prep.+ti*)

Relação	<i>Contigo</i>		<i>Prep.+ti</i>	
	N/T	%	N/T	%
Familiar	27/35	77,1%	37/76	48,7%
Amorosa	27/59	45,8%	108/145	74,5%
Pessoal	3/10	30,0%	18/18	50,0%

Fonte: elaboração da autora

No *corpus* analisado, a variável *relação interpessoal entre missivistas* apresentou um resultado diferenciado para a variação *contigo* ~ *com você*. Enquanto as formas de *prep.+ti* foram predominantes em cartas amorosas, *contigo* foi mais frequente em cartas trocadas entre familiares. Com relação às missivas de teor pessoal, é possível observar equilíbrio entre as estratégias de *prep.+ti* e de *prep.+você*, mas uma predominância da forma *com você* em relação a *contigo*. Dessa maneira, podemos considerar

que *prep.+ti* esteja mais associado a situações de afetividade romântica, enquanto *contigo* é mais frequente em circunstâncias de proximidade dentro do seio familiar. Nossas hipóteses específicas para essa variável não se confirmaram, mas apontam para um possível diferença entre *contigo* e as demais estratégias do paradigma de *tu*, o que era esperado pelas hipóteses gerais.

4.6. O período temporal

Por fim, analisamos como variável independente o período em que a carta foi produzida. Como o somatório de dados referentes à variação *contigo ~ com você* era pequena em relação à quantidade de décadas analisadas, foi necessário agrupar os dados em períodos de tempo maiores. Utilizando como base o padrão de implementação das formas oblíquas observado por Silva (2020), optamos por distribuir os dados nas três fases observadas pela autora.

Com base nos seus resultados, esperamos que, na Fase 3, as formas de *prep.+ti* sejam amplamente substituídas por *prep.+você*, enquanto *contigo* se mantenha maior variação com *com você*. A tabela a seguir apresenta os resultados obtidos tanto para a variação *contigo ~ com você* quanto para os demais sintagmas preposicionados oblíquos de 2SG:

Tabela 7: O período de tempo (valor de aplicação *contigo/prep.+ti*)

Fase	<i>Contigo</i>		<i>Prep.+ti</i>	
	N/T	%	N/T	%
Fase 1 (1870-1899)	14/14	100%	56/60	93,3%
Fase 2 (1900-1939)	37/71	52,1%	105/159	66,6%
Fase 3 (1940-1989)	6/19	31,6%	2/38	5,3%

Fonte: elaboração da autora

Como é possível observar, não foi possível manter um equilíbrio com relação à distribuição de dados por fase, questão já comentada na seção 2 do presente artigo. No entanto, a análise de percentuais confirma nossa hipótese e fornece relevantes apontamentos para o estudo da variação entre as formas oblíquas de 2SG. Na primeira fase, os resultados demonstram um alto predomínio de *contigo* e de *prep.+ti*, com 100% e 93,3% de frequência respectivamente. Na segunda fase, há uma maior aproximação entre os percentuais dos paradigmas de *tu* e *você*. As estratégias de *prep.+ti* alcançaram 66,6% de frequência e *contigo* se aproxima ainda mais de *com você* com percentual de 52,1%. Por fim, na terceira fase, é possível observar que as formas do paradigma de *você* se tornam as mais frequentes e há poucas ocorrências do paradigma de *tu*. Contudo, ao trazer o olhar para elas, se torna perceptível uma diferença entre a forma *contigo* e as demais estratégias preposicionadas do paradigma conservador. O desaparecimento das estratégias de *prep.+ti* em relação a *prep.+você*, com apenas 5,3%, é mais evidente do que observado para *contigo*, que alcança 31,6%. Dessa forma, podemos considerar que, na escrita fluminense, o *ti* preposicionado foi amplamente substituído por *prep.+você* e que a forma morfologizada *contigo* apresenta uma maior resistência a implementação das estratégias preposicionadas oblíquas de 2SG do paradigma inovador.

Considerações finais

O estudo de Silva (2020) apontava para uma possível resistência da forma *contigo* à implementação de formas oblíquas de 2SG do paradigma de *você* em cartas pessoais produzidas no Rio de Janeiro entre fins do século XIX e no decorrer do século XX. Sendo assim, era esperado que, ao analisar os mesmos dados destacando a variação *contigo* ~ *com você* das demais estratégias preposicionadas – *prep.+ti* ~ *prep.+você* – pudéssemos identificar um comportamento diferenciado entre *contigo* e *preposição + ti*.

Os resultados aqui apresentados, com base em percentuais, não nos permitiram traçar completamente esse comportamento diferenciado. Na maior parte das variáveis independentes controladas, os números foram equivalentes. No entanto, com relação à distribuição total de dados e ao sexo do missivista, foi possível observar a oposição *contigo* ~ *com você* levemente mais equilibrada que a oposição *prep.+ti* ~ *prep.+você*. Além disso, ao analisar a *relação interpessoal entre os missivistas*, verificamos que *contigo* foi predominante em cartas trocadas por familiares, enquanto *prep.+ti* foi mais frequente em cartas amorosas e mais afetivas. No entanto, os resultados ainda não elucidam de maneira satisfatória possíveis diferenças entre a forma morfologizada *contigo* e as demais estratégias do paradigma de *tu*.

Lidar com *corpora* históricos oferece algumas limitações que nem sempre nos permite traçar conclusões. Apesar disso, acreditamos que o fato de, no terceiro período de tempo analisado, a forma tônica *ti* preposicionada ter sido mais amplamente substituída pelas demais estratégias do paradigma de *você* do que se observou para a variação *contigo* ~ *com você*, sugira que, em sincronias atuais, essas formas tenham se diferenciado. Os resultados diacrônicos apresentados no presente artigo podem não trazer respostas conclusivas a respeito da variação *contigo* ~ *com você* no decorrer do tempo, mas, sem dúvidas, permitem a elaboração de novas perguntas para análises futuras a respeito da variação entre as formas oblíquas de 2SG no dialeto carioca.

Referências

CÂMARA JUNIOR, Joaquim. Mattoso. *História e estrutura da língua portuguesa*. 2. ed., Rio de Janeiro: Padrão, 1979.

CARVALHO, Bruna Brasil Albuquerque de. “O que você acha do uso de tu?”: a percepção da variação dos pronomes de 2SG no dialeto carioca. Dissertação de Mestrado. Rio de Janeiro: Faculdade de Letras/UFRJ. 2019.

HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE SILVESTRE, Juan Camilo. *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, 2012.

HERNÁNDEZ-CAMPOY Juan Manuel; SCHILLING, Natalie. *The Application of the Quantitative Paradigm to Historical Sociolinguistics: Problems with the Generalizability Principle*. In: HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE-SILVESTRE, Juan Camilo. *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, pp. 63-79, 2012.

LABOV, William. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*. In: SANKOFF, D. et al. (eds.). *Language Variations and Change*. Cambridge: Cambridge University Press, v. 2, n. 2, pp. 135-6, 1990.

LOPES, Célia Regina dos Santos. Retratos da mudança no sistema pronominal: o tratamento carioca nas primeiras décadas do século XX. In: Arnaldo Cortina; Silvia Maria Gomes da Conceição Nasser. (org.). *Sujeito e Linguagem: Séries Trilhas Linguísticas*. Araraquara: Cultura Acadêmica, v. 17, pp. 47-74, 2009.

LOPES, Célia Regina dos Santos; CAVALCANTE, Silvia Regina Oliveira. A cronologia do voçamento no português brasileiro: expansão de você-sujeito e retenção do clítico te. *Linguística* (Madrid), v. 25, pp.30-65, 2011.

LOPES, Célia Regina dos Santos. *Mudança Sintática das Classes de Palavra: Perspectiva Funcionalista, História do Português Brasileiro*, v. 4. São Paulo: Contexto, 2018.

LOPES, Célia Regina dos Santos; RUMEU, Márcia Cristina de Brito. *A identificação dos perfis socioculturais dos redatores de corpora históricos: encaminhamentos metodológicos*. Diadorim (Rio de Janeiro), v. 20, pp. 147-68, 2018.

MARCOTULIO, Leonardo Lennertz. *A preservação das faces e a construção da imagem no discurso político do marquês do Lavradio: as formas de tratamento como estratégias de atenuação da polidez linguística. Volumes I e II*. Dissertação de Mestrado em Língua Portuguesa, Faculdade de Letras, Universidade do Rio de Janeiro. Rio de Janeiro, 2008.

MATEUS, Maria Helena Mira et al. *Gramática da Língua Portuguesa*. Lisboa: Caminho, 2003.

OLIVEIRA, Thiago Laurentino de. *Entre o Linguístico e o Social: Complementos Dativos de 2ª pessoa em Cartas Cariocas (1880-1980)*. Dissertação de Mestrado em Língua Portuguesa, Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2014.

PEREIRA, Rachel de Oliveira. *Pronomes possessivos de segunda pessoa: a variação teu/seu em uma perspectiva histórica*. Tese de Doutorado em Língua Portuguesa. Rio de Janeiro: Faculdade de Letras/UFRJ. 2016.

RAPOSO, Eduardo Buzaglo Paiva et al. *Gramática do Português*, Vol I e II. Coimbra: Fundação Calouste Gulbenkian, 2013.

RUMEU, Márcia Cristina de Brito. *A implementação do 'você' no português brasileiro oitocentista e novecentista: um estudo de painel*. Rio de Janeiro. Tese de Doutorado em Língua Portuguesa, Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2008.

RUMEU, Márcia Cristina de Brito. *A difusão do 'você' pelos contextos sintáticos de complementação e de adjunção*. *Revista Portuguesa de Humanidades*, v. 18-I, pp. 91-114, 2014.

SILVA, Thaissa Frota Teixeira de Araujo. *Lembro de você; preciso de ti: uma análise diacrônica das variantes oblíquas de 2SG na escrita epistolar fluminense*. Dissertação de Mestrado em Língua Portuguesa. Rio de Janeiro: Faculdade de Letras/UFRJ. 2020.

SOTO, Eva Ucy Miranda Sá. *Varição/mudança do pronome de tratamento alocutivo: uma análise enunciativa em cartas brasileiras*. Tese (Doutorado em Linguística) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2001.

SOUZA, Janaina Pedreira Fernandes. *Mapeando a entrada do você no quadro pronominal: análise de cartas familiares dos séculos XIX-XX*. Dissertação de Mestrado em Língua Portuguesa, Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2012.

SOUZA, Camila Duarte de. *Eu te amo, eu lhe adoro, eu quero você: a variação das formas de acusativo de 2ª pessoa em cartas pessoais (1880-1980)*. Dissertação de Mestrado em Língua Portuguesa, Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2014.

WENREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança linguística*. Trad. Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].

EXPRESSIONS DE TEMPO DECORRIDO EM CARTAS PESSOAIS DO ALAGOANO GRACILIANO RAMOS*ELAPSED TIME EXPRESSIONS IN GRACILIANO RAMOS' PERSONAL LETTERS**Pedro Henrique Sousa dos Santos¹**Elyne Giselle de Santana Lima Aguiar Vitória²***RESUMO**

As Expressões de Tempo Decorrido (ETDs) são um fenômeno de recente estudo no português brasileiro, e inédito no português alagoano em sincronia passada. Dentre as análises realizadas, grande parte tem enfoque sincrônico, investigando o português atual em suas variedades. Logo, este trabalho objetiva analisar essa variação com o aporte teórico da Sociolinguística Variacionista (LABOV, 2008 [1972]) e da Sociolinguística Histórica (ROMAINE, 1982) para explicitar os fatores condicionantes do uso desse fenômeno a partir do *corpus* de 112 cartas pessoais do escritor alagoano Graciliano Ramos no início do século XX. Assim, o trabalho justifica-se por ampliar as pesquisas sobre a variedade alagoana do português. Como resultado, foram encontrados 77 dados referentes às variantes “Haver” e “Fazer”, que correspondem a um percentual de 86% (66 dados) para a primeira e 14% (11 dados) para a segunda, o que motiva esse estudo a refletir sobre a ausência das variantes “Ter” e “Estar com” nas cartas. Para análise quantitativa dos dados, utilizou-se o programa RStudio, que proporcionou análises univariadas e resultou na significância das variáveis independentes aspecto ($p < 0,05$), complementizador ($p < 0,001$) e ordem ($p < 0,001$) no condicionamento das variantes das ETDs observadas. Já para análise qualitativa, observaram-se os aspectos linguísticos dos dados encontrados, o que resultou preliminarmente na constituição de hipóteses e novos caminhos de pesquisa para futuros trabalhos que investiguem o português alagoano em sua perspectiva histórica para ampliar os dados acerca desse fenômeno e verificar se a ausência de “Ter” e “Estar com” ocorre em outros *corpora*, além de trazer novas questões sobre as ETDs.

PALAVRAS-CHAVE: Expressões de tempo decorrido. Português alagoano. Cartas pessoais.

ABSTRACT

Elapsed Time Expressions (ETEs) are a phenomenon of recent study in Brazilian Portuguese, and unprecedented in Alagoas Portuguese in synchronic past. Among the analyzes carried out, a large part has a synchronic focus, investigating current Portuguese in its varieties. Therefore, this work aims to analyze this variation with the theoretical contribution of Variationist Sociolinguistics (LABOV, 2008 [1972]) and Historical Sociolinguistics (ROMAINE, 1982) to explain the conditioning factors in the use of this phenomenon from the *corpus* of 112 personal letters of the writer from Alagoas Graciliano Ramos at the beginning of the 20th century. Thus, the work justifies itself by expanding the research on the Alagoas variety of Portuguese. As results, 77 data were found referring to the variants “Haver” and “Fazer”, which correspond to a percentage of 86% (66 data) for the first and 14% (11 data) for the second, which motivates this study to wonder on the absence of variants “Ter” and “Estar com” in the letters. For quantitative data analysis, the RStudio program was used, which provided univariate analyzes and resulted in the significance of the independent variables *aspect* ($p < 0.05$),

¹ Graduando do curso de Letras - Campus Arapiraca e bolsista do Pibic/CNPq/UFAL/FAPEAL, pedro.henrique@arapiraca.ufal.br, <https://orcid.org/0000-0001-7069-275X>.

² Doutora em Linguística e professora da Universidade Federal de Alagoas - Campus Arapiraca, elyne.vitorio@gmail.com, <http://orcid.org/0000-0002-6279-2379>.

complementizer ($p < 0.001$) and *order* ($p < 0.001$) in conditioning the variants of the observed ETEs. For the qualitative analysis, however, the linguistic aspects of the data found were observed, which preliminarily resulted in the constitution of hypotheses and new research paths for future works that investigate Alagoas Portuguese in its historical perspective to expand the data on this phenomenon and verify whether the absence of “Ter” and “Estar com” occurs in other *corpora*, in addition to bringing new questions about ETEs.

KEYWORDS: Elapsed time expressions. Alagoas portuguese. Personal letters.

Introdução

Nos estudos linguísticos, as Expressões de Tempo Decorrido (ETDs) servem para indicar um intervalo de estado de coisas entre dois pontos temporais (AVELAR, 2018, p. 85). Na investigação formal, Avelar (2018) argumenta que as ETDs apresentam base nas sentenças possessivas e existenciais com “Haver”, “Ter” e “Estar com”, apesar de possuírem características próprias, especialmente a ETD formada com “Haver”.

Avelar (2018) reporta que, enquanto as sentenças existenciais com “Haver”, em (1), apresentam caráter oracional, as ETDs com essa variante, em (2), possuem caráter preposicional. Por outro lado, as sentenças existenciais, em (3), e as ETDs, em (4), com “Ter” aparentam comungar do mesmo estatuto oracional, visto que as duas se flexionam e podem abrigar sujeito. Já a variante “Estar com”, em (5), varia com “Ter” nas estruturas possessivas e existenciais pelo caráter transitório da primeira e permanente da última, embora não apresente tal diferença em relação à ETD com “ter”, conforme aponta o autor.

- (1) Há uma criança na sala.³
- (2) Há 10 anos que eu moro aqui.
- (3) Tem um menino na escola.
- (4) Tem um tempo que eu vim ao Rio de Janeiro.
- (5) Está com 15 dias que passei pelo Brasil.

Sendo um fenômeno de recente interesse, há poucos trabalhos descrevendo sua constituição em *corpora* falado ou escrito. Na fala carioca, destaca-se o uso maior de “Haver” e “Ter” (AVELAR, 2011) enquanto, nas falas cearense e alagoana, há também a presença de “estar com” (AVELAR, 2018). Na fala alagoana, há apenas a descrição de Avelar (2012) sobre as ETDs na comunidade Muquém, em União dos Palmares. Nesse sentido, é necessário explorar o comportamento dessa variável a fim de descrever os seus contextos de uso.

Além disso, em uma outra face dos estudos linguísticos, na análise de *corpora* históricos, é ainda mais observável a escassez de trabalhos. Na revisão de literatura, será detectado que o presente trabalho é o primeiro a estudar esse fenômeno em sincronia pretérita em Alagoas, o que justifica por si só a sua importância. Logo, este trabalho contribui para a literatura sobre as ETDs no português, ampliando o conhecimento sobre os processos de variação e possível mudança linguística na variedade alagoana.

³ Exemplos criados para exemplificar o fenômeno em estudo.

Para realizar este estudo, tomou-se como aporte a Teoria da Variação e Mudança (LABOV, 2008 [1972]), que considera a língua como fenômeno dotado de uma heterogeneidade ordenada, para focar as ETDs enquanto fenômeno linguístico variável. Além disso, também foi considerada a Sociolinguística Histórica (ROMAINE, 1982; CONDE-SILVESTRE, 2007; HERNÁNDEZ-CAMPOY; CONDE-SILVESTRE, 2012) ao propor estudar fenômenos variáveis, que ocorreram no passado, em amostra de língua escrita.

As hipóteses do trabalho são as seguintes: i) a frequência da ETD com “estar com” será baixa ou nula, pois, como apontado por Avelar (2018), o seu uso é mais produtivo na fala do que na escrita; ii) o uso da ETD com “ter” será baixo ou nulo, pois é essa a variante não encontrada nos dados de Avelar (2018); e iii) a frequência de ETDs com “haver” será majoritária, visto que é uma forma valorizada na escrita pelas gramáticas normativas.

O objetivo geral da pesquisa é analisar o uso das ETDs com “ter”, “haver”, “estar com” e “fazer” em cartas pessoais do escritor alagoano Graciliano Ramos, produzidas na primeira metade do século XX e organizadas em Ramos (2011). O artigo está segmentado do seguinte modo: a primeira seção apresenta a revisão de literatura acerca dos estudos sobre as ETDs; a segunda seção descreve a metodologia da pesquisa, e, na seção seguinte, são descritos os resultados encontrados. Por fim, são apresentadas as considerações finais destacando as contribuições do estudo para as futuras investigações e para a variedade alagoana.

1. Revisão de literatura

Nesta seção, será feita uma breve revisão sobre o que se tem escrito acerca das ETDs. Para isso, será utilizado o método narrativo de revisão de literatura (VOSGERAU; ROMANOWSKI, 2014), especificamente um misto de revisão narrativa, pois há um mapeamento das pesquisas sem critérios de replicabilidade, uma vez que há poucos trabalhos sobre o assunto; e, também, traços de revisões analíticas, como a integrativa, pois há uma avaliação do que as obras (não) comentam.

A seção de revisão de literatura está dividida em três subseções. Na primeira, intitulada “O que dizem as gramáticas normativas”, selecionaram-se quatro obras normativas de diferentes autores para se ter um panorama dos estudos tradicionais sobre o assunto. Na segunda subseção, “O que dizem as gramáticas descritivas”, analisa-se como três gramáticas descritivas, escritas por linguistas, abordam o objeto de pesquisa. Por fim, em “O que dizem os estudos (socio)linguísticos”, serão analisados, mais especificamente, os estudos (socio)linguísticos sobre o assunto, compreendendo aqui suas interfaces com outras áreas.

1.1. O que dizem as gramáticas normativas

Para a análise do que dizem as gramáticas normativas acerca do fenômeno em estudo, analisaram-se as obras de Cunha e Cintra (2017), Rocha Lima (2011) e Bechara (2014a, 2014b) a fim de obter um panorama sucinto sobre a descrição das ETDs.

Na gramática de Cunha e Cintra (2017, p. 144), foi encontrado um tópico sobre orações sem sujeito em que os autores apresentam uma classificação para verbos impessoais com sentido de decorrência temporal. Os autores exemplificam as expressões com “haver” e “fazer”, mas não comentam sobre o estatuto sintático de cada uma e não abordam as expressões com “ter” e “estar com”. Além disso, Cunha e Cintra (2017, pp. 553-4), quando tratam da sintaxe do verbo “haver”, comentam sobre seu uso como expressão de tempo decorrido, somente apresentando mais alguns exemplos de uso, sem promover a análise sintática dessas estruturas.

Em Rocha Lima (2011), encontrou-se quase a mesma descrição da obra de Cunha e Cintra (2017), abordando, dessa vez, estruturas com os verbos “haver”, “ser” e “fazer”, como mostram os exemplos (6), (7) e (8), retirados de Rocha Lima (2011, pp. 289-90). Também nessa obra não foi constatada a descrição de expressões de tempo decorrido com “ter” e “estar com”. Além dessa descrição, Rocha Lima (2011, p. 488), quando aborda orações impessoais, refere-se mais uma vez ao verbo “fazer” com sentido de decorrência de tempo e comenta alguns exemplos com o uso dessa forma gramatical.

(6) Há grandes poetas no Brasil.

(7) Fazia muito frio naquele mês.

(8) Seriam talvez duas horas da tarde.

Bechara (2014a), quando trata de orações impessoais, esclarece rapidamente que podem ser usados “*haver, fazer e ser* nas indicações de tempo” (BECHARA, 2014a, p. 23). Já em sua outra obra (BECHARA, 2014b), o autor trata os verbos “haver”, “fazer” e “ser” como “*impessoais essenciais*, uma vez que há vários outros que acidentalmente aparecem em construções impessoais mas que tendem, em sua maioria, a ser usados com sujeito próprio (...)” (BECHARA, 2014b, p. 40). Entretanto, do mesmo modo que as obras anteriormente citadas, as de Bechara também não tratam do verbo “ter” e da construção “estar com”.

Cabe nesta seção uma indagação de por que dessas variantes não serem abordadas pelas gramáticas analisadas. Uma resposta seria que esses gramáticos consideram as formas “ter” e “estar com” como marginais ou incorretas. Na obra de Bechara (2014b), o autor comenta que o uso de “ter” no lugar de “haver” constitui “*incorreção, na língua culta*” (BECHARA, 2014b, p. 42). Essa não é, como pode vir a se pensar, uma exclusividade do referido autor. A tradição gramatical sempre tratou as formas com “ter” existencial como incorretas, como observado em Ali (1957), e não chega nem a mencionar o uso de “estar com” com essa função gramatical.

1.2. O que dizem as gramáticas descritivas

No estudo das gramáticas descritivas, foram utilizadas as obras de Neves (2011), Castilho (2010) e Perini (2016). Com esses textos, ter-se-á um panorama mais amplo de como as ETDs são abordadas nas descrições do português brasileiro.

Em Neves (2011), só foi encontrada uma menção às expressões de tempo decorrido, apesar de ser bem proveitosa pelo que diz e o que não diz acerca da variável linguística em estudo. Neves (2011, p. 233) menciona, no tópico sobre locuções adverbiais, que a construção “*HÁ/FAZ, HAVIA/FAZIA + substantivo quantificado*” apresenta sentido temporal. Desse modo, sua análise coaduna com algumas pesquisas sociolinguísticas que serão abordadas na próxima seção. Entretanto, deve-se reconhecer a ausência de uma análise semelhante para o verbo “ter” e a locução “estar com”, mostrando que há escassos trabalhos sobre essas variantes.

Em Castilho (2010), não foi observada nenhuma referência às ETDs. Do mesmo modo, Perini (2016) também não faz referência às expressões de tempo decorrido com “ter”, “haver” e “estar com”. Entretanto, o autor separa uma pequena subseção para tratar dos usos de “vai para” e “fazer” em sentido temporal, semelhante às ETDs analisadas aqui (PERINI, 2016, p. 111), como em (9) e (10), retirados da obra do autor.

(9) Vai para sete anos que eu não vejo a minha irmã.

(10) Faz sete anos que eu não vejo a minha irmã.

Desse modo, ao analisar três gramáticas descritivas conceituadas academicamente, foi possível observar a quase inexistência de análise dessas expressões, que são tema do presente estudo. Assim, é indispensável que haja uma maior descrição e explicação de seus usos no português brasileiro para que se tenha um maior conhecimento acerca da gramática da língua.

1.3. O que dizem os estudos (socio)linguísticos

Para se ter um panorama do que dizem os estudos (socio)linguísticos sobre a variação “haver”, “ter”, “fazer” e “estar com” nas ETDs, analisaram-se os trabalhos de Vieira (2008) e Avelar (2011, 2012, 2018) por serem esses os textos encontrados sobre o assunto e relevantes para a pesquisa.

No trabalho de Vieira (2008), estudaram-se as ETDs com “haver”, “ter” e “fazer” no português brasileiro e europeu partindo de uma perspectiva sociofuncionalista. Nessa pesquisa, foi estudado um *corpus* de ocorrências de ETDs retirado dos CRPC-Lisboa (*Corpus* de Referência do Português Contemporâneo e Português Fundamental), VARPORT (Análise Contrastiva de Variedades do Português), NURC-RJ (Norma Urbana Culta), PEUL (Programa de Estudos sobre o Uso da Língua) e jornais do Brasil e de Portugal. Das 361 ETDs observadas, 308 são referentes ao português brasileiro e 53 referentes ao português europeu.

Nesse texto, Vieira (2008) considera a forma linguística “haver” como distante da noção de verbo predicador e próximo da noção de constituinte de uma expressão temporal, aproximando-se da análise de Avelar (2011, 2012, 2017) sobre essa expressão, como se verá ao longo desta revisão. Já na análise sociolinguística, a autora demonstra que a forma “haver” predominou e foi usada 293 vezes (81%) contra 41 usos de “ter” (11%) e 27 de “fazer” (7%). Além disso, Vieira (2008) reporta que todos os usos de “ter” com função de decorrência temporal são do português brasileiro, o que mostra um

contraste entre as duas variedades da língua. Do mesmo modo, a variante “fazer” é majoritariamente encontrada no Brasil, com apenas duas ocorrências no português europeu. Nesse sentido, o uso de “haver” em Portugal é quase categórico (96,2% contra 78,6% no português brasileiro).

Na investigação sobre as variáveis independentes significativas para a escolha de uma forma ou outra, a autora constatou que a variável escolaridade foi muito significativa, pois mostrava que, quanto mais escolarizado é o sujeito, mais ele tende a usar ETDs com “haver”. Além dessa, as variáveis “modo de organização no discurso”, “ordem de expressão temporal na oração” (semelhante à variável Ordem analisada no presente trabalho) e “configuração estrutural da expressão de tempo decorrido” também se mostraram significativas.

Em Avelar (2011), há uma análise sobre as ETDs na fala carioca, tendo como variantes as formas “ter” e “haver”. Nesse trabalho, o autor realizou uma análise formal e sociolinguística das ETDs, uma vez que a pesquisa empreendida se encaixa na vertente da Sociolinguística Paramétrica. Com um *corpus* de dados falados da cidade do Rio de Janeiro entre as décadas de 70 e 90 do século XX provenientes dos bancos de dados do NURC e do PEUL, Avelar (2011) analisou 282 ETDs, das quais, segundo o texto, 79% são ETDs com “haver” e 21% com “ter”.

Ainda na parte da análise quantitativa, o autor compara o uso dessas ETDs entre falantes com e sem ensino superior e constata que houve um acréscimo de uso das ETDs pelos falantes com ensino superior (de 0% na década de 70 para 23% na década de 90) e que houve uma estabilidade em falantes sem ensino superior. Disso, Avelar (2011) conclui que está havendo uma aproximação das variedades culta e não culta na cidade do Rio de Janeiro, pelo menos nessa variável, e que as variantes “ter” e “haver” estão em estabilidade na comunidade carioca.

Ademais, Avelar (2011) apresenta uma investigação formal das ETDs com “ter” e “haver”, comentando algumas diferenças na configuração sintática de cada variante. Na análise estrutural, o autor analisa o processo de clivagem, a presença do complementizador “que”, a ocorrência do advérbio “atrás”, o uso das variantes dentro de adjuntos adnominais, o preenchimento da posição de sujeito e a proposição de constituintes interrogativos à variante.

Em referência ao contexto de clivagem, o autor mostra que as ETDs com “haver”, quando este está no presente do indicativo em sua forma impessoal, podem ser clivadas, ao contrário das ETDs com “ter”, que se tornam agramaticais, como em (11) e (12)⁴.

(11) Foi [*há* mais de duas semanas] que ela saiu de casa.

(12) *Foi [*tem* mais de duas semanas] que ela saiu de casa.

Já no caso do complementizador “que”, Avelar (2011) percebe que, em sentenças iniciadas pelas ETDs, a estrutura com “ter” exige o uso de um “que” complementizador enquanto, na ETD com “haver”, o uso dele torna-se opcional, isto é, não causa perda de gramaticalidade da sentença, como em (13), (14) e (15).

⁴ Exemplos criados para exemplificar o fenômeno em estudo.

- (13) [*Há* duas semanas] (que) eu não vou à praia.
- (14) [*Tem* duas horas] que eu não saio de casa.
- (15) **[Tem* três dias] eu não vou ao mercado.

Após isso, o autor analisa o contexto da ocorrência do advérbio “atrás”, que favorece o uso da ETD com “haver” e desfavorece o uso da ETD com “ter” porque “itens adverbiais [...] são largamente empregados no interior de sintagmas nominais com interpretação locativa temporal (...)” (AVELAR, 2011, p. 171), como em (16) e (17).

- (16) Eu encontrei meus amigos [*há* oito meses *atrás*].
- (17) ??Eu encontrei meus amigos [*tem* oito meses *atrás*].

Em relação à adjunção, Avelar (2011) nota que ETDs com “haver” ocorrem com frequência dentro de adjuntos adnominais, contexto em que as ETDs com “ter” não ocorrem, a custo de perder seu caráter gramatical, como em (18) e (19).

- (18) Aquela prova [*há* duas horas (*atrás*)] amedrontou os alunos.
- (19) *Aquela prova [*tem* duas horas] amedrontou os alunos.

Além disso, Avelar (2011) analisa os contextos de preenchimento de sujeito e observa que as ETDs com “haver” não licenciam sujeitos em nenhum dos dados coletados, ao contrário das ETDs com “ter”, que licenciam sujeitos pronominais que concordam com o verbo já referido, como em (20) e (21).

- (20) **Eles** [*têm* seis anos] que não aparecem aqui.
- (21) ***Eles** [*há* seis anos] que não aparecem aqui.

Por fim, o autor observa que, no caso da presença de constituintes interrogativos temporais, somente a ETD com “ter” possibilita a anteposição desses termos. Isso, segundo o pesquisador, é devido ao caráter oracional da ETD com “ter”, que, ao contrário da variante “haver”, apresenta a posição [spec, CP] no sintagma para o movimento do constituinte interrogativo, como em (22) e (23).

- (22) **[Quantos anos há]* que você não vê sua irmã?
- (23) *[Quanto tempo tem]* que você não vê sua irmã?

Toda a investigação de Avelar (2011) leva à conclusão de que o elemento “ter” funciona como verbo nas ETDs, enquanto o elemento “haver” funciona como constituinte nominal prepositivo. Logo, a variável ETD apresenta não variantes que se alternam lexicalmente, mas sim duas configurações sintáticas diferentes que se alternam nos usos da língua. Vale ressaltar que, em Avelar (2011), não há tratamento sobre a forma “estar com” como uma das variantes para as ETDs, devido à ausência dessa forma linguística na região do Rio de Janeiro.

Em Avelar (2012), há uma análise das ETDs na fala quilombola da comunidade Muquém, situada em União dos Palmares (Alagoas), onde vivem cerca de 500 habitantes. O autor investigou um *corpus* falado de 14 informantes coletado por Moura (2009) a fim de encontrar como está distribuído o uso das variantes “ter”, “haver” e “estar com” na comunidade.

Após quantificar os dados, Avelar (2012) observou que, em Muquém, “*haver* não varia com *ter*, mas com a locução *estar com* [...]” (AVELAR, 2012, p. 74); a porcentagem de uso de cada variante foi de 39% e 71%, respectivamente. Além disso, vale ressaltar que das 18 ETDs coletadas na comunidade, 7 eram ETDs com “haver”, sendo 6 destas observadas em apenas 1 informante. Logo, o ponto forte do trabalho de Avelar (2012) é mostrar que, no português brasileiro, há outras variantes concorrendo dentro das ETDs além das mais usuais “ter” e “haver”, o que opõe tal variedade ao português europeu.

Já em Avelar (2018), o autor traça três panoramas para dar conta das expressões de tempo decorrido com “ter”, “haver” e “estar com”: um descritivo, um formal e um diacrônico. No panorama descritivo, o autor parte dos dados de Avelar (2011, 2012) para analisar a realidade dessas expressões, como já foi mostrado anteriormente neste texto. Já na parte formal, Avelar (2018) retoma os contextos sintáticos de uso das ETDs reportados, também, em Avelar (2011, 2012) e traça diferenças com relação às sentenças existenciais, visto que nestas o termo “haver” comporta-se como verbo, enquanto, nas ETDs, ele atua como elemento preposicionado. Tal análise coaduna com Vieira (2008), uma vez que a autora reconhece que o verbo “haver” funciona como constituinte de expressão adverbial e não como verbo predicador.

Além disso, outra diferença entre as ETDs e as existenciais refere-se ao uso predominante de “ter” nas orações com indicação de existência, enquanto “haver” é amplamente utilizado na fala e na escrita quando em ETDs. Nesse sentido, a derivação proposta em Avelar (2018, p. 104) para as sentenças possessivo-existenciais não são generalizáveis para as expressões de tempo decorrido.

Por fim, no panorama diacrônico, Avelar (2018) apresenta dados que mostram o uso canônico de ETDs com “haver” na escrita de diferentes gêneros textuais. Aliado a isso, o autor propõe que “ter” e “estar com” podem ter surgido como consequência da redução do paradigma flexional do português brasileiro, que forçou a entrada de verbos licenciadores de sujeitos (AVELAR, 2018, pp. 131-4).

Quanto a “estar com”, Avelar (2018, p. 123) destaca que seu uso é desautorizado até na escrita contemporânea monitorada, logo a tendência é não se encontrar esse padrão frásico nos textos antigos. Uma hipótese para sua aparente estigmatização é a possível origem da forma linguística no contato entre o português brasileiro e as línguas africanas, como aponta Avelar (2012). Desse modo, o uso

de “estar com” consolidou-se na fala de escravos e ex-escravos, que faziam parte das camadas mais subjugadas da sociedade.

Assim, a atual pesquisa justifica-se pela necessidade de entender melhor o comportamento diacrônico das ETDs nas cartas pessoais do escritor alagoano Graciliano Ramos justamente pela carência de material referente à região de Alagoas, em que o uso de “estar com” e “ter” é mais usual do que na região sudeste. Nesse sentido, uma investigação por textos alagoanos pode revelar outra realidade linguística para as expressões de tempo decorrido não apresentada por Avelar (2018).

2. Metodologia

A pesquisa empreendida baseia-se na Teoria da Variação e Mudança (LABOV, 2008), que teoriza o estudo da língua como componente social, ou seja, trabalha com a correlação entre variáveis linguísticas e sociais para explicar a variação e a mudança linguística. Nesse sentido, a Sociolinguística Variacionista, como também é chamada, reconhece a heterogeneidade sistemática das línguas humanas e a inerência do processo de variação, logo assume que seu estudo pode se dar quantitativamente por meio de uma metodologia padronizada de pesquisa.

Para análise de processos variáveis em sincronias passadas, também se faz uso da Sociolinguística Histórica (SH), que é o estudo das relações entre as estruturas linguísticas e sociais ao longo do tempo (ROMAINE, 1982). Logo, tem-se o privilégio de olhar como está o processo de mudança no presente para inferir o passado (CONDE-SILVESTRE, 2007). Por outro lado, como aponta Conde-Silvestre (2007), a SH está restrita à disponibilidade dos registros escritos, que podem apresentar dados fragmentários, escassos e dificilmente vinculáveis à produção real dos falantes, visto que os *corpora* de análise são materiais de épocas passadas, que sobreviveram por acidente (BERLINCK; BARBOSA; MARINE, 2008).

A fim de solucionar possíveis problemas metodológicos, os pesquisadores vêm desenvolvendo instrumentos de pesquisa e ferramentas metodológicas. O princípio do Uniformitarismo (ROSA, 2015), que diz respeito ao fato de fatores que influenciam a variação e a mudança linguística no presente serem os mesmos que influenciaram a mudança no passado, e o princípio do Maximalismo Informacional (HERNÁNDEZ-CAMPOY; CONDE-SILVESTRE, 2012), que se refere ao uso de todos os meios possíveis e razoáveis que possibilitem saber o que ocorreu no passado, auxiliam na resolução dos problemas de validade sócio-histórica e de representatividade. Além disso, os pesquisadores fazem uso de gêneros como cartas, peças teatrais e transcrições de discursos orais, pois são documentos assinados e que apresentam traços de oralidade, fazendo com que haja mais variação em sua escrita (CONDE-SILVESTRE, 2007).

Para tal empreendimento, Romaine (1982, pp. 121-3) aponta que o texto escrito só é visto como invariável e menos relevante por ser comparado ao texto oral. O texto escrito apresenta sua própria configuração estrutural e discursiva, necessitando de uma análise própria. Sendo assim, o espaço de variação nos documentos existe, mas não é igual ao encontrado nos enunciados de fala.

Em relação à descrição das ETDs em sincronias passadas, o presente trabalho analisa 112 cartas pessoais do alagoano Graciliano Ramos organizadas em Ramos (2011). Esses documentos foram

revisados segundo o Novo Acordo Ortográfico da Língua Portuguesa, logo se trata de uma edição modernizada, embora deva se destacar que questões gramaticais foram conservadas. Desse modo, o material de Ramos (2011) se apresenta como um *corpus* de análise relevante por seu autor ser alagoano e, assim, revelar traços dessa variedade.

É tempo de deixar o próprio Graciliano revelar suas relações com o cotidiano e as pessoas com as quais mais de perto conviveu – e isto sem a fragmentação de documentos e sem interpretações passionais. Os futuros estudiosos e biógrafos passam a contar com uma fonte documental direta (RAMOS, 2011, p. 6).

As cartas pessoais de Graciliano (RAMOS, 2011), dos subgêneros “amor”, “amigo” e “familiar”, são endereçadas a diversos destinatários, desde à sua família (pais, irmãs, filhos e cunhado) até ao seu amigo e à sua mulher — destinatária da maioria das cartas. Esses textos foram escritos em um período que se estende de 1910 a 1952, estratificado no presente estudo em 2 períodos, 1910-1928 e 1930-1952, abrangendo grande parte da vida do autor.

Os dados coletados são analisados no programa RStudio (R Core Team, 2021), mais especificamente na *interface* RStudio, por meio dos pacotes estatísticos *ggplot2* (WICKHAM, 2016) e *ggstatsplot* (PATIL, 2018), ferramentas estatísticas disponíveis para análise e discussão dos dados. A variável dependente (VD) *ETD* foi constituída pelas variantes “Haver”, “Ter”, “Estar com” e “Fazer”. Como variáveis independentes (VI), são selecionadas *aspecto*, *especificidade*, *ordem*, *complementizador*, *tempo*, *destinatário* e *período*.

A variável *aspecto* é constituída pelas variantes “Pontual”, um evento específico no tempo, e “Durativo”, um evento em processo no tempo (FERRAREZI JR., 2019). Já a VI *destinatário* foi constituída pelos níveis FA (Família = círculo familiar mais íntimo como Pais, Avós e Irmãos), NE (Namorada/Esposa = Heloísa Ramos), AM (Amigo = Joaquim Pinto da Mota Lima Filho) e FI (Filhos = filhos de Graciliano). A VI *período* foi constituída por dois níveis: I (1910-1928) e II (1930-1952); sua construção se deu pelo ano de escrita da carta.

Já a VI *complementizador* refere-se à presença ou ausência do complementizador “que” entre a *ETD* e a sentença matriz à qual ela se liga. Já a variável *tempo*, constituída pelos níveis Presente, Passado e Particípio, refere-se ao tempo do verbo da sentença matriz e ao seu papel no uso de uma ou outra *ETD*. Em seguida, a VI *especificidade*, baseada nos trabalhos de Vieira (2008, 2020) e formada pelas variantes Específico e Não Específico, intenta analisar o papel do SN ligado ao núcleo das *ETDs*. Por fim, a variável *ordem* constitui-se dos níveis “Início”, “Meio” e “Fim” e baseia-se nas análises de Vieira (2008) sobre a influência da ordem de aparecimento da *ETD* em relação à sentença matriz sobre a escolha de cada *ETD*.

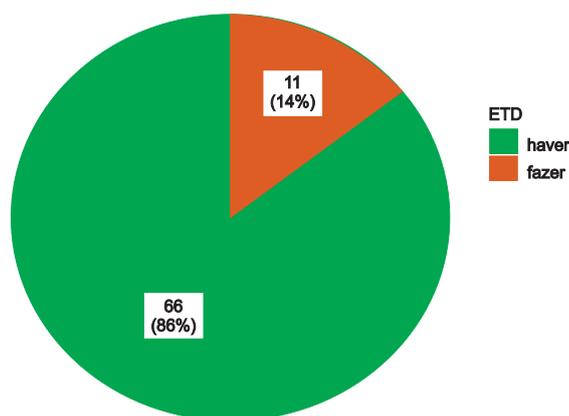
3. Resultados e discussões

O *corpus* analisado apresentou 77 realizações de *ETDs*, das quais 66 são *ETDs* com “haver” e 11 são com “fazer”. Nessa amostra de cartas pessoais de Graciliano Ramos, não foram encontradas realizações com “ter” e “estar com”, conforme mostra o gráfico 1, produzido no pacote *ggstatsplot*

(PATIL, 2018) do programa RStudio (R Core Team, 2021). Nele, é possível observar que, das 77 realizações, “haver” aparece com a maior parte (86%), enquanto “fazer” foi realizado apenas 14%. Esse resultado mostra-se significativo ($\chi^2 = 39,29$ $p < 0,001$), o que aponta para uma prevalência do verbo “Haver” na escrita do século XX.

Gráfico 1: Distribuição das ETDs nas cartas

$$\chi^2_{\text{gof}}(1) = 39.29, p = 3.66e-10, \hat{V}_{\text{Cramer}} = 0.71, \text{CI}_{95\%} [0.48, 0.93], n_{\text{obs}} = 77$$



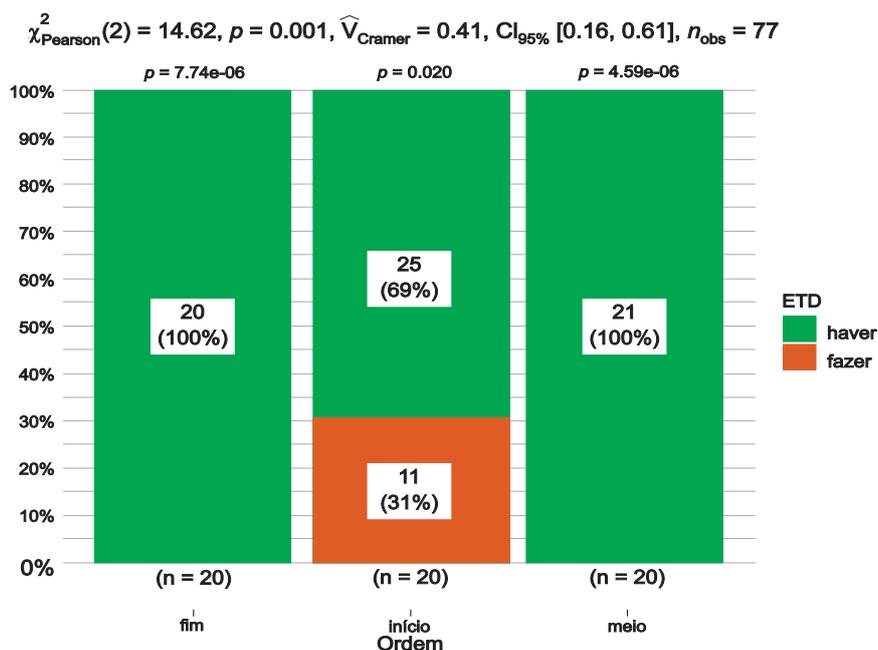
Fonte: elaboração dos autores

Os resultados vão ao encontro das hipóteses iniciais ao não encontrar as variantes “Ter” e “Estar com” na escrita do século XX. É possível que a influência da norma-padrão fosse muito forte naquela época ou, então, que essas duas variantes ainda não estivessem disseminadas na língua, já que Avelar (2012) encontrou a variante “estar com” em uma comunidade alagoana. Já em relação à variação de “Haver” com “Fazer” analisada por Vieira (2008), os dados mostram que “Haver” é muito mais usado na escrita do que “Fazer”, o que indica um processo de variação inalterado ao longo do tempo, em que nenhuma das formas foi suprimida.

Para análise da relevância das variáveis independentes, foram feitos testes de Fisher —tipo de teste de significância (negação da hipótese nula) que indica a relação entre variáveis categóricas dentro de uma tabela de contingência e é adequado para amostras com poucos dados — no RStudio para cada VI (linha), considerando como VD (coluna) a ETD. Nas análises, somente *ordem* ($p = 0.0004159$), *complementizador* ($p = 7.376e-08$) e *aspecto* ($p = 0.04364$) se mostraram estatisticamente relevantes. No caso das VIs *período*, *tempo*, *especificidade* e *destinatário*, não houve relevância estatística, visto que os testes de Fisher resultaram em um valor- $p > 0,05$ para todas. Apesar disso, vale reportar que as VIs *tempo* e *especificidade* podem ser melhor exploradas em *corpora* com mais dados devido ao valor- p das duas ser próximo de 0,05 (0,07943 e 0,09172, respectivamente) e esse valor ser influenciado pelo tamanho amostral.

Em relação à correlação entre o tipo de ETD e a ordem de ocorrência (Gráfico 2), a variante “Fazer” foi realizada apenas no contexto “Início” ($p > 0,02$), como (24) e (25), mas a baixa quantidade de realizações — apenas 11 — dificulta uma afirmação clara — o próprio X^2 é baixo (14,62), apesar da associação ser relevante (V^2 de Cramer⁵ = 0,41). Já a variante “haver”, com 66 realizações, não aparenta ser motivada por algum dos contextos analisados, aparecendo no gráfico com relativamente a mesma frequência no “Início” (25 realizações), “Meio” (21 realizações) e “Fim” (20 realizações), como (26), (27) e (28), respectivamente.

Gráfico 2: Distribuição de uso das ETDs em função da ordem



Fonte: elaboração dos autores

- (24) *Faz quase duas semanas* que não faço nada – nunca estive tão burro. (Palmeira dos Índios, 1914).
- (25) *Faz mais de uma semana* que Marcos Valente e Lima Filho estão na igreja, feito dois malucos, ouvindo umas rezas que nunca mais se acabam. (Palmeira dos Índios, 1914).
- (26) *Há mais de um mês* que aqui estou e creio que me tenho dado bem com o clima do sertão. (Maniçoba, 1911).
- (27) Estive, *há dias*, a palestrar com essa criatura que te prende. (Palmeira dos Índios, 1914).
- (28) A culpa é do padre. Sim, do padre, um padre ordenado por mim *há pouco*, em uma novela que me tomou alguns dias e que há de ter por aí um êxito considerável. (Rio de Janeiro, 1915).

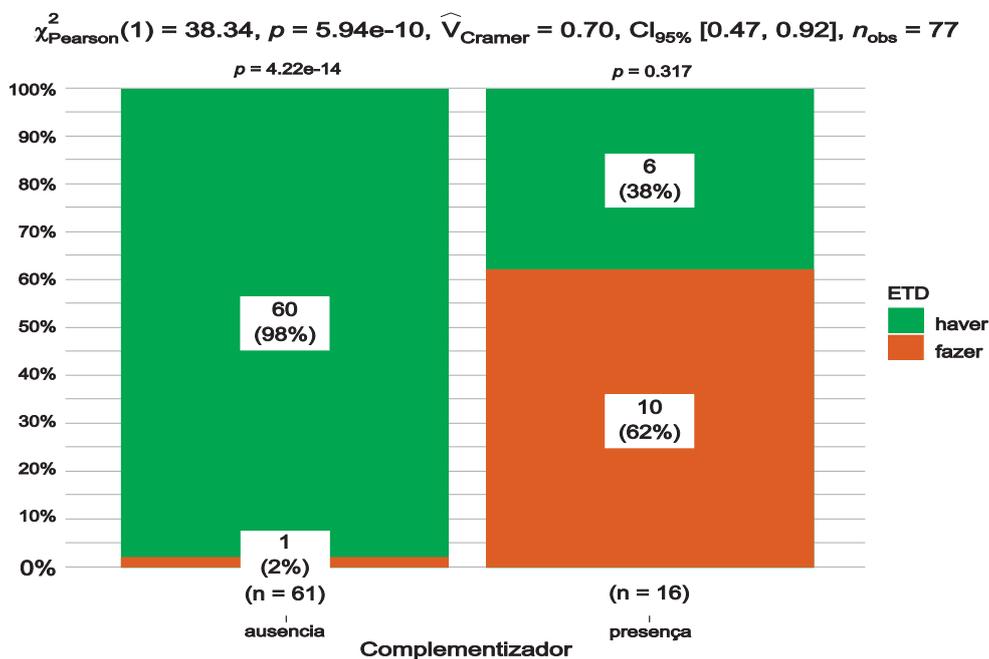
⁵ O teste V^2 de Cramer calcula a associação entre variáveis nominais, com um valor que vai de 0 a 1. Quanto mais próximo de 1, mais associação, e, quanto mais próximo de 0, menos associação.

Os dados apontam que “Haver” possui mobilidade em todos os contextos (“Início”, “Meio” e “Fim”), enquanto “Fazer” foi categórico em “Início”. Esses dados confirmam as conclusões de Vieira (2008) acerca do caráter móvel de “Haver” semelhante a um modificador circunstancial, o que pode ser ressaltado por Avelar (2011, 2012, 2018) ao não encontrar as propriedades verbais nessa ETD. Entretanto, a comparação dos dados sobre “Fazer” fica impossibilitada porque a autora engloba num mesmo bloco os dados de “Ter” e “Fazer”.

Já o gráfico 3 apresenta a relação entre as ETDs e a presença do complementizador “que”, como em (29), (30), e (31). Ao analisar os dados, foi encontrado que a ETD com “fazer” é realizada significativamente mais com a presença do complementizador (10 dos 11 dados encontrados), enquanto a ETD com “haver” apresenta variação na presença ou ausência, apesar de a ausência do complementizador ser muito mais evidente (60 dos 66 dados), o que vai de acordo com Vieira (2008), que reporta a prevalência de “Haver” sem “que” em comparação ao bloco “Fazer/Ter”, que é muito favorecido no contexto de presença do complementizador.

Além disso, outros resultados relevantes são o V^2 de Cramer, que se aproxima de 1 (muita associação), e o valor do qui-quadrado, 38,34, o que mostra uma diferença entre os valores esperados e os observados. Esses dados favorecem a aceitação de que há influência da presença do complementizador “que” na realização de cada tipo de ETD. Assim, observa-se que “Haver” é mais favorecido sem a presença do complementizador “que”.

Gráfico 3: Distribuição do uso das ETDs em função da presença ou ausência do complementizador

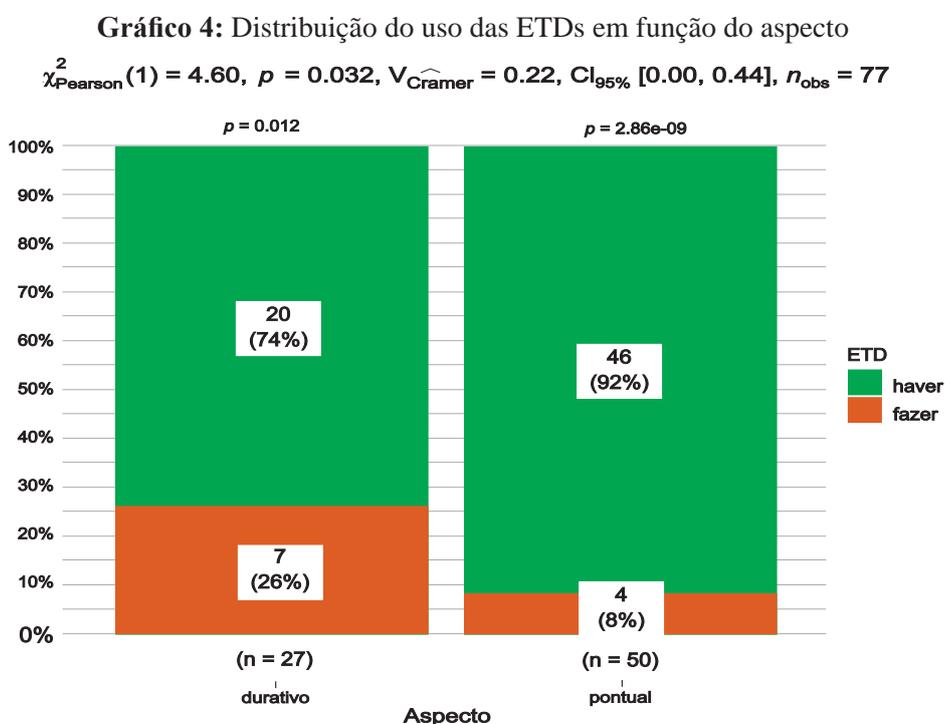


Fonte: elaboração dos autores

- (29) *Há coisa de uma semana* apareceu-me no Café do Rio, com o álbum de uma poetisa, a pedir-me que escrevesse nele uns versos. (Rio de Janeiro, 1915).

- (30) Pinto já não mora com o Rodolfo. *Há quase uma semana que dorme em meu quarto e come em casa do Costa Rego.* (Rio de Janeiro, 1915).
- (31) *Faz bem um mês que o Valdemar me mandou uma carta com a reportagem das festas da paz.* (Palmeira dos Índios, 1932).

Na análise da relação entre as ETDs e a VI *aspecto* (Gráfico 4), como (32), (33), (34) e (35), observa-se uma preferência significativa ($p < 0,001$) do aspecto pontual na ETD com “haver” – 92% versus 74% no aspecto durativo, enquanto, na ETD com “fazer”, há um aumento de dados no aspecto durativo – 26% em relação ao aspecto pontual – 8%, o que pode indicar uma tendência das ETDs com “fazer” aparecerem em contextos de eventos em curso no tempo. A pouca quantidade de dados por célula na ETD com “fazer”, entretanto, enviesava a interpretação do teste de qui-quadrado, pois ele precisaria de pelo menos cinco valores por célula.



- (32) Creio que v. ainda não tinha recebido uma que lhe mandei *há alguns dias*, com dinheiro. (Palmeira dos Índios, 1932).
- (33) *Há dois anos* você estava em Pilar, comendo bagre. (Palmeira dos Índios, 1932).
- (34) Uma crise de chateação medonha: *faz mais duma semana* que não consigo fabricar nada, julgo-me uma besta e não me animo a sair de casa. (Rio de Janeiro, 1937).
- (35) *Faz poucos dias* que te escrevi. (Rio de Janeiro, 1914).

No enunciado (32), observa-se que o aspecto é pontual, visto que a ação de mandar as cartas foi um evento num ponto específico do tempo, enquanto o dado (33) denota um evento contínuo (você estava em Pilar durante um tempo e não num momento específico). Por fim, (34) apresenta o aspecto durativo, visto que o período de ócio de Graciliano permanece por vários dias, e (35) tem o aspecto pontual, pois o autor se refere a um ponto específico no tempo em que escreveu à outra pessoa (nesse caso, sua irmã Leonor).

Considerações finais

O presente trabalho objetivou descrever a variação sintática existente nas ETDs a fim de verificar como essa variável se comporta em cartas pessoais escritas por um alagoano ilustre na primeira metade do século XX. Nesse sentido, analisou-se um *corpus* escrito de 112 cartas pessoais de Graciliano Ramos, visto que o autor é nascido e criado na região de Alagoas.

Os resultados obtidos mostram apenas as realizações das ETDs com “haver” e “fazer”, com a predominância da variante “haver” – 86% *versus* 14% da variante “fazer”, sem menção às variantes “ter” e “estar com”, que apareceram nas amostras alagoanas atuais (AVELAR, 2012). Essa diferença do século XX para o século XXI mostra-se relevante ao indicar uma possível entrada de variantes linguísticas inovadoras provindas da fala não padrão e de comunidades afrodescendentes, como aponta Avelar (2012).

As variáveis *aspecto*, *ordem* e *complementizador* mostraram-se relevantes estatisticamente e revelaram contextos de aparecimento das ETDs, com a variante “fazer” sendo mais frequente nos seguintes contextos: contexto “início”, presença do complementizador “que” e aspecto durativo. Apesar da baixa quantidade de dados, já há um panorama preliminar de como se comportam as ETDs em variados contextos sintáticos escritos, o que pode direcionar pesquisas futuras e incrementar dados sobre a história do português brasileiro.

Novos estudos precisam ampliar a quantidade de dados acerca desse fenômeno para analisar se essa realidade da escrita de Graciliano é semelhante à de outras cartas pessoais de Alagoas. Outras pesquisas podem mostrar também se “Ter” e “Estar com” são traços da língua falada (e avaliar a influência da norma-padrão nessa escolha) ou se aparecem em outros textos escritos, assim como analisar os contextos de aparecimento dessas variantes. Além disso, a análise de outros *corpora* pode revelar novos contextos de uso das variantes ou confirmar/refutar as conclusões apresentadas no trabalho, como a influência da ordem e do aspecto no uso das ETDs e na presença do complementizador “que” ligado a “Haver”.

Referências

ALI, Said. *Dificuldades da Língua Portuguesa*. 5. ed. Rio de Janeiro: Livraria Acadêmica, 1957.

AVELAR, Juanito Ornelas. Expressões de Tempo Decorrente com *ter* e *haver* na fala carioca. *Diadorim*, Rio de Janeiro, v. 8, n. 1, pp. 161-80, 2011.

- AVELAR, Juanito Ornelas. Expressões Possessivo-existenciais de Tempo Decorrente na Fala dos Quilombolas de Muquém. *Stockholm Review of Latin American Studies*, Estocolmo, n. 8, pp. 65-82, 2012.
- AVELAR, Juanito Ornelas. Sentenças Possessivas e Existenciais. In: CYRINO, Sonia; MORAIS, Maria A. Torres (Coords.). *Mudança Sintática do Português Brasileiro: perspectiva gerativista*. São Paulo: Contexto, 2018.
- BERLINCK, Rosane de Andrade; BARBOSA, Juliana Bertucci; MARINE, Talita de Cássia. Reflexões Teórico-metodológicas sobre Fontes para o Estudo Histórico da Língua. *Revista da ABRALIN*, v. 7, n. 2, pp.169-95, jul./dez. 2008.
- BECHARA, Evanildo. *Gramática Fácil*. Rio de Janeiro: Nova Fronteira, 2014a.
- BECHARA, Evanildo. *Lições de Português pela Análise Sintática*. 19. ed. Rio de Janeiro: Nova Fronteira, 2014b.
- CASTILHO, Ataliba Teixeira. *Nova Gramática do Português Brasileiro*. São Paulo: Contexto, 2010.
- CONDE-SILVESTRE, Juan Camilo. *Sociolingüística Histórica*. Madrid: Gredos, 2007.
- CUNHA, Celso; CINTRA, Lindley. *Nova Gramática do Português Contemporâneo*. 7. ed. Rio de Janeiro: Lexikon, 2017.
- FERRAREZI JR., Celso. *Semântica*. São Paulo: Parábola, 2019.
- HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE-SILVESTRE, Juan Camilo. *The Handbook of Historical Sociolinguistics*. UK: Blackwell Publishing Ltd, 2012.
- HERNÁNDEZ-CAMPOY, Juan Manuel; SCHILLING, Natalie. The Application of the Quantitative Paradigm to Historical Sociolinguistics: problems with the generalizability principle. In: HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE-SILVESTRE, Juan Camilo. *The Handbook of Historical Sociolinguistics*. UK: Blackwell Publishing Ltd, 2012.
- LABOV, William. *Padrões Sociolingüísticos*. São Paulo: Parábola, 2008.
- MOURA, Denilda. *Resquícios de Palmares: o que uma comunidade quilombola nos diz*. Maceió: EDUFAL, 2009.
- NEVES, Maria Helena Moura. *Gramática de Usos do Português*. 2. ed. São Paulo: EDUNESP, 2011.
- PATIL, Indrajeet. *Visualizations with statistical details: The 'ggstatsplot' approach*. PsyArxiv, 2018. Doi:10.31234/osf.io/p7mku.
- PERINI, Mário. *Gramática Descritiva do Português Brasileiro*. Petrópolis: Vozes, 2016.
- R Core Team. *R: A language and environment for statistical computing*. Áustria: R Foundation for Statistical Computing, 2021. Disponível em: <https://www.R-project.org/>.
- RAMOS, Graciliano. *Cartas*. 8. ed. Rio de Janeiro, Record, 2011.
- ROCHA LIMA, Carlos Henrique. *Gramática Normativa da Língua Portuguesa*. 49. ed. Rio de Janeiro: José Olympio, 2011.

ROMAINE, Simone. *Socio-Historical Linguistics: its status and methodology*. New York: Cambridge University Press, 1982.

ROSA, Eliane. Sociolinguística Histórica. *Revista de Letras*, Curitiba, v. 17, n. 21, jul./dez. 2015.

SILVA, Rosa Virgínia Mattos. A Arte de Fazer o Melhor Uso de Maus Dados: estudos diacrônicos no português antigo. *Anais do 1º Encontro de Estudos Românicos*, v. 1, 1988. DOI: <http://dx.doi.org/10.17851/2447-0538.1.0.68-84>.

VIEIRA, Marcia dos Santos Machado. Haver, Ter ou Fazer na Expressão de Tempo Decorrido. In: RONCARATI, Claudia; ABRAÇADO, Jussara (orgs.). *Português Brasileiro II: contato linguístico, heterogeneidade e história*. 19. ed. Rio de Janeiro: EDUFF, 2008.

VOSGERAU, Dilmeire Sant'Anna Ramos; ROMANOWSKI, Joana Paulin. Estudos de revisão: implicações conceituais e metodológicas. *Rev. Diálogo Educ.*, Curitiba, v. 14, n. 41, pp. 165-89, 2014.

WICKHAM, Hadley. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag, 2016. Disponível em: <https://ggplot2.tidyverse.org>.

GRAMATICALIZAÇÃO DO VERBO VISAR

GRAMMATICALIZATION OF THE VERB 'VISAR'

Geovana Portela de Moura¹

Célia Márcia Gonçalves Nunes Lôbo²

Lúcia Regiane Lopes-Damasio³

RESUMO

Este artigo resulta de uma pesquisa desenvolvida no âmbito do Projeto para a História do Português Brasileiro – Mato Grosso (PHPB-MT) e teve como objetivo principal estudar o processo de mudança linguística das construções com o verbo *visar*. Neste estudo, foi mapeada a trajetória de gramaticalização das construções com o verbo *visar*, por meio de uma descrição analítica do comportamento sintático, semântico e pragmático do item à luz dos pressupostos teóricos de Gramaticalização, na perspectiva da Linguística Funcional. As ocorrências com o objeto dessa pesquisa foram catalogadas e analisadas a partir de uma amostra panocrônica, extraídas de jornais escritos e publicados no estado de Mato Grosso, nos séculos XIX, XX e XXI, a fim de captar a variação e comprovar a mudança histórica do fenômeno em estudo. Conclui-se que *visar* está num processo de gramaticalização, em que parte do uso mais concreto, como verbo pleno, significando “olhar”, para um uso mais abstrato, na formação de perífrase (*visar* + infinitivo) com marcação de volição.

PALAVRAS-CHAVE: Linguística funcional. Gramaticalização. Verbo *visar*.

ABSTRACT

This article is the result of a search carried out within the scope of the Project for the History of Brazilian Portuguese – Mato Grosso (PHPB-MT) and its main objective was to study the process of linguistic change in constructions with the verb *visar*. In this study, the grammaticalization trajectory of constructions with the verb *visar* was mapped, through an analytical description of the syntactic, semantic and pragmatic behavior of the item in the light of the theoretical assumptions of Grammaticalization, from the perspective of Functional Linguistics. The occurrences with the object of this research were cataloged and analyzed from a panchronic sample, extracted from newspapers written and published in the state of Mato Grosso, in the 19th, 20th and 21st centuries, in order to capture the variation and prove the historical change of the phenomenon under study. It is concluded that *visar* is in a grammaticalization process, in which it starts from a more concrete use, as a full verb, meaning “to look”, to a more abstract use, in the formation of periphrasis (*visar* + infinitive) with volition marking.

¹ Mestra em Estudos Linguísticos pela Universidade Federal do Mato Grosso. Docente da Secretaria de Educação de Mato Grosso, prof.giportela@gmail.com, <https://orcid.org/0000-0002-1539-8512>.

² Doutora em Estudos Linguísticos pela Universidade Federal de Goiás. Docente do Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, celiamarciagn@gmail.com, <https://orcid.org/0000-0001-8811-814X>.

³ Doutora em Linguística pela Universidade Estadual Paulista - Júlio de Mesquita Filho. Docente da Universidade Estadual Paulista Júlio de Mesquita Filho, l.damasio@unesp.br, <https://orcid.org/0000-0002-9058-3566>.

KEYWORDS: Functional linguistics. Grammaticalization. Verb *visar*.

Introdução

Esta pesquisa se respalda no entendimento de língua como unidade concreta, histórica e em constante “fazimento”, assim como apregoa Coseriu (1979). Nosso estudo fundamenta-se, sobretudo, nos moldes do funcionalismo desenvolvido pela vertente americana, que tem como principais nomes os de linguistas como Givón (1979; 1983), Hopper (1991), Hopper e Traugott (1993), entre outros. Ademais, amparamo-nos nos postulados de autores pertencentes ao grupo funcionalista alemão (LEHMANN, 1988; HEINE *et al.*, 1991), uma vez que discorrem especificamente sobre os verbos, objeto de análise deste trabalho.

Utilizaremos, desta corrente linguística, o arcabouço teórico da Gramaticalização (GR, daqui em diante), definido, genericamente, como um processo linguístico no qual um item/construção lexical adquire caráter gramatical, ou, se gramatical, passa a ser mais gramatical, com o objetivo de descrever o comportamento sintático, semântico e pragmático de *visar*, com o significado de “ter por objetivo”, em processo de mudança.

Nossa investigação amparou-se pelas seguintes perguntas de pesquisa: 1) Qual é o processo histórico de mudança que confirma a descrição sincrônica dos usos de *visar*? 2) O princípio da unidirecionalidade⁴ é validado nesse processo de mudança perpassado por *visar*? 3) Quais os contextos que propiciam as mudanças de categoria e de sentido de *visar*?

Nesse sentido, os processos de GR do verbo *visar* foram analisados numa perspectiva panocrônica (conjugação de dados sincrônicos e diacrônicos). O resultado dessa análise será exposto na seção 5 deste artigo. Antes, porém, se faz necessário apresentar os pressupostos teóricos e metodológicos que subsidiam tal análise, o que é disposto nas seções de 1 a 4.

1. Breves considerações sobre gramaticalização

A gramaticalização consiste em um tipo especial de mudança em que, a depender da abordagem (clássica ou construcional), se reconhece quando um item/construção lexical se torna um item/uma construção gramatical, ou quando um item/construção gramatical se torna ainda mais gramatical, podendo mudar de categoria sintática, receber propriedades funcionais na sentença, adquirir alterações semânticas e fonológicas (cf. HOPPER, 1991; TRAUOGOTT e TROUSDALE, 2013).

Para Hopper e Traugott (1993), os itens lexicais são utilizados para relatar ou descrever as coisas, as ações e as qualidades; já os itens gramaticais são aqueles que estabelecem relações entre nomes (função das preposições); articulam os discursos (função dos conectivos); designam se coisas ou seres do texto já foram ou não identificados (atribuição dos pronomes e artigos); e revelam se essas coisas ou seres estão próximos ou distantes do falante ou do ouvinte (papel desempenhado pelos

⁴ A unidirecionalidade prevê que as mudanças linguísticas ocorrem num continuum, “[...] operadas sempre da esquerda para a direita e, nesse caso, de categorias cognitivas mais próximas do indivíduo, [+concretas], para categorias cognitivas mais distantes do indivíduo [-concretas]” (GONÇALVES *et al.*, 2007, p. 40).

demonstrativos). Nesse sentido, uma forma gramaticalizada é aquela que um item/construção lexical assume quando passa a desempenhar funções de um item/construção gramatical.

Nessa mesma perspectiva, Heine *et al.* (1991) apresentam os conceitos: fontes e gramaticais. Os conceitos fontes referem-se a elementos caracterizados por lexemas, que possuem significação própria e codificam objetos concretos, processos, localizações e dêixis. Os conceitos gramaticais estão ligados a elementos mais abstratos, codificados por formas não lexicais, como auxiliares, partículas, clíticos, afixos etc., e caracterizam-se por serem determinados na estrutura de uso.

Numa visão diacrônica, o processo de GR assume, necessariamente, um caráter gradual, no qual o estudo destina-se a explicar como as formas gramaticais aparecem e como passam a exercer funções na língua. Numa perspectiva sincrônica, assume o objetivo de identificar os graus de gramaticalidade que uma forma linguística desenvolve, estando, portanto, o caráter gradual também presente nessa perspectiva.

Abordando sobre esse assunto, Neves (2004, p. 118) assegura que a GR pode ser compreendida numa posição pancrônica, na qual “acentua a interdependência entre sistema linguístico e uso, entre natureza fluida da gramática e a importância da história para a compreensão da gramática sincrônica”, sendo esse o enfoque que escolhemos para o estudo do fenômeno de GR eleito. A complementação entre as duas perspectivas de análise, mais do que observar a relação entre sistema e uso, permite comprovar um entendimento de língua concreta e histórica, que se encontra em constante “fazimento”, nos termos de Coseriu (1979).

2. Heine: a mudança de pleno para auxiliar

É pertinente apresentarmos as contribuições teórico-metodológicas de Heine (1993, pp. 58-66) para o paradigma teórico da GR, no que tange, especificamente, às motivações para a formação de verbos auxiliares, objeto deste estudo. Sete estágios são propostos por esse autor para a GR de verbos, a saber:

1º estágio: o verbo tem a sua raiz no concreto, significado lexical pleno, e seus objetos também se referem a objetos concretos.

2º estágio: dá-se o início do processo de auxiliarização, em construções em que aparecem complementos nominais ou complementos formados por verbos não-finitos. Neste estágio, para Heine, os itens, em processo de GR, associam-se sempre a outros itens que são classificados, de forma equivocada, como verbos. As seguintes características costumam ser observadas nesses casos: (i) os itens verbais podem se associar a verbos não-finitos ou terem complementos nominais; (ii) pode não ser identificada a identidade do sujeito para os dois verbos; (iii) várias construções concorrem na complementação do verbo, como as de infinitivo, de gerúndio ou de particípio; (iv) a oração com verbo finito também pode ser o complemento do verbo.

3º estágio: o verbo, em processo de GR, neste estágio, passa a desempenhar outras funções, podendo, dessa forma, ter referentes não-humanos, isso porque eliminam-se as restrições de seleção

e os verbos passam a desempenhar funções como as de tempo, modo ou aspecto. Unido ao verbo principal, passa a constituir uma unidade semântica, tornando-se conhecido como semiauxiliar.

4º estágio: neste estágio, os verbos em GR perderam sua categoria, ou seja, foram decategorizados e apresentam as seguintes características: (i) perdem a capacidade de formar imperativos, de ser nominalizados ou de se apassivar; (ii) os núcleos de seus complementos já não são os nomes; (iii) na formação de perífrase, associam-se apenas à forma nominal não-finita.

5º estágio: nesta etapa, o item já não é mais classificado gramaticalmente como verbo, sofre erosão, perde o estatuto de palavra e substância fonológica e/ou morfológica. Neste estágio, portanto, as seguintes características podem ser percebidas: (i) não são classificados como pertencentes à categoria dos verbos; (ii) separadamente, perdem a possibilidade de serem negados; (iii) ocupam uma posição fixa dentro da oração; (iv) em algumas situações, podem ser um híbrido, como marcador gramatical, resultante dos vestígios de quando eram verbos; (v) podem ser cliticizados e/ou sofrerem erosão:

6º estágio: já se encontra fixado firmemente, mas possui, ainda, vestígios de sua origem. Passa a ser visto como clítico ou afixo, sendo acompanhado de um verbo, denominado como principal.

7º estágio: estágio final da GR, em que o item perde toda e qualquer característica de verbo e torna-se um marcador gramatical com a forma de um afixo flexional sem tom ou acento distintivo próprio.

Geralmente, os verbos podem apresentar características de mais de um estágio numa determinada fase da GR. Passar por todos os estágios não é um requisito para consolidar o processo, uma vez que, dificilmente, isso ocorre. Heine (1993) classifica os verbos, nestes estágios da GR, da seguinte forma: a) nos 1º e 2º estágios: lexemas, verbos plenos; b) no 3º estágio: quase-auxiliares, semi-auxiliares, concatenativos; c) nos 4º e 5º estágios: auxiliares; d) no 6º estágio: auxiliares ou afixos; e) no 7º estágio: afixos ou flexões. A proposta de Heine subsidiou a nossa análise e o entendimento do processo de GR do verbo *visar*.

3. Mecanismos do processo de GR

Em relação aos mecanismos e motivações que envolvem o processo da GR, vale ressaltar: (i) aqueles associados à mudança semântica, na qual estão contemplados os processos metafóricos na perspectiva de Heine *et al.* (1991) e Sweetser (1988, 1991); o princípio da iconicidade; e o processo metonímico; bem como (ii) aqueles relacionados à mudança morfossintática, em que se destacam os conceitos de reanálise e analogia, atuando no processo de mudança gramatical.

A iconicidade é um princípio geral e, por isso, aparece como explicação para vários fenômenos linguísticos. As relações icônicas, segundo Casseb-Galvão (1999), contrariam o princípio da arbitrariedade proposto por Saussure, pois há, no signo icônico, uma relação motivada. Dessa forma, Hopper e Traugott (1993) definem a iconicidade como a propriedade de similaridade entre um item e outro. Neves (1997) completa esse entendimento, afirmando que ela é um princípio que considera

a relação **forma e função**, ou **código e mensagem** na linguagem humana, que aconteceria de forma não-arbitrária.

Quanto ao entendimento da metáfora e da metonímia como mecanismos que ocorrem na mudança semântica, partindo da concepção semântico-cognitivista conforme Sweetser (1988) e Heine *et al.* (1991), compreendemos que a metáfora opera nos domínios conceituais e paradigmáticos, é um mecanismo motivador da GR que se associa à abstratização de significados, ou seja, atua diretamente no processo de deslizamento do domínio lexical para gramatical, ou de gramatical para mais gramatical, pois atua como veículo na transferência de significados de um domínio mais concreto para um mais abstrato.

Aproxima-se da metáfora o princípio da iconicidade, já que a metáfora é icônica, uma vez que sua atuação se baseia na semelhança ou no compartilhamento semântico entre o significado inicial e o derivado. Já a metonímia ocorre em situações mais específicas, no eixo sintagmático, e promove uma extensão gradual do significado numa direção única e contínua.

Ambas, metáfora e metonímia, devem ser compreendidas atuando de forma complementar. Num trajeto unidirecional, sinalizam o aumento da abstratização do item/construção na passagem do menos para o mais gramatical. A metáfora resolve um problema de representação, e a metonímia, problemas de informatividade e relevância na comunicação.

No processo de GR, encontramos, também, outros mecanismos envolvidos: a reanálise, que permite a criação de formas gramaticais, atuando, assim como a metonímia, no eixo sintagmático, pois uma forma reanalisada passa a pertencer a uma categoria diferente de sua origem; e a analogia, que, como a metáfora, atua no eixo paradigmático, uma vez que a inovação que acontece naquela já está prescrita no sistema.

Partindo dessas observações, podemos entender que a GR envolve mecanismos que atuam nos diferentes níveis da linguagem, resultando na abstratização do item/construção.

4. Universo da pesquisa e procedimentos metodológicos

A elaboração deste artigo provém das discussões e das análises de resultados mais amplas constantes em Moura (2016). Para tal trabalho, adotamos como *corpus* os jornais publicados no estado de Mato Grosso. Nossos dados são baseados em dois recortes temporais: (i) sincrônico, século XXI; e (ii) diacrônico, séculos XIX e XX. Para cada recorte, uma forma de coleta do material linguístico ocorreu. Na composição do recorte do *corpus* diacrônico, a coleta ocorreu na Superintendência de Arquivo Público de Mato Grosso (SAP/MT), localizada em Cuiabá.

Nesse lugar, nosso trabalho foi amparado no acervo microfilmado que, além dos jornais que utilizamos, é composto por relatórios de Presidentes, de Governadores de Mato Grosso, mensagem de governo, avisos imperiais, livros do Banco do Estado de Mato Grosso (BEMAT), cartórios do 1º, 2º e 5º ofícios, coletoria de Cuiabá, correspondências, gazeta oficial e diário oficial, inventários e herança. O acervo dos jornais em microfilme é formado por edições de jornais pertencentes a Mato

Grosso desde 1840 a 1929, ou seja, antes da divisão do Estado, ocorrida em 11 de outubro de 1977 pelo presidente Ernesto Geisel, momento que se passou a ter também Mato Grosso do Sul. Essa informação torna-se relevante, pois, na escolha do material para a coleta do *corpus*, tivemos que prestar atenção à cidade a qual o jornal pertencia.

Foram pesquisados, ao todo, 16 jornais, dos quais oito constituíram o nosso *corpus*, a saber: “Themis Matogrossense” (1840), “A Gazeta Cuyabana” (1847 e 1848), “Echo Cuiabano” (1850), “O Mattogrosso” (1873), “O Liberal” (1871-1878), “A Imprensa” (1911), “O Republicano” (1920 e 1950), “A Cruz” (1930-1969). A escolha por esses jornais se deu pelo fator datação, com alternância de uma década entre as suas respectivas publicações.

Já em relação ao recorte do *corpus* sincrônico, ao todo, 06 jornais foram utilizados, a saber: “24 Horas News”, “Agora MT”, “Diário de Cuiabá”, “Gazeta Digital”, “G1/TVCA”, “Primeira Hora”. A escolha desses periódicos deve-se à sua relevância no estado de MT. Desses 06 jornais, obtivemos 405 ocorrências extraídas do gênero notícia, encontradas nas mais diversas seções.

A descrição e a análise das ocorrências do verbo *visar*, em seus contextos de uso, amparam-se numa conjugação das abordagens qualitativa e quantitativa, em perspectiva funcional, a partir de critérios semânticos e pragmáticos:

- (i) traços semânticos do sujeito: se [+animado][+humano], [-animado][-humano] ou [+animado][-humano];
- (ii) traços semânticos dos complementos internos: [afetado] ou [alvo];
- (iii) uso do *visar*: [+abstrato], [-concreto] ou [+concreto].

Além desses critérios, adotamos o de frequência proposto por Bybee (2007). Essa autora sugere dois métodos de contagem da frequência, denominados como frequência *token* e *type*. A primeira, diz respeito à frequência de ocorrência do item investigado, já a segunda refere-se aos padrões de uso desse item, especificados mediante os critérios de análise.

Para Bybee (2007, p. 338), o aumento da *frequência type* constitui um indício de GR de um item, uma vez que os itens gramaticais apresentam frequência alta, ao contrário dos lexicais. A autora afirma que a repetição frequente de uma forma tem papel importante na mudança, pois: (i) a frequência de uso leva ao enfraquecimento da força semântica pela habitualidade; (ii) à mudança fonológica através das reduções e fusões; (iii) ao enfraquecimento ou à perda de associação do item em GR com outros; (iv) à perda da transparência semântica, ocasionando a mudança de sentido do item; e (v) à perda de autonomia do item/construção, tornando-o mais enraizado na língua.

As abordagens qualitativa e quantitativa percorreram, basicamente, as seguintes etapas principais: (i) levantamento das ocorrências de *visar* em textos jornalísticos pertencentes ao gênero notícia, em contexto mato-grossense e em perspectiva sincrônica e diacrônica; (ii) descrição e análise do funcionamento sintático, semântico e pragmático dos usos do verbo *visar*, em dados do português

contemporâneo e dos séculos XIX e XX, com o controle de suas frequências *token* e *type*; (iii) a partir de (ii), proposição de um quadro ilustrativo dos diferentes usos desse item, ordenados do mais concreto até os mais abstratos, de acordo com um *cline*⁵ de aumento de gramaticalidade e GR, que enfatiza o aspecto da categoricidade de formas em GR; (iv) a partir de (iii), confirmação do princípio da unidirecionalidade e, também, dos contextos que teriam propiciado as mudanças de categoria e sentido; e (v) proposta de relações entre os resultados das descrições sincrônica e diacrônica e o processo de GR de *visar*.

Para esse trabalho, reportamo-nos a um levantamento etimológico de “visar” e, com base nessa pesquisa, percebemos que o verbo *visar* tem a sua origem numa acepção mais concreta, “dirigir o olhar para lançar algo numa direção”, e, a partir dela, adquire outros sentidos, menos concretos, como “ter por objetivo”, revelando usos cada vez mais abstratizados, como podemos observar nos exemplos de (01) a (05), abaixo, em que o primeiro representa um uso de *visar* mais concreto e o último, por conseguinte, um mais abstrato:

- (01) As crianças visavam a tela do cinema atentas.
- (02) O arqueiro concentrado visava o alvo.
- (03) O caçador visava com a arma.
- (04) O time visava ao título de campeão.
- (05) O projeto visa beneficiar a sociedade.

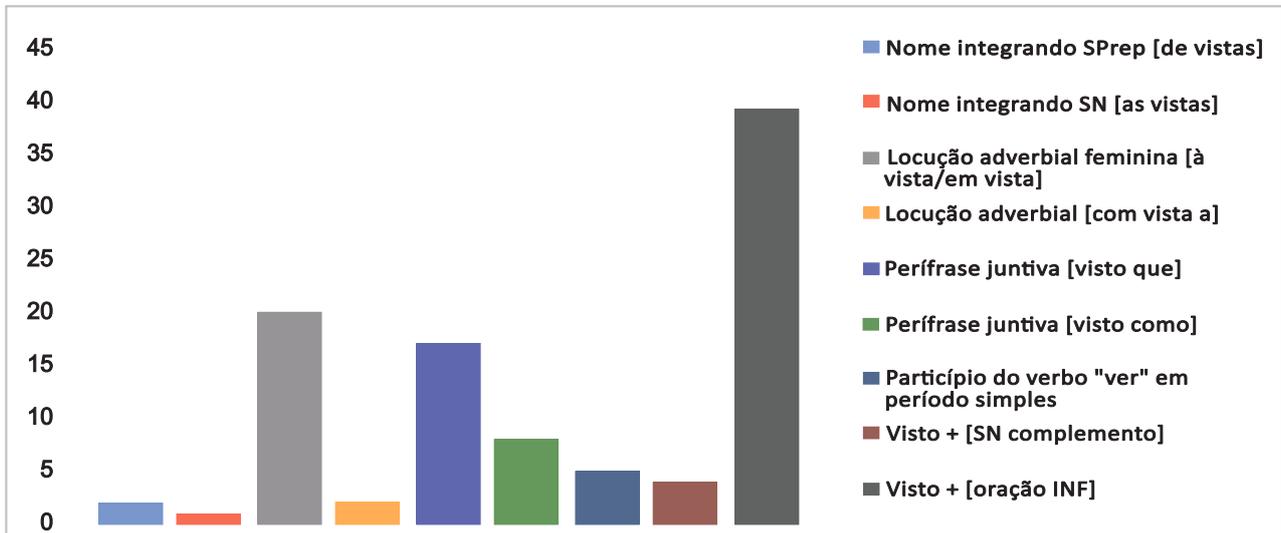
Em (01), *visar* aparece no sentido de “olhar fixamente”, em (02) e (03) adquire o sentido de “mirar”. Nos exemplos em (01) e (02), *visar* é um verbo que recruta, como complemento interno, um objeto direto, ou seja, trata-se de um verbo transitivo direto. Em (03), temos um exemplo de *visar* como um verbo intransitivo, acompanhado de um adjunto de natureza instrumental. Em (04), é um verbo transitivo indireto, recrutando um objeto indireto como argumento interno, com acepção de “ter por objetivo, almejar”. Por fim, em (05), *visar* se apresenta na construção, *visar* + *infinitivo*, como um verbo auxiliar, agregando sentido de futuridade. Essa construção deriva de seu uso com argumento interno oracional, no qual apresenta o sentido de “ter por objetivo, meta”.

A pesquisa à amostra dos jornais mato-grossenses do período entre 1840 a 1911 revelou baixa frequência de usos do verbo *visar*, apenas 3. Diante desse resultado, voltamos o nosso olhar para construções que traziam termos relacionados à sua etimologia. Nessa perspectiva, chamamos a atenção para os usos, destacados, do verbo *ver*; relacionados aos de *visar* pela proximidade de sentido, tais como: (i) dirigir à vista ou olhar fixamente para; e (ii) mirar. Esses usos mostraram-se recorrentes,

⁵ *Cline* ou *continuum* é a indicação de que a mudança ocorre de forma unidirecional e gradual. Pressupõe, única e exclusivamente, a noção de categorias não discretas, que refletem a gradualidade sincrônica e diacrônica da mudança experimentada via GR, eliminando, portanto, qualquer espécie de separação clara entre itens/construções lexicais e gramaticais, suas categorias e funções (LOPES-DAMASIO, 2008, p. 20).

em diferentes codificações sintáticas, no *corpus* investigado, nesse período que as ocorrências com *visar* foram poucas, em que as ocorrências de *visar* foram encontradas com baixa frequência. A seguir, o gráfico 1, apresenta a sistematização dessas frequências:

Gráfico 1: Frequência de ocorrências de *visto*



Fonte: elaboração das autoras

Comparando as ocorrências de *visto* com as de *visar*, encontradas no mesmo recorte temporal, em meados do século XIX e início do século XX, percebe-se a grande diferença na frequência de uso, uma vez que a de *visto* ultrapassa 90% e *visar* fica abaixo de 10%.

A comparação entre as frequências *tokens* desses dois itens, no mesmo recorte temporal, corrobora a afirmação de Bybee (1999) sobre o valor da frequência no processo da GR. Em outras palavras, podemos sugerir que, nesse intervalo, o processo de mudança de *visar* ainda não estava implementado, o que encontra justificativa, ainda, na natureza das três ocorrências de *visar*, encontradas nesse ínterim, associadas ao sentido mais concreto e à categoria de verbo pleno, assim como pode ser verificado em (06), mais adiante.

A correlação das acepções de *visar* com os usos diacrônicos de *visto* sugere-nos a hipótese de que, no momento em questão, os usos mais gramaticalizados de *visto*, especificamente como em: “Pelas razões apresentadas na discussão ao Snr. Presidente por á votos, se a questão era objeto de deliberação, foi resolvido pela negativa **visto estar** já providenciado pela Lei Provincial accima citada” [TMT-01-04-1840], apresentam coerência categórico-funcional com aqueles que serão constatados em relação a *visar*, em período subsequente. Em outras palavras, tanto *visto* quanto *visar* ocorrem em contexto de orações infinitivas, formando uma construção em que sua função é de auxiliar. Dada essa coerência, explica-se a frequência elevada da construção com *visto* em detrimento da frequência da com *visar*, no recorte temporal.

5. Os *types* de “*visar*” numa abordagem panocrônica

Com o objetivo de ilustrar os estágios de GR, empreendemos uma apresentação dos usos do verbo *visar* em padrões funcionais, de acordo com os resultados analíticos em recorte diacrônico e sincrônico. Segundo Bybee (2003), essa metodologia garante o controle mais adequado da produtividade/frequência dos usos, bem como, a partir desse controle, possibilita enxergar possíveis indícios de GR.

5.1. *Visar* 1: [SN (+animado/+humano) *visar* (VTD) (olhar) SN (afetado)]

O *Visar* 1 é um verbo pleno, pois é núcleo de predicação. Esse verbo é transitivo direto, ou seja, não há preposição entre ele e o seu complemento. Seleciona dois argumentos obrigatórios: um argumento externo, sujeito agente, com os traços [+animado] e [+humano] que exerce uma atividade: “olhar”, tendo assim, um sentido [+concreto]; e um argumento interno, o complemento afetado (aquilo que é visado/olhado) que é um sintagma nominal, assim como podemos perceber no exemplo:

- (06) Um superior como o que fica debuxado, no quadro precedente e por entre cujas sombras não deixarão os leitores de **visar** o vulto angelico, majestoso e imponente, desse ancião venerável, no qual vê o clero cuiabano seu chefe, os súbditos obedece, adorão, idolatrão em cumprem sem exame e sem reflexão, as ordens e determinações, por que sabem que ellas não podem deixar de trazer o cunho da justiça [OL-19-02-1874].

Essa é a única ocorrência do *Visar* 1 encontrada em nosso *corpus* diacrônico, representando 1,42% do total das ocorrências de *visar* nesse recorte. Destacamos, portanto, que não constatamos ocorrências desse *type* nos dados investigados em perspectiva sincrônica, o que já nos permite sugerir que, com o desenvolvimento de outras novas funcionalidades, com base em novos traços semântico-sintáticos, esse primeiro uso, fonte, já não se reflete no português contemporâneo, obviamente, no recorte (tipo de texto) que investigamos.

5.2. *Visar* 2: [SN (+animado/+/-humano) *visar* (VTI) (ter como objetivo) SP (alvo)]

Este *type*, assim como o anterior, tem o seu uso prescrito pelos dicionários e gramáticas da Língua Portuguesa. Vejamos abaixo a ocorrência deste *type* encontrada em nosso *corpus* diacrônico:

- (07) O pai de família sendo exemplo, há de admoestar com amor, mostrando que **visa** ao bem dos seus. [AC-02-11-1958]

Das 70 ocorrências de construções diacrônicas com *visar*, 3 pertencem a este *type*, num total de 4,28%. Não foi encontrada, na sincronia, nenhuma ocorrência referente a esse *type*.

5.3. *Visar* 3: [SN (+animado/+/-humano) *visar* (VTD) (dirigir a vista para) SN (afetado)] e [SN (+animado/+/-humano) *visar* (VTD) (ter como objetivo) SN (alvo)]

Neste *type*, encontra-se a co-ocorrência de sentidos de *visar*: [+concreto] e [-concreto]. Os usos de *visar* significando “dirigir a vista para” representam uma acepção que está diretamente

ligada ao mundo bio-psíquico-social, então, mais concreta⁶. Classificamos neste *type*, também, a característica [-concreto] ao sentido do *visar*, significando, assim, “ter com objetivo”. Na diacronia, foram encontradas 4 (5,71%) ocorrências em configurações com sujeito [+animado] e [+/-humano] e SN complemento alvo – portanto, sem preposição entre o verbo e seu complemento –, tal como podemos observar abaixo:

- (08) Reuniram-se, na última noite de maio, trabalhadores com Dom Jorge Marcos, Bispo de Santo André, para lançar, no Sindicato dos Metalúrgicos, as bases da Frente Nacional Trabalhista Cristã, que **visa** uma nova liderança sindical, a fim de combater os grupos extremistas e pelegos sindicais. [AC-30-06-1960]

Já no *corpus* sincrônico, 07 ocorrências enquadram-se neste *type*, ou seja, apenas 1,73%. Essas aparecem em apenas dois dos seis jornais que constituem a amostra sincrônica investigada, por isso, o uso pode ser interpretado como uma escolha estilística do escrevente:

- (09) O objetivo do atleta mato-grossense vai além de seu aperfeiçoamento e preparação física, o atleta **visa** a expansão do judô, que para ele é muito mais que um esporte. [GD/Ev03-02-2014]

Neste *type*, encontramos, também, *visar* desempenhando uma função ambígua, na qual o item estudado apresenta o sentido mais concreto de “olhar para”, e, concomitantemente, pode, igualmente, sugerir o significado “ter como objetivo”, menos concreto. Em *Visar 3* e, especialmente, nessa ambiguidade, estão a comprovação histórica da hipótese de mudança de sentido no uso de *visar*. Vejamos:

- (10) [...] e dissuadindo-os de quaisquer infundados preconceitos quanto aos objetivos do Governo, que não **visa**, com a indagação projectada, o levantamento de novos impostos, nem a collecta de elementos que facilitem o sorteio militar. [OR-02-12-1920]

O verbo *visar*, nessa ocorrência, apresenta usos ambíguos, significando “olhar para e/ou ter por objetivo”. Heine e Reh (1984) denominam esse processo de *Split*. Dessa forma, tanto podemos interpretar: (i) [...] aos objetivos do governo, que não **olha**, com a indagação projectada, o levantamento de novos impostos [...]; (ii) [...] aos objetivos do governo, que não **tem por objetivo**, com a indagação projectada, o levantamento de novos impostos [...].

Como argumento externo, *visar* recruta o SN “o governo”, constitutivo do SP “aos objetivos do governo”, que pode ser classificado como [+animado] e [+/-humano], uma vez que se entenda por “governo” todas aquelas pessoas que governam – ou seja, há aí, ainda, características de extensão contextual, relativa ao traço semântico da animacidade. Esse sentido ambíguo ocorreu em 5 construções todas encontradas no *corpus* diacrônico, perfazendo, assim, 7,14% das ocorrências diacrônicas.

⁶ Em nosso *corpus* (sincrônico e diacrônico), não encontramos, em *Visar 3*, exclusivamente o sentido mais concreto, porém, esse uso é atestado em gramáticas e dicionários, como observa-se no exemplo retirado de Abreu (2006): “O contratorpedeiro Seafish manobrou o canhão de ré, visando a linha d’água do navio inimigo”. Percebe-se que *visar* aparece no sentido de dirigir o olhar para fazer pontaria, empregado no gerúndio, recrutando, como argumento externo, o SN “O contratorpedeiro Seafish” e, como argumento interno, o SN “a linha d’água do navio inimigo”.

Sweetser (1991) afirma que, na diacronia, o sistema metafórico é responsável pela emergência de formas polissêmicas, ou seja, a metáfora é responsável pela transferência de significados, na qual parte do significado concreto “olhar para” contribui para um significado mais abstrato “ter como objetivo”. Esse salto de significado se dá por meio da contiguidade, ou seja, a partir de contextos que permitem a constatação de ambiguidades por proximidade conceitual, via metonímia, capaz de captar a gradualidade desse processo de mudança semântica que acompanha e é imprescindível para o processo de GR de *visar*.

No caso do nosso objeto de estudo, *visar*, significando “olhar, mirar”, mais próximos à experiência física do falante, sendo assim mais concreto, passa por um estágio de polissemia, no qual acontece a mudança semântica que resulta em outra acepção do verbo, a de “pretender, almejar, ter por objetivo”, associada ao contexto do ato de fala, portanto, mais abstrata.

O gatilho dessa mudança acontece quando *visar* tem a acepção de “mirar”, como no seguinte exemplo: “(i) O arqueiro **visou** o alvo e atirou”. Percebe-se, a partir desta frase, com referentes pertencentes ao contexto concreto, que para “mirar”, faz-se necessário “olhar”, de maneira fixa, o alvo com o intuito de atingi-lo com alguma coisa, no caso do exemplo, a flecha. Através da manipulação conceitual, metafórica e metonímica, a aproximação de significados transfere para o verbo a ideia de projeção futura, uma vez que a ideia de “olhar” contida em *visar* depreende de um olhar sempre com um foco, uma meta.

Na categoria processo, temos *visar* significando “olhar”, mais concreto; na categoria tempo, sua acepção é de futuro, menos concreto; e em qualidade, ele será mais abstrato, significando “ter por objetivo”.

Sendo assim, *visar*, significando “mirar” (exemplo i), a partir de usos metafóricos, desenvolve uma nova acepção, significando polissemicamente “olhar / ter por objetivo” em decorrência dos contextos de uso. Com o tempo, porém, a partir da frequência de uso de *visar*, o uso mais concreto pode dar lugar ao novo, aquele mais abstrato ou as duas formas poderão conviver sem necessariamente uma ser extinta.

5.4. *Visar* 4: [SN (+animado/+humano) *visar* (VTD) (ter por objetivo) Oração (alvo)]

A estrutura deste *type* é quase igual à do anterior. A diferença consistirá no complemento oracional que, neste caso, não aparece preposicionado. Apenas dois dos dicionários pesquisados trazem essa possibilidade de uso, em que, com a acepção “ter como objetivo”, pode ser usada com essa regência. Porém, trazem esse uso, ainda, como nota da regência da acepção transitiva indireta, ou seja, como uma variação daquela. Vejamos, a seguir, a ocorrência extraída do *corpus* diacrônico, um exemplo prototípico desse *type*:

- (11) Não ha lei alguma o creando, ato siquer do Intendente autorizando a sua fundação; é um fructo espurjo da cachola do tenente Mello que só **visa** arrancar do povo mais alguns mil réis de cada rêz abatida. [OR-07-03-1920]

Já o exemplo, seguinte, diferencia-se do anterior, pois apresenta materiais intercalados entre o predicado e o argumento interno.

- (12) O operoso deputado Paes de Oliveira **visa**, por via do presente projecto de lei, limitar o numero de pedido de reconsideração no fôro administrativo [AC-12-10-1930]

A explicativa intercalada entre o verbo e seu objeto oracional, mostra-nos que ainda há um afastamento, tanto semântico quanto sintático entre esses dois elementos, o que indica não se tratar de uma construção “visar + infinitivo” altamente cristalizada.

As ocorrências que preenchem as características de *Visar 4*, na diacronia, correspondem a 5,71%, dentre as quais, formas, como a apresentada acima, compõem 33,33%, são exemplos não prototípicos deste *type* uma vez que apresentam material intercalado entre o predicado e o objeto.

Na sincronia, foram constatadas 07 ocorrências deste *type*, equivalentes a 1,73% de todas as ocorrências de *visar* no *corpus* nessa perspectiva:

- (13) Diamantino **visa** explorar seu charme histórico e cultural. [DC/Tur-15-09-2014]

Temos, nessa frase, um SN “Diamantino”, exemplo de oração com grau fraco de entrelaçamento do sujeito, pois apresenta os traços [+explicitude do sujeito] e [-correferencialidade]. Segundo Lehmann (1988), esse parâmetro indica que, quanto maior for o entrelaçamento dos sujeitos, no complexo oracional, maior o compartilhamento de elementos e mais as sentenças encontram-se num grau forte de integração sintática. A oração “explorar seu charme histórico e cultural” é o complemento interno de *visar*. A forma não finita dessa oração, como complemento do predicado, tal como já apresentamos em outras ocorrências, pode indicar, segundo Lehmann (1988), a GR do verbo da matriz. Em relação a essa ocorrência o argumento externo “Diamantino” recebe, também de acordo com o contexto, os traços de [+animado] e [-humano].

Neste *type*, assim como vimos em *Visar 3*, apareceram usos que co-ocorrem (4 ocorrências), ou seja, no mesmo trecho de texto, o autor utiliza mais de uma forma de empregar o verbo em estudo. Em outras palavras, além deste *type* (*Visar 4*), ele também emprega o *Visar 8* (como veremos mais adiante). Entendemos que este *type* enquadra-se no segundo estágio proposto por Heine (1993), em que *visar* associa-se a verbos não finitos e o sujeito de ambos os verbos é o mesmo.

5.5. *Visar 5*: [SN (-animado/-humano) *visar* (VTI) (ter por objetivo) SP (alvo)]

A partir deste *type*, o sujeito será classificado como [-animado] e [-humano]. Cezario (2001, p. 88), especificamente sobre o traço [-animado], assevera que “um ser inanimado como sujeito da cláusula subordinada aumenta a integração entre os eventos codificados na cláusula principal e na subordinada”, uma vez que os referentes inanimados não são capazes, normalmente, de quebrar a cadeia tópica, ou seja, a atenção não é desviada. Sendo assim, o tópico pode ser representado por uma anáfora zero, fato que dificilmente ocorrerá quando se tratar de referentes animados. Heine *et al.* (1991), em relação ao que esse traço revela sobre o processo GR, afirmam que, se existirem duas

categorias e a diferença entre elas for a animacidade do sujeito, aquela que for [-humano] poderá ser considerada mais gramaticalizada que aquela categoria com o sujeito [+humano]. Utilizamos essa compreensão para diferenciarmos os padrões subsequentes, que apresentam esse traço, dos anteriores.

Sendo assim, este *type* apresenta os traços semânticos do argumento externo: [-animado] [-humano], o que faz com que esse *type* se encontre no terceiro estágio do processo de GR proposto por Heine (1993). Para o autor, isso ocorre devido às restrições de seleção adquiridas pelo verbo. O predicado, num sentido mais abstrato, com a acepção “ter por objetivo”, aqui, une-se ao seu argumento externo, passando a constituir uma unidade semântica. Os complementos do verbo, neste *type*, são SPs alvos da ação, tal como exemplificam as ocorrências diacrônicas abaixo:

- (14) A construção de Brasília **visa** a esse objetivo: levar para o centro do país atividades que desenvolvam a sua economia e integram na comunhão brasileira de dois terços do território nacional, ainda que inexploradas>>, declarou o presidente da República em recente discurso. [AC-02-02-1958]
- (15) [...] por-quanto na prática essa criação se desviara inteiramente de seus nobres fins, para **visar** unicamente ao interesse pecuniário do inspetor geral das aulas, autor de dito regulamento, e um dos professores do referido lycêo (...) [OL-09-05-1878]

Em (14), temos uma ocorrência prototípica desse *type*. Dois aspectos merecem destaque nesse contexto: (i) percebe-se, após esse SP, a oração “levar para o centro do país atividade que desenvolvam [...]”. Essa contiguidade co-textual revela alguns traços de GR em curso: além de as duas orações compartilharem o mesmo sujeito e o mesmo tempo verbal, caracterizando, respectivamente, menor explicitude do sujeito e maior correferencialidade, a ocorrência contígua das duas construções – SP e oração reduzida de infinitivo – revela uma proximidade que favorecerá a constituição de padrões mais gramaticalizados, conforme apresentaremos na sequência, favorece a mudança via metonímia e reanálise categorial; (ii) o SN, argumento externo, “A construção de Brasília”, que funciona como sujeito do predicado constituído por *visar*, não permite uma interpretação em que é possível derivar os traços de animacidade e humanidade, tal como propusemos em padrões anteriores, o que também indicia traços do processo de mudança semântico-categorial de *visar*.

A ocorrência (15) não é prototípica de *Visar 5*, porque apresenta o advérbio “unicamente” entre o predador e o seu objeto, distanciando-os e mostrando, assim, estágio anterior de GR em comparação com o exemplo contido em (15). Das 7 ocorrências pertencentes ao *corpus diacrônico* encontradas neste *type*, 57,17%, equivalente a 4 ocorrências, apresentam o advérbio como material interveniente; ou seja, diferente do que veremos no *Visar 6*, neste, a construção prototípica é menos utilizada.

Dentre as ocorrências coletadas na sincronia, 12 revelaram as características deste *type*, equivalendo a 2,96% do total das ocorrências de *visar* nessa perspectiva:

- (16) Projeto **visa** à revogação de aumento. [...]O parlamento acrescenta que, além de vetar o aumento da taxa, o novo projeto também **visa** revogar a elevação da alíquota da energia rural, que passou de 29% para 30%. [DC/Pol19-04-2012]

Este *type* coexiste com outra construção de *visar*, a saber, *Visar 8*, conforme apresentaremos na sequência, comprovando que, na GR, a mudança não é abrupta, as construções coexistem e podem permanecer à disposição das necessidades comunicativas do escrevente/usuário da língua. Percebe-se como as construções “visa à revogação”/ “visa revogar” são utilizadas em alternância pelo sujeito, e, além disso, como, num mesmo recorte sincrônico, temos refletidos estágios distintos da mudança experimentada pelo item, em direção à construção que caracterizará o *type 8*, em que o complemento preposicional SP, formado pela preposição “a” + o SN “a revogação”, cujo nome, de natureza verbal, aparece também complementado, passa a funcionar como complemento de natureza verbal, propriamente dita, “visa + revogar”. Com isso, identificamos que um possível elo fundamental para a passagem desse *type 5* para a construção mais gramaticalizada é exatamente a condição de se ter um nome verbal que é substituído por um verbo, dentro de uma construção que se cristaliza com o tempo.

5.6. *Visar 6*: [SN (-animado/-humano) *visar* (VTD) (ter como objetivo) SN (alvo)]

Temos aqui como sujeito um SN com os traços [-animado] e [-humano], *visar* com o sentido [+abstrato] de “ter por objetivo” e o seu complemento é um SN alvo. Trata-se, portanto, de um uso em que, além da perda dos traços de animacidade e humanidade do sujeito, constata-se também a perda do uso da preposição, em que verbo e complemento ligam-se, transitivamente, de forma direta. Foram encontradas 14 ocorrências desta construção, correspondentes a 20% de todos os usos de *visar* na diacronia. Vejamos, na ocorrência, abaixo, um uso prototípico deste *type*:

- (17) A estatística empreendida **visa** a demonstração da nossa força, do nosso trabalho e do nosso progresso, em benefício da pátria comum. [OR-02-12-1920]

Neste *type*, em todas as ocorrências em que foi observada a presença de material intercalado, entre predicado e complemento sua natureza é adverbial. Como no exemplo:

- (18) A criação deste pólo tecnológico **visa**, principalmente, a geração de oportunidades e empregos, com o desenvolvimento mais acelerado das atividades produtivas e com a geração de postos de trabalho pelo surgimento de novas empresas”, externa o autor do projeto. [AMT/Pol-02-09-2014]

Esse aspecto reporta-se à questão da iconicidade relativa ao princípio da proximidade: quanto mais distante, estruturalmente, está uma oração da outra ou de seu complemento, também assim se encontram no aspecto semântico. Além disso, quanto mais gramaticalizada é uma construção, menor é a probabilidade de ocorrências de material linguístico que “quebre” sua unidade enquanto construção cristalizada.

5.7. *Visar* 7: [SN (-animado/-humano) *visar* (VTI) (ter por objetivo) Oração(alvo)]

Aqui, encontraremos as seguintes características: o sujeito será [-animado] e [-humano], *visar* terá um uso [+abstrato], com sentido de “ter por objetivo” e, como um verbo transitivo indireto, recrutará uma oração precedida de preposição para completá-lo. Essa construção aparece 7 vezes no *corpus* diacrônico, dando-nos um total de 10% de todas as ocorrências encontradas na diacronia. Vejamos o exemplo prototípico desse *type*:

- (19) A medida da direção da referida empresa **visa** a atender centenas de pedidos de ligação. [AC-19-11-1960]

Na sincronia, foram constatadas 18 ocorrências deste *type* no *corpus*, equivalendo a uma porcentagem de 4,44%. Vejamos:

- (20) A solicitação **visa** a verificar a qualidade dos serviços prestados pela contratada. [GD/Econ-30-12-2013]

5.8. *Visar* 8: [SN (-animado/-humano) *visar* (VTD) (ter como objetivo) Oracional (alvo)]

Neste *type*, as construções apresentam as características de um sujeito composto por um SN [-animado] e [-humano], com *visar* tendo um uso [-concreto] e recrutando uma oração como objeto direto, ou seja, sem o uso da preposição. Trata-se, pois, de uma relação transitiva em que o uso da preposição deixa de ser observado, em contexto similar ao do *type* anterior.

Foram constatadas 24 ocorrências de *Visar* 8, no recorte diacrônico, correspondentes a 34,28% de todo o *corpus*, o que permite afirmar que este *type* representa a escolha mais utilizada pelos escreventes na amostra investigada, nessa perspectiva. Vejamos um exemplo:

- (21) A campanha **visa** realizar o ideal cristão da família, solucionar problemas a ela referentes e defende-la dos perigos e ameaças que a cercam [AC-03-03-1957].

Esse exemplo é a forma prototípica deste *type*. Vale destacar que, apesar de termos considerado esses como os usos mais prototípicos do *type* em questão, não estamos diante da construção “*visar* + infinitivo”, conforme mostraremos em dados da sincronia atual, uma vez que há a possibilidade de coordenação de mais de uma estrutura oracional, tal como exemplificado anteriormente. Na diacronia, temos, no entanto, a forma que abre caminho para o desenvolvimento da construção a partir da cristalização via frequência de uso.

Este *type* também apresentou usos, menos prototípicos, com material intercalado, em todos os casos o advérbio, entre o verbo e o seu complemento oracional, e com o próprio verbo *visar* encaixado em relativa; essas formas totalizaram 12,5% de todas as encontradas.

Já na sincronia, das 405 ocorrências encontradas desse verbo, 291, equivalentes a 71,85%, enquadram-se neste formato, ou seja, a maioria das ocorrências:

- (22) Folhetos com dicas de prevenção estão sendo distribuídos. A operação “Saque Seguro” começou no ano de 2011 **visa** dar maior segurança a população neste período em que ocorre pagamento dos funcionários públicos e aposentados, onde ocorre maior circulação de pessoas com valores neste período em toda Capital. [AMT/Des>Pol-30-04-2014]

Temos, nessa oração, um uso prototípico de *Visar* 8, no qual o SN “A operação ‘Saque Seguro’” funciona como argumento externo ao predicado verbal e a oração “dar maior segurança à população neste período” como argumento interno, não preposicionado. Percebe-se o sujeito com menor grau de explicitude, com maior correferencialidade e constituído por uma anáfora zero, porém possível de ser resgatado no contexto, constituído, por sua vez, pela coordenação por justaposição, além de marcado pelas características semânticas [-animado] e [-humano]. Ainda sobre o critério de entrelaçamento de Lehmann (1988), percebemos, também, que há correlação entre a oração principal e a subordinada quanto ao aspecto modo-temporal “visa-dar”.

Correlacionamos este *type* ao quarto estágio de GR de verbos de Heine (1993), uma vez que nessa construção não há a possibilidade de formação de imperativos ou de apassivação, os núcleos dos complementos não são mais nomes e, devido à formação de perífrase, o verbo subordinado, na maioria dos casos, aparecerá no infinitivo. É exatamente esse *type* que se cristaliza na construção “visar + infinitivo”, no uso considerado mais gramaticalizado. Veja a ocorrência, a seguir, que se diferencia do prototípico:

- (23) Denominado “AmbientAÇÃO”, o programa da AL/MT envolve os servidores do poder Legislativo Estadual em ações que **visam** reduzir, reutilizar e reciclar dentro do próprio órgão, que conta como uma oficina de reciclagem. [24HN/Esp-29-11-2012]

Acima, temos o exemplo do SN “ação”, como argumento externo da construção relativa em que *visar* encontra-se encaixado. Esse verbo é, por sua vez, complementado, internamente, pela oração “reduzir”, que se coordena a “reutilizar” e “reciclar”. Encontramos 77 ocorrências, em nosso *corpus*, de oração encaixada em relativa, o que corresponde a 26,46% de todas as ocorrências deste *type*.

- (24) Evento **visa** que educador olhe criança em sua perspectiva. [24HN/Esp-05-11-2012]

Na ocorrência acima, temos o sujeito formado pelos SNs “Evento”, [-animados] e [-humanos], e como complemento de *visar* a oração desenvolvida “que educador olhe criança em sua perspectiva”. Esse tipo de uso foi encontrado apenas em duas ocorrências, totalizando 0,68% de todas as ocorrências de *Visar* 8. Vale enfatizar que somente nesse *type* encontramos orações desenvolvidas completando o verbo *visar*, representando um grau de entrelaçamento menor.

- (25) Conforme Pedro Pelizer, inspetor da Polícia Rodoviária Federal, a operação durante o carnaval, **visa** não só reduzir o número de acidentes de trânsito nas rodovias. [G1/MT>Not-20-02-2012]

Nota-se, nesse exemplo, que a construção é precedida de “não só”, único caso observado no *corpus* (0,34%). Como já dito anteriormente, a ocorrência desses materiais indica menor grau de GR da cláusula, uma vez que, semântica e estruturalmente, as orações principal e subordinada encontram-se mais distantes, ou seja, a construção ainda permite a interveniência de material intercalado, o que indica seu “processo de cristalização em curso”.

- (26) Faltando um dia para o início da Copa do Mundo no Brasil, manifestações são organizadas em Cuiabá. Duas delas acontecem nesta sexta-feira (13) em pontos distintos. Uma tem como objetivo protestar contra as obras inacabadas e a outra **visa** lutar por um país mais justo e menos desigual. [GD/Cid-11-09-2014]

Nota-se uma anáfora pronominal no SN sujeito “a outra”, referindo-se ao substantivo “manifestações”, e, como argumento interno, a oração “lutar por um país mais justo e menos desigual”. Na oração, “Uma tem como objetivo protestar contra as obras inacabadas”, o escrevente utiliza o substantivo “objetivo” e, na próxima oração, para evitar a repetição da palavra, opta por *visar*, evidenciando, assim, o valor semântico com que o escrevente emprega o verbo em estudo.

- (27) O objetivo **visa** dar celeridade e qualidade no atendimento médico às pessoas que procuram uma Policlínica, Unidade de Pronto Atendimento (UPA) ou Pronto-Socorro (PSMC). [GD/Pol-15-10-2011]

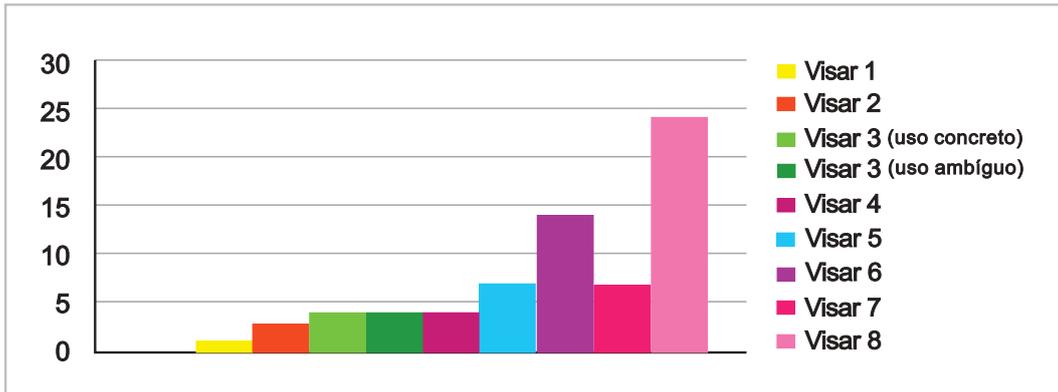
Nessa construção, o sujeito “o objetivo” e o *visar*, significando “ter como objetivo”, mostram sentidos que se repetem. É exatamente esse contexto, em que o sentido mais gramaticalizado de *visar* já se encontra presente no SN sujeito, que permite a emergência de um outro sentido e de uma outra funcionalidade para a construção em que no verbo se encontra, a saber, a marcação de futuridade na perífrase em formação “visa dar”. Trata-se de um exemplo de “desbotamento semântico”⁷ de um sentido “ter como objetivo” que se soma a outro, o de “futuridade”, num uso mais gramaticalizado. Nesse caso, o complemento de *visar*, a oração “dar celeridade e qualidade no atendimento médico às pessoas que procuram uma Policlínica (...)”, tem seu verbo, na forma infinitiva, migrando, num processo de reanálise, que permite a emergência da perífrase somada à marcação de futuridade.

5.9. Aspectos gerais sobre os usos de *visar*

A seguir, nos gráficos 2 e 3, apresentamos a quantidade de ocorrências encontradas em cada *type* apresentado nas subseções anteriores:

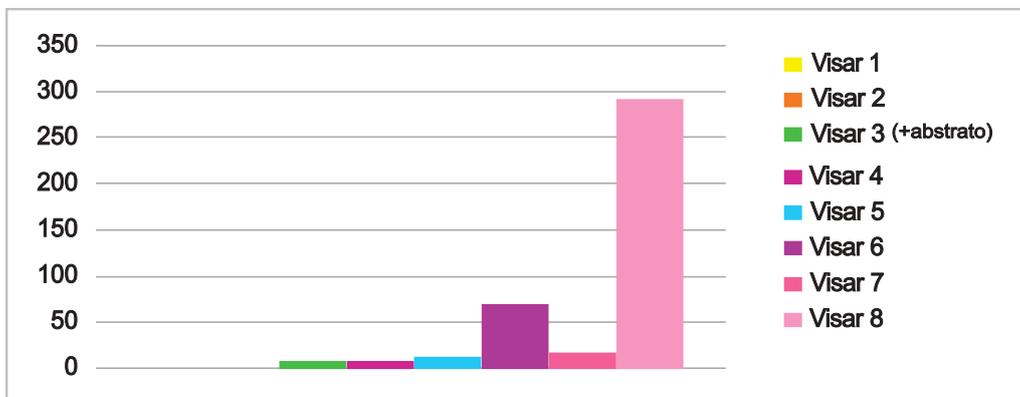
⁷ Característica de desvanecimento e abstratização semântica envolvidos nos processos de GR, em que um conceito fonte, que carrega consigo o significado completo, ao final do processo de GR, encontra-se reduzido, isto é, as especificidades semânticas de sua fonte tornam-se esvaziadas ou desbotadas (HEINE *et al.*, 1991).

Gráfico 2: Ocorrências de *visar* na diacronia



Fonte: elaboração das autoras

Gráfico 3: Ocorrências de *visar* na sincronia



Fonte: elaboração das autoras

Ao analisarmos os gráficos 2 e 3, percebemos, a partir dos números das ocorrências neles apresentados, o que Haiman (1991 *apud* HOPPER; TROUGOTT, 1993) afirma sobre o processo de GR, salientando que esse é um processo de rotinização da língua. Uma forma ou combinação de formas ocorre, no discurso, com frequência crescente e, começando como uma forma não usual de fazer ou reforçar um ponto do discurso, passa a ser um meio usual e não marcado de desempenhar esse papel.

Dessa maneira, a frequência com que tais expressões ocorrem será um fator que determina se a forma passa ou não a ser considerada gramatical pela comunidade de fala. Bybee (2010), nesse mesmo prisma, afirma que, quanto maior for a frequência de um item/construção, mais ele terá a probabilidade de ampliar o seu espectro funcional. Comparando os dois gráficos, notadamente conseguimos perceber a assertiva dessa autora sobre a frequência de uso, pois é a partir do aumento do uso da construção que as mudanças ocorrem, ou seja, os resultados da análise aqui apresentados mostram que a mudança funcional está atrelada ao aumento das ocorrências do verbo *visar*.

Considerações finais

Neste trabalho, realizamos uma análise sintático-semântico-pragmática do item *visar* num caminho de mudança de seu estatuto categorial, saindo de um verbo pleno e adquirindo traços de perífrase na marcação de intencionalidade/volição, perpassada por uma projeção de tempo futuro.

Utilizamos como *corpus*, para tal investigação, os jornais publicados no estado de Mato Grosso em dois recortes temporais, a saber: sincrônico, século XXI, e diacrônico, séculos XIX e XX. A complementação entre esses dois recortes, perspectiva pancrônica, nos possibilitou, não somente observar as polissemias e multiplicidade de formas e funções apresentadas pelo objeto de estudo, como, também, comprová-las na história da língua e traçar um trajeto de mudança percorrido por *visar*. Sendo assim, entendemos que as mudanças que ocorrem na língua provêm das necessidades comunicativas, é no falante/escrevente a partir de suas necessidades e intenções de comunicação que surgem as mudanças.

A partir do levantamento etimológico de *visar* e dos usos catalogados e analisados, a explicação cognitiva para o processo e os usos de *visar* identificam-se com o fato de esse verbo, com o sentido de “ter por objetivo”, ter surgido da palavra francesa *viser*, tendo o primeiro registro no ano de 1160, com o significado especificamente de “dirigir o olhar em direção a um alvo para lançar algo nesta direção”.

Correlacionamos a essa etimologia, o que Meillet 1948 [1912] registrou, em seus estudos sobre o *bleaching* semântico: a fonte, aqui representada pela forma e significado originais de *visar*, mais concretos (*olhar = mirar um alvo = objeto concreto específico*), reflete uma ação humana elementar, “olhar”, e de, a partir desse olhar, “mirar um alvo”, também concreto. Esse significado fonte dá origem ao significado da forma-alvo, “ter por abjetivo”, pois entendemos que, quando há algo a ser almejado, volta-se o olhar, o desejo de conquistar tal coisa e, para isso, lança-se à ação para obtê-la. Há, portanto, a passagem de um sentido mais concreto e básico, para outro mais abstrato e cognitivamente complexo, por meio de mudanças fundamentais nos aspectos estruturais e semânticos do contexto de usos do item, até alcançar a construção “*visar* + infinitivo”, em co-textos em via de cristalização e altamente recorrentes, que contribuem para o sentido ainda mais abstrato de volição, apontando para uma possível inferência de futuridade.

Partindo desse deslizamento semântico, as ocorrências de *visar*, classificadas em oito *types*, foram analisadas e distribuídas num *continuum* entre os seus diferentes usos e acepções, partindo do mais concreto (*Visar 1*) em direção ao mais abstrato (*Visar 8*). Nos quais *visar*, no primeiro estágio, é um verbo pleno, tendo um significado concreto que recruta como argumento externo um sujeito agente [+humano] e [+animado] e um argumento interno que recebe a ação. Já no último, o item tem significado abstrato e recruta como sujeito nomes [-humanos] e [-animados] e se une a orações no infinitivo, formando uma construção com marcações de volição em contexto que permite inferência de futuridade.

Nesse percurso, do mais gramatical para o mais abstrato, *visar* passa por várias pequenas mudanças, como apresentamos na seção de análise, comprovando que as mudanças na língua não são abruptas, são graduais. Destaca-se, nesse longo trajeto, o *Visar 3*, ponto em que se observa a ambiguidade de sentido, *type* em que encontramos a atuação da metáfora e da metonímia, responsáveis

pela transferência de significados. Observamos que o significado mais concreto de *visar*, “dirigir a vista”, contribuiu para um significado mais abstrato “ter como objetivo”. Entendemos que esse salto metafórico de significado se deu por meio da contiguidade, ou seja, por meio da ação da metonímia, que é capaz de captar a gradualidade desse processo de mudança semântica, em contextos que propiciam a ambiguidade e a transferência de significado.

Dessa forma, *visar*, significando “olhar, mirar”, mais concreto, mais próximo à experiência física do falante, a partir do aumento da frequência de uso, passa por um estágio de polissemia, tendo como resultado uma outra acepção do verbo, a de “pretender, almejar, ter por objetivo”, associada ao contexto do ato de fala, portanto, mais abstrata. Consideramos esse significado de *visar* mais abstrato uma vez que ele tem carácter cognitivo [-concreto] e recruta como sujeito palavras marcadas pelas características [-animado] e [-humano].

Percebemos, por meio desta pesquisa, que são as relações que o predicado estabelece com os seus argumentos que determinarão o estatuto sintático-semântico dos verbos. Sendo assim, cria-se um processo de mudança em cadeia: o que subsidia a distinção em dois tipos semânticos distintos do verbo, [-concreto] e [+abstrato], é uma informação, também de natureza semântica, que vem do sujeito.

Sobre o complemento recrutado por *visar*, aparece um SV que integra uma oração reduzida de infinitivo, de acordo, portanto, com o processo de dessentencialização, segundo Lehmann (1988), em que o verbo perde a função de principal da oração e passa a exercer função de auxiliar com forte marcação de volição e possível inferência de tempo futuro, a partir do “desbotamento” de significado, no qual um sentido “ter como objetivo” soma a outro, o de “volição”, num uso mais gramaticalizado. Dessa forma, percebe-se, através do percurso de mudança, que o item em estudo forma uma nova construção, documentada pela frequência de uso na escrita/fala.

Referências

- ABREU, Antônio Suárez. *Gramática mínima: para o domínio da língua padrão*. São Paulo: Ateliê Editorial, 2006.
- BYBEE, Joan. *Frequency of Use and the Organization of Language*. New York: Oxford University Press, 2007.
- BYBEE, Joan. Mechanisms of Change Grammaticalization: the role of frequency. In: JOSEF, B.; JANDA, R. (eds.) *The Handbook of historical linguistics*. Oxford, UK: Blackwell Publishing, 2003, pp. 602-23.
- BYBEE, Joan. Usage-based Phonology. In: DARNELL, Michael *et al.* (ed.) *Functionalism as formalism in linguistics*. v.1. John Benjamins Publishing, 1999. pp. 211-42.
- CASSEB-GALVÃO, Vânia Cristina. *O achar no português no Brasil: um caso de gramaticalização*. 1999. p. 167. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, SP, 1999.
- CEZARIO, Maria Maura. *Graus de integração de cláusulas com verbos cognitivos e volitivos*. 2001. p. 217. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.
- COSERIU, Eugenio. *Sincronia, diacronia e história: o problema da mudança linguística*. Rio de Janeiro: Presença, 1979.

- GIVÓN, Talmy. *On understanding grammar*. New York: Academic Press, 1979.
- GIVÓN, Talmy. *Topic continuity in discourse: a quantitative cross-language study*. Philadelphia, PA: John Benjamins, 1983.
- GONÇALVES, Sebastião Carlos Leite *et al.* (orgs.). *Introdução à gramaticalização*. São Paulo: Parábola, 2007.
- HEINE, Bernd. *Auxiliaries: Cognitive forces and grammaticalization*. Oxford: University Press, USA, 1993.
- HEINE, Bernd *et al.* From cognition to grammar: Evidence from African Languages. In: TRAUGOTT, Elizabeth Closs; HEINE, Bernd. *Approches to Grammaticalization*. v. 1. Amsterdam/ Filadélfia: John Benjamins Publishing Company, 1991, pp. 149-88.
- HEINE, Bernd; REH, Mechthild. Grammaticalization and reanalysis. In: *African Languages*. Hamburg: Helmut Booke Verla, 1984.
- HOPPER, Paul. On some principles of gramaticalization. In: TRAUGOTT, Elizabeth Closs; HEINE, Bernd. *Approches to Grammaticalization*. v. 1. John Benjamins Publishing Company. Amsterdam/ Filadélfia. 1991.
- HOPPER, Paul; TRAUGOTT, Elizabeth Closs. *Gramaticalization*. Cambridge: Cambridge University Press, 1993.
- LEHMANN, Christian. Towards a typology of clause linkage. In: HAIMAN, John e THOMPSON, Sandra A. (eds.). *Clause combining in gramar and discourse*. Amsterdam e Philadelphia: j. Benjamins, 1988. Disponível em: www.christianlehmann.eu/publ/linkage.pdf. Acesso em: 23 out. 2014.
- LOPES-DAMASIO, Lúcia Regiane. *A emergência do marcador discursivo Assim sob a óptica da Gramaticalização: um caso de multifuncionalidade e (inter)subjetivação*. 2008. p. 247. Dissertação (Mestrado em Estudos Linguísticos) - Instituto de Biociência, Letras e Ciências Exatas, UNESP/ São José do Rio Preto, 2008.
- MEILLET, Antoine. *Linguistique Historique et Linguistique Générale*. Paris: Librairie Ancienne Honoré Champion, 1948 [1912].
- MOURA, Geovana Portela de. *Visa gramaticalizar: o processo de gramaticalização de visar*. 2016. p. 165. Dissertação (Mestrado em Estudos da Linguagem) - Programa de Pós-Graduação em Estudos de Linguagem, Universidade Federal de Mato Grosso Cuiabá, 2016.
- NEVES, Maria Helena de Moura. *A gramática funcional*. São Paulo: Martins Fontes, 1997.
- NEVES, Maria Helena de Moura. Uma introdução ao funcionalismo: proposições, escolas, temas e rumos. In: CRHISTIANO, Maria Elizabeth A.; SILVA, Camilo Rosa; HORA, Demerval da. (orgs). *Funcionalismo e gramaticalização: teoria, análise e ensino*. João Pessoa: Idéia, 2004, pp. 13-28.
- SWEETSER, Eve. *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*. New York: Cambridge University Press, 1991.
- SWEETSER, Eve. Grammaticalization and semantic bleaching. In: AXMAKER, Shelley; JAISSER, Annie; SINGMASTER, Helen (orgs.). *General Session and Parasession on Grammaticalization*. Berkeley Linguistics Society, 1988. pp. 389-405.
- TRAUGOTT, Elizabeth Closs; TROUSDALE, Graeme. *Constructionalization and constructional changes*. Oxford: Oxford University Press, 2013.

DA EMERGÊNCIA DE FORMAS COMPOSTAS PARA A EXPRESSÃO DO ANTEPRETÉRITO: CONSTRUTOS MORFOSSINTÁTICOS E SEMÂNTICOS

ON THE EMERGENCE OF COMPOUND FORMS TO THE EXPRESSION OF THE 'ANTEPRETÉRITO': MORPHOSYNTACTIC AND SEMANTIC CONSTRUCTS

Márluce Coan¹

Francisco José Gomes de Sousa²

Laila Cavalcante Romualdo³

RESUMO

Em busca de indícios de emergência da forma composta de antepretérito, este artigo apresenta motivações morfofossintático-semânticas que conduziram ao processo de gramaticalização de *haver/ter* como auxiliares mais participio verbal. Via análise de dados provenientes de 1.680 cantigas medievais galego-portuguesas, a emergência do processo de gramaticalização de tempos compostos (*haver/ter* mais participio) é retratada mediante quatro parâmetros de análise, dois deles referentes à seleção argumental de sujeito e de objeto, um acerca do tipo de combinação verbal e, ainda, outro para verificação de marcas morfológicas da forma participial, para que fosse possível detectar estágios de gramaticalização atrelados a princípios como estratificação, divergência, persistência, especialização e decategorização. Dessa guisa, vimos que as ocorrências evidenciaram, predominantemente, falta de coincidência entre sujeitos de *haver/ter* e participio; noção de posse metafórica do objeto; combinação entre *haver/ter* muito mais com verbos que não contradizem a noção de posse e neutralização (masculino/singular) da estrutura participial. Ademais desses construtos contextuais que possibilitaram a gramaticalização de estruturas compostas de *haver/ter* no imperfeito mais participio, outros tempos verbais combinados com o participio, a depender do contexto, também manifestam a noção temporal de passado anterior a outro passado nas cantigas, quais sejam: presente do indicativo, pretérito perfeito do indicativo e pretérito imperfeito do subjuntivo seguidos de participio.

PALAVRAS-CHAVE: Haver/ter. Participio. Antepretérito. Lírica profana. Galego-português.

ABSTRACT

In search of evidence of the emergence of the compound form of the *antepretérito* (past before), this article presents morphosyntactic-semantic motivations that led to the grammaticization process of the verbs *haver/ter* as auxiliaries plus verbal participles. Through the analysis of data from 1,680 medieval Galician-Portuguese songs, the emergence of the grammaticization process of compound tenses (*haver/ter* more participle) is portrayed through four analysis parameters, two of them referring to the argument selection of subject and object, one about the type of verbal combination and yet another to verify morphological marks of the participial form, so that it would be possible to detect grammaticization stages linked to principles

¹ Professora da Universidade Federal do Ceará, coanmalu@ufc.br, <https://orcid.org/0000-0001-7809-8624>.

² Graduado em Letras pela Universidade Federal do Ceará, francisco.jose.letas@gmail.com, <https://orcid.org/0000-0001-6115-0107>.

³ Graduada em Letras pela Universidade Federal do Ceará, lailacavrom@gmail.com, <https://orcid.org/0000-0001-6923-7442>.

such as layering, divergence, persistence, specialization and decategorization. This way, we noticed that the occurrences showed, predominantly, a lack of coincidence between subjects of *haver/ter* and participle; notion of metaphorical possession of the object; combination of *haver/ter* much more with verbs that do not contradict the notion of possession and neutralization (masculine/singular) of the participial structure. In addition to these contextual constructs that enabled the grammaticalization of structures composed of *haver/ter* in the imperfect tense plus participle, other verb tenses combined with the participle, depending on the context, have also manifested the temporal notion of past before another past in the songs, namely: *presente do indicativo* (present tense), *pretérito perfeito do indicativo* (simple past tense) and *pretérito imperfeito do subjuntivo* (subjunctive imperfect tense) followed by participle.

KEYWORDS: *Haver/ter*. Participle. *Antepretérito*. Profane lyric. Galician-Portuguese.

Introdução: proposta e perspectivas analíticas

Analisamos, nesta pesquisa, as formas de codificação de antepretérito, nos termos de Bello (1984 [1847]), ou seja, as formas de expressão de um tempo passado anterior a outro passado. Esse uso, mais frequentemente expresso pela forma composta do mais-que-perfeito na atualidade, subjaz da noção de que essa estrutura tem a mesma função que a forma simples, cujas razões estariam assentadas em três critérios, conforme Ilari (1997): (i) as perífrases de “*ter*” mais *particípio* comutam com as formas simples do verbo (cf. *tinha feito/fizera*); (ii) é impossível atribuir um mesmo sentido ao verbo “*ter*” quando principal e quando auxiliar; (iii) há indicação de uma única ação expressa por “*ter*” mais a base verbal, atribuída a um único sujeito.

A preferência pela forma composta, conforme Fiorin (1996), decorreria do fato de a forma analítica ter a possibilidade de expressar a relação de anterioridade e o aspecto perfectivo (concluso), enquanto a forma sintética representaria apenas a relação de anterioridade. Tal tendência/preferência pela forma composta, ao passar do tempo, pode ser constatada, também, em pesquisas realizadas por Gonçalves (1993), Coan (1997, 2003), Martins (2010, 2011), Freitag, Oliveira e Coan (2018) e Coan, Lima e Sampaio (2019).

Em busca de indícios da emergência de tal estrutura composta (*haver/ter* mais *particípio*), optamos por analisar as cantigas trovadorescas galego-portuguesas, as quais, segundo Vasconcelos (1904, p. 780), “falam eloquentemente a favor da afinidade primitiva de lusitanos e galaicos”. Monteagudo (2012) também observa que foram os cancioneiros *galego-portugueses* que registraram as mais antigas manifestações literárias de nosso idioma. Trata-se de produções de um período em que a língua falada se desenvolve como língua literária por excelência. Ademais, conforme Lorenzo (1995), já na época medieval, as formas compostas têm muitíssimo uso.

Ao investigarmos vestígios do processo de gramaticalização de *haver/ter* mais *particípio* como forma de expressão do antepretérito na lírica profana galego-portuguesa, detectamos que tal função não era expressa somente por *haver/ter* no imperfeito do indicativo mais *particípio*, como ilustrado em (01). Esse valor de antepretérito foi observado em combinações outras com *haver/ter* no presente e no pretérito perfeito do indicativo, bem como no pretérito imperfeito do subjuntivo mais *particípio*, conforme ilustram, respectivamente, os exemplos de (02) a (04).

- (01) Que muito m'el **HAVIA JURADO** / que me nom visse, mais, a Deus grado, / par Deus, donas, aqui é já. {Cantiga de Amigo: O meu amigo, que mi dizia, linhas 4-6}
- (02) Se sabedes novas do meu amado, / aquel que mentiu do que mi **HÁ JURADO?** / Ai Deus, e u é? {Cantiga de Amigo: - Ai flores, ai flores do verde pino, linhas 10-12}
- (03) Do que mi **HOUVE JURADO**, / pois mentiu per seu grado, / sanhuda lh'and'eu. {Cantiga de Amigo: Ai madre, bem vos digo, linhas 4-6}
- (04) Que alongad'eu ando d'u iria, / se eu **HOUVESSE AGUISADO** d'ir i, / que viss'a dona que veer querria, (...) {Cantiga de Amor: Que alongad'eu ando d'u iria, linhas 1-7}

Para garantir uma análise mais bem fundamentada, separamos essas construções verbais de *haver/ter* mais particípio neutro (masculino/singular), que codificam um passado anterior a outro passado, em blocos analíticos, aos quais acrescentamos outro destinado somente para ocorrências de *haver/ter* mais particípio flexionado no feminino e/ou plural, conforme exemplo (05).

- (05) E por levá'la quitaçom dobrada / se [me] queixou; e catei u jazia / eno padrom, e achei que **HAVIA** / de todo bem sa quitaçom **LEVADA**; (...) {Cantiga de Escárnio e Maldizer: Cirola vi [eu] andar-se queixando, linhas 15-18}

De posse dessas estruturas, analisamos seus contextos de ocorrência, especificamente por meio de fatores morfossintáticos e semânticos. Verificamos, primeiramente, se os sujeitos de *haver/ter* eram idênticos, em contraposição à seleção de sujeitos distintos (manutenção do significado original de posse de *haver/ter* e do significado adjetival da forma participial), já que sujeitos idênticos poderiam indicar um primeiro construto para a emergência da forma composta. Seguindo a perspectiva sintática, observamos o segundo argumento (objeto) de *haver/ter*, perspectiva alargada para o nível semântico, quando averiguamos a noção de posse codificada por esse objeto de *haver/ter*, com o intuito de detectar, para além da ideia de posse material, posse metafórica (quando o objeto representa posse inalienável) ou mesmo ausência de argumento. Ainda no âmbito semântico, consideramos as combinações entre *haver/ter* mais particípio, com o fito de flagrar se suas ocorrências viriam adjuntas a participios que contradissem a noção de posse ou que contivessem lexema idêntico aos dos verbos *haver/ter* (como em: *havia perdido*; *tinha tido*, respectivamente). Não poderia ficar por menos o exame das marcas morfológicas que indicam concordância entre o particípio e o argumento de *haver/ter* ou tendência ao uso do particípio neutro (masculino/singular).

Estando prospectada a investigação, seguem-se os procedimentos metodológicos, na seção segunda, e dois blocos analíticos nas seções terceira e quarta. Especificamente nesta última, correlacionamos os resultados aos princípios de gramaticalização propostos por Hopper (1991): *estratificação*, *divergência*, *especialização*, *persistência* e *de categorização*. Consideramos mais oportuno, em se tratando do tema em pauta, proceder desse modo a apresentar separadamente e primeiramente o referencial teórico. Cremos que a visão em conjunto, em se tratando de fenômenos históricos, serve para validar resultados, além de conferir profundidade à pesquisa, já que é com base

em nossas teorias que produzimos conhecimento sobre o passado. De acordo com Lass (1980), é o ato intelectual humano que engatilha provas do passado em sequências coerentes.

1. Procedimentos metodológicos: contextualização e *corpus*

Os dados sob análise foram coletados a partir da leitura de um conjunto de 1680 cantigas medievais galego-portuguesas, denominadas de profanas ou de corte, disponíveis no site: <https://cantigas.fcsh.unl.pt> (LOPES *et al.*, 2011). Tal base de dados online, resultado do projeto *Littera* (Edição, atualização e preservação do património literário medieval português), inclui a totalidade das cantigas medievais dos cancioneiros galego-portugueses, bem como as respectivas imagens dos manuscritos, para conferência de dados. As cantigas provêm de três recolhidas trovadorescas, quais sejam: o *Cancioneiro da Ajuda*, o *Cancioneiro da Biblioteca Nacional* e o *Cancioneiro da Biblioteca Vaticana*. Ademais, os textos estão divididos em diversos gêneros, como cantigas de amor, de amigo, de escárnio e maldizer e outros gêneros menores, a exemplo das cantigas de *tenção*, de *espúria* e as de *sirventês moral*.

O corpus em questão refere-se à produção feita durante o período denominado de galego-português, ou, ainda, português arcaico, antigo ou trovadoresco (conforme, respectivamente, Cuesta, Leite de Vasconcelos, Lindley Cintra e Serafim da Silva Neto *apud* Castro, 1988, p. 12), período que se inicia em torno do ano 1200 e vai até cerca de 1385 a 1420 (em torno dos séculos XIII a XV). Insere-se, portanto, no contexto histórico do período medieval.

Dessa amostra, obtivemos 41 dados em que foram detectadas as formas *haver/ter* mais participio passado com acepção de antepretérito. Pode parecer curioso o fato de haver somente 41 dados em 1680 cantigas, mas estamos em viés analítico verticalizado, considerando-se uma função específica, a de antepretérito, mais distante temporalmente em relação ao momento de fala, portanto, menos usual na amostra sob análise, cujas referências temporais mais se voltam ao centro dêitico. Relativamente à forma, estamos em busca de indícios de gramaticalização, assim sendo não são quaisquer usos de *haver/ter* e de verbos no participio que nos interessam, ou seja, não os analisamos isoladamente, mas combinados em um mesmo enunciado. Agrega-se a isso o fato de estarmos diante de um processo de gramaticalização, mais especificamente de sua emergência, portanto, poucos usos são mesmo esperados, já que uma estrutura não se gramaticaliza abruptamente. Todo processo é gradual. Da mesma forma que tratar de um fenômeno residual tem seu mérito por mostrar sua resistência ao longo do tempo, também o tem o tratamento de estágios primeiros de um processo que veio a consolidar-se em português (uso da forma composta para expressão do antepretérito). Por um ou outro viés, há uma considerável empreitada de retratar parcelas históricas, cujo reflexo pode ser visto na atual configuração gramatical do português.

2. Expressão de antepretérito por estruturas compostas: *tipos de construção, competição entre “haver” e “ter” e ocorrências por tipo de cantiga*

Os dados de *haver/ter* mais particípio expressando antepretérito, provenientes de cantigas medievais galego-portuguesas, estão divididos em cinco grupos, os quais ilustramos a seguir. Note-se que tais usos se referem a uma situação passada anterior a outra também passada expressa no trecho sob análise: (i) “dado havia” em relação a “acordou”; (ii) “houvesse aguisado” em relação a “iria”; (iii) “há jurado” em relação a “foi”; (iv) “tive guisado” em relação a “foi” e (v) “houvesse pagada” em relação a “ia”.

I. *Haver/ter* no imperfeito do indicativo mais particípio neutro:

(06) Ca, se per seu grado foss', al seria;/ mais daquesto nunca m'enfingirei,/ ca hoje verdadeiramente o sei/ que per seu grado nunca mi o daria;/ mais, u estava coidando em al, / deu-mi um gram peid'e foi-lhi depois mal, / u s'acordou que mi o **DADO HAVIA**. {Cantiga de Escárnio e Maldizer: Moitos s'enfingem que ham gaanhado, linhas 8-14}

II. *Haver/ter* no imperfeito do subjuntivo mais particípio neutro:

(07) Que alongad'eu ando d'u iria, / se eu **HOUVESSE AGUISADO** d'ir i, / que viss'a dona que veer querria, / (que nom visse! ca por meu mal a vi!) / de que m'eu mui sem meu grado parti / e mui coitad'; e foi-s'ela sa via / e fiquei eu, que mal dia naci! {Cantiga de Amor: Que alongad'eu ando d'u iria, linhas 1-7}

III. *Haver/ter* no presente do indicativo mais particípio neutro:

(08) Foi-s'agora meu amig'e por en / **HÁ**-mi **JURADO** que polo meu bem / me quis e quer mui melhor doutra rem; / mais eu bem creio que nom est assi, / ante cuid'eu que moira el por mi / e eu por el, em tal hora o vi. {Cantiga de Amigo: Foi-s'agora meu amig'e por en, linhas 1-6}

IV. *Haver/ter* no pretérito perfeito do indicativo mais particípio neutro:

(09) Depoi'lo **TIV**'eu **GUISADO** /que s'el foi daqui sanhudo, /e atendi seu mandado / e non'o vi, e perdido / é comigo, e alá x'ande / sanhud'e nom mi o demande; / quant'el quiser, atant'ande / sanhud'e nom mi o demande. {Cantiga de Amigo: Assanhou-s'ó meu amigo, linhas 17-24}

V. *Haver/ter* em qualquer tempo mais particípio em feminino/plural:

(10) Vejo-vos jazer migo muit'aguada, / Luzia Sánchez, porque nom fodo nada; / mais se eu vos per i **HOUVESSE PAGADA**, / pois eu foder nom posso, peer-vos-ia. {Cantiga de Escárnio e Maldizer: Luzia Sánchez, jazedes em gram falha, linhas 7-10}

Na tabela (1) abaixo, apresentamos o quantitativo de cada uma das formas verbais ilustradas acima. Como mostram os dados, há mais casos de *haver/ter* no presente do indicativo mais particípio (15 dados, totalizando 36,6% da amostra) do que de imperfeito e perfeito do indicativo mais particípio (ambos com 9 dados), usos mais esperados, tendo em vista que foi o uso de *haver/ter* no imperfeito do indicativo que se gramaticalizou como pretérito mais-que-perfeito composto, ademais de ser o uso de

perfeito mais participio uma estrutura corrente no período medieval. Conforme observa Silveira Bueno (1995), o perfeito anterior (*haver/ter* no pretérito perfeito mais participio), cuja vitalidade foi evidente entre o final do século XII até meados do século XVI, deixou de ser utilizado no português clássico.

Tabela 1: Correlação entre forma verbal e expressão de antepretérito

Forma verbal	Aplicação/Total	Percentual
<i>Haver/ter</i> no imperfeito do indicativo mais participio neutro	9/41	22.0
<i>Haver/ter</i> no imperfeito do subjuntivo mais participio neutro	4/41	9.7
<i>Haver/ter</i> no presente do indicativo mais participio neutro	15/41	36.6
<i>Haver/ter</i> no perfeito do indicativo mais participio neutro	9/41	22.0
<i>Haver/ter</i> em qualquer tempo mais participio feminino/plural	4/41	9.7

Fonte: elaboração dos autores.

Após o mapeamento das formas verbais, observamos qual dos dois verbos (*haver* ou *ter*) era mais frequente, pois, em português, diferentemente de outras línguas românicas, o auxiliar preferido é *ter*, em vez de *haver*. Embora se considere que essa substituição de *haver* por *ter* vem da Idade Média, nos séculos XIII a XIV, a forma *default* era o verbo *haver*, fato que se confirma pelos resultados apresentados na tabela (2): *haver* aparece em 36 dados, o que corresponde a 87,8% da amostra, sobrepondo-se fortemente à ocorrência do verbo *ter* (5 dados, totalizando 12,2% da amostra).

Tabela 2: Variação entre “haver” e “ter” na expressão de antepretérito

Forma verbal	Perífrase de imperfeito do indicativo mais participio neutro	Perífrase de imperfeito do subjuntivo mais participio neutro	Perífrase de presente do indicativo mais participio neutro	Perífrase de perfeito do indicativo mais participio neutro	Perífrase verbal com participio em feminino/plural	Total
<i>Haver/Ter</i> <i>Haver</i>	9	3	12	8	4	36
<i>Ter</i>	0	1	3	1	0	5
<i>Total</i>	9	4	15	9	4	

Fonte: elaboração dos autores.

Os dados mostraram-se relativamente equilibrados entre os tipos de cantiga, como se pode confirmar na tabela (3), havendo mais ocorrências em cantigas do gênero de escárnio e maldizer (14 dados). Em seguida, foram observados 11 dados relativos às cantigas de amigo e 10 relativos às cantigas de amor. Gêneros menores somaram 6 dados.

Tabela 3: Correlação entre forma verbal e tipo de cantiga

Forma verbal Tipo de cantiga	<i>Haver/ter</i> no imperfeito do indicativo mais particípio neutro	<i>Haver/ter</i> no imperfeito do subjuntivo mais particípio neutro	<i>Haver/ter</i> no presente do indicativo mais particípio neutro	<i>Haver/ter</i> no perfeito do indicativo mais particípio neutro	<i>Haver/ter</i> em qualquer tempo mais particípio feminino/ plural	Total
Amor	1	2	6	1	0	10
Amigo	2	0	4	5	0	11
Escárnio/ Maldizer	4	2	4	1	3	14
Outros	2	0	1	2	1	6
<i>Total</i>	9	4	15	9	4	

Fonte: elaboração dos autores.

3. Emergência e gramaticalização de estruturas compostas para a expressão do antepretérito: *restrições morfossintáticas e semânticas*

Para detectar contextos propícios ou restritivos à gramaticalização das estruturas compostas por *haver/ter* mais particípio, investigamos os dados mediante quatro parâmetros: dois referentes à seleção argumental de *haver/ter* (primeiro e segundo argumentos), um referente à combinação entre *haver/ter* e particípio e outro referente às marcas morfológicas de concordância no particípio. Os resultados de cada análise estão detalhados na sequência, por meio de resultados quantitativos, expostos nas tabelas de (4) a (7). Para cada parâmetro analítico, são especificados os fatores que o integram, bem como são apresentados exemplos advindos da amostra. Ademais, tecemos correlações aos pressupostos teóricos (princípios de gramaticalização) e contrapomos nossos resultados ao estado da arte sobre o tema.

Pautamos nossas considerações sobre os resultados em pressupostos funcionalistas, especificamente nos princípios de gramaticalização propostos por Hopper (1991): estratificação (emergência e coexistência de camadas), divergência (permanência da forma lexical como elemento autônomo e como forma gramaticalizada), especialização (estreitamento de escolhas formais), persistência (traços do significado lexical e detalhes da história lexical de um item podem ser refletidos em sua nova distribuição gramatical) e decategorização (perda ou neutralização de marcas morfológicas e/ou privilégios sintáticos característicos das categorias lexicais).

3.1. Seleção argumental de HAVER/TER (primeiro argumento)

Trata-se, a partir desse parâmetro, dos sujeitos expressos em estruturas com *haver/ter* mais particípio. Observamos, especificamente, se possuem o mesmo sujeito, ou seja, se o primeiro argumento de *haver/ter* equivale ao primeiro argumento do particípio, fato que indicaria a emergência de uma forma composta por seleção argumental idêntica (Ele tinha conquistado um castelo/uma

aldeia/dois castelos/duas aldeias. – Ele possuía e ele conquistou). No entanto, quando *haver* e *ter* são verbos plenos, podem selecionar objetos, os quais, por sua vez, seriam os “sujeitos” participiais em estrutura verbo-nominal, como em: ‘Ele tinha um castelo conquistado/ uma aldeia conquistada/ dois castelos conquistados/ duas aldeias conquistadas.’ – Ele possuía algo que foi conquistado. Assim, os dados foram agrupados de acordo com as seguintes proposições: (a) o sujeito de *haver/ter* coincide com o sujeito do particípio (*O cavaleiro havia [x] voltado*, por exemplo); (b) o sujeito de *haver/ter* não coincide com o sujeito de uma estrutura predicativa de particípio (*Helena havia bom conselho recebido*, por exemplo) ou (c) não há seleção de primeiro argumento pelos verbos *haver/ter* (*Havia chovido*, por exemplo).

Não houve, na amostra, casos sem primeiro argumento; houve, entretanto, considerável diferença, conforme tabela (4), entre estruturas nas quais há coincidência argumental, como no exemplo (11), e estruturas em que não houve coincidência, conforme exemplo (12).

- (11) Joam Bol’anda mal desbaratado /e anda trist’e faz muit’aguizado, /ca perdeu quant’**HAVIA GANHADO** /e o que lhi leixou a madre sua: (...) {Cantiga de Escárnio e Maldizer: Joam Bol’anda mal desbaratado, linhas 1-4}
- (12) mais, u estava cuidando em al, / deu-mi um gram peid’e foi-lhi depois mal, / u s’acordou que mi o **DADO HAVIA**. {Cantiga de Escárnio e Maldizer: Moitos s’enfingem que ham gaanhado, linhas 12-14}

Tabela 4: Correlação entre forma verbal e seleção de primeiro argumento na expressão de antepretérito

Forma verbal / Primeiro argumento	<i>Haver/ter</i> no imperfeito do indicativo mais particípio neutro	<i>Haver/ter</i> no imperfeito do subjuntivo mais particípio neutro	<i>Haver/ter</i> no presente do indicativo mais particípio neutro	<i>Haver/ter</i> no perfeito do indicativo mais particípio neutro	<i>Haver/ter</i> em qualquer tempo mais particípio feminino/plural	Total
Coincidência	1	2	3	1	0	7
Sem coincidência	8	2	12	8	4	34
Sem argumento	0	0	0	0	0	0
Total	9	4	15	9	4	

Fonte: elaboração dos autores

A análise da seleção argumental mostra-nos um estágio bem inicial do processo de gramaticalização, pois, de acordo com Benveniste (1966), um dos contextos favoráveis ao desenvolvimento das formas compostas a partir de *haver* é a igualdade de sujeito, ou seja, o sujeito do particípio é igual ao sujeito de *haver*. Desse modo, as formas *haver/ter* e particípio tendem a aproximar-se, fundindo-se em significação temporal de antepretérito.

3.2. Seleção argumental de HAVER/TER (segundo argumento)

Investigamos, na sequência, se *haver* e *ter* já ocorrem sem argumento interno, ou seja, sem um objeto (2º argumento) indicativo de posse material ou, havendo tal objeto, se representaria posse inalienável, como etapa intermediária do processo de gramaticalização (posse material > posse metafórica > ausência do objeto possuído). Dessa guisa, os fatores sob análise estão assim distribuídos: a) o 2º argumento indica posse material (*Havia o castelo cercado*, por exemplo); (b) o 2º argumento indica posse inalienável (*Havia um ensinamento recebido*, por exemplo) ou (c) não há um candidato para a posição de 2º argumento (*[Já fazia quase um século] que ele havia nascido*, por exemplo).

Conforme nos mostra a tabela (5), foram detectados nove casos em que *haver/ter* indicam posse material, conforme exemplo (13). Entretanto, os números brutos evidenciam maior quantidade de casos em que *haver/ter* expressam posse inalienável, conforme exemplo (14), uma etapa intermediária do processo de mudança em direção à gramaticalização da forma composta. Houve, ainda, sete casos em que não há um candidato para a posição de 2º argumento, conforme exemplo (15), fator que evidencia ainda mais a gramaticalização.

- (13) E por levá'la quitaçom dobrada / se [me] queixou; e catei u jazia / eno padrom, e achei que **HAVIA** / de todo bem sa quitaçom [recibo de pagamento] **LEVADA**; (...) {Cantiga de Escárnio e Maldizer: Cirola vi [eu] andar-se queixando, linhas 15-18}
- (14) El andava trist'e mui sem sabor, / come quem é tam coitado d'amor, / e **PERDUD'**[HÁ] o sem e a color, / pero / quando me viu, disse-mi assi: / "Ai senhor, ide rogar mia senhor, / por Deus, que haja mercee de mi." {Cantiga de Amigo: O voss'amig', amiga, vi andar, linhas 7-12}
- (15) Que alongad'eu ando d'u iria, / se eu **HOUVESSE AGUISADO** d'ir i, / que viss'a dona que veer querria (...) {Cantiga de Amor: Que alongad'eu ando d'u iria, linhas 1-7}

Tabela 5: Correlação entre forma verbal e seleção de segundo argumento na expressão de antepretérito

Forma verbal / Segundo argumento	<i>Haver/ter</i> no imperfeito do indicativo mais particípio neutro	<i>Haver/ter</i> no imperfeito do subjuntivo mais particípio neutro	<i>Haver/ter</i> no presente do indicativo mais particípio neutro	<i>Haver/ter</i> no perfeito do indicativo mais particípio neutro	<i>Haver/ter</i> em qualquer tempo mais particípio feminino/plural	Total
Posse material	3	0	2	1	3	9
Posse inalienável	5	2	10	7	1	25
Sem argumento	1	2	3	1	0	7
<i>Total</i>	9	4	15	9	4	

Fonte: elaboração dos autores.

Na seleção do segundo argumento de *haver/ter*, vemos, outrossim, predominância da noção de posse inalienável, como em (14), mas também casos em que não há um candidato para objeto de *haver/ter*, como em (15). Corroborando as observações de Olbertz (1993) e Ribeiro (1996), também consideramos esses usos como indícios de emergência da forma composta de expressão do antepretérito. Para as autoras, a mudança começa quando *haver/ter* passam a ser usados para posses inerentes, mas o primeiro passo para a gramaticalização é refletido em construções em que não há um candidato para a posição de segundo argumento, reduzindo, portanto, o significado de *haver/ter* como um predicado de dois lugares, expressando uma relação de posse entre o primeiro e o segundo argumentos, como ocorre em (13). Pereira (1923, p. 463) observa que, a partir do século XVI, em alguns contextos, *ter* e *haver* foram esvaziando-se de sentido (sentido de posse). Desse modo, foram pouco a pouco entrando na categoria dos auxiliares, mas conservando seu valor original, quando não se achavam em conjunção com o particípio passado e com o infinitivo. Nossos achados remontam tal emergência para um período anterior, pois os dados são dos séculos XIII a XV.

3.3. Tipo de combinação entre HAVER/TER e particípio

Averiguamos as combinações entre *haver/ter* e particípio, considerando-se três critérios: (a) combinação com particípio cujo significado contradiz a noção de posse (*tinha perdido*, por exemplo); (b) combinação de *haver/ter* com eles mesmos (*havia havido/tinha tido*, por exemplo) ou (c) combinação com outro tipo verbal (*havia escrito*, por exemplo). Foram detectadas nove combinações que contradizem a noção de posse, conforme exemplo (16), ao passo que a maioria dos casos (32 dados) mostra combinações de *haver/ter* com outro tipo de verbo, como ilustrado em (17). Não houve, entretanto, combinação de *haver/ter* com eles mesmos, conforme se observa na tabela (6) abaixo.

- (16) Ante me quis leixar perder o sem / por vós, senhor; des i soub'alongar / meu bem, que era em mi a morte dar, / e quis que já sempre eu vivess'assi, / em gram coita como sempre vivi, / e que m'**HOUVESSE PERDUDO** meu sem. {Cantiga de Amor: Que pret'esteve de me fazer bem, linhas 7-12}
- (17) Que amor tam astroso e tam delgado, / quen'ó **TEVESS[E]** um ano **SOTERRADO!** {Cantiga de Escárnio e Maldizer: Ai amor, amore de Pero Cantone, linhas 21-26}

Tabela 6: Correlação entre forma verbal e tipo de combinação na expressão do antepretérito

Forma verbal Tipo de Combinação	<i>Haver/ter</i> no imperfeito do indicativo mais particípio neutro	<i>Haver/ter</i> no imperfeito do subjuntivo mais particípio neutro	<i>Haver/ter</i> no presente do indicativo mais particípio neutro	<i>Haver/ter</i> no perfeito do indicativo mais particípio neutro	<i>Haver/ter</i> em qualquer tempo mais particípio feminino/ plural	Total
Contradição	4	1	1	2	1	9
Igual	0	0	0	0	0	0
Outro	5	3	14	7	3	32
<i>Total</i>	9	4	15	9	4	

Fonte: elaboração dos autores.

No que diz respeito ao tipo de combinação entre os verbos, apesar de a maioria das ocorrências evidenciarem uma combinação com um tipo verbal diferente de *haver/ter* ou que não contradiz a noção de posse, conforme ilustramos em (17), flagramos combinações que contradizem tal noção, como demonstramos em (16), estando em evidência mais um indício de gramaticalização da forma composta. De acordo com Olbertz (1993), se *haver/ter* denotam uma relação de posse entre o primeiro e o segundo argumentos, não podem estar combinados com particípio cujo significado contradiz a noção de posse (*perder/dar*, por exemplo) nem com eles mesmos, o que justificaria a falta de ocorrência de estruturas como “*ter tido*” e “*haver havido*”, já que, quando passam a aceitar tais combinações, o elemento marginal do predicado (particípio) torna-se mais proeminente, assim quanto mais o particípio ganha proeminência, menos proeminente fica o outro verbo.

3.4. Marcas morfológicas no particípio

Examinamos, ainda, marcas morfológicas indicativas de concordância entre o particípio e o argumento de *haver/ter*, de acordo com quatro critérios, para detectar traços de gramaticalização da forma composta: (a) presença de marcas de gênero feminino e/ou número plural (*Havia as graças recebidas, por exemplo*); (b) forma neutra (masculino/singular) em concordância com o argumento selecionado por *haver/ter* (*Havia o mal feito, por exemplo*); (c) forma neutra (masculino/singular) sem concordância com o argumento selecionado por *haver/ter* (*Havia a desonra recebido, por exemplo*) ou (d) supressão das marcas morfológicas finais do particípio (*Havia feit’isso, por exemplo*).

Conforme a tabela (7), há, na amostra, dois casos em que há marca de feminino, conforme exemplo (18), dois em que há marca de plural, como em (19), e seis casos de supressão, como em (20), quando não é possível saber se o autor estaria usando a forma do particípio com marcas explícitas de feminino e/ou plural, ou se já estaria utilizando a forma neutra. Essa forma neutra foi detectada na maioria dos casos, havendo equilíbrio entre a forma neutra em concordância (16 casos), conforme exemplo (21), e sem concordância (15 casos), conforme exemplo (22), o que também aponta para o processo de gramaticalização da forma composta.

- (18) Vejo-vos jazer migo muit'aguada, / Luzia Sánchez, porque nom fodo nada; / mais se eu vos per i **HOUVESSE PAGADA**, / pois eu foder nom posso, peer-vos-ia. {Cantiga de Escárnio e Maldizer: Luzia Sánchez, jazedes em gram falha, linhas 7-10}
- (19) Que panos perdi de peso / e outros bem bastoados / que m'**HAVIAM** já **MANDADOS!** {Cantiga de Escárnio e Maldizer: Dom Foão, em gram cordura, linhas 13-15}
- (20) Um dia que vi mia senhor / quis-lhe dizer o mui gram bem / que lh'eu quer'e como me **TEM / FORÇAD'**e preso seu amor; / e vi-a tam bem parecer / que lhe nom pude rem dizer. {Cantiga de Amor: Um dia que vi mia senhor, linhas 1-6}
- (21) E se mais d[e] oito dias nom som / que de mia senhor foi alongado, / forte preito **TENHO COMEÇADO**, / pois m'oito dias foi tam gram sazom! {Cantiga de Amor: Estes com que eu venho preguntei, linhas 19-22}
- (22) (...) pois me deu por en mui grand'afã e desej'e cuidado que **HOUVI** dela, poila vi, **LEVADO**, / [e] per que viv', amigos, na maior / coita do mundo, ca, mao pecado!, / sempr'eu houve por amar desamor; / de mia senhor tod'este mal me vem, / [e] al me fez peor, ca me fez quem / servo servir e nom seer amado por en; (...) {Cantiga de Amor: Muitos me dizem que servi doado, linhas 8-14}

Tabela 7: Correlação entre forma verbal e marcas morfológicas na expressão do antepretérito

Forma verbal Marcas Morfológicas	<i>Haver/ter</i> no imperfeito do indicativo mais particípio neutro	<i>Haver/ter</i> no imperfeito do subjuntivo mais particípio neutro	<i>Haver/ter</i> no presente do indicativo mais particípio neutro	<i>Haver/ter</i> no perfeito do indicativo mais particípio neutro	<i>Haver/ter</i> em qualquer tempo mais particípio feminino/ plural	Total
Feminino/ Plural	-	-	-	-	4	4
Neutro (Masculino/ Singular – Com concordância)	5	3	6	2	-	16
Neutro (Masculino/ Singular – Sem concordância)	2	1	5	7	-	15
Supressão	2	0	4	0	-	6
Total	9	4	15	9	4	

Fonte: elaboração dos autores

Quanto às marcas morfológicas, notamos que o processo de gramaticalização já se encontrava em andamento, haja vista que a maioria dos casos apontaram para as formas neutras do particípio (masculino/singular). Todavia, detectamos o particípio em sua estrutura adjetival, uma vez que os dados nos mostraram formas no feminino e no plural. A propósito dos achados, os dados em concordância ainda eram mais usuais na época medieval (localizamos 20 dados), mas foram, pouco a pouco, neutralizando-se, conforme evidenciamos via uso do particípio neutro sem concordância (15 dados).

Ademais dos três fatores (neutro com concordância, neutro sem concordância e morfologicamente marcado por feminino ou plural), houve necessidade de instituímos mais um fator para contextos como o de (20), em que não era possível identificar marcas morfológicas da terminação participial, somando-se 6 dados neste caso. Ressaltamos que, em relação ao fator das marcas morfológicas, houve casos em que não foi possível reconhecer a concordância explicitamente, uma vez que, como mostra o exemplo (23), o segundo argumento aparecia em forma pronominal, impossibilitando a identificação do gênero referente ao objeto. Inferimos, nesses casos, a concordância da forma participial pelo tipo de cantiga, haja vista que, nas amostras em que isso aconteceu, o gênero do eu lírico seria uma característica inerente à tipologia do texto, como ocorre em (23), dado proveniente de uma cantiga de amigo (cantiga em voz feminina).

- (23) Foi-s'agora meu amig'e por en / **HÁ**-mi **JURADO** que polo meu bem / me quis e quer mui melhor doutra rem; (...) {Cantiga de Amigo: Foi-s'agora meu amig'e por en, linhas 1-6}

Cumpramos destacar que Mattos e Silva (2001) observara também, no português arcaico, mais usos de particípio concordando com o complemento de *haver/ter*, embora a variação já aparecesse documentada, difundindo-se a interpretação de estrutura de tempo composto do século XV para o século XVI. Atesta também Lorenzo (1995) que, na época medieval, o particípio geralmente concorda com o objeto, mas *ter* pode já ser visto como auxiliar em alguns casos. Além da perda de concordância, as duas formas aparecem lado a lado, e, quando há uma palavra no meio, conforme Olbertz (1993), é geralmente um clítico (uma palavra de pouca proeminência prosódica).

3.5. Articulação entre resultados e princípios de gramaticalização

Detalhados e aplicados os fatores sob análise, nossos dados demonstram indícios de gramaticalização que se coadunam aos princípios propostos por Hopper (1991): *estratificação, divergência, especialização, persistência e decategorização*. Sendo a gramática um sistema adaptativo (DU BOIS, 1984), é natural que ocorra a “atribuição de um caráter gramatical a uma palavra outrora autônoma” (MEILLET, 1965 [1912], p. 131); neste caso, em particular, *haver/ter* plenos passam a ocorrer como meros auxiliares de um particípio que encerra a noção verbal. Segundo Ribeiro (1996, p.367), “pode-se considerar que a perda de restrição do traço mais transitivo dos particípios é evidência clara da reanálise de *haver* como verbo auxiliar”. As construções em que o particípio é adjetival ficam, portanto, menos frequentes, mas não são eliminadas da língua. O que geralmente ocorre é

ambiguidade, ou seja, duas leituras são possíveis, por exemplo, “...*havia escondido o tesouro*” ou “...*havia o tesouro escondido*”, tema que consideramos relevante para desdobramento da pesquisa.

De acordo com Givón (1991), no processo de gramaticalização, as mesmas unidades de codificação podem ser usadas para diferentes funções (uma forma – *haver/ter* mais particípio para duas funções: antepretérito e posse mais objeto possuído caracterizado por particípio adjetival). O mais comum, entretanto, é ocorrer uma modificação estrutural para o desempenho de uma nova função (mesmo sujeito, ausência de objeto possuído, combinação com particípio que contradiz a noção de posse e particípio neutro, no caso em pauta neste artigo). Essas observações acoplam-se à perspectiva de Heine e Reh (1984), para quem a gramaticalização é um processo no qual há perdas, por exemplo, de complexidade semântica, significação pragmática, liberdade sintática e substância fonética. Ademais, podemos aludir ao conceito de gramaticalização proposto por Traugott e König (1991), como um processo histórico em que itens lexicais adquirem novo status morfossintático, codificando novas ou diferentes relações.

Dos resultados, destacamos, inicialmente, a coexistência de camadas: (i) usos de *haver/ter* plenos e auxiliares; (ii) usos do particípio neutro e com marcas explícitas de concordância e (iii) variação entre *haver/ter*. Tais achados evidenciam, portanto, o princípio da *Estratificação*. Paralelamente, observa-se o princípio da *Divergência*, já que *haver/ter* permanecem no sistema, como elementos autônomos (por seleção de dois argumentos – sujeito e objeto) e como formas gramaticalizadas (casos em que houve a perda do segundo argumento, provocando, por conseguinte, a coincidência entre sujeitos de *haver/ter* e particípio). Paulatinamente, também, amplia-se a noção de posse: de bens adquiríveis materiais para posse de bens ou qualidades adquiríveis imateriais, até chegar a estruturas em que a posse é uma qualidade inerente, não transferível do possuidor, conforme Mattos e Silva (1995) e Osório (2008).

O princípio da *Especialização* pode ser ilustrado por estreitamento de escolhas formais: ao encaminharem-se para a expressão temporal da forma composta de mais-que-perfeito, *haver/ter* passaram a comportar-se, também, como auxiliares de outros verbos, especializando-se para a construção de tempos compostos. Mesmo especializando-se para a codificação de tempos compostos, observa-se *Persistência*, pois traços do significado lexical e detalhes da história lexical persistem, tendo em vista ocorrências como: “*tinha cercado os inimigos/ tinha os inimigos cercados*”.

Por fim, observamos *Decategorização*, ou seja, perda ou neutralização de marcas morfológicas flexionais do particípio, além da falta de concordância entre as formas participiais e o segundo argumento (direcionando-se ao uso neutro, no masculino/singular), bem como perda de privilégios sintáticos característicos das categorias lexicais, especificamente em relação à seleção argumental, pois *haver/ter* passaram a não mais selecionar um segundo argumento indicando posse material. Primeiramente ocorreu metaforização (selecionavam argumento correspondente a posse inalienável); por fim, passaram a não mais selecionar um segundo argumento, processo que conduziu, paralelamente, ao esvaziamento semântico de *haver/ter*, deslocando-se a proeminência ao particípio, o que consolidou a formação de tempos compostos.

Considerações finais

Em perspectiva analítica do presente ao passado, observamos que o uso de *haver/ter* mais particípio como forma composta foi consolidado gradualmente, razão pela qual investigamos um recorte temporal, o do galego-português, em busca de indícios de emergência dessa forma composta. Assim, situamo-nos em duas vias analíticas. Primeiramente, voltamos ao passado para iluminar o presente. Como observa Faraco (1998, p. 121), “o atual estado de coisas teve uma gênese”; se buscamos “no ontem a raiz do hoje”, retrazendo “o fluxo histórico”, tornam-se mais compreensíveis configurações gramaticais atuais. Paralelamente, por fixarmo-nos em uma sincronia passada, datada e documentada, situamo-nos também na perspectiva da Linguística Histórica *lato sensu*, nos termos de Mattos e Silva (1999).

Todo esse panorama expande os estudos gramaticais do Português, especialmente em relação à história da configuração das estruturas compostas. Os parâmetros de análise propostos mostram indícios do processo gramaticalização, retratando seus primórdios, desde contextos favoráveis à implementação da estrutura composta de antepretérito até aqueles que se configuraram como contextos de resistência das reverberações das formas compostas. Nossas constatações propiciam, portanto, mais conhecimento sobre processos de mudança linguística.

Referências

- BELLO, Andrés. *Gramática de la lengua castellana destinada al uso de los americanos*. Madrid: EDAF, [1847] 1984.
- BENVENISTE, Emile. Estrutura das relações de pessoa no verbo. In: BENVENISTE, Emile. *Problemas de Lingüística Geral I*. São Paulo: Editora Nacional, 1966.
- CASTRO, Ivo. *Sete Ensaios sobre a obra de J. M. Piel*. Lisboa: Instituto de Linguística da Faculdade de Letras de Lisboa, 1988, p. 201.
- COAN, Márluce. *Anterioridade a um ponto de referência passado: pretérito (mais-que-) perfeito*. 1997. 183 f. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 1997.
- COAN, Márluce. *As categorias Tempo, Aspecto, Modalidade e Referência na significação dos pretéritos mais-que-perfeito e perfeito: correlação entre função(ões)-forma(s) em tempo real e aparente*. 2003. 238 f. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, 2003.
- COAN, Márluce; LIMA, Ester Vieira; SAMPAIO, Mariana Freire. Um retrato do pretérito mais-que-perfeito de 1887 a 2012. *D.E.L.T.A.*, v. 35-2, pp. 1-26, 2019.
- DUBOIS, John. Competing Motivations. In: HAIMAN, John (org.). *Typological Studies in Language*. Amsterdam: John Benjamins Publishing Co, 1984. pp. 229-40.
- FARACO, Carlos Alberto. *Lingüística histórica*. São Paulo: Ática, 1998.
- FIORIN, José Luiz. *As Astúcias da Enunciação: as categorias de pessoa, espaço e tempo*. São Paulo: Ática, pp. 127-255, 1996.

FREITAG, Raquel Meister Ko.; OLIVEIRA, Josane Moreira; COAN, Márluce. Formas simples e perifrásticas do verbo em relação ao domínio tempo-aspecto-modalidade. In: CASTILHO, Ataliba; LOPES, Célia Regina (orgs.). *Mudança sintática das classes de palavra: perspectiva funcionalista* (Coleção história do Português Brasileiro, Vol. IV). 1. ed. São Paulo: Contexto, 2018. pp. 186-239.

GIVÓN, Talmy. *Functionalism and grammar: a prospectus*. University of Oregon, 1991.

GONÇALVES, Carlos Alexandre Victorio. Falara-se mais-que-perfeito: estudo presente do tempo pretérito. *ALFA*, v. 37, São Paulo, pp. 135-42, 1993.

HEINE, Bernd; REH, Mechthild. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske, 1984.

HOPPER, Paul. On Some Principles of Grammaticization. In: TRAUGOTT, Elizabeth Closs; HEINE, Bernd. (eds.). *Approaches to Grammaticalization*. Amsterdam/Philadelphia: John Benjamins Publishing Co, 1991.

ILARI, Rodolfo. *A expressão do tempo em português: expressões da duração e da reiteração, os adjuntos que focalizam eventos, momentos estruturais na descrição dos tempos*. São Paulo: Contexto, 1997.

LASS, Roger. *On explaining language change*. New York: Cambridge, 1980.

LOPES, Graça Videira et al. *Cantigas Medievais Galego Portuguesas* [base de dados online]. Lisboa: Instituto de Estudos Medievais, FCSH/NOVA, 2011. Disponível em: <http://cantigas.fcs.unl.pt>. Acesso em: 13 jan. 2021.

LORENZO, Ramón. Galegische Koine. In: HOLTUS, Günter; METZELTIN, Michael; SCHMITT, Christian. (eds.). *Lexikon der Romanistischen Linguistik*. Volume II. Tübingen: Max Niemeyer Verlag, 1995, pp. 649-79.

MARTINS, Kellen Cozine. A variação entre o pretérito mais-que-perfeito simples e composto em textos jornalísticos. *Diadorim*, v. 8, pp. 15-30, 2011.

MARTINS, Kellen Cozine. *A expressão variável da anterioridade a um ponto de referência passado na escrita midiática*. 2010. 131 f. Dissertação (Mestrado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

MATTOSE SILVA, Rosa Virgínia. A sócio-história do Brasil e a heterogeneidade do português brasileiro. *Boletim da ABRALIN*, v. 17, pp. 73-85, 1995.

MATTOS E SILVA, Rosa Virgínia. Orientações atuais da Lingüística Histórica brasileira. *D.E.L.T.A.*, v. 15, Número Especial, pp. 147-66, 1999.

MATTOS E SILVA, Rosa Virgínia. *O português arcaico: morfologia e sintaxe*. São Paulo: Contexto, 2001.

MEILLET, Antoine. L'Évolution des Formes Grammaticales. In: *Linguistique Historique et Linguistique Générale*. 6o. ed. Paris: Honoré Champion, 1965. pp. 130-48.

MONTEAGUDO, Henrique. A Galiza e o espaço linguístico-cultural de expressão portuguesa. In: LOBO, Tânia; CARNEIRO, Zenaide; SOLEDADE, Juliana; ALMEIDA, Ariadne; RIBEIRO, Silvana (orgs.). *ROSAE: linguística histórica, história das línguas e outras histórias* [online]. Salvador: EDUFBA, 2012. pp. 51-64.

OLBERTZ, Hella. The grammaticalization of Spanish haber plus participle. In: MARLE, Jaap. van. *Historical linguistics*. Amsterdam: John Benjamins Publishing Co, 1993. pp. 243-63.

OSÓRIO, Paulo. Linguística histórica e história da língua. Aportações teóricas e metodológicas. *I SIMELP*. USP, 2008. Disponível em: <http://simelp.fflch.usp.br/sites/simelp.fflch.usp.br/files/inline-files/S401.pdf>

PEREIRA, Eduardo Carlos. *Gramática Histórica*. São Paulo: Editora Monteiro Lobato & cia, 1923.

RIBEIRO, Ilza. A formação dos tempos compostos: a evolução histórica das formas ter, haver e ser. In: ROBERTS, Ian; KATO, Mary Aizawa. (orgs.) *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. Unicamp, 1996. pp. 343-86.

SILVEIRA BUENO, Francisco. *A Formação histórica da língua portuguesa*. Rio de Janeiro: Livraria Acadêmica, 1955.

TRAUGOTT, Elizabeth Closs; KÖNIG, Ekkehard. The Semantics-Pragmatics of Grammaticalization Revisited. In: *Approaches to Grammaticalization*, v. 1, Philadelphia: John Benjamins Publishing Co, 1991.

VASCONCELOS, Carolina Michaëllis. *Cancioneiro da Ajuda*. Vol. II, Lisboa, Imprensa nacional - Casa da Moeda (reimpressão da edição de Halle, 1904), 1990.

A CONSTRUÇÃO X-ADA: UM CASO DE CONSTRUCIONALIZAÇÃO

THE X-ADA CONSTRUCTION: A CASE OF CONSTRUCTIONALIZATION

João Carlos Tavares¹

RESUMO

Os objetivos do presente artigo são, em primeiro lugar, discutir o processo de construcionalização de X-ada, apresentando o que essa construção herda da categoria de origem e no que inova, e, em segundo lugar, descrever morfossemanticamente as construções analisadas. Para esses fins, apenas os dois padrões X-ada mais antigos da língua serão analisados, a saber, X-ada_{ATO DE...} e X-ada_{GOLPE}. Esse trabalho toma por base os principais pressupostos da Linguística Cognitiva (LANGACKER, 1987; 2008) e das abordagens construcionais (BOOIJ, 2010; TRAUGOTT; TROUSDALE, 2013). Argumenta-se, ao longo deste texto, que a construção X-ada, de caráter derivacional, tem sua origem no particípio, caracterizando um processo de construcionalização (TRAUGOTT; TROUSDALE, 2013). Para a descrição morfossemântica das construções analisadas, toma-se por base a proposta de Tavares da Silva (2020), cujo cerne é a identificação do significado esquemático da construção morfológica.

PALAVRAS-CHAVE: Construcionalização. Sufixo -ada. Morfologia diacrônica.

ABSTRACT

The objectives of this article are, firstly, to discuss the constructionalization process of X-ada, presenting what this construction inherits from the origin category and in what it innovates, and, secondly, to morphosemantically describe the analyzed constructions. For these purposes, only the two oldest X-ada patterns of the language will be analyzed, namely, X-ada_{ATO DE...} and X-ada_{GOLPE}. This work is based on the main assumptions of Cognitive Linguistics (LANGACKER, 1987; 2008) and constructional approaches (BOOIJ, 2010; TRAUGOTT; TROUSDALE, 2013). It is argued throughout this text that the derivational X-ada construction has its origin in the participle, characterizing a process of constructionalization (TRAUGOTT; TROUSDALE, 2013). For the morphosemantic description of the analyzed constructions, the proposal of Tavares da Silva (2020) is taken as a basis, whose core is the identification of the schematic meaning of the morphological construction.

KEYWORDS: Constructionalization. -ada suffix. Diachronic morphology.

Introdução: objetivos, referencial teórico e corpus utilizado

A partir de dados históricos e com base nos conceitos teóricos da Linguística Cognitiva (LANGACKER, 1987, 2008) e de modelos construcionais (BOOIJ, 2010; TRAUGOTT; TROUSDALE, 2013), o presente artigo tem por objetivo discutir o processo de construcionalização de X-ada, apresentando o que essa construção herda da categoria de origem e no que inova. Para esse fim, apenas os dois padrões mais antigos da língua serão analisados, a saber, X-ada_{ATO DE...} e X-ada_{GOLPE}.

Argumenta-se, ao longo deste texto, que a construção X-ada tem sua origem no particípio, herdando deste o caráter delimitativo e a *sumatividade interna*, nos termos de Becker (2014). Como

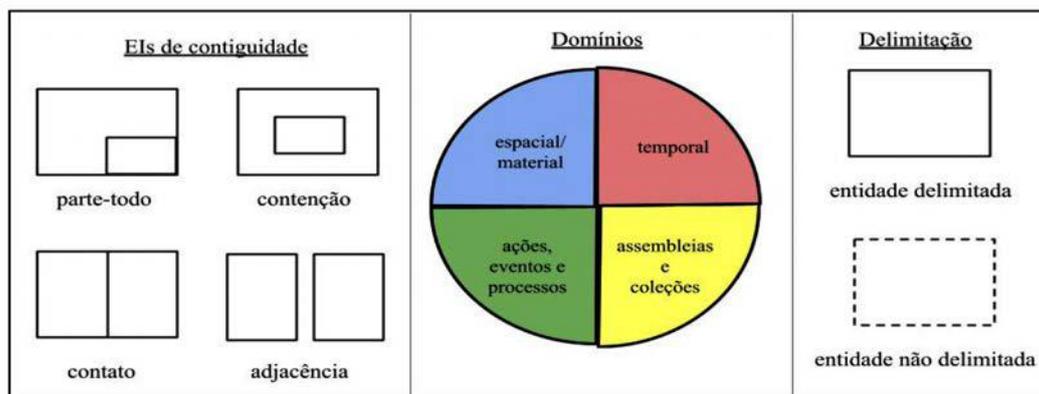
¹ Universidade do Estado do Rio de Janeiro, tavares.jct@gmail.com, <https://orcid.org/0000-0002-2646-3981>.

inovações, passa a figurar como construção derivacional prototípica, ao contrário do participípio (construção mais próxima da flexão prototípica). Tais mudanças caracterizam um processo de construcionalização, nos termos de Traugott e Trousdale (2013).

Por limites de espaço, optou-se por não fazer uma detalhada seção destinada à apresentação do arcabouço teórico que sustenta as análises. Para suprir minimamente essa falta, alguns conceitos serão definidos à medida que surgirem e/ou se fizerem necessários, seja no corpo do texto seja em nota de rodapé. Por ora, o referencial teórico pode ser resumido nos quatro tópicos abaixo:

- ◆ Para a compreensão das propriedades herdadas, são caras as noções langckerianas de escaneamento sequencial e sumário, bem como sua proposta acerca da definição semântica das classes de palavras e da distinção entre perfectivo e imperfectivo (LANGACKER, 2008).
- ◆ Para o tratamento da mudança, assume-se aqui a distinção entre mudança construcional e construcionalização (TRAUGOTT; TROUSDALE, 2013).
- ◆ Para o tratamento das estruturas linguísticas em perspectiva construcional, assume-se as noções de construção², esquema³ e subesquema⁴.
- ◆ Para a descrição morfossemântica das construções analisadas, toma-se por base a proposta de Tavares da Silva (2020), para quem os afixos denominais do português podem ter sua contraparte semântica descrita a partir da combinação das noções de esquema imagético, domínio conceptual e estado de delimitação (delimitado vs. não delimitado). Cada uma dessas noções é representada conforme as imagens a seguir:

Figura 1: esquema imagético, domínio conceptual e estado de delimitação



Fonte: Tavares da Silva (2020, p. 46)

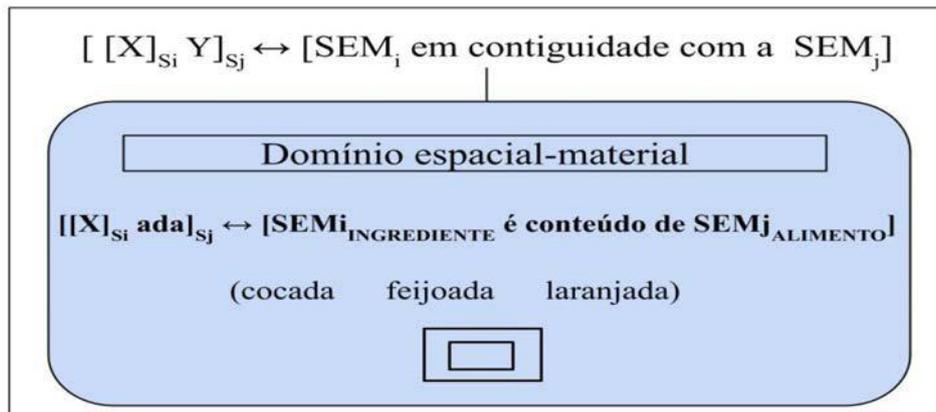
² **Construção** pode ser definido como um pareamento entre forma e significado-função (LANGACKER, 1987; CROFT, 2001; GOLDBERG, 1995, 2006), em que significado-função pode reunir informações semânticas, pragmáticas, discursivas, textuais e sociais.

³ **Esquema** pode ser entendido como uma estrutura conceitual de alto nível que organiza a experiência anterior e nos ajuda a interpretar novas situações (GURECKIS; GOLDSTONE, 2010, p. 725), sintetizando, a partir da abstração dos componentes formais e semânticos mais estáveis e importantes, nossa experiência linguística de contato com um grupo de palavras. Em esquemas, o que é variável pode ser abstraído em forma de *slots* a serem preenchidos, ao passo que o que é recorrente pode ser fixado na construção. Todo esquema é uma construção, mas nem toda construção é um esquema.

⁴ Um **subesquema** é uma instancia que sinaliza uma hierarquização, isto é, uma especificação mais detalhada e subordinada a um esquema mais alto (BOOIJ, 2010). Todo subesquema é um esquema, mas nem todo esquema é um subesquema.

O autor propõe que o significado das construções morfológicas deve ser descrito esquematicamente e não apenas por meio de paráfrases. Assim, **a identificação do significado esquemático da construção morfológica é o cerne dessa proposta**. A construção X-ada_{ALIMENTO}, por exemplo, é descrita em Tavares da Silva (2021)⁵, como segue:

Figura 2: O subesquema X-ada_{ALIMENTO}



Fonte: Tavares da Silva (2021, p. 112)

No topo da rede, está o esquema geral dos denominais, que reflete a estrutura esquemática capaz de abarcar não só todas as construções X-ada, mas também construções com outros sufixos denominais (-eiro(a), -ário, -agem, -al, -aria). Essa construção prevê que a relação entre a entidade designada pela palavra base e pela derivada é sempre de contiguidade (laranja-laranjeira [PARTE-TODO], água-aquário⁶ [CONTENÇÃO], canela-caneleira [CONTATO]).

Logo abaixo, está o subesquema $[[X]_{si} \text{ ada}]_{sj} \leftrightarrow [SEMi_{INGREDIENTE} \text{ é conteúdo de } SEMj_{ALIMENTO}]$. Nas palavras X-ada_{ALIMENTO}, tanto a base quanto o derivado são substantivos concretos. Sendo base e derivado substantivos concretos, o domínio envolvido é o material-espacial. A base designa o principal ingrediente, que pode ser o único (laranja) ou o mais saliente (feijão), e o derivado uma iguaria culinária, que pode ser comida ou bebida. O esquema de imagem é o de contenção (ingredientes são conteúdos). As palavras entre parênteses são apenas para exemplificar algumas formações sancionadas por essa construção. Portanto, esse é o significado esquemático dessa construção: uma relação base-derivado ancorada no domínio material-espacial e no esquema imagético de contenção.

O corpus utilizado para esta pesquisa se constitui de 422 palavras, sendo 222 com a acepção “ato de X” e 200 com a acepção “golpe”. Esses dados foram extraídos do dicionário Houaiss (2002) e dos trabalhos de Becker (2014) e de Takahashi (2014). Como esses dois autores também se valeram

⁵ Por questão de espaço, não será feita, neste texto, uma revisão detalhada acerca do modelo proposto por Tavares da Silva (2020). Para maiores detalhes, indico as seguintes leituras: *A semântica dos sufixos denominais* (TAVARES DA SILVA, 2020) e *Breve estudo sobre as construções denominais X-ada* (TAVARES DA SILVA, 2021).

⁶ A primeira acepção dessa palavra ainda em latim é “reservatório de água; local que contém água” (*aqua-aquarium*).

do dicionário Houaiss, praticamente todas as palavras do corpus vêm dessa fonte, com poucas delas providas de textos medievais, coletadas dos referidos trabalhos.

Para entender a gênese das construções X-ada_{ATO DE...} e X-ada_{GOLPE}, traçou-se uma cronologia estruturada com base nas datações encontradas no dicionário Houaiss (2002) e no dicionário etimológico de Cunha (1982). No entanto, tomar por base abonações de dicionários implica lidar com alguns percalços. Em primeiro lugar, a data apresentada nos dicionários é referente à data do documento mais antigo em que a palavra foi atestada e não ao surgimento da palavra na língua. Além disso, há um grande número de palavras sem data de abonação. Dos 422 dados analisados, 147 não apresentam datação em nenhum dos dicionários consultados (cf. tabela 1). Outro problema está relacionado à polissemia. Raras são as vezes em que o dicionário apresenta uma data de abonação para cada significado da palavra, a exemplo de ‘pedrada’ “1. (XIII) ação de lançar uma pedra; 2. (1679) golpe com pedra; 3. (1665) Derivação: sentido figurado. palavra ou gesto que atinge a dignidade ou a honra de alguém; insulto, ofensa. **Etimologia:** pedra + -ada; ver petr-”

Tabela 1: Distribuição dos dados por grupos morfossemânticos.

	Com datação	Sem datação	Total
Ato de X	156	66	222
Golpe	119	81	200

Fonte: elaboração do autor

Contudo, se, por um lado, esse terreno se mostra arenoso, por outro lado, é bastante confiável, pois, se conjugadas a outros elementos, as datações apresentadas em dicionários podem fornecer um panorama geral acerca das trajetórias morfossemânticas, permitindo que se chegue a conclusões plausíveis e coerentes.

Toda a análise é feita com base em dados de língua portuguesa, pois, até o momento, carece-se de dados ou fontes que comprovem a existência do sufixo -ata latino com as mesmas acepções do -ada português.

O surgimento de uma nova construção implica herança de propriedades das sua(s) forma(s) gatilho(s) ao mesmo tempo em que adquire propriedades inteiramente inovadoras que lhe são próprias. Desse modo, entender o surgimento da construção X-ada impele tratar, mesmo que brevemente, do participio, pois o processo se deu da seguinte forma: participio > deverbais > denominal. Assim, as próximas seções são destinadas a discutir essa trajetória.

1. De -ata a -ada: mudança construcional e construcionalização do latim ao português

Vários filólogos, gramáticos e linguistas convergem ao remontar a origem de -ada ao feminino (ou neutro plural) do participio passado latino -ata (MEYER-LÜBKE, 1894; ROHLFS, 1969;

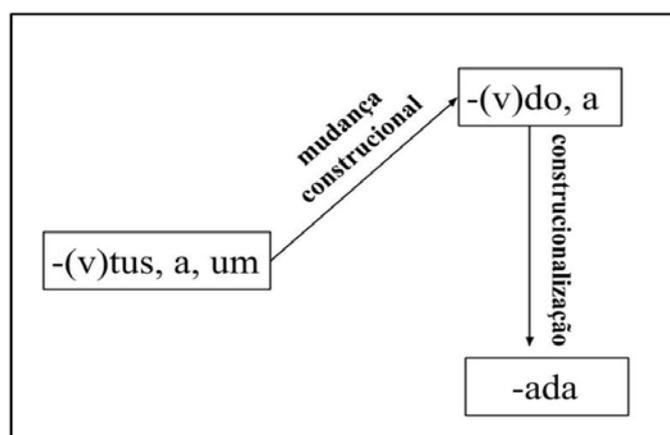
TEKAVČIĆ, 1972; SAID ALI, 1964; CUNHA, 1982; HOUAISS, 2002; SOLEDADE, 2004, 2020; BECKER, 2014; TAKAHASHI, 2014; apenas para citar alguns). Mas, ao contrário do particípio presente, que sofreu recategorização (GONÇALVES; TAVARES DA SILVA, 2020), o particípio passado permanece na língua portuguesa. Sendo assim, coexistem a construção participial, com suas propriedades verbo-nominais, e a construção X-ada, historicamente derivada da primeira.

A noção de que a mudança nem sempre resulta em desuso da forma mais antiga em detrimento da forma inovadora tem sido ressaltada nos estudos de base funcionalista, sobretudo em pesquisas sobre gramaticalização. Mas, numa perspectiva construcional, é apenas em *Constructionalization and Construction changes* (TRAUGOTT; TROUSDALE, 2013) que esse e outros aspectos da mudança linguística ganharam tratamento mais apropriado, amparados pelos conceitos de *mudança construcional* e *construcionalização*.

Os autores definem construcionalização como a criação de nova forma e novo significado, ou seja, é a formação de novos tipos de nós, que têm nova sintaxe ou morfologia e novo significado codificado na rede linguística de uma população de falantes. (TRAUGOTT; TROUSDALE, 2013, p. 22). Acrescentam ainda que a construcionalização é acompanhada por mudanças no grau de esquematização, de produtividade e de composicionalidade. Já a mudança construcional é “uma mudança que afeta uma dimensão interna de uma construção. Não envolve a criação de um novo nó” (TRAUGOTT; TROUSDALE, 2013, p. 26)⁷. Isso significa que, na mudança construcional, uma ou mais estruturas internas são modificadas. Pode haver uma mudança apenas na fonologia ou na morfossintaxe e não no significado-função, por exemplo.

Com base nessa distinção, pode-se afirmar que o particípio português é resultado de mudanças construcionais ocorridas no particípio passado latino, ao passo que a construção X-ada é resultado de construcionalização, conforme representado na figura a seguir.

Figura 3: mudança construcional do particípio e construcionalização de -ada



Fonte: elaboração do autor

⁷ A constructional change is a change affecting one internal dimension of a construction. It does not involve the creation of a new node. (TRAUGOTT & TROUSDALE, 2013, p. 26)

É possível, no entanto, que o processo de construcionalização de X-ada remonte ao latim medieval, como sugerem os trabalhos dos filólogos consultados (cf. discutido na próxima seção). Porém, até o momento, os dados e as fontes disponíveis não são suficientes para comprovar tal hipótese e o presente estudo toma por base apenas dados da língua portuguesa, suas etimologias e suas datações.

1.1. O particípio português: um caso de mudança construcional

O particípio passado latino (doravante Ppss) caracteriza-se por ser uma categoria híbrida, apresentando tanto propriedades verbais quanto nominais.

Assim como os nomes, apresentava as categorias gramaticais de caso, de número e de gênero, comportando-se morfossintaticamente como adjetivos de primeira classe (ex.: *amatus, a, um sum* “sou amado(a)”). Decorrentemente, já no latim, é possível encontrar alguns Ppss gramaticalizados em verdadeiros adjetivos e, por isso, muitos deles aparecem nos dicionários como adjetivos de primeira classe, a exemplo de *perītus, a, um*⁸, derivado do particípio de *perēō*. Podiam também, como é comum à classe dos adjetivos, substantivar-se.

No que tange às suas características verbais, ao contrário do particípio presente, que era ativo, o Ppss atualizava a voz passiva: “Cumpra ainda notar que o particípio presente só tem a forma ativa e o particípio passado, só a passiva” (FARIA, 1958, p. 463). Isso se dava porque o Ppss era formado a partir do tema de *perfectum*, ou seja, aspecto perfectivo e, sendo assim, “apresentava basicamente um sentido resultativo, indicando uma ação concluída” (TAKAHASHI, 2014, p. 336).

Além das propriedades de tempo (passado) e aspecto (perfectivo), o Ppss podia receber pronome sujeito ou objeto. Nesse aspecto, o comportamento do particípio português é menos verbal que o das outras formas nominais (infinitivo e gerúndio), pois um particípio nunca pode receber um pronome tônico ou átono como sujeito ou objeto, nem pode ser seguido de um clítico, afastando-se, neste ponto, de outras línguas românicas, como o italiano, por exemplo (cf. 1a). Assim, a frase em (1b) é agramatical em português:

(1a) Le informazioni datemi sono incomplete.

(1b) *As informações dadas-me são incompletas.

Com relação à constituição morfológica, o Ppss é formado a partir de um sufixo *-to-, que se une ao tema de *perfectum* (ERNOUT, 1914). Essa associação gera tanto o adjetivo verbal quanto a flexão perifrástica da voz passiva nos tempos de *perfectum*. No que tange à funcionalidade, o formador de particípio em latim tem comportamento mais próximo dos afixos flexionais, pois apresenta algumas das principais características da flexão.

⁸ *perītus, a, um* - 1. que tem a experiência de, experimentado; 2. Versado, instruído, perito, hábil em.

(2) *Hunc sortita⁹ locum miserae sunt ossa puellae*

(Os ossos da pobre mulher foram lançados ao acaso neste lugar)

Com base em (2), pode-se ver que a marca de participio: (a) depende de um contexto sintático (relevância sintática), como a voz passiva analítica; (b) *-tus, a, um* não é cabeça semântica, pois a interpretação parte da base e a marca de participio apenas agrega propriedades gramaticais (a acepção “lançar ao acaso” está na base); (c) não é cabeça morfológica, pois o gênero é determinado pelo sujeito ou substantivo com o qual concorda (o participio *sortita* concorda em gênero, número e caso com *ossa*).

Em suma, o participio português, portanto, se afastou mais da sua origem latina que as outras formas nominais, se aproximando mais do nome se comparado ao infinitivo e ao gerúndio, que apresentam mais propriedades verbais que aquele. Por sua própria natureza híbrida, não tardou em gerar palavras que funcionassem como autênticos adjetivos e substantivos, o que permanece até hoje.

Em comparação ao participio latino, de todas as propriedades elencadas, o participio português perde a capacidade de receber argumentos e de expressar tempo, já que “em si não marca o tempo da situação” (TRAVAGLIA, 1985, p. 200). A marcação morfológica de caso não deve ser considerada, pois foi mudança que ocorreu na língua como um todo.

Pode-se afirmar, então, que o participio passado português difere do participio latino por ter sofrido mudanças construcionais ao longo do tempo, sendo as mais perceptíveis a mudança fonológica (*-atus(a) > -ado(a)*), a não marcação de tempo e a impossibilidade de receber pronome sujeito ou objeto. Pode-se, ainda, afirmar que o Ppss latino e o participio português são a mesma construção, que se manteve na língua por milênios.

É essa construção participial, ao que tudo indica, que deu origem aos nomes de ação em *-ada*. Meyer-Lübke (1894), bem como filólogos posteriores (ROHLFS, 1969; TEKAVČIĆ, 1972), defende a hipótese de que a mudança de Ppss a sufixo derivacional tem como ponto de partida a elipse do substantivo em construções Subs+Adj¹⁰, o que resultou em posterior substantivação do adjetivo e consequentes alterações da função semântica dessa forma.

Segundo esses autores, o Ppss ocorria como adjetivo em expressões sintagmáticas como *pecunia collecta* (dinheiro coletado), *causa defensa* (causa defendida) *peccata remissa* (pecados perdoados) dentre outras. Com a eliminação do substantivo por elipse, o adjetivo participial foi, posteriormente, reanalisado como substantivo, o que acarretou perda do caráter perfectivo e da voz passiva¹¹.

⁹ Participio de **sortior, iris, iri, itus sum**. v. dep. intr. Tirar à sorte.

¹⁰ Dado interessante é a etimologia de ‘levada’, que, segundo José Pedro Machado (1972), vem do latim *levata (aqua)* ‘água levada’, sugerindo a elipse do substantivo, o que vai ao encontro da referida hipótese.

¹¹ Cabe destacar que a proposta de mudança categorial por elipse do substantivo também é sugerida para outros afixos. Esse é o caso do sufixo *-eiro(a)*, cuja origem remonta ao sufixo latino *-arius, a, um*, formador de adjetivos, que passou a *-arius*, um formador de substantivos (cf. VIARO, 2011), a exemplo de *Servus coquinarius* (“o servo que trabalha na cozinha”) > *coquinarius* (“aquele que trabalha na cozinha”) e *Vir ferrarius* (“o homem que trabalha com ferro”) > *ferrarius*

Forma latina:	<i>pecunia collecta</i>	>	\emptyset <i>collecta</i>	>	<i>collecta</i>	
	o dinheiro coletado		>	o coletado	>	a coleta

Para esses autores, portanto, a reanálise do adjetivo participial fez com que -ada, perdesse suas propriedades verbais (voz e aspecto) e assumisse o sentido/função de nomes de ação. Assumindo, então, a origem participial da construção X-ada, a questão que se coloca é: o que essa construção herda da sua categoria de origem e no que inova? A próxima seção se destina a esse debate.

1.2. A construção X-ada_{ATO DE...}: do participípio ao formador de nomes de ação

Como já exposto, o sufixo -ada formador de substantivos tem sua origem no participípio. No que tange a propriedades morfosintáticas, o sufixo -ada formador de nomes de ação ('pegada', 'batucada') figura como afixo derivacional, pois, ao contrário do participípio, deixa de ter relevância sintática e passa a ser a cabeça morfológica, pois determina a categoria lexical da palavra (sempre um substantivo) e o gênero (sempre feminino). Passa também a ser a cabeça semântica da construção, pois a interpretação parte do afixo (ex.: olhada "ato de olhar").

Com relação às propriedades semânticas, é preciso partir do significado esquemático do participípio para averiguar o que -ada herda dessa construção e o que são inovações adquiridas. Não será feita aqui ampla descrição morfossemântica da construção participial, limitando-se esta seção às questões mais relevantes e suficientes para entender o processo de construcionalização -(v)do, a > X-ada. Assim, serão apresentados e discutidos os conceitos de *aspecto* e de *escaneamento*, no intuito de argumentar que o efeito de delimitação e a sumatividade interna das construções X-ada são heranças/decorrem do aspecto perfectivo (propriedade presente na construção participial) e do tipo de escaneamento imposto à classe dos não verbos (cf. LANGACKER, 2008).

A diferença entre tempo e aspecto é sutil, pois ambas fazem referência à temporalidade da situação, a um conceito de tempo. Não à toa, Travaglia (1985) faz a distinção entre **tempo** e **TEMPO** definidos nos seguintes termos: **tempo**: categoria verbal (correspondente às épocas: passado, presente, futuro); **TEMPO**: ideia geral e abstrata de tempo sem consideração de sua indicação pelo verbo ou qualquer outro elemento da frase. O autor ressalta que tanto **tempo** quanto **aspecto** são categorias de TEMPO,

entretanto as duas não se confundem pois: **tempo** situa o momento de ocorrência da situação a que nos referimos em relação ao momento da fala como **anterior** (passado), **simultâneo** (presente) ou **posterior** (futuro). É uma categoria dêitica. [...] aspecto não é uma categoria

("aquele que trabalha com ferro"). Assim, o adjetivo *ferrarius*, que podia ser parafraseado como "relativo ao ferro", após a eliminação do substantivo no complexo Subs+Adj, passou a ter a paráfrase equivalente a "aquele que trabalha com ferro", além de sofrer mudança categorial, passando a substantivo. A mudança categorial por elipse é, inclusive, um processo bastante comum e acontece ainda hoje em português como dão mostra as reanálises telefone celular > celular, forno microondas > microondas, empregada doméstica > doméstica, guarda de segurança > segurança, dentre tantas outras.

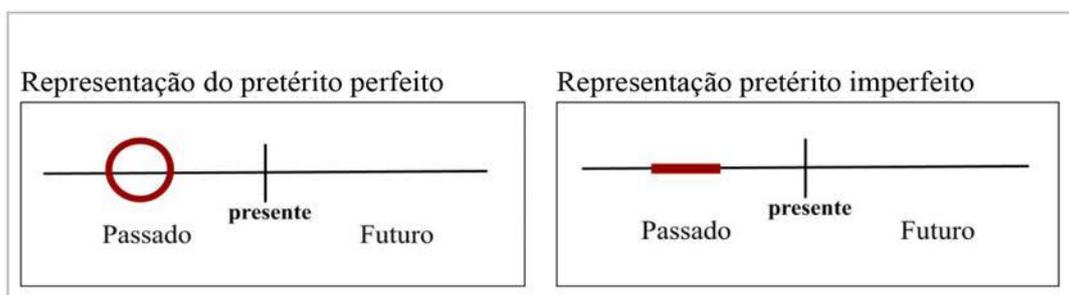
dêitica. Aspecto são as diferentes maneiras de ver a constituição temporal interna da situação, sua duração. [...] podemos dizer que tempo é “um TEMPO externo à situação” e o aspecto é “um TEMPO interno à situação”. (TRAVAGLIA, 1985, p. 52)

Para ilustrar bem essa distinção, podemos tomar como exemplo o pretérito perfeito e o pretérito imperfeito em português. Essas duas formas verbais expressam o momento passado tomando o presente como ponto de referência (cf. figura 4); a diferença entre as formas é aspectual, pois estão relacionadas a diferentes perspectivas que se pode ter em relação à constituição temporal interna da situação.

(3a) Antônio *ouviu* música o dia todo. (pret. perf. /aspecto perfectivo)

(3b) Antônio *ouvia* música, quando saiu de casa. (pret. imperf. /aspecto imperfectivo)

Figura 4: Distinção entre pretérito perfeito e imperfeito



Fonte: elaboração do autor

Em (3a), a situação passada é vista como um todo, sem ênfase nas suas fases constitutivas. Na imagem à esquerda, o traço interno ao círculo representa o intervalo em que a situação ocorreu, ao passo que o círculo perfilado representa a ênfase que é dada ao todo da situação. Esse aspecto é conhecido como perfectivo. Já na imagem à direita, a situação passada é vista a partir da sua estrutura interna, com ênfase nas suas fases constitutivas. O traço mais fortemente marcado na linha do tempo representa o perfilamento que é dado ao próprio intervalo em que a situação ocorreu, ou seja, à estrutura interna. Esse aspecto é conhecido como imperfectivo.

Travaglia define o perfectivo como

caracterizado por **apresentar a situação como completa**, isto é, em sua totalidade. O todo da situação é apresentado como **um todo único**, inalisável, com começo, meio e fim englobados juntos. Não há tentativa de dividir a situação em suas fases de desenvolvimento. É como se a situação fosse vista de fora, em sua globalidade. (TRAVAGLIA, 1985, p. 96)

No âmbito da Linguística Cognitiva, Langacker (2008, p. 147) afirma que o contraste semântico entre perfectivo e imperfectivo é exatamente a propriedade de delimitação, sendo os

perfectivos delimitados no tempo, enquanto os **imperfectivos não delimitados especificamente**¹². Sendo então o particípio um marcador de aspecto perfectivo (TRAVAGLIA, 1985, p. 199), possui, portanto, caráter delimitativo e totalizador. Em outras palavras, o particípio põe ênfase (perfila) no todo e não nas fases constitutivas.

Outra característica das formas participiais é o tipo de escaneamento. *Escaneamento* (LANGACKER, 2008) refere-se à nossa capacidade de rastrear uma relação através do tempo ou no espaço. Há dois tipos de escaneamento: escaneamento sequencial e escaneamento sumário. O primeiro equivale a rastrear mentalmente um evento que se desenrola ao longo do tempo. Langacker (2008) exemplifica com a imagem de uma bola rolando um declive abaixo, salientando que, em nossa visualização em tempo real desta ocorrência, vemos a cada instante a bola em apenas uma posição durante todo o processo. Nossa compreensão dessas várias etapas na linha do tempo depende da nossa habilidade de acessar esses estados componentes na sequência precisa (e, por isso, escaneamento sequencial) de sua manifestação temporal.

Somos, no entanto, capazes de visualizar eventos de outra maneira, sobrepondo mentalmente cada etapa do processo, resultando na sua ativação simultânea. O resultado final é que todos os estados componentes são simultaneamente ativos e disponíveis. Uma boa analogia para diferenciar o escaneamento sequencial do sumário é entender aquele como um pequeno filme e este como uma fotografia de múltipla exposição (LANGACKER, 2008).

O autor relaciona o tipo de escaneamento às classes gramaticais. Sendo assim, o verbo perfila um processo, definido esquematicamente como uma relação escaneada sequencialmente em sua evolução ao longo do tempo (LANGACKER, 2008, p. 147). Apesar de sua base verbal, particípios não são eles mesmos verbos e normalmente perfilam relações não processuais¹³ (escaneamento sumário). Então o ajuste mínimo provocado pela participialização é a imposição do escaneamento sumário.

Em suma, particípios possuem caráter delimitativo e totalizador (propriedade relacionada ao aspecto perfectivo), bem como ativação simultânea dos estados componentes (fruto do escaneamento sumário). A proposta do presente trabalho, portanto, é a de que, do ponto de vista esquemático, a construção X-ada deverbal herda da participial seu caráter delimitativo e totalizador, bem como a simultaneidade dos estados/elementos componentes. Becker (2014) já havia percebido o caráter delimitativo das formações X-ada, além de destacar o que chama de *sumatividade interna*.

Quando se passa em revista os subpadrões de -ada, compreende-se que **o efeito de delimitação conceitual é primordial na “ação” de -ada, mas que não consegue explicar por completo o “funcionamento” do sufixo. Não deve passar despercebido que -ada veicula um matiz coletivo particular que podemos chamar de “sumatividade interna”**. Observando a leitura

¹² Essa caracterização parece ser consensual entre os linguistas, mesmo entre os de inflexão divergentes. Numa abordagem gerativa, por exemplo, Iatridou, Anagnostopoulou & Izvorski (2003) afirmam que os traços subjacentes às morfologias perfectivas e imperfectivas são [+delimitado] e [-delimitado], respectivamente.

¹³ **Relações processuais** são aquelas em que a evolução temporal é essencial para sua caracterização e reconhecimento. O verbo é uma expressão que perfila um processo. Já na **relação não processual**, não há evolução temporal. Adjetivo, advérbio, preposição, infinitivo e particípio são todos caracterizados como não processuais.

deverbal de “ação e efeito de V” (como *chamada*), a acepção de “golpe com N” (*martelada*), de “coletivo de N” (*meninada*) ou de “duração prolongada” (invernada), **em todos os casos nos defrontamos com unidades externamente delimitadas, mas que dispõem de uma estrutura interna complexa composta de várias fases (portanto, “sumatividade”).** Percorramos outra vez a lista das “leituras”: **1.** A *chamada* (e inclusive *chegada*) é um evento que pressupõe uma fase preparatória, uma acumulação suficiente de fases interativas (no caso de chegada são, melhor dizendo, fases prévias ao ponto culminante), uma fase final e, enfim, uma fase posterior (ou “resultativa”). **2.** A acepção de “golpe com N” (*martelada*) pressupõe também um evento com estrutura interna que não pode ser introduzida pela base que designa um nome de instrumento (martelo). [...] **3.** O subpradrão “coletivo” (*meninada*) se refere a conjuntos que se compõem de indivíduos (logo, dispõem de uma mesma estrutura interna complexa). **4.** O subgrupo “duração prolongada” (invernada) focaliza inteiramente a estrutura interna de uma unidade temporal (daí o efeito de prolongamento). Por fim, podemos postular que **o sufixo -ada produz um efeito duplo: gera unidades delimitadas que pressupõem uma estrutura interna (complexa). O traço [+delimitado], assim como o matiz de “sumatividade”** (que se apresenta como uma manifestação da noção de “coletividade”) **constituem a semântica abstrata do sufixo -ada** que desenrola o seu potencial nas diferentes leituras e acepções. (BECKER, 2014, pp. 124-5, grifos meus)

Pode-se afirmar, portanto, que o *insight* do autor acerca do caráter delimitativo e da sumatividade interna das construções X-ada é, na verdade, herança do significado esquemático do participípio. A delimitação tem origem no aspecto perfectivo das formas participiais e a sumatividade interna resulta do escaneamento sumário.

No entanto, isso não explica a mudança categorial de -ada. Em outras palavras, isso apenas nos mostra que tanto o participípio quanto a construção X-ada deverbal apresentam propriedades semânticas comuns que, conforme a presente proposta, foram herdadas daquele por esta. Mas não explica a passagem de -ada formador de participípio a formador de nomes de ação; logo, não explica também a decorrente mudança semântica de *relação* a *coisa*¹⁴.

Como explanado, a participialização implica mudança no tipo de escaneamento (sequencial > sumário), mas isso por si só não implica uma mudança de perfil. Para que um conteúdo conceptual seja enquadrado como substantivo – que é o caso das construções X-ada deverbais (‘a entrada’, ‘a arrancada’ etc.) –, é necessária uma mudança no perfil para uma coisa, que pode ser tanto um participante no processo original ou então uma reificação conceitual desse processo em si.

A palavra ‘vendedor’ pode exemplificar o primeiro caso. O verbo ‘vender’ expressa um evento que apresenta evolução temporal, ou seja, é um processo. Além disso, vender é um processo que se caracteriza por apresentar 2 participantes: aquele que vende e a coisa vendida¹⁵. O sufixo -dor gera,

¹⁴ Esquemáticamente, substantivo é uma expressão que perfila uma **coisa**, entendida como termo técnico para designar qualquer produto da *reificação conceitual*. Já **relação** é o termo técnico utilizado para se referir ao tipo de perfilamento de classes básicas não substantivos, como verbo, adjetivo e advérbio (ver LANGACKER, 2008, p. 93-108).

¹⁵ Para fins de simplificação, a análise feita exclui o terceiro participante da cena de vender, a saber, aquele que compra, ou seja, para quem é vendido.

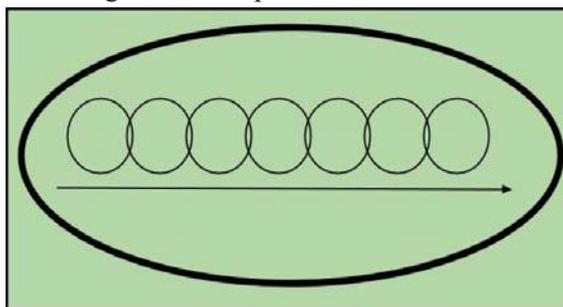
portanto, uma mudança de perfil na palavra base, focalizando/perfilando um dos participantes – aquele que vende. Já o segundo caso – reificação conceitual do processo em si – ocorre por metonímia.

Também é bastante comum para as mesmas formas [infinitivos e participios] funcionar gramaticalmente como substantivos. [...] Esta extensão para uso nominal é bastante simples, dada a descrição do CG de categorias de base. Coisas e relações não processuais representam um agrupamento natural uma vez que ambos interpretam uma situação de forma sumária. Conseqüentemente, a visão holística imposta pela marcação de infinitivo ou de participio é um dos dois fatores essenciais envolvidos na derivação de um substantivo a partir de um verbo. O outro fator requisito é uma mudança no perfil de uma relação para uma coisa. **Suponha-se, então, que um infinitivo ou participio devem ser submetidos a uma tal mudança. Se não houver nenhuma marcação adicional para sinalizá-la, a mesma forma vai perfilar uma coisa e, portanto, ser classificada como um substantivo.** Isto não é nem implausível nem improvável. **Uma mudança implícita no perfil nada mais é que o fenômeno linguístico onipresente conhecido como metonímia** (LANGACKER, 2008, p. 119, grifos meus).

Ora, os substantivos deverbais em -ada, ao tempo das primeiras formações, não são frutos do acréscimo de um sufixo a uma base verbal. Afinal, -ada não formava substantivos, mas participios. Essa mudança de perfil, então, só pode ter se dado por metonímia. Importante frisar que tal análise se concilia com a proposta de filólogos como Meyer-Lübke (1894), para quem o ponto de partida da mudança do Ppss a sufixo derivacional foi a elipse do substantivo, ou seja, um processo metonímico, logo, uma mudança implícita no perfil. Essa análise então encontra respaldo nas pesquisas de cunho histórico.

Foram essas formações, frutos de processos metonímicos, que fomentaram a emergência da construção $[[X]_v \text{ ada}]_s \leftrightarrow \{\text{ato/ação de } x\}$. A imagem a seguir, representa o significado esquemático dessa construção.

Figura 5: significado esquemático dos deverbais X-ada



Fonte: elaboração do autor

O círculo indica uma ação/processo qualquer, ao passo que sua sequência representa o escaneamento sumário, ou seja, a simultaneidade das várias etapas. A seta indica a transcorrência desse processo no tempo. A elipse em destaque indica o perfilamento que recai sobre o todo (caráter delimitativo), ficando subfocalizadas suas partes constitutivas e seu caráter temporal. Uma vez que

as fases internas foram sumarizadas, elas funcionam como uma unidade, como fotografia de múltipla exposição, para usar aqui a analogia proposta por Langacker. Pela base verbal e, sobretudo, por expressar um nome de ação, o domínio que ancora esse subesquema é o das ações/eventos/processos (cor verde) (cf. TAVARES DA SILVA, 2020).

Em suma, o nome deverbal em -ada designa uma ação atemporalizada e com ênfase no processo em si¹⁶. É a partir do significado esquemático da construção deverbal que será analisada a construção denominal, pois, como será discutido na próxima seção, os deverbais tiveram importante papel na emergência da construção X-ada_{GOLPE}.

1.3. A construção denominal X-ada_{GOLPE}

Os padrões de formação de palavras não emergem do nada e partem sempre de uma forma específica ou um pequeno grupo de duas ou três palavras que podem receber o nome de forma(s)-gatilho. Na ausência de dados que nos aponte a(s) forma(s)-gatilho, é preciso, ao menos, identificar o processo que deu origem às primeiras formações do padrão, como foi o caso já discutido da elipse do núcleo do SN e consequente reanálise do adjetivo participial e de sua estrutura interna. Sendo assim, nesta seção, argumenta-se que as formas-gatilho que impulsionaram o surgimento do padrão X-ada_{GOLPE} foram provavelmente as palavras ‘lançada’ e ‘porrada’.

No corpus analisado, foram encontradas apenas três palavras cujas etimologias registradas remontam ao latim: ‘levada’ (1067), ‘cilada’ (1344) e ‘parada’.

levada (1067)

ato ou efeito de levar

1. torrente de água que se desvia de um rio para mover moinhos, fábricas etc., para irrigar terrenos, para alimentar reservatórios etc.
2. queda-d’água em que a água ger. escorre por entre pedras
3. construção de terra, pedra etc. que se destina a represar as águas dos rios, para que sejam utilizadas no abastecimento, na agricultura etc.; açude

Etimologia: segundo JM, do lat. levata (aqua) ‘água levada’; ver llev(i)-; f.hist. 1067 levata, 1258 Levada top., sXV levada

cilada (1344)

1. ação de surpreender o inimigo ou a caça; emboscada
2. Derivação: por metonímia.
lugar encoberto, ger. junto ao caminho, onde se fica de tocaia para aguardar o inimigo ou a caça
3. ação que visa iludir, lograr; ardil, estratégia, armadilha, arapuca
4. perfídia, traição

Etimologia: lat.medv. da península Ibérica, segundo Corominas, celata,ae ‘id.’, fem.substv. de celatus,a,um, part.pas. de celare ‘esconder’, prov. através do esp. celada (1140); f.hist. 1344 celada, sXIV cilada, sXIV çuada, 1603 silada

¹⁶ Cabe ressaltar que, por metonímia, nomes de ação podem designar um dos participantes, gerando polissemia: ‘pousada’ (ato de pousar ou local do pouso), ‘picada’ (ato de picar ou o ferimento, efeito da picada), ‘mijada’ (ato de mijar ou a quantidade de urina vertida em cada micção), dentre outros.

parada (1536)

1. ação ou efeito de parar
2. lugar em que se pára, esp. os pontos de transportes públicos
Ex.: a próxima p. fica perto do parque
3. interrupção de uma atividade ou movimento; paralisação, suspensão
Ex.: um problema mecânico provocou a p. nos trens do metrô

Etimologia: lat. vulg. *parata, f. substv. do part. pas. de paráre ‘esforçar-se para obter’, donde ‘obter, alcançar, adquirir com dinheiro, comprar’; ver para-; a1536 é a data para ‘lugar em que se pára’ e 1672 é a data para ‘ação de parar’.

No período que vai do séc. IX ao XII, das 13 palavras X-ada, 9 são deverbais, cuja acepção mais geral é acolhida pela paráfrase “ato/ação de X”. As outras quatro são ‘lançada’, ‘porrada’, ‘canada’ e ‘jugada’. Fica nítido então que até o século XII, há apenas um padrão verbal e alguns germens do que será, a partir do século XIII, o subesquema X-ada_{GOLPE}¹⁷.

No padrão formado a partir de bases verbais, as palavras mais antigas são ‘achada’ (952) “ato ou efeito de achar” e ‘pousada’ (959) “ato ou efeito de pousar; pouso”. A palavra mais antiga do corpus, no entanto, é ‘lançada’ (886) “1. golpe dado com a lança; 2. ferimento feito por lança”, de “lança + -ada”, ou seja, um denominal, conforme etimologia do dicionário Houaiss. Em Cunha (1982), tanto ‘lançada’ quanto ‘lançar’ estão dentro do verbete ‘lança’, ambos com a datação do séc. XIII. O que se pode afirmar para além das informações encontradas no verbete é que tanto ‘lançar’ quanto ‘lança’ já existiam em latim.

lançar - lat. lanceo, as, avi, atum, are ‘manejar a lança, jogar a lança’; observe-se que lançar, originalmente ‘jogar a lança’, sofreu ampl. de signf. (segundo JM já no lat. vulg.), passando a designar ‘jogar qualquer objeto, arremessar, atirar’, tomando para si, deste modo, os sentidos do v. lat. jactare, v. freq. de jacere ‘jogar, atirar, arremessar, lançar’ (ver jact-), e mesmo afastando-se do signf. primário; ver lanç-; f. hist. sXIV alçar, sXIV lançã, sXIV lãçey, sXIV lancou, sXV lãçado, sXV lançades, sXV lamçou, sXV llançar. (HOUAISS, 2002)

lança - lat. lancea, ae ‘lança, arma dos suevos, arremessão, dardo, venábulo’, prov. de orig. céltica; ver lanç-; f. hist. sXV llança, sXV lãça, sXV lamça. (HOUAISS, 2002)

Sendo assim, tanto é possível que ‘lançada’ tenha vindo da substantivação do particípio de lançar, quanto que tenha vindo de lança + ada. Diante da incerteza, o que se torna relevante é que há uma ambiguidade com relação à natureza da base, que, juntamente com outros casos ambíguos (cf. quadro 1), parece ter propiciado a emergência de um padrão denominal (nesse caso, X-ada_{GOLPE}) e impulsionado sua prolificidade.

¹⁷ Na verdade, a análise dos dados sugere que ‘lançada’ e ‘porrada’ foram as formas-gatilho do padrão X-ada_{GOLPE}, ao passo que ‘canada’ e ‘jugada’ foram o ponto de partida para o padrão X-ada_{QUANTIDADE QUE CABE EM...} que sanciona palavras como ‘braçada’ (XIII) “quantidade de coisas ou de determinada coisa que se abarca cingindo com os braços” e ‘tonelada’ (XIV) “tonel cheio; conteúdo de um tonel”.

Como já mencionado, antes do século XIII, há apenas duas palavras datadas com acepção “golpe”, lançada (886) e porrada (1139)¹⁸. A proliferação da acepção “golpe” começa a aumentar no século XIII (7 dados – dentada, espadada, palmada, orelhada, pancada, punhada, setada), justamente quando começa a aumentar o número de possíveis ambiguidades nas formações deverbiais.

Quadro 1: Casos de possíveis ambiguidades na relação base-derivado

X-ada “ato de X”	Verbo e nome relacionados ¹⁹
cavada (1272)	cavar (1272) cava (XIII)
tornada (XIII)	tonar (XIII) torno (s/data)
caçada (XIII)	caçar (XIII) caça (XIII)
pesada (XIII)	pesar (XIII) peso (XIII)
cantada (XIII)	cantar (XIII) canto (XIII)
jogada (XIII)	jogar (XIII) jogo (XIII)

Fonte: elaboração do autor

Todas as palavras da coluna esquerda no quadro 1 são deverbiais, segundo o dicionário Houaiss (2002). Mas, para o falante, a relação base sufixo nem sempre é clara. Não se pode prever que todos os falantes inequivocamente vão estabelecer exatamente a mesma relação entre base e derivado para uma dada palavra. A analisabilidade²⁰, muitas vezes, depende do quão proeminente e ativo está determinado conceito na memória do falante. Além disso, não é um fenômeno preciso e de manifestação idêntica para todos. O trabalho de Viaro (2007) dá mostras dessa imprecisão, patente nas respostas obtidas em um teste por ele aplicado.

A partir de 120 palavras formadas com o sufixo -eiro, coletadas num corpus de textos, num teste informal, um informante de nível superior, doutor em Letras, afirmou não conseguir entender ou detectar a base em 38 casos (31,7%). Em alguns casos, a derivação assume posições interessantes: afirmou que gelo → geladeira. De fato, a memória depende da frequência usual e, para o informante, gelo é mais associável, por estar mais rapidamente presente na memória, do que gelar, verbo defendido, por exemplo, por um linguista que postulava a derivação deverbal a partir do particípio gelado para esses casos em -eiro. Outras situações que causaram dúvida no informante foram as seguintes derivações: engenho → engenheiro, cruz → cruzeiro, enfermo → enfermeiro, carta → carteira, barro → barreira, solto → solteiro, cocho → cocheiro, travessa → travesseiro, tesoura → tesoureiro, carne → carneiro, caldo → caldeira, trinchar → trincheira. De fato, outros informantes

¹⁸ **porrada** (1139): pancada, bordoadada. *Etimologia*: *porra + -ada, embora os registros de datação disponíveis não corroborem a der.; ver porr-; f.hist. 1139 porrada, sXIV porrada. ***porra** (1209): 1. Diacronismo: antigo. clava com ponta redonda e reforço de ferro; 2. pedaço de pau; porrete, cacete.

¹⁹ A exceção de ‘caça’ e ‘torno’, todas as demais palavras já existiam em latim, mesmo que com significado relativamente distinto do atual.

²⁰ A analisabilidade se relaciona ao grau com que os falantes reconhecem – e tratam distintamente – as partes componentes de uma palavra e/ou um esquema (Cf. BYBEE, 2010). Portanto, está diretamente relacionada à identificação de unidades (reconhecimento de formas linguísticas).

com certeza não teriam dúvidas em alguns (como enfermo → enfermeiro), ou não veriam casos de derivação em outros (como em cruzeiro), ou proporiã outras palavras simples em outros (talvez coche → cocheiro, tesouro → tesoureiro). Subjetivismos semelhantes se instauram em propostas teóricas, em que testes semelhantes são aplicados a si mesmo. O informante encontrou, ainda, situações, que apresentava como palavras simples outras palavras derivadas ou ainda palavras que necessitariã muitas mudanças fonéticas para justificar a derivação. Isso mostra que a palavra-base que vem para a consciência do falante nativo é a mais proeminente na memória, definida, entre outras coisas, com base em sua usualidade. (VIARO, 2007, p. 6-7)

Cabe destacar ainda um outro fator importante, a saber, o caráter hiponímico entre as acepções “ato de...” e “golpe...”. Tomando como exemplo as palavras ‘pedrada’ (XIII)²¹ “ação de lançar uma pedra” e ‘picada’ (1269) “ato ou efeito de picar(-se); machucado ou ferida produzida por objeto pontiagudo”, vê-se naquela um denominal com acepção típica das palavras “ato de x” e nesta um deverbais com semântica muito próxima a das palavras que designam golpe, ou seja, de uma ação que causa algum tipo de ferimento ou machucado. Essa fronteira, de fato, não é muito bem delimitada e isso se dá pelo caráter hiponímico da acepção ‘golpe’ em relação a noção “ato de x”, pois todo “golpe” é um “ato/ação de...”.

‘Lançada’ e ‘porrada’, então, provavelmente foram as palavras gatilho para as demais formações com a acepção “golpe”, que, por analogia, a tomaram como modelo. Desse modo, as quatro molas propulsoras para a emergência do subesquema X-ada_{GOLPE} foram (a) a analogia com as formas já existentes (‘lançada’ e ‘porrada’), (b) a grande proximidade entre as noções “ato de x” e “golpe com x” (caráter hiponímico), (c) a potencial ambiguidade da relação base-derivado não só em ‘lançada’, mas também em (d) outras palavras X-ada deverbais igualmente ambíguas com relação à natureza da base, se verbal ou nominal. Essa ambiguidade²² foi fundamental para reforçar a possibilidade de uma derivação [X]_{Substantivo} → X-ada_{substantivo} e consequente emergência do padrão X-ada_{GOLPE}, o primeiro padrão X-ada denominal da língua.

Importante atentar também que essa hipótese destaca o fato de que o surgimento de um novo padrão é sempre complexo e sempre envolve vários fatores. Não parece que a simples existência de bons candidatos à forma-gatilho seja suficiente para explicar a emergência do padrão denominal, pois ‘lançada’ e ‘porrada’ podem até explicar as formações denominais posteriores, mas fica sem resposta o porquê de essas duas palavras terem bases nominais se ainda não existia um padrão X-ada

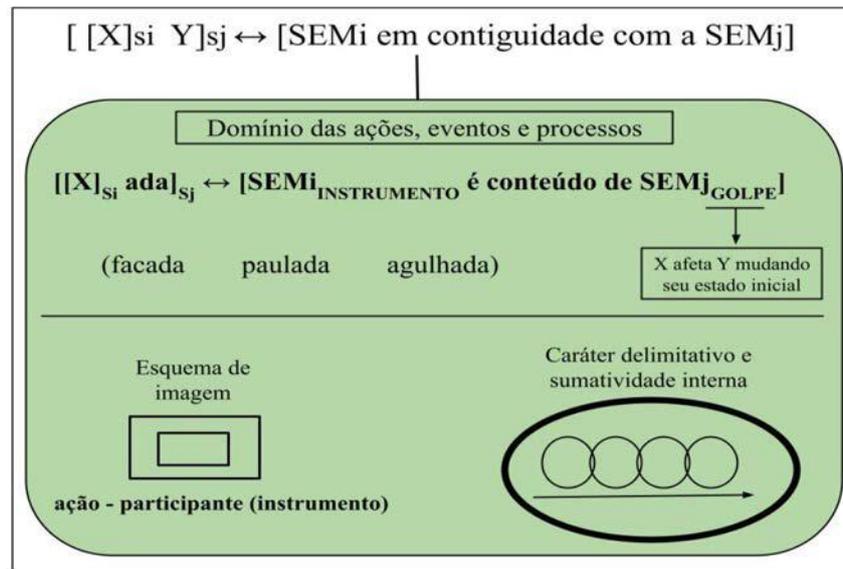
²¹ Segundo Houaiss (2002), a primeira acepção de ‘pedrada’ é “ação de lançar uma pedra”, datando a acepção “golpe” de 1679.

²² Os estudos sobre gramaticalização destacam a ambiguidade como um estágio fundamental para a mudança linguística, o que ficou conhecido pelo termo *overlap model*, o que pode ser exemplificado a partir das três sentenças a seguir: (a) Ele vai para casa falar com Paulo. (b) Ele vai falar com Paulo. (c) Vai chover. Vê-se que a sentença (b) é ambígua com relação ao significado de deslocamento no espaço (Ele vai até Paulo para falar com ele) ou tempo futuro (Ele falará com Paulo). Já em (c), não há qualquer ambiguidade, sendo a única leitura possível a de tempo futuro. Não parece inconsistente a aplicação da noção de *overlapping* à análise que o(s) falante(s) faz(em) da constituição interna das palavras e construções morfológicas como uma etapa importante em muitos casos de mudança. Afinal, seja em construções sintáticas, por meio de inferência contextualmente localizada, seja em palavras, por meio da análise de sua constituição interna (que independe de contexto), o que está em jogo é uma possível ambiguidade na análise que o(s) falante(s) faz(em) da construção.

denominal na língua. Assim, a confluência dos fatores elencados de (a) a (d) (e talvez de outros ainda obscuros), parece ser mais coerente e bastante plausível.

Diante do que foi exposto e com base na proposta de Tavares da Silva (2020), a matriz esquemática do padrão X-ada_{GOLPE} pode ser representada conforme a figura a seguir:

Figura 6: O subesquema X-ada_{GOLPE}



Fonte: elaboração do autor

No topo da rede, está o esquema geral dos denominais (cf. seção 1). Logo abaixo, está o subesquema $[[X]_{si} \text{ ada}]_{sj} \leftrightarrow [SEMi_{INSTRUMENTO} \text{ é conteúdo de } SEMj_{GOLPE}]$, em que X deve ser entendido como participante instrumento, conteúdo da ação de “golpear”²³. X-ada_{GOLPE} herda dos nomes de ação o caráter delimitativo e a sumatividade interna. Sendo um nome que expressa um “ato de...”, herda também a semântica de ação. Portanto, o domínio que ancora esse subesquema é o das ações/eventos/processos.

Como propriedades inovadoras, estão a base nominal e a ideia de uma ação mais específica, genericamente parafraseada como “golpe”. A base das construções X-ada_{GOLPE} é um substantivo concreto que designa o instrumento; o derivado, por sua vez, é um substantivo abstrato que designa a ação praticada com o instrumento. Nos deverbiais, a ação expressa é inteiramente dependente da base verbal e, conseqüentemente, vários tipos de ação podem ser expressos (de calar, de pousar, de entrar etc..). Já nos denominais, uma vez que substantivos concretos não expressam ação, essa noção precisa estar codificada no significado da construção, além de ser necessariamente uma ação específica (nesse caso a de “golpe”), já que não pode ser extraída da base. Isso faz com que haja uma restrição no que concerne à semântica da base, que precisa ser um nome que possa se compatibilizar com essa

²³ A especificação GOLPE, no entanto, deve ser entendida de modo bem amplo, como uma ação pela qual uma entidade X afeta outra Y, modificando seu estado inicial. As palavras ‘dentada’ e ‘palhetada’, por exemplo, não se encaixam muito bem na ideia comum que fazemos de golpe, mas sem dúvida são ações em que uma entidade afeta outra causando-lhe alguma modificação.

noção de “efetuar um golpe (ou algo do tipo)”; daí serem todos substantivos concretos que designam entidades passíveis de funcionar como instrumento.

Uma ação pode ter uma ampla gama de participantes, que pode ser um agente, um paciente, uma localização, seu tempo ou um instrumento. A relação entre ação e um participante pode ser entendida em termos de contenção (ao contrário de uma relação do tipo Evento e Subevento, em que este pode ser entendido como parte daquele – EI PARTE-TODO).

Esse é o significado esquemático dessa construção: uma relação base-derivado ancorada no domínio das ações/eventos/processos e no esquema imagético de contenção, em que a base é sempre o instrumento de um golpe, conceptualizado como uma ação atemporalizada e com ênfase no processo em si (caráter delimitativo), cujas fases constitutivas são sumarizadas (escaneamento sumário).

Considerações finais

O presente texto enfocou o processo de construcionalização de X-ada, sob a proposta de que o caráter delimitativo e a sumatividade interna próprios dessa construção são heranças do significado esquemático do particípio, sua categoria de origem. Com isso, a distinção entre mudança construcional e construcionalização proposta por Traugott e Trousdale (2013) assume aqui caráter fundamental para demonstrar e embasar teoricamente a coexistência de duas construções na língua: o particípio português, fruto de mudança construcional do particípio passado latino, e as construções X-ada, resultado de construcionalização a partir do particípio.

Esta pesquisa também teve o intuito de demonstrar que padrões não surgem arbitrariamente nem são produções *ad hoc*. Estamos em consonância, portanto, com Traugott e Trousdale (2013, p. 2), para quem “o locus da mudança é o constructo, uma instância de uso”. Estamos nos harmonizando também com postulações mais antigas, a exemplo das defendidas pelos difusionistas, que advogavam que uma mudança ocorre inicialmente em algumas palavras e se propaga para outras, podendo atingir todo o léxico ou parte dele. Embora estes se referissem especificamente à mudança sonora, não tardou para que o modelo da difusão lexical fosse estendido a outros fenômenos.

Sendo assim, no que tange ao estudo da morfologia diacrônica, assumir que a mudança começa em itens lexicais específicos implica a identificação da(s) forma(s)-gatilho que impulsionaram o surgimento de um novo padrão ou, pelo menos, dos fenômenos que estiveram por trás dessa emergência. Argumentamos que as formas-gatilho para a emergência de X-ada_{GOLPE} foram provavelmente ‘lançada’ e ‘porrada’. Já os nomes de ação deverbais, ao que tudo indica, têm sua origem em processos metonímicos, ocorridos ainda no latim medieval, de elipse do núcleo do SN e conseqüente reanálise do adjetivo participial. Identificada(s) suas origens, cabe investigar o que cada subesquema herda de sua categoria “mãe” e no que inova.

O subesquema X-ada_{ATO DE...} herda do particípio o caráter delimitativo e a sumatividade interna. Mantém-se também a base verbal. Como principais inovações, está o comportamento derivacional, com perda da relevância sintática e mudança da posição da cabeça morfológica e semântica (construção de cabeça à direita), passando a determinar o gênero da palavra (sempre feminino).

O subesquema X-ada_{GOLPE}, por sua vez, herda dos deverbais o caráter delimitativo, a sumatividade interna, a estrutura esquemática de imagem (EI DE CONTENÇÃO), e a especificação para gênero feminino (cabeça à direita). Por expressar nomes que denotam ações, mantém-se também a ancoragem no domínio das ações/eventos/processos.

No entanto, passa a selecionar bases substantivas. Além disso, considerando a trajetória participio > deverbais > denominal, percebe-se que, à medida que o padrão vai se afastando da construção de origem, vai também perdendo aplicabilidade. Se o participio se aplica a todos os verbos, a construção denominal X-ada_{GOLPE} passa a ser bem mais restrita, se anexando apenas a nomes que são passíveis de funcionar como instrumento. O padrão deverbais, por sua vez, estaria justamente numa zona intermediária em termos de aplicabilidade.

Referências

- CROFT, William. *Radical Construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press, 2001.
- CUNHA, Antônio Geraldo. *Dicionário etimológico da língua portuguesa*. Rio de Janeiro: Lexikon, 1982.
- BECKER, Martin. O sufixo -ada em português - aspectos semânticos e diacrônicos. In: VIARO, M. E. (org.). *Morfologia Histórica*. São Paulo: Cortez, 2014. pp. 118-53.
- BYBEE, Joan. *Language, usage and cognition*. New York: Cambridge University Press, 2010.
- BOOIJ, Geert. *Construction Morphology*. Oxford: Oxford University Press, 2010.
- ERNOUT, Alfred. *Morphologie historique du latin*, Paris, ed. Klincksieck, 1914.
- FARIA, Ernesto. *Gramática superior da língua latina*. Biblioteca Brasileira de Filologia, n. 14. Livraria Acadêmica, 1958.
- GOLDBERG, Adele. *Constructions: a construction grammar approach to argument structure*. Chicago; London: The University of Chicago Press, 1995.
- GOLDBERG, Adele. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press, 2006.
- GONÇALVES, Carlos Alexandre Victorio; TAVARES DA SILVA, João Carlos. *Sobre o estatuto de -nte: evidência de um continuum flexão-derivação para a mudança morfológica do latim ao português*. LABORHISTÓRICO, Dossiê “Descrições morfológicas das línguas românicas: abordagens históricas”, v. 6, n. 1, 2020.
- GURECKIS, Todd; GOLDSTONE, Robert. Schema. In: HOGAN, P. C. (ed.). *The Cambridge Encyclopedia of the Language Sciences*. Cambridge: Cambridge University Press, 2010.
- HOUAISS, Antônio. *Dicionário Digital da Língua Portuguesa - versão 1.0*. Rio de Janeiro: Objetiva, 2002.

- IATRIDOU, Sabine; ANAGNOSTOPOULOU, Elena; IZVORSKI, Roumyana. Observations about the form and meaning of the perfect. In: ALEXIADOU, A.; RATHERT, M.; VON STECHOW, A. (eds.). *Perfect Explorations*. Berlin: Mouton de Gruyter, 2003. pp. 153-205.
- LANGACKER, Ronald Wayne. *Foundations of cognitive grammar*, v. 1: theoretical prerequisites. Stanford, Calif.: Stanford University Press, 1987.
- LANGACKER, Ronald Wayne. *Cognitive Grammar: a basic introduction*. Oxford: Oxford University Press, 2008.
- MACHADO, José Pedro. *Dicionário Etimológico da Língua Portuguesa*, v. 2, Lisboa, Confluência, 1972.
- MEYER-LÜBKE, Wilhelm. *Grammatik der Romanischen Sprachen: Formenlehre*. Leipzig: O. R. Reisland, 1894. v. 2.
- ROHLFS, Gerhard. *Grammatica storica della lingua italiana e dei suoi dialetti: Sintassi e formazione delle parole*. Torino: Einaudi, 1969. v. 3.
- SAID ALI, Manuel. *Gramática histórica da língua portuguesa*. São Paulo: Melhoramentos, 1964.
- SOLEDADE, Juliana. *Semântica morfolexical: contribuições para a descrição do paradigma sufixal do português arcaico*, v. 2, 575 f. Tese (Doutorado em Letras e Linguística) — Instituto de Letras, Universidade Federal da Bahia, Salvador, 2004.
- SOLEDADE, Juliana. *Esquemas construcionais no português arcaico: um estudo sobre X-ada1, X-ada2, X-ado, X-do, X-da*. LABORHISTÓRICO, v. 6, pp. 41-56, 2020.
- TAKAHASHI, Mônica Yuriko. Mudanças semânticas do sufixo -ada. In: VIARO, M. E. (org.). *Morfologia histórica*. São Paulo: Cortez, 2014.
- TAVARES DA SILVA, João Carlos. *A semântica dos sufixos denominais*. Curitiba, Appris, 2020.
- TAVARES DA SILVA, João Carlos. Breve estudo sobre as construções denominais X-ada. In: SOLEDADE, J.; GONÇALVES, C. A.; SIMÕES NETO, N. (org.). *Morfologia Construcional: avanços em língua portuguesa*. Salvador: EDUFBA, 2021.
- TEKAVČIĆ, Pavao. *Grammatica storica dell'italiano: lessico*. Bologna: Il Mulino, 1972. v. 3.
- TR AUGOTT, Elizabeth Closs; TROUSDALE, Graeme. *Constructionalization and Constructional Changes*. Oxford, 2013.
- TRAVAGLIA, Luiz Carlos. *O Aspecto Verbal em Português*. Uberlândia: UFB, 1985.
- VIARO, Mário Eduardo. Estudo diacrônico da formação e da mudança semântica dos sufixos derivacionais eiro/eira na Língua Portuguesa. In: MASSINI-CAGLIARI, G. et al. (ed.). *Trilhas de Mattoso Câmara e outras trilhas: Fonologia, Morfologia, Sintaxe*. São Paulo: Editora Cultura Acadêmica, 2007. pp. 45-84.
- VIARO, Mário Eduardo. *A derivação sufixal do português: elementos para uma investigação semântico-histórica*, 2011. 220 f. Tese de Livre-docência para o Departamento de Letras Vernáculas e Clássicas, área de Filologia e Língua Portuguesa. Universidade de São Paulo. São Paulo, 2011.

HOMENAGEM À NOSSA MIRIAM LEMLE*HOMAGE TO OUR MIRIAM LEMLE**Isabella Lopes Pederneira¹*

Um número sobre estudos gramaticais sincrônicos e diacrônicos não poderia ficar sem a presença de Miriam Lemle, por isso presto minha homenagem póstuma a ela através deste número.

Miriam foi daquelas figuras centrais, pilares e raras da Faculdade de Letras da UFRJ.



Professora Miriam Lemle, na sacada da Ca'Foscari, Universidade de Veneza, em 2013.

Homenageá-la na apresentação desse número da Revista Linguística da UFRJ é uma honra. Seu legado é imensurável, e minha apresentação será a partir de uma visão particular sobre ela, e não, propriamente, a partir de sua atuação como Professora Titular Emérita de Linguística da Universidade Federal do Rio de Janeiro. A opção por essa visão advém do fato de ter conhecido a Miriam quando ela estava a poucos

anos de sua aposentadoria compulsória. Desse modo, há outras testemunhas oculares de um número maior de elementos e detalhes de seu percurso acadêmico.

Fui a última de suas orientandas que ela acompanhou em todas as etapas acadêmicas. Nos últimos anos, estive ao lado dela quase diariamente. Com e como ela, tornei-me gerativista. Há muitos momentos marcantes em nossos inúmeros encontros, em que se misturavam objetivos acadêmicos e pessoais. Em dezembro de 2018, dia em que estava sendo homenageada através do livro “O apelo das árvores”, grande e merecida homenagem realizada pelos professores Alessandro Boechat e Andrew Nevins, coincidiu com o dia em que soubemos que eu ocuparia uma vaga entre os professores da Universidade Federal do Rio de Janeiro. Seu alívio e satisfação foram na mesma proporção que os meus. Ela sabia o tamanho de sua participação nessa conquista.

Comecei quase pelo fim, mas agora vou voltar um pouco no tempo, cerca de dezesseis anos, que foi o tempo que tive oportunidade de conciliar meus dias com os dela. As idas e voltas textuais confirmam nossa resistência à linearidade, representada nas arquiteturas sintáticas. Em janeiro de 2004, quando ingressei em seu grupo de estudos, através da querida professora Aniela Improta França,

¹ Departamento de Letras Vernáculas/PPG Linguística - Universidade Federal do Rio de Janeiro (UFRJ), isbellapederneira@letras.ufrj.br, <https://orcid.org/0000-0002-5884-8472>.

também ex-orientanda e parceira de trabalho da Miriam por longos anos, deparei-me, pela primeira vez, com a imagem da tão conhecida Professora Miriam Lemle. Tive medo de trocar uma palavra sequer com ela, embora aquele sorriso simpático de que me lembro até hoje já estivesse lá. Meus laços específicos e individuais com ela foram de iniciativa dela, em janeiro de 2005, já que eu não tinha ousadia suficiente para propor um trabalho com Miriam Lemle. Eu soube que estava recebendo uma proposta grandiosa na minha vida. Deitei contente naquela noite, excitada com a oportunidade. Ela havia descoberto que eu era aluna de latim e sempre teve interesse em estudos diacrônicos também, ainda que sua produção não seja exatamente voltada para essa área de estudos.

Daquele dia em diante, Miriam me acolheu, me ensinou e me conduziu não só na linguística, mas também na vida comum. Em nossa mesa de trabalho, misturavam-se árvores sintáticas, chocolates para dar energia e deixar inteligente (na falta do café), dicionários e papéis, além de choros, gritos e muita risada e planejamento em todas as nossas manhãs, tardes – e também noites, quando os prazos apertavam. Particularmente, Miriam estava sempre com lindas e modernas bolsas, que ela dizia que comprava na praia, mas lá dentro, além de muitos utensílios perdidos, havia a magia da linguística, que ela conseguia enxergar em notícias jornalísticas, por vezes, banais. O brilho em seus olhos e a energia do conhecimento que brotavam de um título daquele boletim diário de notícia eram momentos inspiradores, que fazem parte da minha formação. Emocionava e impulsionava o quanto ela insistia em um mundo constituído apenas de verdades e estruturas linguísticas.

A ousadia que estava nos acessórios da Miriam não estava restrita aos colares e bolsas. Miriam ousava nos trabalhos acadêmicos também, o que tornava nossas pesquisas sempre mais árduas, pois a literatura não era tão fácil de ser encontrada, às vezes nem havia mesmo. Essa é uma marca de sua personalidade, certamente conhecida entre aqueles que conviveram com ela. Talvez tenha sido essa a característica que motivou que um colega a descrevesse como “rock and roll”. Ela foi responsável por inaugurar muitos campos de estudo em linguística gerativa no Brasil, por exemplo. Além da própria linguística chomskyana, Miriam adentrou temas polêmicos e arriscados em seu início, como a interface com a neurociência da linguagem e os próprios estudos em modelos construcionistas de Gramática Gerativa.

Miriam, do início ao fim, além de ousada, foi surpreendente. Sua presença, sempre forte, decidida e decisiva, coloca-nos a dificuldade em lidar com sua ausência física, mesmo passados mais de dois anos de sua partida inesperada. Miriam tinha um jeito próprio, peculiar, até mesmo de arrastar os pés pelo corredor H do terceiro andar da Faculdade de Letras e de abrir a porta da sala H-308, de modo que era difícil não identificar sua iminente chegada para nossos encontros, sempre um pouquinho atrasada. Atrasada no horário, mas nunca na ideia e no pensamento. Deste modo, ela tornava aqueles 30 minutos de espera um passo adiante nos estudos. O processo de realização de um trabalho era tão importante quanto seu fim. Ainda ressoam suas gargalhadas durante a produção de nosso eternamente inacabado artigo sobre “piadas linguísticas”. Eternamente inacabado, até nos reencontramos no “Cosmos”, tenho certeza. Lá, a gente recomeça do início, dá umas boas risadas e, finalmente, terminamos. Muitos artigos inacabados, ao longo do tempo, consigo terminar, mas não esse. Esse eu não quero que seja sem ela.

Como todos sabem, Miriam Lemle foi uma das maiores linguistas do Brasil, tendo sido precursora, entusiasta e defensora da linguística gerativa. É autora de inúmeros artigos, capítulos de livros e livros em linguística gerativa. O que poucos sabem é que, quando eu estava iniciando os estudos, antes mesmo de conhecê-la, enviei um corajoso e-mail para Noam Chomsky e uma das perguntas que fiz a ele foi uma indicação de um bom linguista para estudar gerativismo no Brasil. Sua resposta foi Miriam Lemle. Esse dado mostra parte de sua grandiosidade. Miriam era minimalista em seus entusiasmos, mas é certo que homenageá-la em um número da Revista Linguística da UFRJ, proveniente do Programa de Pós-Graduação de que ela tinha tanto orgulho e pelo qual ela tanto lutou, com garra e dedicação exclusiva (quase literalmente, em seu caso), vai deixá-la feliz e satisfeita.

Apesar de ser adepta a uma vida de poucas euforias, foi uma mulher complexa, forte, de uma sofisticação simples nas palavras, lia e escrevia com a visão clara das hierarquias sintáticas das palavras e das sentenças. Sem muita ortodoxia religiosa, segundo ela, a prova de que Deus existe é a linguagem humana, mais propriamente a sintaxe. Ela dizia que era uma arquitetura tão perfeita, que só poderia ser obra de Deus. Alguns poderiam argumentar que nesse contexto residiria sua verdadeira ortodoxia: a linguística gerativa. Controverso ou não, seguiu em defesa desta corrente teórica até 2020, sempre percorrendo caminhos aventureiros, seja na vida particular, cheia de esportes radicais, seja na profissional. Temas de estudo muito comportados, como disse, não eram seus prediletos. Era preciso sempre criatividade para construir uma carreira junto dela. Houve riscos e risadas. Extremista nos sentimentos, ia dos gritos aos sorrisos, sempre com sinceridade no olhar. Uma ingenuidade peculiar e encantadora diante da vida, e uma perspicácia linguística raríssima.

Esta homenagem é, portanto, muito pertinente, já que Miriam Lemle ajudou na construção do que hoje é o Programa de Pós-Graduação em Linguística da UFRJ. Esteve entre os primeiros professores que formavam os primeiros linguistas do Brasil. Miriam Lemle, mais do que uma grande pesquisadora, foi uma professora orgulhosa. Importante ressaltar que ela não fazia questão de qualquer rótulo nobre. Quando conheci a Miriam, ela já era reconhecida e premiada, mas só soube disso por informações de segunda mão. Ela nunca mencionou, por exemplo, que havia sido uma pesquisadora A1 do CNPq (nem sabia direito o que era) ou que era Professora Titular ou que havia sido presidente da ABRALIN. As claras hierarquias ficavam reduzidas somente às construções arbóreas. Seu orgulho era por ter entendido algo importante, lá nos anos 70, quando descobria a recursividade e a Gramática Universal. Ela me pedia até para não permitir salas em seu nome após sua partida. O que ela me fez prometer mesmo é que a linguística gerativa resistiria.

Suas demonstrações de afeto eram comedidas. Estavam em inúmeras trocas de e-mails e telefonemas diários, num “oi, adoro quando abro a porta e você já está aqui”, num “tô com saudade, Isabella”, quando eu estava trabalhando em Teresópolis e conseguia ir pouco ao Fundão, ou numa chamadinha discreta, agradecendo pela minha presença diária. Às vezes, não era tão comedida assim: ela me dizia que eu não deveria me mudar para Petrópolis, porque subir e descer a serra diariamente poderia me causar problema cardíaco e eu acabar morrendo disso antes da hora. No fim de tantas discussões em torno da minha moradia, entendi que ela achava que eu estaria menos presente no Fundão, e era isso que ela temia.

Ela foi daquelas acadêmicas formadas em um tempo que passou. Prezava a humildade do conhecimento verdadeiro, a grandiosidade de um saber sem grandes performances produtivistas. Estava sempre percorrendo caminhos abstratos, onde ela encontrava as construções arbóreas. Teórica e poliglota, era fácil, para ela, encontrar padrões gerais na linguagem. Tendo sido bilíngue português e italiano, ela reconhecia derivação morfológica com significado regular no par nome/verbo janela/defenestrar, o que nos causava grandes discussões. Ela era teimosa, mas não insistia em equívocos. Apesar da aquisição do italiano como língua materna e de ter nascido em Roma, tinha orgulho de ser brasileira, naturalizadíssima. Sua história familiar, em meio a uma das tragédias da humanidade, foi decisiva para que tivéssemos alguém como ela, que lutasse por direitos sociais acima tudo e pela educação pública, sobretudo pela UFRJ. Estava aflita com o discurso anti-indigenista que nos ronda no momento, de modo que ficou muito feliz com seu último trabalho em parceria com colegas de longa data do Programa de Pós-Graduação de Linguística. Ela estava entusiasmada, anotava em seu caderninho os encontros para desenvolver um trabalho muito importante que enfatiza aspectos da Gramática Universal em línguas indígenas brasileiras, que une, portanto, o início de seu percurso no Museu Nacional, com estudos em línguas indígenas brasileiras, e Gramática Universal, uma verdade absoluta para ela. Seu vigor no auge de seus mais de oitenta anos a permitiu estar presente no Centro do Rio de Janeiro, em 2019, para lutar pela educação pública brasileira, na chuva, mas plena de energia e olhar crítico e preciso.

O ano de 2019 foi agitado para ela. Além de artigos e livros em meio a algumas orientações, tivemos uma aventura na Quinta da Boa Vista num domingo qualquer para promover ciência num momento sombrio e, finalmente, começava a ter forma, nossa maneira de contribuir para a educação básica através da Gramática Gerativa: submetemos, juntas, o projeto que hoje coordeno com outras colegas, mas que, sem ela, não existiria. Minha maneira particular de demonstrar eterna gratidão é defendendo a educação e desenvolvendo estudos e projetos em Gramática Gerativa. A agitação de seu último ano foi merecida e desejada. Aposentadoria absoluta não era um horizonte vislumbrado e vislumbrável por ela. Sequer gostava de férias.

Diante de tantas facetas de uma única mulher, é difícil, na verdade, definir em uma palavra quem foi Miriam Lemle. Corajosa, talvez. Verdadeira, sincera e independente, certamente. Deixou dois filhos biológicos e, por enquanto, quatro netos, mas muitos, muitos herdeiros de sua linguística, da teoria ao experimentalismo. Estou dentre uma destas orgulhosas herdeiras de um conhecimento complexo e honroso.

Seu maior legado iniciou-se no som da queda do *Syntactic Structure*, jogado em sua mesa por um estagiário americano de passagem no Museu Nacional. Ela fazia questão das onomatopéias desse encontro dela com sua história. Nas vezes em que ela compartilhou esse evento conosco, era possível ouvir, junto com ela, o som do livro batendo na mesa. Era uma cena real, com toques teatrais, que dão o devido relevo à importância daquele momento para ela. Ali começava, além da história dela, a história da linguística gerativa no Brasil, parte da história do Programa de Pós-Graduação de Linguística da UFRJ e ... a minha história.