

**36/1 (2022)**

revista do centro de estudos humanísticos

# diacrítica

Psicolinguística

*Psycholinguistics*



A *Diacrítica* é uma revista científica, de cariz multidisciplinar, dedicada aos estudos literários, culturais, linguísticos e artísticos. É editada pelo Centro de Estudos Humanísticos da Universidade do Minho (CEHUM), desde 1986, e subsidiada pela Fundação para a Ciência e a Tecnologia. Desde janeiro de 2017 é publicada em formato eletrónico, com periodicidade trimestral.

A revista está registada com o ISSN 0870-8967 (formato em papel) e 2183-9174 (formato eletrónico) e está licenciada com uma Licença Creative Commons CC BY-NC.

*Diacrítica is a multidisciplinary scientific journal devoted to literary, cultural, linguistic and artistic studies. It is edited by the Center for Humanistic Studies of the University of Minho (CEHUM) since 1986 and funded by the Portuguese Foundation for Science and Technology. Since January of 2017 it is published in electronic format, on a quarterly basis.*

*The journal is registered with ISSN 0870-8967 (paper version) and 2183-9174 (electronic version) and is distributed under Creative Commons CC BY-NC license.*

#### **Editor responsável / Editor-in-chief**

Orlando Grossegeisse

#### **Direção da Revista / Journal Management**

Cristina Flores; Margarida Pereira; José Eduardo Silva (Direção do CEHUM)

#### **Comissão científica / Scientific Committee**

Abel Barros Baptista (Universidade Nova de Lisboa); Antónia Coutinho (Universidade Nova de Lisboa); António Branco (Universidade de Lisboa); Ana Brito (Universidade do Porto); Augusto Soares da Silva (Universidade Católica Portuguesa); Bernard McGuirk (University of Nottingham); Clara Rocha (Universidade Nova de Lisboa); Conceição Paiva (Universidade Federal do Rio de Janeiro); Eduardo Paiva Raposo (University of California); Fátima Oliveira (Universidade do Porto); Fernando Cabo Aseguinolaza (Universidad de Santiago de Compostela); Graça Rio-Torto (Universidade de Coimbra); Helder Macedo (King's College); Helena Buescu (Universidade de Lisboa); Ivo Castro (Universidade de Lisboa); João de Almeida Flor (Universidade de Lisboa); José Luís Cifuentes Honrubia (Universitat d'Alacant); José Luís Rodrigues (Universidade de Santiago de Compostela); Jürgen M. Meisel (Universität Hamburg / University of Calgary); Maria Alzira Seixo (Universidade de Lisboa); Maria Irene Ramalho (Universidade de Coimbra); Maria João Freitas (Universidade de Lisboa); Maria Manuela Gouveia Delille (Universidade de Coimbra); Mary Kato (Universidade de Campinas); Nancy Armstrong (Brown University); Rui Marques (Universidade de Lisboa); Susan Bassnett (University of Warwick); Susan Stanford Friedman (University of Wisconsin-Madison); Tomás Albaladejo Mayordomo (Universidad Autónoma de Madrid); Vita Fortunati (Università di Bologna); Vítor Aguiar e Silva (Universidade do Minho).

**Diacrítica, volume 36, número 1 (2022)**

*Psicolinguística*

ISSN: 2183-9174

DOI: <https://doi.org/10.21814/diacritica.36.1>

Website: <https://revistas.uminho.pt/index.php/diacritica>

Editores: Maria do Carmo Lourenço-Gomes (CEHUM), Armanda Costa (FLUL/CLUL) e Juliana Novo Gomes (CEHUM)

Revisão: Ana Cristina Silva, Orlando Grossegeesse e Cristina Flores

**Revisores científicos:** Aida Cardoso (Centro de Linguística da Universidade de Lisboa); Aléria Lage (Faculdade de Letras da Universidade Federal do Rio de Janeiro); Alina Villalva (Faculdade de Letras da Universidade de Lisboa); Álvaro Iriarte Sanromán (Centro de Estudos Humanísticos da Universidade do Minho - CEHUM); Ana Isabel Mata (Faculdade de Letras da Universidade de Lisboa); Ana Madeira (Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa); Anabela Leal de Barros (Escola de Letras, Artes e Ciências Humanas, Universidade do Minho); António José Gonçalves de Freitas (Centro de Estudos Humanísticos da Universidade do Minho - CEHUM); Arlette Verhaeghe (Faculdade de Psicologia da Universidade de Lisboa); Candido Oliveira (Centro Federal de Educação Tecnológica de Minas Gerais-CEFET-MG); Cristina Maria Moreira Flores (Centro de Estudos Humanísticos da Universidade do Minho - CEHUM); Déborah González Martínez (Universidade de Santiago de Compostela); Dina Alves (Centro de Linguística da Universidade de Lisboa/Escola Superior de Saúde do Instituto Politécnico de Setúbal); Eduardo Kenedy (Instituto de Letras da Universidade Federal Fluminense); Elisângela Teixeira (Departamento de Letras Vernáculas da Universidade Federal do Ceará); Gabriela Matos (Centro de Linguística da Universidade de Lisboa); Íris Susana Pires Pereira (Centro de Investigação em Educação, Universidade do Minho); Isabel Barahona da Fonseca (Faculdade de Psicologia da Universidade de Lisboa); Isabel Falé (Departamento de Humanidades da Universidade Aberta); Isabel Margarida Duarte (Faculdade de Letras da Universidade do Porto); Isabel Pereira (Faculdade de Letras da Universidade de Coimbra); Joana Batalha (Centro de Linguística da Universidade Nova de Lisboa); João Veríssimo (Faculdade de Letras da Universidade de Lisboa); Leticia Maria Sicuro Correa (Departamento de Letras da Pontifícia Universidade Católica do Rio de Janeiro); Luciana Mendes (Universidade Federal do Rio de Janeiro); Lurdes Moutinho (Centro de Línguas Literaturas e Culturas da Universidade de Aveiro); Manuela Ambar (Centro de Linguística da Universidade de Lisboa); Margarita Correia (Faculdade de Letras da Universidade de Lisboa/CELGA-ILTEC); Maria Serena Felici (Università degli studi Internazionali di Roma - UNINT); Mariagrazia Russo (Università degli studi Internazionali di Roma - UNINT); Marília Uchôa (Departamento de Letras-Libras da Universidade Federal do Rio de Janeiro); Marina Rosa Ana Augusto (Departamento de Estudos da Linguagem da Universidade Estadual de Campinas); Nélia Alexandre (Faculdade de Letras da Universidade de Lisboa); Norberto Pereira (Centro de Linguística da Universidade de Lisboa); Paulo José Tente da Rocha Santos Osório (Departamento de Letras da Universidade da Beira Interior); Rafael Minussi (Departamento de Letras da Universidade de São Paulo); Rosa Lídia Coimbra (Universidade de Aveiro); Rui Marques (Faculdade de Letras da Universidade de Lisboa); Sandra Fernandes (Faculdade de Psicologia da Universidade de Lisboa); Sandro Rodrigues da Fonseca (Instituto de Letras da Universidade Federal do Rio Grande do Sul); Tjerk Hagemeyer (Faculdade de Letras da Universidade de Lisboa).

## PSICOLINGUÍSTICA

- 1-5 **Palavras introdutórias**  
Maria do Carmo Lourenço-Gomes, Armanda Costa, Juliana Novo Gomes
- 6–36 **Eye tracking sentences in language education**  
Marcus Maia
- 37–65 **Differences between stereotypical gender and definitional gender in pronominal antecedent retrieval in Brazilian Portuguese**  
Michele Calil dos Santos Alves
- 66–88 **Leitura de sentenças em alemão por falantes de hunsriqueano: acesso lexical não seletivo e coativação sintática**  
Bernardo Kolling Limberger
- 89-107 **Aprendizagem de dependências não adjacentes por adultos em uma língua artificial: evidências experimentais**  
Cristina Name, Késsia Henrique
- 108-132 **A aquisição de clíticos em Português Europeu L2 e a Hipótese de Reconfiguração dos Traços**  
Ronan Pereira
- 133-162 **Aceitabilidade de sentenças de tópico-comentário em PB e em inglês: um estudo psicolinguístico**  
Lorrane da Silva Neves Medeiros Ventura
- 163-182 **Dificuldades de leitura de estudantes universitários com TDAH: um estudo da influência da memória de trabalho na compreensão leitora**  
José Ferrari Neto, Gustavo Lopez Estivalet, Priscilla de Albuquerque Almeida

## RECENSÕES

- 183-187 **Roepert, T.; Maia, M.; Pilati, E. Experimentando Linguística na Escola: conhecimento gramatical, leitura e escrita. São Paulo: Pontes, 2020**  
Heloisa Maria M. Lima Salles
- 188-192 **Recensión de Martín Aizpuru, Leyre (2020). La escritura cancelleresca de Fernando III, Alfonso X, Sancho IV y Fernando IV. Estudio paleográfico y gráfico-fonético de la documentación real de 1230 a 1312. Peter Lang, 364 pp**  
Pedro Dono López

## VARIA

- 193-224 **Paisagem linguística da “Zona Velha” da cidade do Funchal: considerações históricas e sociolinguísticas**  
Bruna Pereira
- 225-237 **Empréstimos do Italiano em nomes empresariais em Portugal**  
Rosa Lúcia Coimbra, Lurdes de Castro Moutinho
- 238-259 **Instituições de formação de professores como comunidades de escrita: estudo com base nos títulos dos relatórios de Mestrado**  
Luís Filipe Barbeiro, José António Brandão Carvalho, Célia Barbeiro

## PALAVRAS INTRODUTÓRIAS

### INTRODUCTORY WORDS

Maria do Carmo Lourenço-Gomes\*  
mclgomes@elach.uminho.pt

Armanda Costa\*\*  
armandacosta@letras.ulisboa.pt

Juliana Novo Gomes\*\*\*  
juliana.novo.gomes@elach.uminho.pt

•

A faculdade da linguagem permite-nos pensar, estruturar o pensamento, estabelecer rotinas de comunicação e criar discursos e textos surpreendentes pelo seu valor retórico, estético, ou pela sua eficácia na comunicação. Para tal há recursos bem identificados pelas Ciências da Linguagem e da Cognição: um conhecimento linguístico mental que integra uma componente computacional que opera sobre objetos de natureza variada (fonológicos, morfológicos ou sintáticos) e um léxico cujos itens são especificados em várias dimensões. A gramática (o sistema computacional e o léxico) é a condição para a aquisição da linguagem e está na base da comunicação verbal entre falantes de uma mesma língua. Potencia também a aquisição de mais do que uma língua e a comunicação interlínguas porque possui propriedades transversais às várias línguas humanas.

Produzir e compreender enunciados orais ou escritos são comportamentos assentes nesse conhecimento linguístico estruturado, complexo e partilhado, mas que precisam de condições neurofisiológicas e de capacidades cognitivas suficientes para serem bem-sucedidos. Ora o foco de estudo da Psicolinguística é exatamente o uso da linguagem balizado pelas propriedades linguísticas dos sistemas envolvidos e pelas capacidades neuropsicofisiológicas necessárias ao sucesso da comunicação. Apoia-se nas teorias sobre a gramática para fazer predições sobre o processamento da linguagem e contribui para a validação empírica ou infirmação de algumas das suas hipóteses. A Psicolinguística cria interfaces com a psicologia cognitiva e as neurociências; parte dos seus pressupostos sobre o funcionamento do cérebro e da mente humana; usa as suas ferramentas de

---

\* Centro de Estudos Humanísticos (CEHUM), Escola de Letras, Artes e Ciências Humanas (ELACH), Universidade do Minho (UMinho), Braga, Portugal. ORCID: 0000-0001-9003-0438

\*\* Faculdade de Letras/ Centro de Linguística, Universidade de Lisboa (FLUL/CLUL), Lisboa, Portugal. ORCID: 0000-0003-4255-5779

\*\*\* Centro de Estudos Humanísticos (CEHUM), Escola de Letras, Artes e Ciências Humanas (ELACH), Universidade do Minho (UMinho), Braga, Portugal. ORCID: 0000-0001-6834-1432

investigação para, através da observação de comportamentos e da captação da atividade neural, inferir processos cognitivos inerentes ao uso da linguagem.

Na vasta área dos estudos psicolinguísticos emergem domínios especializados. O processamento da informação linguística (da conceptualização à produção, da percepção à elaboração de um significado mental) toma relevo. Investigam-se os efeitos das propriedades do sistema linguístico de cada língua e da natureza neurocognitiva nos vários níveis de organização da gramática. Entre outros, estuda-se o processamento lexical (organização mental do léxico, o impacto da estrutura interna da palavra (fonológica ou morfológica) no seu reconhecimento auditivo e visual, o acesso lexical; estuda-se o processamento sintático (ordem de palavras, complexidade estrutural e ambiguidade). Estudam-se efeitos de alterações da memória de trabalho e da atenção, efeitos da ativação de estratégias metacognitivas na resolução de problemas de processamento. Analisa-se o impacto das perturbações da linguagem na comunicação. Aprofunda-se o estudo dos processos de leitura, identificando-se as condições necessárias à sua aprendizagem, à melhoria da compreensão e da fluência; estudam-se perturbações na leitura e as suas manifestações na dislexia, na perturbação da hiperatividade e défice de atenção, nas perturbações auditivas. Analisam-se efeitos de idade no processamento linguístico: declínio de capacidades de processamento, alterações no acesso lexical, redução do *reading span*. Investiga-se sobre o processamento em condição bilingue e plurilingue.

A investigação psicolinguística é de carácter experimental. O trabalho laboratorial procede segundo planos rigorosamente detalhados e com apoio em metodologias de investigação adequadas à obtenção de dados comportamentais, fisiológicos ou neurais, que abrem janelas para a compreensão da relação linguagem-mente-cérebro.

Este volume da *Diacrítica*, dedicado à investigação psicolinguística, tem a língua portuguesa como referência e reúne um conjunto de estudos que se distribuem por alguns dos domínios acima referidos.

“Eye Tracking Sentences in Language Education”. Marcus Maia demonstra o uso educacional da técnica de rastreamento ocular através da apresentação de um conjunto de dados trabalhados com estudantes em oficinas de leitura no 3º Ciclo do Ensino Básico (Ensino Fundamental, no Brasil) e em cursos de licenciatura e de pós-graduação. Entremado ao relato de ricas experiências com os estudantes em aulas de gramática, sintaxe gerativa e psicolinguística, o autor discute e reflete sobre os benefícios da aplicação em sala de aula do conhecimento científico sobre processamento da leitura, advindo da tecnologia de rastreamento ocular, e sobre metacognição.

“Differences between Stereotypical Gender and Definitional Gender in Pronominal Antecedent Retrieval in Brazilian Portuguese”. Michele Alves investiga sobre o processamento de cadeias correferenciais em que, para a atribuição de um antecedente a um pronome, há diferentes fontes de informação linguística em competição: princípios estritos da gramática que determinam as condições estruturais que permitem a ligação de um antecedente a um pronome (Princípio B) e informação morfossintáctica e pragmática que marcam o género gramatical e o género de estereótipo. Em dois estudos de eye-tracking em que participam nativos adultos do PB, avalia-se o peso dos diferentes fatores que contribuem para o estabelecimento da cadeia anafórica e discute-se o contributo e a

precedência das restrições estipuladas pela gramática teórica contra a marcação morfológica e informação não linguística.

“Leitura de Sentenças em Alemão por Falantes de Hunsriqueano: Acesso Lexical Não Seletivo e Coativação Sintática”. Bernardo Limberger estuda o acesso lexical em contexto de frase, tendo como alvo o hunsriqueano (*Hunsrückisch*), uma variedade linguística do alemão usada no Brasil. A partir de uma tarefa de compreensão de frases em alemão, e com base na precisão e nos tempos de respostas, o autor analisa e discute o compartilhamento de representações ortográficas e fonológicas dos cognatos e de representações sintáticas, comparando o desempenho de falantes de hunsriqueano, multilíngues (português e alemão) e de falantes também multilíngues (português, alemão e inglês), mas não falantes de qualquer língua minoritária, em estruturas com variações no verbo (cognato *vs.* não cognato) e no tempo verbal (*Perfekt vs. Präsens*).

“Aprendizagem de Dependências Não Adjacentes por Adultos em Uma Língua Artificial: Evidências Experimentais”. Cristina Name e Késsia Henrique trazem evidências da capacidade de aprendizagem de dependências não adjacentes, nas quais há coocorrência de marcadores morfossintáticos e/ou fonológicos, por falantes adultos e, portanto, com L1 bem consolidada. As autoras examinam a aptidão dos participantes do estudo para identificar, abstrair e generalizar novos padrões de combinação das dependências não adjacentes em uma língua artificial. Além da contribuição que oferecem à investigação sobre o tema tratado, as autoras combinam técnicas de coleta de dados *online* e *offline* na experiência usada, e dão ênfase a questões metodológicas importantes, procurando contribuir também neste domínio da investigação.

“A Aquisição de Clíticos em Português Europeu L2 e a Hipótese de Reconfiguração dos Traços”. Ronan Pereira apresenta um estudo sobre a aquisição de PE como L2 por falantes nativos de espanhol, com controlo de propriedades da gramática relativas à distribuição dos clíticos nas duas línguas. Colocam-se hipóteses sobre a ativação do conhecimento linguístico da L1 no processamento de L2, e usa-se a variável nível de proficiência em L2 para verificação do contributo de diferentes estádios de conhecimento da língua alvo na reconfiguração dos traços que presidem à distribuição dos clíticos nas duas línguas em confronto. Foram usadas tarefas de completamento de frases e de julgamento de aceitabilidade para, respetivamente, avaliar o uso e aceitação de contextos de próclise e ênclise em contextos estruturais distintos.

“Aceitabilidade de Sentenças de Tópico-Comentário em PB e em Inglês: Um Estudo Psicolinguístico”. Lorrane Ventura investiga sobre o processamento de frases com diferentes tipos de topicalização: topicalização por movimento e topicalização com tópicos gerados na base, por falantes adultos nativos de PB e de inglês e nativos de PB fluentes em inglês, com recurso a tarefa de julgamento de aceitabilidade sobre frases apresentadas auditivamente. Há um duplo objetivo: recolher dados empíricos que confirmem a já clássica tipologia de línguas com proeminência de sujeito ou de tópico e que esclareçam a hipótese de o PB se comportar como língua mista; compreender o que se passa no processamento de uma L2 que se distingue estruturalmente da L1 relativamente a estruturas de topicalização.

“Dificuldades de Leitura de Estudantes Universitários com TDAH: Um Estudo da Influência da Memória de Trabalho na Compreensão Leitora”. José Neto, Gustavo

Estivalet e Priscilla Almeida apresentam uma investigação sobre perturbações na leitura numa população clínica em particular: sujeitos adultos diagnosticados com défice de atenção e/ou hiperatividade. Através da aplicação de 3 tarefas experimentais – *teste Cloze*, *Tarefa de Atenção Linguística* e *Reading Span Test* – compararam os dados da população alvo e de um grupo de controlo, de forma a avaliar relações entre variações na capacidade da memória de trabalho e da atenção linguística na compreensão na leitura.

Há ainda duas recensões críticas. A primeira, do livro de Roeper, Maia e Pilati – *Experimentando Linguística na Escola: Conhecimento Gramatical, Leitura e Escrita*, da autoria de Heloisa Salles, aborda tópicos caros à investigação linguística e psicolinguística e ao seu impacto na educação. A segunda, da tese de doutoramento da Professora Leyre Martín Aizpuru – *La Escritura Cancilleresca de Fernando III, Alfonso X, Sancho IV y Fernando IV. Estudio Paleográfico y Gráfico-Fonético de la Documentación Real de 1230 a 1312*, e da autoria de Pedro Dono López, salienta a contribuição da obra no campo da linguística histórica.

A secção *Varia* integra três estudos, um na área da sociolinguística, outro na área da lexicologia e ainda outro que versa sobre literacia académica.

“Paisagem Linguística da ‘Zona Velha’ da Cidade do Funchal: Considerações Históricas e Sociolinguísticas”. Bruna Pereira evidencia a influência da língua inglesa, reflexo da presença e ocupação britânica, na Zona Velha da cidade do Funchal, a partir de um estudo de caso que conta com uma ampla análise de fotografias de sinais linguísticos em espaços públicos da localidade. A autora discrimina e discute o impacto desta influência, considerando os planos sociopolítico, cultural e económico, bem como o fenómeno turístico.

“Empréstimos do Italiano em Nomes Empresariais em Portugal”. Rosa Coimbra e Lurdes Moutinho, sob uma perspetiva sincrónica, analisam marcas lexicais de proveniência italiana no português atual em um largo corpus de nomes de empresas em Portugal. Na análise e discussão dos dados obtidos, as autoras consideram categorias morfológicas, distribuição por ramo de atividade e distribuição geográfica das empresas e deixam revelar o encontro de línguas e culturas nos estrangeirismos e o enriquecimento do léxico que advém deste encontro.

“Instituições de Formação de Professores como Comunidades de Escrita: Estudo com Base nos Títulos dos Relatórios de Mestrado”. Luís Filipe Barbeiro, José António Brandão Carvalho e Célia Barbeiro analisam e discutem a construção de títulos de relatórios de mestrado de formação de professores, reunidos em um corpus. Identificam características como número de elementos, estruturas sintático-semânticas, estratégias discursivas, palavras-chave e saliência dada às dimensões pedagógica e de investigação. Da análise, os autores deixam indicações das diferenças encontradas entre instituições na construção dos títulos desses trabalhos e sugerem que essas diferenças indicam que “as instituições se constituem como comunidades locais, nas quais predominam certas escolhas”.

Na variedade dos tópicos, questões de investigação, discussões teóricas e metodológicas e nas muitas reflexões deixadas no conjunto de artigos temáticos, o volume exhibe a riqueza e a contribuição da Psicolinguística nas suas múltiplas interfaces.

Juntamente com as duas resenhas e os trabalhos da secção *Varia*, o volume mantém o cariz multidisciplinar da Revista *Diacrítica*.

## EYE TRACKING SENTENCES IN LANGUAGE EDUCATION

### RASTREAMENTO OCULAR DE PERÍODOS NA EDUCAÇÃO LINGUÍSTICA

Marcus Maia\*  
maiamarcus@gmail.com

The present study reports and discusses the use of eye tracking qualitative data (static and dynamic gaze plots and heatmaps) in reading workshops in a middle school and in Generative Syntax and Sentence Processing courses at the undergraduate and graduate levels. Both endeavors take the sentential level as the proper object to be metacognitively explored in language education in order to develop innate science forming capacity and knowledge of language. In both projects non-discrepant qualitative eye tracking data collected and quantitatively analyzed in psycholinguistic studies carried out in LAPEX-UFRJ (Experimental Psycholinguistics Laboratory of the Federal University of Rio de Janeiro) were displayed to students as a point of departure, triggering discussions. Active, problem-solving based methodologies were employed with the objective of stimulating student participation. The article also discusses the importance of developing full literacy, epistemic vigilance and intellectual self-defense in an infodemic world.

**Keywords:** Eye tracking. Sentence. Reading. Language Education. Metacognition.

O presente estudo relata e discute a utilização de dados qualitativos de rastreamento ocular (mapas de fixação estáticos e dinâmicos e mapas de calor) em oficinas de leitura no ensino fundamental e em cursos de Sintaxe Gerativa e de Processamento de Frases na graduação e pós-graduação. Ambos os empreendimentos tomam o nível sentencial como o objeto apropriado a ser explorado metacognitivamente na educação linguística, a fim de desenvolver a capacidade inata de formação científica e o conhecimento da linguagem. Em ambos os projetos, dados de rastreamento ocular qualitativos não discrepantes coletados e analisados quantitativamente em estudos psicolinguísticos realizados no LAPEX-UFRJ (Laboratório de Psicolinguística Experimental da Universidade Federal do Rio de Janeiro) foram apresentados aos alunos como ponto de partida, gerando discussões. Foram utilizadas metodologias ativas baseadas em resolução de problemas, com o objetivo de estimular a participação dos alunos. O artigo também discute a importância de desenvolver o letramento pleno, a vigilância epistêmica e a autodefesa intelectual em um mundo infodêmico.

**Palavras-chave:** Rastreamento ocular. Período. Leitura. Educação linguística. Metacognição.

•

---

\* Department of Linguistics, College of Letters, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. ORCID: 0000-0002-1583-3334.

## 1. Introduction

Eye tracking studies have taught us important lessons about the visual scanning of static and dynamic scenes and about reading processes. Since Javal (1879) pioneer stereoscope studies, we have learned that the human eyes do not glide smoothly on the lines of texts, but intermingle stops and jumps, foveally fixating areas of one to two degrees for accurate reading. Modern eye tracking has carefully established the relevance of eye movements in the investigation of the interfaces between language processing, attention, vision, and oculomotor control, spanning basic perceptual information on grapheme recognition and high-level cognitive processes.

Rayner (1998) gives a thorough overview of the research in the last two decades of the 20<sup>th</sup> century. Schroeder, Hyöna and Liversedge (2015) summarize the state of the art in eye tracking and outline the main research lines in studies presented in a special issue of the *Journal of Cognitive Psychology*, with special attention to developmental aspects of reading. We have learned from Rayner's review that there were three eras in eye tracking studies. The first followed Javal and Huey (1908) research until around 1920, establishing many basic facts concerning eye tracking, such as saccadic suppression, saccadic latency, and the perceptual span. The second era had a more applied focus during the behaviorist mainstream period and, as it could be expected due to the prevailing antimentalism, did not teach us much about cognitive processes. The third era started in 1970 and is responsible for major technological advancements and the development of theories of language processing. It was during this period that influential ideas such as the eye-mind hypothesis were proposed, discussed, and partially refuted. Just and Carpenter (1980) initially formulated the strongest version of the hypothesis, stating that "there is no appreciable lag between what is fixated and what is processed" (*cf.* p. 331). If this was strictly true, wherever our gaze is our attention and mind processing would also be, in a strong cognitive version of WYSIWYG.<sup>1</sup> Even though this strong formulation had the potential to provide researchers with a powerful rationale to study and make inferences about visual processing and directly correlate fixation latencies and saccadic movement patterns to attentional and cognitive processes, subsequent studies have disproved the strong version of the eye-mind hypothesis in the light of covert attention and gaze contingency facts. Wright and Ward (2008) reviews the literature on the orientation of attention in detail, including weaker versions of the eye-mind hypothesis, which seem to be more widely accepted nowadays.

Eye tracking has also had many applications ever since its inception with Laval in the turn of the 19<sup>th</sup> to the 20<sup>th</sup> century. Laval himself had an ophthalmological laboratory at the Sorbonne and, in addition to his pioneer original research on eye movements during reading, he had an applied interest in the study of strabismus and glaucoma. Duchowski

---

<sup>1</sup> According to the Oxford Dictionary, WYSIWYG (/ˈwɪziwɪɡ/ WIZ-ee-wig), is an acronym for What You See Is What You Get, referring to a system in which an editing software allows content to be edited in a form that resembles its appearance when printed or displayed as a finished product, such as a printed document, web page, or slide presentation.

(2002) discusses eye tracking applications in the fields of Neuroscience, Psychology, Engineering, Marketing/Advertising, in what he calls the fourth era of eye tracking, adding to Rayner's typology alluded above. This fourth era would be characterized exactly by the emergence of several interactive applications in many different areas.

An area in which eye tracking has been explored is the Human Factors Engineering (*cf.* Kovesdi, Spielman, LeBlanc & Rice 2017), pertaining to measurement, evaluation, and control of human visual behavior in different settings. In a world in which information collection is an important asset in marketing research with sophisticated systems to detect consumers' preferences, the detailed information on human visual behavior, such as the frequency of fixations and saccades, the progressive and regressive scan path with fixation durations in precise areas of the visual scanning process and even pupil diameter and blink rates, offers a very accurate tool to monitor citizens. If, on the one hand, smarter cities have been envisaged with the remote eye tracking capture of citizens' gaze, allowing hands free operation in cell phones, elevators, ATM bank machines and even including algorithms capable of detecting and preventing shoplifting behavior in stores, on the other hand, neo-totalitarian surveillance systems by means of mass monitoring of citizens with the objective of control and manipulation should not be discarded. Chomsky's 1986 characterization of Plato's and Orwell's problems<sup>2</sup> would also offer in our view a proper framework to analyze the different applications of the eye tracking technology. If the educational possibilities may be mindboggling, for example, in the coupling of eye tracking and virtual reality systems (VR), a burgeoning field (*cf.* Clay, König & König 2019), allowing preliminary explorations of VR 3D spaces of different natures, providing potentially transforming VR sensorial experiences yet to be fully explored in all fields of knowledge, the use of such systems for manipulation and control must be surveyed and in itself controlled in democratic societies.<sup>3</sup> In this sense there is even urgency that universities and other scientific research institutions investigate and develop educational uses of the eye tracking technologies, creating expertise, and offering knowledge and democratic alternatives to surveillance and control as intellectual self-defense to societies. In July 2020, during the pandemic, the World Health Organization called the First WHO Infodemiology Conference intended to contribute "to understand, measure and control infodemics" (WHO 2020)<sup>4</sup>, the overabundance of information, especially low-quality manipulative information (fake news), people are massively exposed to in the world. The WHO evaluated that as the pandemic requires actions of epidemiological vigilance, the infodemic requires actions of epistemic vigilance. It is our intention that the educational proposals we advance in this article may take advantage of the eye tracking technology and contribute to improve language awareness, develop

---

<sup>2</sup> Chomsky (1986) characterizes the two most important intellectual problems which, in his view, deserve serious consideration and people's attention. Plato's problem is defined by the question "how can we know so much, if we have so little evidence?" and is discussed in terms of the poverty of stimulus problem in language acquisition. Orwell's problem is defined by the question "how can we know so little, if we have so much evidence?" and is discussed in terms of political, religious or corporate manipulation in societies.

<sup>3</sup> Morais (2017) discusses the role of literacy in democracies, which he analyzes as actually pseudo-democracies where the public opinion is shaped by ideology and political marketing interests. In such contexts, scientists would have a collective responsibility in promoting full literacy, enhancing free, argumentative and critical thinking, inspired by humanist values.

<sup>4</sup> [https://www.who.int/health-topics/infodemic#tab=tab\\_1](https://www.who.int/health-topics/infodemic#tab=tab_1)

language literacy, and enhance epistemic vigilance (*cf.* Sperber *et al.* 2010) at schools and universities, so students and society at large can be better equipped to face the challenges posed by infodemics.

The objective of the present article, therefore, is of course not to fully review the vast eye tracking basic and applied literature as a whole, but to focus on studies which report the educational use of eye tracking in order to set the foundations to allow the presentation and discussion of our own educational experiences with the use of qualitative analyses of eye tracking heatmaps and gazeplot videos, presentially in sentence reading workshops in a middle school, and later, in the remote classrooms we have been teaching linguistics in Brazil during the pandemic. In order to support both the reflections on the use of the eye tracking technology in sentence reading workshops in the middle school and in undergraduate and graduate linguistics classes at the university level, we need to discuss the concepts of sentence and metacognition, which we will take up, respectively, in sections 2 and 3 of the article, before we discuss the project in middle school in section 4 and the activities at the university in section 5. Section 6 will offer the final considerations of the article.

## 2. The case for the sentence

The study of the sentence is the focus of three subfields of Linguistics and Psycholinguistics that present issues that have been investigated in Brazil for more than two decades by the author of this paper: Grammatical Theory, Sentence Processing and Experimental Syntax. As we reviewed in Maia (2014a), the tensions between these areas can be better framed by the proposal of a single cognitive system, measured at different levels (*cf.* Lewis & Phillips 2015). From a non-reductionist standpoint, this same single system can be viewed from different related, but ontologically autonomous angles: a more theoretical one, which describes its general properties (Grammatical Theory), another that studies the system's implementation in relation to other cognitive functions, such as memory and attention and in situations of ambiguity and depth of analysis (Sentence Processing). The third angulation – intermediate to the first two – would be Experimental Syntax – the proposal of off-line and on-line investigation of the computation of grammar in processing, abstracting it, however, from mnemonic, attentional as well as uncertainty and depth of analysis. This present article explores grammatical (and ungrammatical) constructions in these three areas from an educational perspective, always taking the sentence as a central category.

Gallego and Chomsky (2020) reviews some of the basic properties of sentence generation and takes into consideration the advantages of incorporating some of these ideas in secondary education, as they may prove useful to develop what Chomsky (1988) has called science forming capacity. The science forming capacity is the innate ability to formulate and entertain questions that we do not know the answer to. It is an immensely complex skill, derived from many other auxiliary skills. Honda and O'Neil (1993) propose that the innate science forming capacity can be triggered by properties of the formulation of linguistic theories and methodologies which can be constituted as epistemic tools to be used in the development of basic reasoning skills – observation of

empirical data, generalizations, formulation of hypotheses, comparison and evaluation of explanatory theories. In Brazil, Pilati (2017) pioneered proposed a new understanding of the role of grammar in linguistic education. Pilati's *Gramaticoteca* includes the building and the use of sets of concrete manipulable materials actively constructed by teachers and students to learn different grammar topics based on Generative Linguistics principles.

Noam Chomsky, the main proponent of Generative Linguistics in the mid-fifties, has published himself a book called *On Miseducation* (2004), in which necessary educational reforms are discussed in the light of social change and the modern democracy crises. In more recent work, Gallego and Chomsky (2020), in addition to the presentation of generative grammar key ideas, such as the sentence compositional operation MERGE<sup>5</sup>, also discuss the role of grammar in educational programs in terms of reduction of contents and focus on small objects. In Roeper, Maia & Pilati (2020), we suggest that the sentence would be the proper small object to be explored in language education. Phrase compositional recursive merge has been shown to be directly related to a type of active constructive imagination called Prefrontal Synthesis (PFS). Vyshedskiy (2019) proposes that "PFS is defined neurobiologically as the conscious, purposeful process of synthesizing novel mental images from two or more objects stored in memory" (*cf.* p.5), and states that Chomskyan Merge, defined as a process of combining any two syntactic objects to create a new one could be taken as the linguistic equivalent of PFS. According to Vyshedskiy (2019), research has shown that children who were not exposed to recursive merge may have PFS disabilities. Vyshedskiy, Mahapatra and Dunn (2017) reviewed research that demonstrate that "individuals deprived of recursive conversations until puberty performed poorly in all PFS tests, both verbal and nonverbal, despite focused multi-year post-pubertal rehabilitation efforts" (*cf.* p. 30). Therefore, it seems only natural that sentential merge processes, including recursion, would be an important small object to be explored in language education.

However, despite the very clear potential of the sentence in language education, Scott (2009), in her "A case for the sentence in reading comprehension", points out that:

I will make the case that sentence comprehension is a culprit for some readers and is commonly overlooked when thinking about improving reading comprehension and content knowledge. If a reader cannot derive meaning from individual sentences that make up a text, that is going to be a major obstacle in text-level comprehension. This statement seems so obvious that it is all the more puzzling that so little attention has been paid to sentence parsing as a component of reading comprehension. The word sentence does not even appear among the recommended domains of best practice instruction taken up by the National Reading Panel (2000) and popularized as the five topic headings of phonemic awareness, phonics, fluency, vocabulary, and text comprehension. Nor are sentences discussed at all under the topic of text comprehension. (*cf.* Scott 2009, p. 184)

In Brazil, the negligence towards the sentence in language education is unfortunately also prevailing. Lobato (2003) in a lecture published in Pilati, Vicente, Naves and Lima-Salles (2014) proposes that taxonomic grammatical teaching should be

---

<sup>5</sup> Gallego and Chomsky (2020) define MERGE as "a simple compositional operation, ..., which applies to two objects,  $\alpha$  and  $\beta$ , to yield a third one,  $K$ , which we can represent as  $[K \alpha, \beta]$ " (*cf.* p. 5).

abandoned at the schools but that a new perspective of grammatical teaching has to remain, making explicit reference to the sentential level. Lobato ponders that the same abstract grammar underlies words, phrases, clauses, sentences, and texts and points out that the human mind would have to be antieconomic if there had to be different principles to each of these levels. She recognizes that there are principles that may go beyond the limits of the sentence, but these principles cannot be different in nature of the principles that govern sentences (*cf.* p. 25).

As a matter of fact, in Brazil, there has been an important proposal to focus on the articulation of clauses at the sentential level in language teaching. However, this proposal has been little explored at the schools, which in general adopt textbooks where larger texts of different genres are most often chosen as the main object of study. Othon Moacyr Garcia's book *Comunicação em Prosa Moderna – aprendendo a escrever, aprendendo a pensar*<sup>6</sup>, on the other hand, argues that the logical analysis of clauses in sentences and sentences in paragraphs would be fundamental to developing reading and writing skills. Garcia's book does not only discuss theoretical issues, but mainly proposes a variety of exercises of clause articulation that have the potential to shed light on the important notion of **perspective**, established by the syntactic hierarchy and semantic autonomy of the main clause, especially in adjective clause modification and adverbial clause contexts in complex subordinate sentences, which are also compared to flatter coordination processes. Therefore, the central syntactic processes of coordination, adjunction and embedding are actively practiced in Garcia's manual so that students do have a proper small object (*cf.* Gallego & Chomsky 2020) to become aware of core structural properties of sentential organization and, in a next step, of the organization of sentences in paragraphs. Even though the psycholinguistics of reading concept of **gist**<sup>7</sup> is not explicitly mentioned in the book, Garcia's notion of syntactic and semantic perspective does the job quite well. The book also explores other central notions in the study of sentences, such as viewpoint, emphasis, parallelism, and even prosody notions as rhythm and cadence. Maia (2006) further elaborates on Garcia's ideas and exercises, but as evaluated for the US<sup>8</sup> in Scott (2009), it is regrettable that also in Brazil there are few language manuals and textbooks with a real focus on the sentence.

---

<sup>6</sup> Communication in Modern Prose: learning to write learning to think. ([https://www.amazon.com.br/Comunica%C3%A7%C3%A3o-prosa-moderna-Moacyr-Garciaebook/dp/B01695S0LU/ref=tmm\\_kin\\_swatch\\_0?\\_encoding=UTF8&qid=&sr=](https://www.amazon.com.br/Comunica%C3%A7%C3%A3o-prosa-moderna-Moacyr-Garciaebook/dp/B01695S0LU/ref=tmm_kin_swatch_0?_encoding=UTF8&qid=&sr=) )

<sup>7</sup> In Psycholinguistics, gist has been defined as “the residue of linguistic input that remains as time passes”. The concept has been broadened to “encompass nonlinguistic as well as linguistic, patterns of sense that are present in incoming information, especially the crucial background data that accompany thinking, reasoning and learning tasks. (...) Gist is not simply the ghost of degraded verbatim information. Instead it consists of memory representations that are extracted in parallel with the encoding of such information, and it is thought to function independently of verbatim information once it has been stored”. (*cf.* Braynerd & Reyna 2012).

<sup>8</sup> In the US, Tankersley (2003, 2005) discuss and propose practices focusing on the articulation of clauses in sentences as part of literacy strategies.

### 3. Metacognition

The simplest definitions of metacognition are generally phrased as “thinking about thinking” or “cognition about cognition”. The term, however, covers phenomena that are anything but simple. As evaluated by Mahdavi (2014), since its introduction by the developmental psychologist John Flavell in the seventies (*cf.* Flavell 1979), the term has become an umbrella for the most different purposes in a wide range of areas. Metcalfe (2008) discusses the importance of self-reflective cognition in the evolution of human consciousness and reviews investigations of its presence in primates and even in rats, to conclude that only humans would really have metacognition, which seems to have evolved relatively recently in the species, as an evolutionary advantage. According to Metcalfe, most claims that other non-human animals would have metacognition could be ruled out as “mere conditioning contingencies” (*cf.* Metcalfe 2008, p. 18). Having acquired this unique capacity of projecting levels of representation beyond the immediate here and now experience, humans would be free “to escape from being stimulus bound, and allow self-control of their learning, and actions” (Metcalfe 2008, p.1). The human metacognitive evolutionary advantage would also extend to auto-noetic consciousness, the capacity to recursively introspect on one’s own subjective experience through time. As pointed out by Metcalfe (2008), “being able to reflect on past occurrences also has its own adaptive value, freeing such an animal from the constraints of the stimulus, and allowing more rational, adaptive future responding” (*cf.* Metcalfe 2008, p. 5).

As reviewed in Maia (2019a), Hauser, Chomsky and Fitch (2002) argue that the basic property of the human faculty of language would be strictly the recursive operation, which has recently appeared in human phylogeny. This ability to manipulate symbols recursively, embedding and re-embedding them in a creative and unlimited way, to produce and understand utterances never said or heard before, could be thought of as related to the emergence of the capacity of metacognitive representation. According to De Villiers & De Villiers (1999), cognitive development and language acquisition by the child are directly dependent on sophisticated metacognitive faculties such as the Theory of Mind (ToM), the ability to attribute mental states and points of view which are different from our own to others. As discussed in Amaral, Maia, Nevins and Roeper (2018, pp. 5–7), when answering a second-order belief question such as “what did the girl say she bought?” children would have up to a certain age a tendency to respond with what the girl actually bought rather than what she said she did. Obviously, the two things could differ, as the girl could have said, for example, that she bought a toy car, when in fact she would have bought a doll. The attribution of false belief would depend both on metacognitive development and on the full development of recursive subordinate embedding, which according to Roeper & Oseki (2018) would be the final stage of the acquisition path, whose initial stage would be the simple juxtaposition of items. Interestingly, metacognitive abilities seem to be directly relevant to research such as Vyshedskiy’s (2019), briefly reviewed in the previous section, which has shown that children who were not exposed to recursive merge may have PFS disabilities. In this sense, linguistic metacognitive practices as we discuss in this article may be central to language, reasoning and imagination development. In addition to its relevance to ToM and recursion,

metacognition has also been analyzed as comprising different skills which are relevant in developmental psychology research about empathy and perspective taking<sup>9</sup>, including not only their characterization but also disabilities. Metacognitive complex abilities have thus also been object of clinical research on autism (*cf.* Frith & Happe 1999) and schizophrenia (*cf.* Coulacoglou & Saklofske 2017).

The aim of this section is, however, not to fully survey the heterogenous literature on metacognition, but to focus on some concepts and practices more directly related to the topic of the current article, namely, the use of eye tracking as an epistemic tool in language education. In both educational endeavors we describe in this article, we use qualitative eye tracking data as an aid to help students to “go meta”, as put by Mahdavi (2014). Going meta with eye tracking has the potential not only to help students become aware of their own on-line reading properties but also offer them interesting data to exercise their science forming capacity, inspecting, and contrasting gaze plots and heatmaps to make hypotheses, establish variables, make inferences and draw conclusions. In these lines, Mason, Pluchino and Tornatora (2015) proposed the EMME (Eye movement modelling example) model, based on experiments carried out with 64 seventh-year elementary school students in Italy. Considering that modern eye trackers are non-intrusive, allowing not only the capture of fixations and saccades, but also the display of results in the form of videos, the authors hypothesized that the viewing of eye tracking videos, modeling the behavior of more competent readers, could help less proficient readers. The results obtained were discussed by the authors based on theoretical frameworks for the development of self-regulatory observational processes (*cf.* Groenendijk, Janssen, Rijlaarsdam & van den Bergh 2013), demonstrating that the procedures used effectively contributed to refining students' awareness of their strategies for reading texts and images, which were positively impacted.

In the metacognitive workshops we discuss in the present article, we have proposed to start with epilinguistic activities, that is, activities with initial inductive adherence to the playful exploration of data, which are distinguished from “meta” activities that deductively depart from broader analytical categories. According to Garcia and Sista (2020), the concept of epilinguistic activity was first proposed in Culioli (1968), based on the Theory of Predicative and Enunciative Operations. The Brazilian linguist Carlos Franchi, in an article from 2002 and, mainly, in his 2006 book, proposes that these epilinguistic activities deserve to be further explored in language education, in which metalinguistic activities tend to predominate. According to Franchi: “we call epilinguistic activity this practice that operates on language itself, compares expressions, transforms them, experiments with new modes of construction, canonical or not, plays with language, invests linguistic forms with new meanings (Franchi 2006, p. 97).<sup>10</sup>

In the following sections 4 and 5, both the objectives of using eye tracking to improve self-awareness and self-regulation of students of their own reading properties

---

<sup>9</sup> Perspective taking is directly relevant to the clause articulation practices discussed in section 2 and also described with examples in sections 4 and 5 of this current article.

<sup>10</sup> “...chamamos de atividade epilinguística a essa prática que opera sobre a própria linguagem, compara as expressões, transforma-as, experimenta novos modos de construção canônicos ou não, brinca com a linguagem, investe as formas linguísticas de novas significações (Franchi 2006, p. 97).

and to trigger their science forming capacity will be exemplified in projects developed in a middle school and at the university level.

#### 4. Psycholinguistics and metacognition at the school

In 2014, my Experimental Psycholinguistics Laboratory (LAPEX/UFRJ) and Aniela Improta França's neuroscience of language laboratory (ACESIN/UFRJ) participated as founding members of the creation of *Rede Nacional de Ciência para a Educação* – a national network of scientists in Brazil intended to discuss and develop means of translating epistemologies and methodologies used in scientific research in all fields to impact all the basic education disciplines at elementary, middle and high school levels. LAPEX and ACESIN then created Laboratory LER to develop a first project with the objective to translate linguistic concepts and methodologies to a public middle school in Rio de Janeiro.<sup>11</sup> The project had as inspiration the “triggering science forming capacity” concepts and practical proposals (cf. Honda & O’Neil 1993, 2008), as mentioned above. The fundamental principles of Chomskyan sentence-based linguistics were also assumed (cf. Gallego & Chomsky 2020), specially the innatism hypothesis. Knowledge of language is part of the human nature, a genetic endowment of the species. Therefore, people may have deep intuitions and insights on their languages, including also metacognitive inspections. The project subscribed to active methodologies, problem-solving based education, student protagonist decolonizing pedagogies as used in the *3Grau Indígena project*, and other previous educational experiences.<sup>12</sup> The whole story of the three phases of the project in 8<sup>th</sup> grade middle school classes in a public school in Rio de Janeiro is presented in Maia (2019b), *Psicolinguística & Metacognição na Escola*, in which the conceptual foundations, the diagnostics, the different workshops and the reassessment phases are fully described and discussed. The project was presented to and approved by the school community, including administrators, teachers, students and parents. In brief, three groups of 20 students each were tested, namely, the workshop group, the middle school control group and the university control group. ANOVAS and t-tests were performed, showing statistically significant differences, as indicated in Graphs 1 and 2. The students in the workshop group took part in weekly extra active, problem-solving based sessions, whereas the students in the middle school control group received only their regular classes at the school.

---

<sup>11</sup> Professor Aniela Improta França (UFRJ) and I founded a new laboratory sharing resources of her Neuroscience of Language EEG Laboratory (Acesin - <http://www.acesin.letas.ufrj.br/>) and the eye tracking psycholinguistics laboratory [LAPEX - [https://pt.wikipedia.org/wiki/Laborat%C3%B3rio\\_de\\_Psicolingu%C3%ADstica\\_Experimental\\_\(LAPEX/UFRJ\)](https://pt.wikipedia.org/wiki/Laborat%C3%B3rio_de_Psicolingu%C3%ADstica_Experimental_(LAPEX/UFRJ))]. This new laboratory was also joined by laboratories from UERJ, the Rio de Janeiro state university and was called LER (<http://www.ler.letas.ufrj.br/>) and was designed to be dedicated to reading studies using neurolinguistic and psycholinguistic methodologies, such as EEG and eye tracking.

<sup>12</sup> Maia, Franchetto, Storto and Sandalo (2002) present and discuss their proposals to use active decolonizing methodologies and problem-solving based practices to prepare indigenous teachers as researchers when they were consultants and teachers in the 3Grau Indígena Project ([https://pt.wikipedia.org/wiki/Faculdade\\_Ind%C3%ADgena\\_Intercultural](https://pt.wikipedia.org/wiki/Faculdade_Ind%C3%ADgena_Intercultural)), the first program of higher education of indigenous teachers in Brazil (MEC/UNESCO).

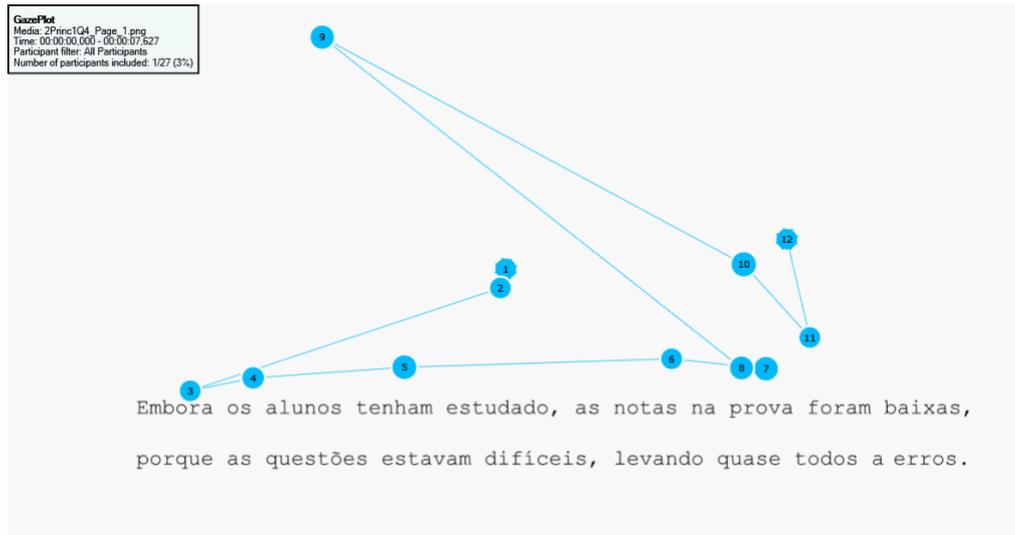
One of the major findings obtained in psycholinguistic and neurolinguistic tests in the diagnostics phase of the project showed the precocious disengagement in the computation of clauses during sentence reading. Full details of methodology and statistical analyses are presented in the Maia (2019b) book, but it is worth mentioning that the disengagement in reading was consistently captured in EEG, eye tracking, self-paced reading and cloze tests carried out by LER laboratory members during the diagnostics semester phase. All tests captured a remarkably strong effect of abandoning the reading to speed up to the interpretation questions in typical reading questionnaire exercises which are common at the schools. As this disengagement effect could also be qualitatively captured and displayed in the form of static and dynamic eye tracking fixation gaze plots or heatmaps, which are very intuitive and user friendly, among the several workshops developed and applied during the treatment year, we decided to explore eye tracking as a metacognitive educational tool, envisaging the future possibility of integrating the technique in school routines, as a reading microscope in school science labs.

Another important finding of the first phase of the project concerned the average fixation latencies and lookback rates on the main clause of the sentences as compared to the subordinate clauses. Proficient university undergraduate students would display higher rates than middle school students in these two on-line measures as well as higher accuracy rates in interpretation questions. The lookback rates found in the diagnostic eye tracking tests run with middle and university students were in line with Hyöna, Lorch and Kaakinen (2002) findings. The authors argue that lookbacks would be indicative of the integrative effort of the information read in order to solve comprehension difficulties. Based on the examination of these second pass refixations in the reading of texts, the authors were able to establish strategies of reading, differing linear readers from structuring selective readers, who are capable to identify the gist in reading more efficiently than the linear readers.

The eye tracking workshops developed during the treatment phase of the project were restricted to one of the two 8<sup>th</sup> grade classes (workshop group), as the other class was taken as a control group, not taking part of the workshop sessions, which took place every week during a one-hour recess time for two months. Inspired by problem-solving based teaching methodologies (*cf.* Duch, Groh & Allen 2001), and supported by the intuitive nature of the qualitative eye tracking reading fixations videos and heatmaps, the workshops sought to trigger science forming capacities of systematic observation, comparison, contrast, variables identification, generalization, explanation, hypotheses forming, inference, etc. Students were encouraged to explore all eye tracking parameters without much previous information on eye tracking measures, which were easily discovered by them. For example, in the first session we showed them actual readings of sentences such as below<sup>13</sup>:

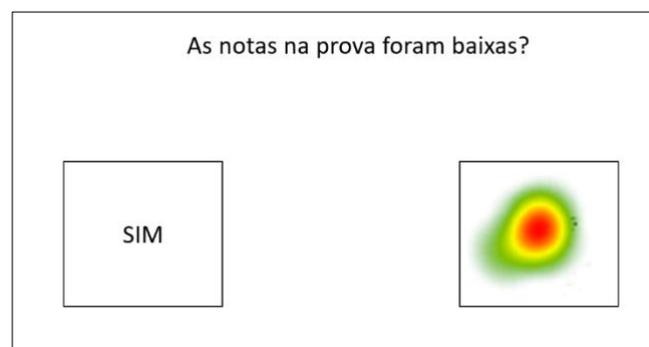
---

<sup>13</sup> The Brazilian Portuguese four clause sentence displayed in Video 1 could be translated into English as “Even though students had studied, test grades were low because the questions were too hard, leading almost everyone to errors”.



**Video 1. Middle school student reading disengagement.**  
(Click on the image to watch the video)

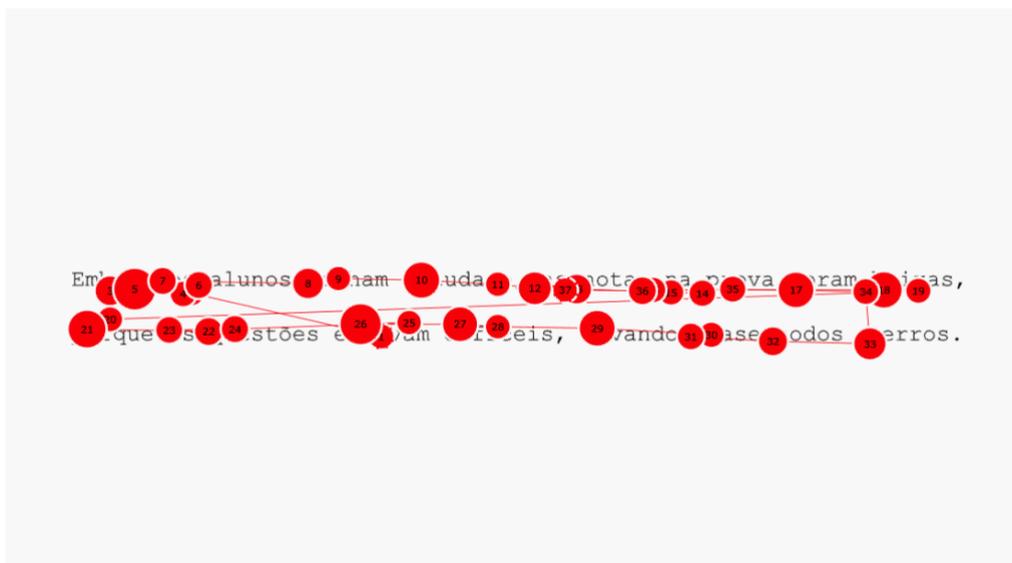
Workshop participants were very motivated to appreciate and analyze eye-tracking gaze plot videos and heatmaps of readings, collected during the first phase of the project both in the two 8<sup>th</sup> grade classes at the school, and at the university, with undergraduate students, taken as a second control group. Participants could immediately identify the early disengagement in the reading patterns in Video 1, and could also predict that the interpretation answer would be wrong, as actually illustrated by the heatmap in Figure 1<sup>14</sup>:



**Figure 1. Heatmap of Answer to Interpretation Question of Sentence in Video 1**

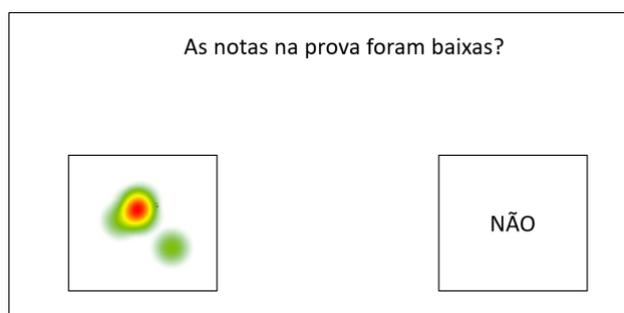
Exercising innate comparison and contrast capacities, students were usually highly motivated to discover and point out very naturally (and cheerfully) relevant eye tracking parameters in a proficient reading as illustrated in Video 2:

<sup>14</sup> The heatmap shows the answer of the student who read the sentence displayed in Video 1 to the question “were the test grades low?”. The sentence stated in its main clause that the test grades were low (*cf.* footnote 12), but the student incorrectly fixated the square with NÃO “no”, rather than the square with SIM “yes”, making a mistake.



**Video 2. University student proficient reading.**  
(Click on the image to watch the video)

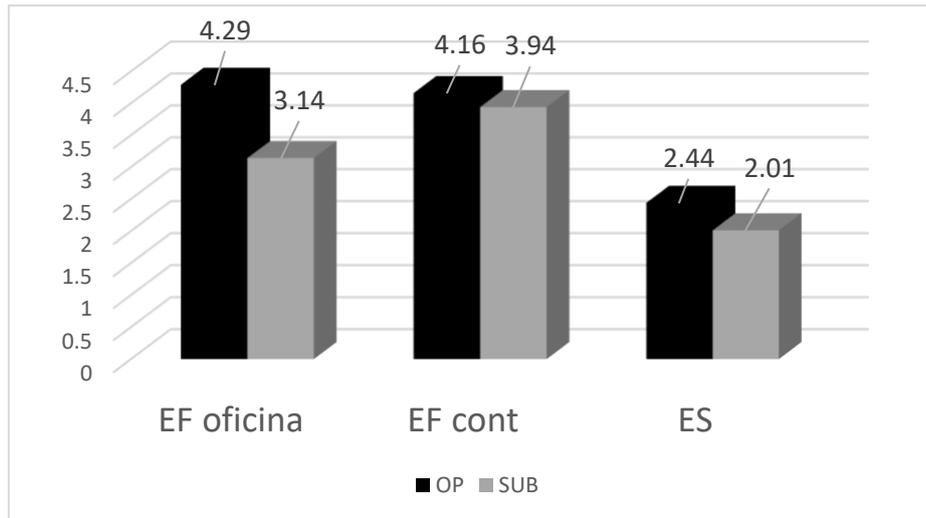
Students could also predict that a proficient reading as in Video 2 would yield a correct answer to its interpretation question, as illustrated in the heatmap in Figure 2:



**Figure 2. Heatmap of Answer to Interpretation Question of Sentence in Video 2**

Finally, in the last semester of the project several post-tests were applied, including eye tracking experiments which found significant improvement in reading patterns in the students of the middle school class which took part in workshops as compared both to control middle class and higher education students, as indicated in Graph 1<sup>15</sup> below and respectively illustrated by prototypical gaze plot videos 3, 4 and 5 in the next section.

<sup>15</sup> Graph 1 displays post-test main clause total fixation durations in black and subordinate clauses in gray for the three groups of subjects, namely, the workshop group, the middle school control group and the university control group.

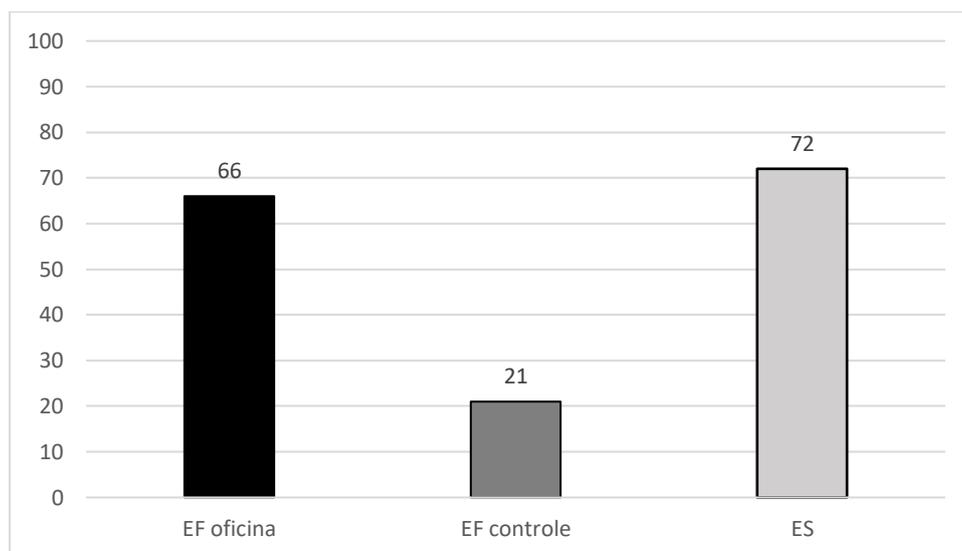


**Graph 1. Post-test Total Fixation Durations (ms) in Main and Subordinate Clauses in Three Groups of subjects: the workshop group (EF oficina), the middle school control group (EF cont) and the university control group (ES).**

Fonte: Maia (2019b, p. 92).

Graph 1 shows that the middle school group who took the workshops now displayed significantly longer latencies on main clauses than on the average of subordinate clauses, in contrast with the middle school control group who did not show significant differences in fixation between main and subordinate clauses. The last two columns in Graph 1 shows that university students read sentences faster but display significant differences in their reading of main and subordinate clauses.

Graph 2 displays accuracy rates in the interpretation questions, demonstrating how gist identification significantly improved for the workshop group who are now closer to the higher education group, having significantly improved accuracy rates in relation to the 8<sup>th</sup> grade class who did not take the workshops.



**Graph 2. % of gist identification in workshop (EF oficina), middle (EF) and university (ES) control groups.**

Fonte: Maia (2019b, p. 98).

In the next section, prototypical comparative videos 3, 4 and 5 are illustrative of the fixation and saccadic movement patterns across the three group of subjects. As they were also used in Linguistics classes at the university they will be shown in the next section of the article.

## 5. Eye tracking in the teaching of linguistics

The objective of this section is to demonstrate and discuss the use of eye tracking examples in the reading of different constructions in undergraduate and graduate linguistics classes I have taught at the Federal University of Rio de Janeiro in 2020 and 2021. The practices allowed classes with active participation of the groups, under the hypothesis that the study of syntactic theory, experimental syntax and sentence processing, including, in addition to syntactic issues, epistemological and methodological issues, would benefit significantly from quali-quantitative analyses of prototypical gaze plots and heatmaps from sentences in our eye tracking corpora, which were examined in group discussions during the courses.

In these classes not only quantitative data were discussed, but classes would generally start with the display of videos showing qualitative data, such as gaze plots and heatmaps from several studies in Psycholinguistics and Experimental Syntax that we had already developed in our laboratory.<sup>16</sup> The videos usually triggered discussions with students about linguistic and psycholinguistic issues, such as the reading of sentences for gist task used in the project described in the previous section, but also included garden-path sentences, syntactic and semantic anomalies, the filled-gap effect, island effects, direct and indirect cause, and recursive constructions, among other topics.

If the reader of the present article would like to experience how these problem-solving based classes would happen, let me suggest that the gaze plots and heatmaps illustrating each section be comparatively inspected before the analyses proposed in the text are read. Examine the fixation patterns as well as the progressive and regressive saccadic movements and try to make hypotheses concerning representation and processing issues for each construction. We will offer short briefings of the grammar and processing questions at stake and the analytical threads explored in usually cheerful 'think aloud in group' sessions, which would usually include many replays of the videos during discussions to compare observations and analyses.

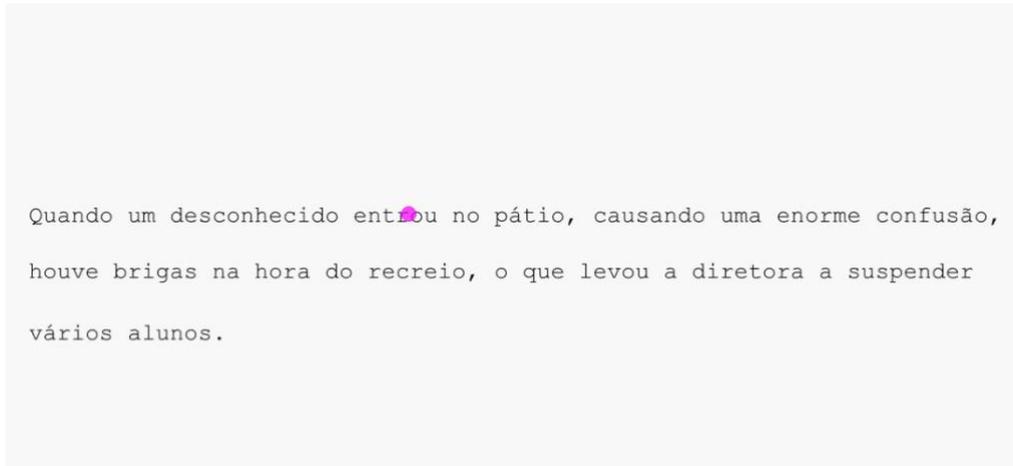
### 5.1. Gist identification data

Let us start with the post-test gaze plots illustrative of the graphs which end section 4 above. Video 3 is a prototypical middle school control group sentence reading; Video 4

---

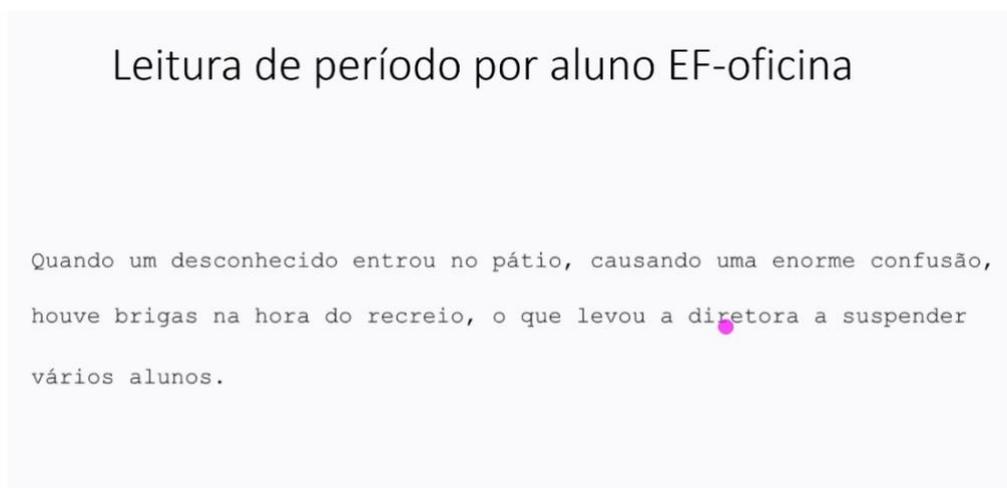
<sup>16</sup> The Laboratory of Experimental Psycholinguistics (LAPEX) at the Federal University of Rio de Janeiro has been running eye tracking experiments since 2007. During the years of 2020 and 2021, due to the pandemic, no new eye tracking tests could be implemented and run, but we could still resort to our existing eye tracking data and make active use of them in classes. These were exploratory sessions which usually yielded lively participation and enthusiastic approval from students, but at this point they have not been thoroughly evaluated yet, as suggested they should be by an anonymous Diacrítica reviewer.

is a prototypical middle school workshop group reading and Video 5 is a prototypical Higher Education sentence reading.



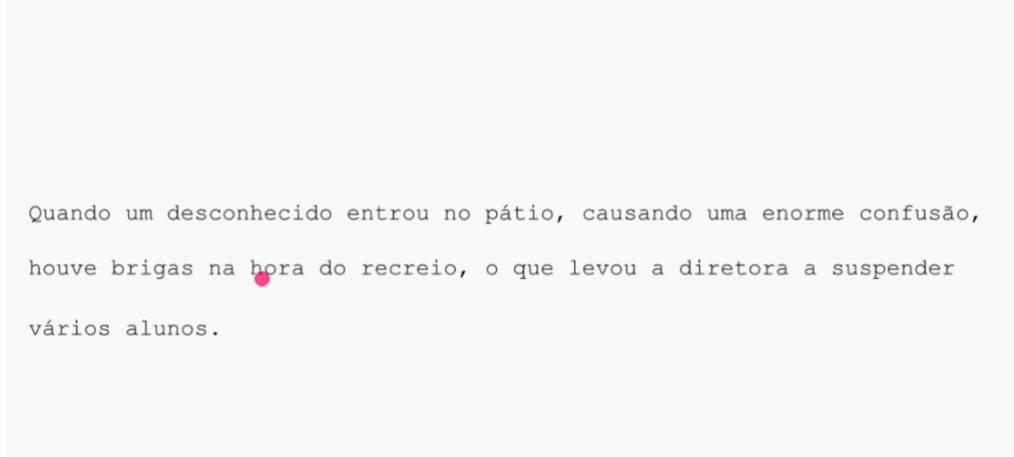
**Video 3. Middle school control group reading.**  
(Click on the image to watch the video)

Students would evaluate the reading patterns of the dynamic gaze plot in Video 3 and would generally infer low proficiency in gist identification, as there was early disengagement during sentence reading. Even though the first clause was reread, reading was abandoned right after the first pass reading of the second clause. On the other hand, Video 4 illustrates the reading of a middle student who took part in the eye tracking workshops and of other activities in the project treatment phase. Linguistics students could easily detect a selective structural reader (*cf.* Hyöna *et al.* 2002) who looks back to the critical gist clause both during first and second pass readings:



**Video 4. Middle school workshop group reading.**  
(Click on the image to watch the video)

Finally, the comparison of the same sentence read by a university student reveals that her overall reading speed is faster than the average middle school student, as indicated in Graph 1, in section 4. Higher Education subjects in average also proficiently return to previously read critical materials in their search for the gist, as illustrated in Video 5, below:



**Video 5. Higher Education control group.**  
(Click on the image to watch the video)

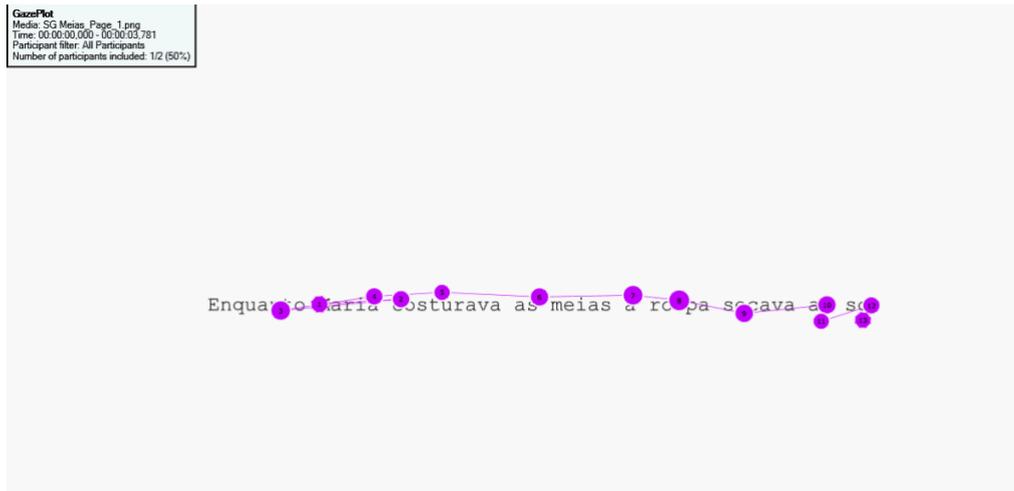
## 5.2. Sentence processing data: late closure

Sentence Processing principles such as Late Closure – LC (Frazier & Rayner 1982; Ribeiro 2012) were also presented and discussed with the active participation of undergraduate and graduate students, based on the visualization of typical garden-path readings which violate the LC principle, in contrast to readings which do not violate it.<sup>17</sup> The discussion about the psychological reality of punctuation as a cue to separate syntactic constituents could also benefit from the presentation of the eye tracking gaze plots. Let us, first, look at the relevant qualitative eye tracking data, comparing videos 6, 7 and 8<sup>18</sup>:

---

<sup>17</sup> These gaze plots were prototypical examples of a study reported in Ribeiro (2012).

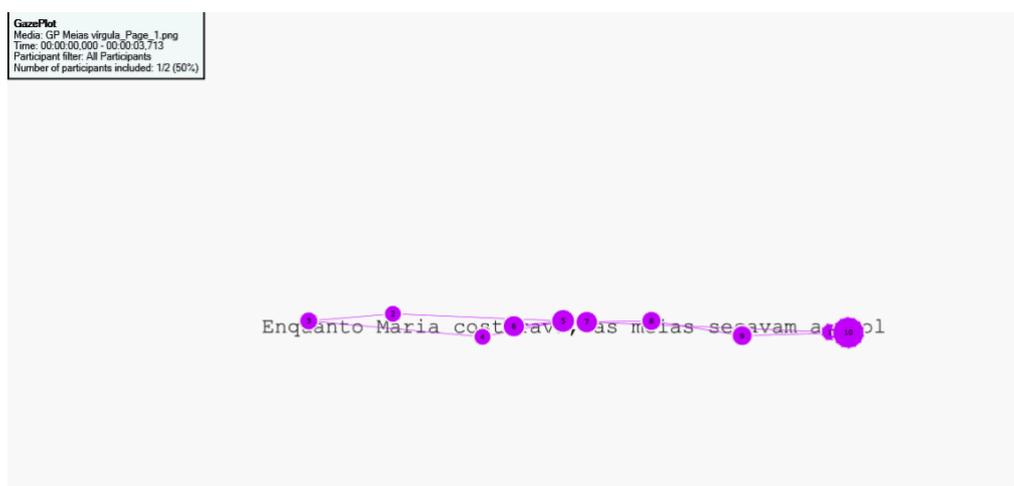
<sup>18</sup> Video 6 sentence displays the BP equivalent to the English sentence While Maria sewed her socks the clothes dried in the sun; video 7 BP garden path sentence can be translated as While Maria sewed her clothes dried in the sun; video 8 BP sentence includes a comma which avoids the garden path in the sentence in video 7 and can be translated as While Maria sewed, her clothes dried in the sun.



**Video 6. Sentence without a garden-path.**  
(Click on the image to watch the video)



**Video 7. Garden-path sentence.**  
(Click on the image to watch the video)



**Video 8. Garden-path avoided by comma.**  
(Click on the image to watch the video)

The videos posed interesting problems to be solved by the students. While video 6 was quickly read requiring only progressive fixations, video 7 had several regressions to the critical area of the garden-path, namely the region in which the *NP as meias* “the socks” is initially attached as the object of the verb *costurava* “sewed” and needed to be reanalyzed as the subject of the verb *secavam* “dried”. In video 8, the comma avoided the garden-path. Even undergraduate students who had had no previous classes about the garden-path theory or the late closure principle were able not only to establish the facts comparing the readings of the three sentences but could also deduce the basic tenets of garden-path theory and elaborate about the usage of commas as syntactic markers.

### 5.3. Sentence processing and experimental syntax data: filled gap effect data

Maia (2014b) investigated Filled Gap Effect (FGE) constructions (*cf.* Stowe 1986) such as in Example 1 in comparison with controls as in Example 2 in Brazilian Portuguese, through eye-tracking and self-paced reading experiments:

(1)

Que livro o professor escreveu a tese sem ler antes?

Which book the teacher wrote the thesis without reading before

“Which book<sub>i</sub> did the teacher write the thesis without reading t<sub>i</sub> before?”

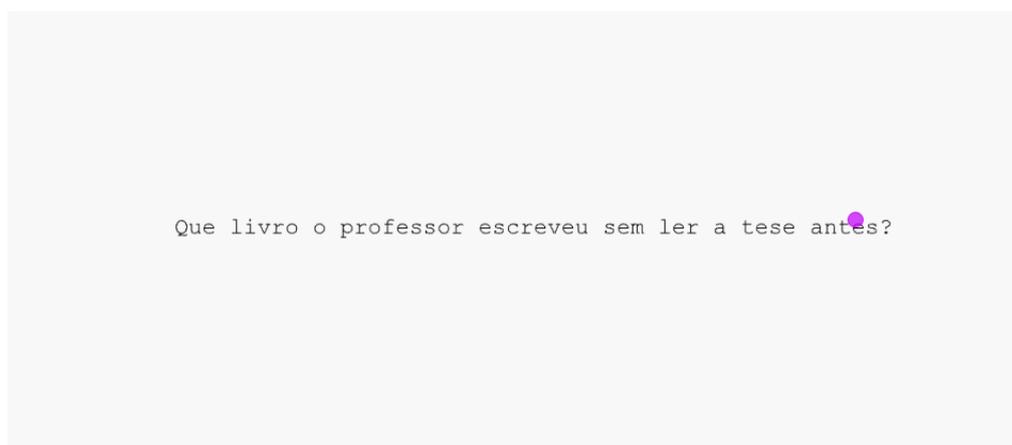
(2)

Que livro o professor escreveu sem ler a tese antes?

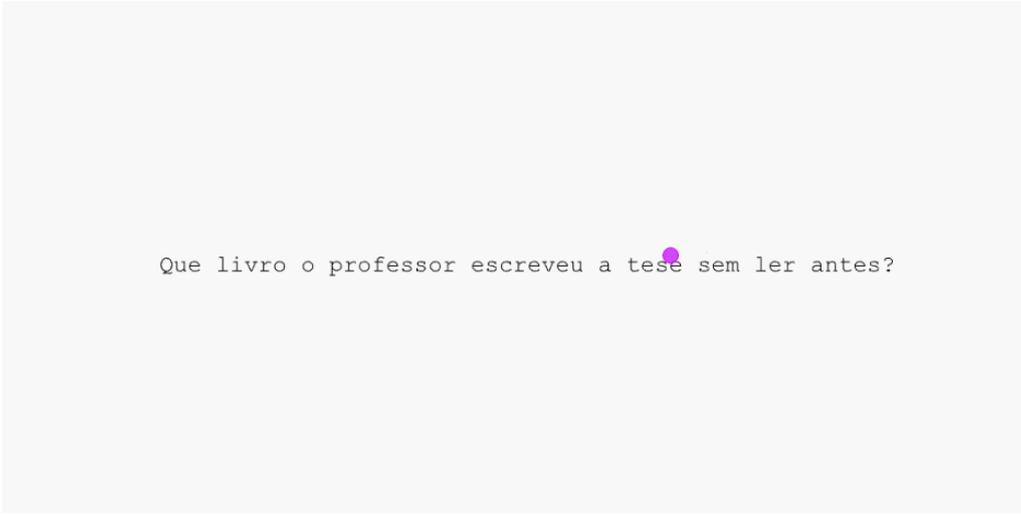
Which book the teacher wrote without reading the thesis before

“Which book<sub>i</sub> did the teacher write t<sub>i</sub> without reading the thesis before?”

Typical non-discrepant qualitative data of the eye tracking readings of sentences in these two conditions, such as videos 9 and 10 below, obtained in the study reported in Maia (2014b), were showed, and discussed in graduate Experimental Syntax and Sentence Processing classes in 2020 and in 2021. As proposed above, the readers of this present article are also invited to inspect and try to comparatively analyze the gaze plots:



**Video 9. Sentence without an FGE.**  
(Click on the image to watch the video)



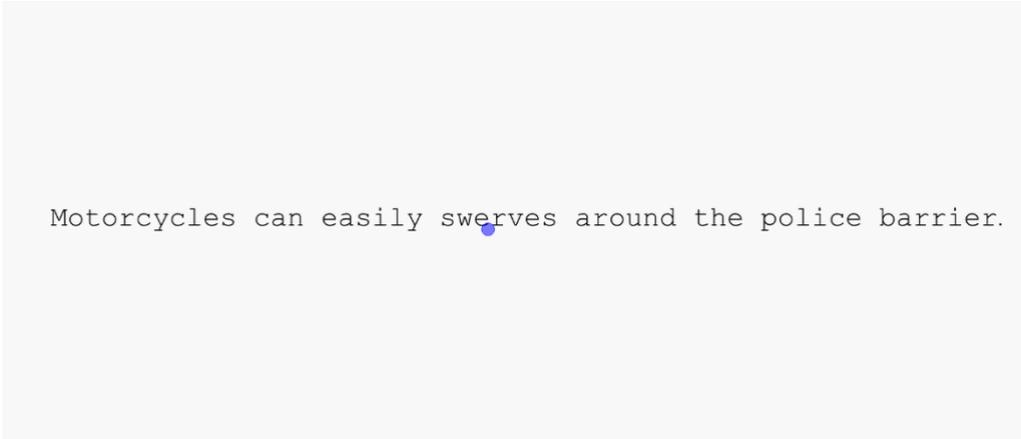
Que livro o professor escreveu a tese sem ler antes?

**Video 10. FGE sentence.**  
 (Click on the image to watch the video)

As the reader may have observed in Video 9, the fixation point is in the center of the screen when the sentence first appears, and then immediately moves to the left to start the reading of the sentence. Fixations then move forward along the sentence up to the end when there is a selective regressive movement to the WH-phrase *que livro* “which book”, in the beginning of the sentence. In Video 10, on the other hand, the fixation point also starts at the center of the screen and moves left to start reading the sentence. However, the sentence reading pattern is more costly now. There is a long fixation in first pass reading on the NP *a tese* “the thesis”, and there are also second pass regressive movements to this same NP, during the reading of the sentence. These patterns were actively established by the students early in the class and explanatory hypotheses were advanced along two lines: (i) the Filled Gap Effect sentence processing principle and (ii) the immediate satisfaction of grammatical conditions in the experimental syntax framework. As indicated in the introduction of the current paper, the Sentence Processing and the Experimental Syntax specialties are related but should not be properly taken as isomorphic. The examination of quantitative and qualitative data on the Filled Gap Effect was a relevant point of departure to substantiate the epistemological discussion concerning these two levels of analyses, namely the parsing algorithm and the grammatical assignment of theta-role to the WH-Phrase *que livro* “which book”, in an A-bar position (*cf.* Wagers & Phillips 2009), which would require the postulation of a gap as a first resort, which would be trivial in Example 1, but costly in Example 2, since there is an NP (*a tese*) filling the object position where the Wh-Phrase could be interpreted as first resort.

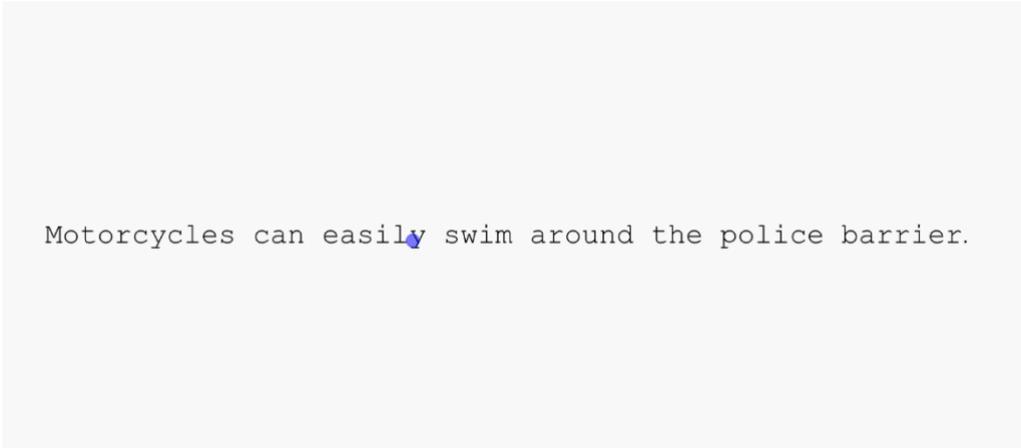
#### 5.4. Syntactic and semantic anomalies data

Data as the dynamic eye tracking gaze plots<sup>19</sup> in videos 11 and 12 below were used to trigger discussions in a graduate course about syntactic and semantic anomalies and their detection in reading comprehension.



Motorcycles can easily swerves around the police barrier.

**Video 11. Syntactic anomaly.**  
(Click on the image to watch the video)



Motorcycles can easily swim around the police barrier.

**Video 12. Semantic anomaly.**  
(Click on the image to watch the video)

Maia & Nascimento (2020) investigated to what extent the reading of sentences in Brazilian Portuguese may differ in relation to grammatical and semantic anomalies and how these sentences are evaluated for their good formation, taking as control syntactically and semantically well-formed sentences. The analyses of the on-line data suggested that sentences that present anomalous number agreement features in the object noun phrase are more costly in terms of fixation and re-fixation patterns than the semantic anomalous sentences, even though they may be better accepted by educated speakers in the off-line

---

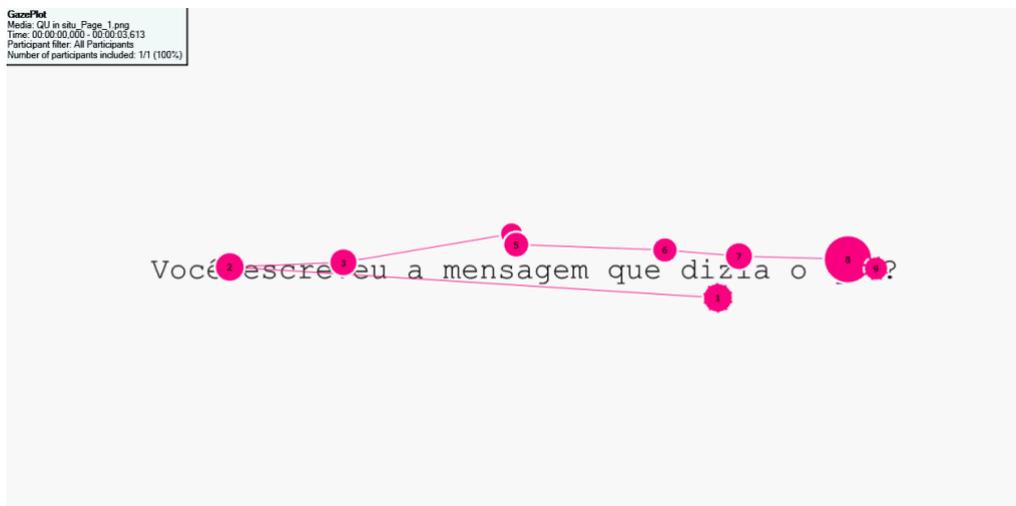
<sup>19</sup> These English gazeplot data are part of a pilot study carried out with native New Zealand English speakers in the Laboratory of Experimental Psycholinguistics (LAPEX/UFRJ) in 2019. The data go in the same direction as the BP data collected in 2019 and analyzed and published in Maia and Nascimento (2020) in which syntactic anomalies were shown to trigger first pass regressions in contrast with semantic anomalies which trigger second pass regressions.

measure than sentences with semantic anomalies. It was argued that the results provide evidence in favor of structural models of sentence processing. These issues were discussed both in relation to the BP and the English data, offering an interesting opportunity to explore the syntax/semantics interface in terms of language architecture. In the English data in video 11, it is clearly noticeable that the subject-verb agreement anomaly triggered first pass regressions to the anomaly area, whereas the semantic anomaly in video 12 only triggered second pass regressions to the anomalous s-selection area.

### 5.5. Subjacent and in situ WH

Maia, Moura and Souza (2016) report an eye tracking experiment investigating the filled gap effect (FGE) in subjacent and non-subjacent constructions<sup>20</sup> in Brazilian Portuguese in a 2x2 design in which the plausibility of the verb complement s-selection is also tested. The results suggest that relative clause syntactic island configurations are quickly perceived by the parser, preventing both the postulation of the FGE, and the plausibility evaluation of the verb complement s-selection. In non-subjacent conditions, the FGE is instantiated, and the semantic evaluation occurs only subsequent to parsing decisions, not avoiding the FGE, as predicted by syntax-first processing models.

Videos 13 and 14 were shown to students in a graduate course in Experimental Syntax before Maia *et al.* (2016) was read and discussed in class.



**Video 13. Wh-phrase in situ.**  
(Click on the image to watch the video)

<sup>20</sup> According to Chomsky (1973, 1977) subjacency is a locality constraint to movement. For example, a relative clause is an island, that is a syntactic context which blocks movement.



**Video 14. Wh-phrased moved out of Relative Clause Island.**  
(Click on the image to watch the video)

Video 13 shows the reading of a BP sentence in which a wh-phrase is in situ:

- (3)  
 Você escreveu a mensagem que dizia o quê?  
 You wrote the message that said what  
 “Did you write the message that said what?”

In Video 14, the wh-phrase was moved to the left periphery of the same sentence:

- (4)  
 O que você escreveu a mensagem que dizia?  
 What you wrote the message that said  
 “What did you write the message that said?”

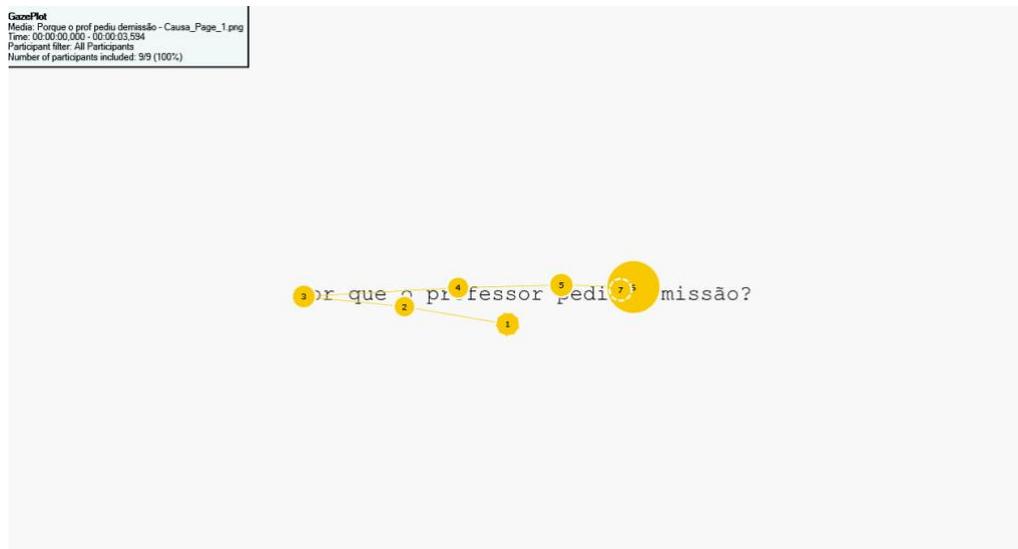
Students readily noticed that while the sentence in video 13 was read with only 9 fixations and no regressions, sentence in video 14 required 23 fixations and many second pass refixations. These striking differences motivated the active participation of the students in interesting discussions concerning the relative clause island constraint.

## 5.6. Inference data

As discussed in the introduction and in section 2 of the current paper, the case for the sentence proposed in Scott (2009) should also be considered in Brazilian education, where the sentential level has not been so explored as higher levels of the textual organization, which seem to be the mainstream topic in school textbooks. In this subsection of section 5 in which we explore the qualitative use of eye tracking gaze plots at the university level, we would like to argue that even macrostructural processes such as presuppositional and inferential analyses are already available at the sentential level

and can be productively assessed and discussed in terms of problem-solving methodologies in classes.

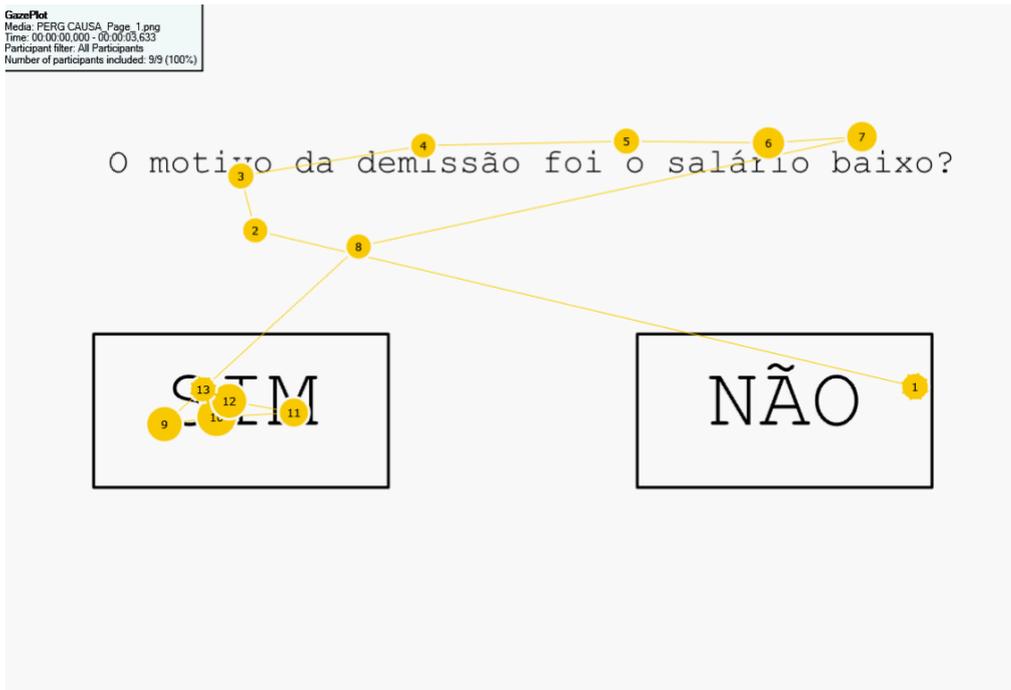
Let us start with the relevant data, involving direct and indirect causality.



**Video 15. Initial question in direct causality condition.**  
(Click on the image to watch the video)

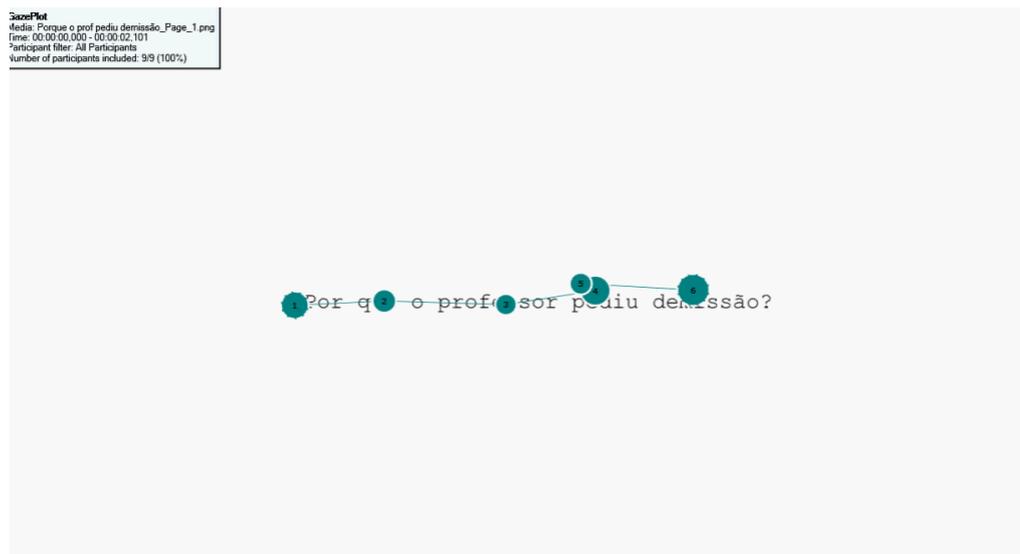


**Video 16. Direct causality sentence.**  
(Click on the image to watch the video)

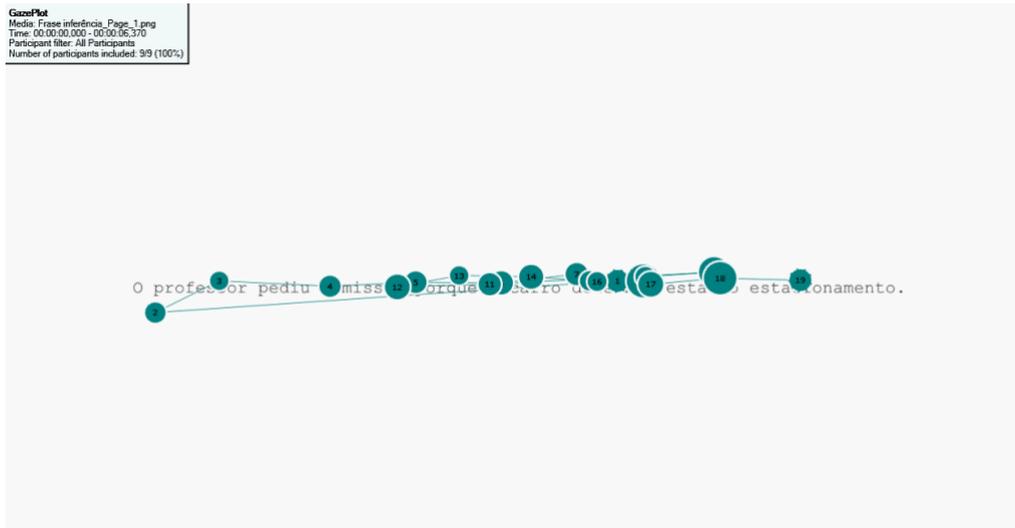


**Video 17. Final question in direct causality condition.**  
 (Click on the image to watch the video)

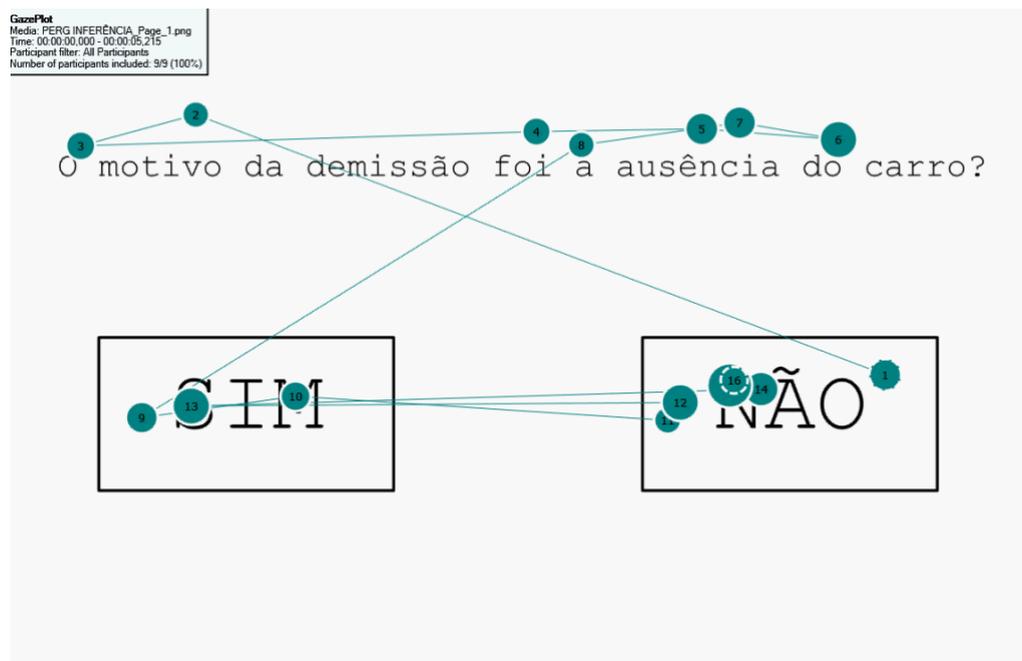
As discussed in Santos (2019) and in Roeper, Maia and Pilati (2020), direct or explicit causality, as illustrated above, seem to be processed more easily than indirect or implicit causality, which requires the establishment of presuppositions and inference. If we compare the three gaze plots above with the three gaze plots below, as we have done in undergraduate and graduate classes at UFRJ, we find a very rich material to discuss a wide variety of theoretical and methodological issues with a clear impact on the development of science forming capacity.



**Video 18. Initial question in indirect causality condition.**  
 (Click on the image to watch the video)



**Video 19. Indirect causality sentence.**  
 (Click on the image to watch the video)



**Video 20. Final question in indirect causality condition.**  
 (Click on the image to watch the video)

The comparison of the initial question in both the direct and indirect causality conditions does not yield any significant difference, as expected, since they are exactly the same stimulus in both direct and indirect causality conditions and have no previous context:

(5)

Por que o professor pediu demissão?

Why the teacher quit

“Why did the teacher quit?”

However, the readings of sentences with direct and indirect causality clearly show differences in fixation numbers and latencies, as well as in regression patterns, as students usually very accurately pointed out.

(6)

O professor pediu demissão porque o salário dele não dava para as despesas.

The teacher quit because the salary of him not enough for the expenses

“The teacher quit because his salary was not enough for the expenses”

(7)

O professor pediu demissão porque o carro dele não está no estacionamento.

The teacher quit because the car of him not is in the parking lot

“The teacher quit because his car is not in the parking lot”

In Example 6, there is a very direct cause relation between the main clause idea of quitting a job and the fact that the salary was not enough. In Example 7, on the other hand, there is only indirect causation requiring inferential reasoning in a series of presuppositional steps, which students would formulate, for instance, as:

- The teacher would probably drive to work every day.
- The teacher would usually park his car at the parking lot.
- The teacher would never miss a day.
- The fact that his car is not in the parking lot suggests that he might have quit.

Accordingly, requiring more reasoning steps than direct causality, the indirect causality gaze plot displays about 50% more fixations, also requiring second pass regressions to be properly processed. Interestingly, the final question also yields differences. While the direct causality sentence question is answered in a very straightforward fashion, the indirect causality sentence question is answered with gaze patterns which suggest doubt, as the gaze moves back and forth, scanning both answer possibilities, before finally deciding for the right answer.

### 5.7. PP direct recursion vs. PP indirect recursion

Finally, we would like to give an illustration of the use of heatmaps in graduate classes exploring the recursive embedding and juxtaposition of Prepositional Phrases in BP. Maia (2016) reported an eye-tracking sentence picture matching experiment comparing the processing of coordination (direct recursion) and embedding (indirect recursion) of PPs in BP, predicting that self-embedding would be more costly than conjoining. Results confirmed the prediction. Before reading and discussing this paper in class, students were shown heatmaps as illustrated in Figures 3 and 4, obtained in the study:

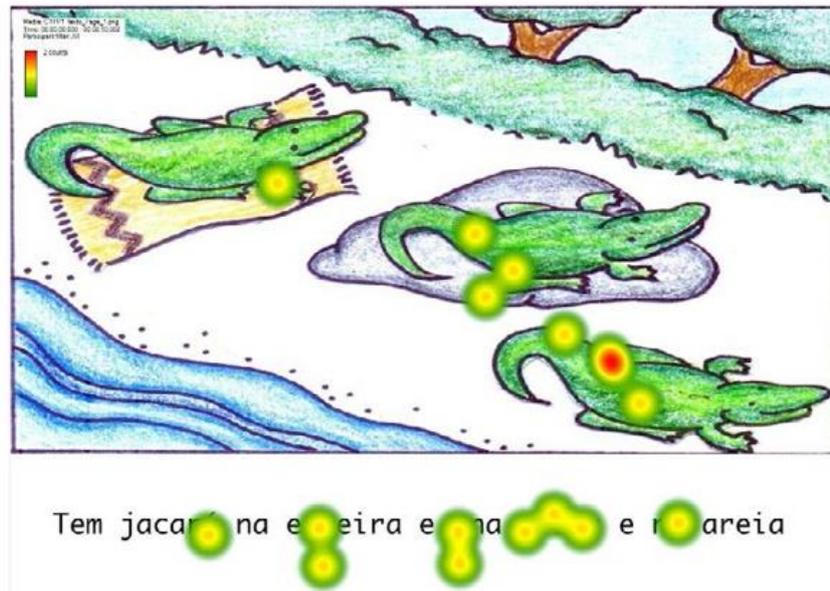


Figure 3. Heatmap of PP conjoined sentence *Tem jacaré na esteira e na pedra e na areia* “There is an alligator on the mat and on the rock and on the sand”

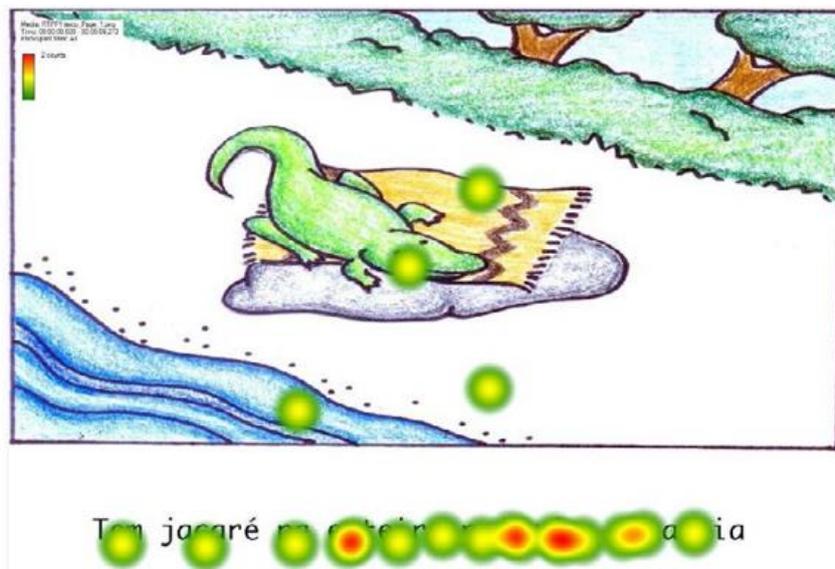


Figure 4. Heatmap of PP embedded sentence *Tem jacaré na esteira na pedra na areia* “There is an alligator on the mat on the rock on the sand”

The students very intuitively identified that the green areas in the heatmaps indicated less fixation durations, the yellow areas had less fixation durations than the red areas, which had the highest fixation durations in reading. Students also concluded that fixation patterns in the reading of the PP conjoining condition illustrated in Figure 3 tended to be less costly than the reading of the PP embedding condition in Figure 4, in which the red areas indicated higher fixation durations than in Figure 3. These observations led to interesting discussions about juxtaposition as being a flatter process than hierarchical embedding. The class analyzed these patterns as an indication of the difficulty inherent to the embedding process, which would not be the default process, if compared to the conjoining process.

## 6. Final remarks

As we have reviewed in the present paper, a metacognitive ability which came to be known as ‘theory of mind’ has been initially assessed as having first appeared in non-human primates to become later highly developed in humans, allowing individuals to infer others’ mental states and emotions. However, these earlier findings that other non-human animals would have metacognition were later ruled out as mere contingency effects (*cf.* Metcalfe 2008).

If, on the one hand, the appearance of a unique metacognitive ability would have allowed the first hominids to transcend the limits of environmental stimuli, developing consciousness of others’ mental states and self-reflexivity to become aware of themselves, on the other hand, a new challenge would have been posed: How reliable are others as sources of information? According to Sperber *et al.* (2010), to try to avoid being misinformed and deceived, humans must develop epistemic vigilance. If metacognition can be thought of as a secondary level of representation which could be fully established through hierarchical structuring made possible by language, epistemic vigilance on metacognition would be even recursively more complex, adding a third level of representation to cognition. This multilayered view of human cognition would be related to what Chomsky (2016) has called the basic property of language – the recursive capacity that allows the generation of an infinite array of hierarchically structured expressions. That being true, the question of how to develop epistemic vigilance in an infodemic world would be properly tackled through language reasoning which could then be thought of in terms of educational metacognitive exercises of recursion and inference, in addition to other structures, as we have sought to demonstrate in the paper.

Finally, we have made a case here that the sentence is the proper “small object” to start developing fundamental microstructural abilities and we have also presented our initial attempts to explore eye tracking as a metacognitive tool in education, in order to develop students’ language awareness and science forming capacity in lively active classes which may offer breakthroughs to counter infodemics, to be further explored in a post-pandemic new normal at the schools and universities.

**Financing:** This article and several experiments carried out at Lapex/UFRJ benefitted from CNPq (Brazil Research Council) grants received by the author as a 1B researcher.

**Thanks:** The ideas presented in this paper have benefitted from many interesting discussions in live talks during several remote events the author has participated, in 2020 and 2021. Taking the risk of unintended omissions, I would like to mention the 36 Encontro da Associação Portuguesa de Linguística – APL, the ABRALIN Linguists on-line event, the Third Experimental Portuguese Linguistics Workshop (University of Toronto), the III and IV Encontro Nacional sobre o Ensino de Língua Portuguesa (ENELP), the 73 Reunião Anual da Sociedade Brasileira para o Progresso da Ciência (SBPC), the 68 Seminário do Grupo de Estudos Linguísticos do Estado de São Paulo (GEL), Maia & Gomes 2020 presentation at the Language Acquisition Research Center (LARC-UMass) and the 12th International Society of Applied Psycholinguistics Congress (ISAPL).

## References

- Amaral, L., Maia, M., Nevins, A., & Roeper, T. (2018). *Recursion across domains* (1st ed.). Cambridge: Cambridge University Press.
- Braynerd, C. J., & Reyna, V. F. (2012). Domains of Fuzzy Trace Theory. In M. L. Howe & R. Pasnak (Orgs.), *Emerging themes in Cognitive Development Theory: Volume I – Foundations* (pp. 50–93). New York: Springer-Verlag.
- Chomsky, N. (1973). Conditions on transformations. In S. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 232–286). New York: Holt, Rinehart and Winston.
- Chomsky, N. (1977) On Wh-Movement. In P. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal Syntax* (pp. 71–132). New York: Academic Press.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, N. (1988). *Language and problems of knowledge: The Managua Lectures*. Cambridge: The MIT Press.
- Chomsky, N. (2004). *On miseducation*. Lanham: Rowman & Littlefield Publishers.
- Chomsky, N. (2016). *What kind of creatures are we?*. New York: Columbia University Press.
- Clay, V., König, P., & König, S. (2019). Eye tracking in Virtual Reality. *Journal of Eye Movement Research*, 12(1). <https://doi.org/10.16910/jemr.12.1.3>
- Coulacoglou, C., & Saklofske, D.H. (2017). *Psychometrics and psychological assessment*. Cambridge: Academic Press.
- Culioli, A. (1968). La formalisation en linguistique. *Cahiers pour l'Analyse*, 9, 106–117.
- De Villiers, J.G., & De Villiers, P.A. (1999). Language development. In M. Bornstein & M. Lamb (Eds.), *Developmental psychology: An advanced textbook* (2nd ed., pp. 313–373). Mahwah: Lawrence Erlbaum.
- Duch, B. J., Groh, S. E., & Allen, D. E. (Eds.). (2001). *The power of problem-based learning*. Sterling, VA: Stylus.
- Duchowski, A. T. (2002). A breadth-first survey of eye tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4), 455–470. <https://doi.org/10.3758/bf03195475>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906–911.
- Franchi, C. (2002). Linguagem – atividade constitutiva. *Revista do GEL* (número especial), 37–74.
- Franchi, C. (2006). *Mas o que é mesmo ‘gramática’?*. São Paulo: Parábola.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements and the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14 (2), 178–210. [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1)
- Frith, U., & Happe, F. (1999). Theory of mind and self-consciousness: What is it like to be autistic? *Mind and Language*, 14(1), 82–89. <https://doi.org/10.1111/1468-0017.00100>
- Gallego, A.J., & Chomsky, N. (2020). The Faculty of Language: A biological object, a window into the mind, and a bridge across disciplines. *Revista de la Sociedad Española de Lingüística*, 2020.
- Garcia, M. G., Sisle, H.C. (2020). Atividades epilingüísticas e práticas pedagógicas. *Revista de Educação (PUCCAMP)*, 25, 1–16. <https://doi.org/10.24220/2318-0870v25e2020a4904>
- Groenendijk, T., Janssen, T., Rijlaarsdam, G., & van den Bergh, H. (2013). Learning to be creative. The effects of observational learning on students’ design products and processes. *Learning and Instruction*, 28, 35–47. <https://doi.org/10.1016/j.learninstruc.2013.05.001>
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve?. *Science*, 298(5598), 1569–1579. <https://doi.org/10.1126/science.298.5598.1569>
- Honda, M., & O’Neil, W. (1993). Linguistic inquiry in the science classroom: "It is science, but it's not like a science problem in a book" (MIT Occasional Papers in Linguistics, No. 6). Cambridge: MITWPL.

- Honda, M., & O' Neil, W. (2008). *Thinking linguistically: A scientific approach to language*. Oxford: Blackwell Publishers.
- Huey, E.B. (1908). *The psychology and pedagogy of reading*. New York: Macmillan.
- Hyöna, J., Lorch, R. F. Jr., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*, 44–55. <https://doi.org/10.1037/0022-0663.94.1.44>
- Javal, E. (1879). Essai sur la physiologie de la lecture. *Annales D'Oculistique, 82*, 242–253.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading from eye fixations to comprehension. *Psychological Review, 87*(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kovesdi, C., Spielman, Z., LeBlanc, K., & Rice, B. (2017). Application of eye tracking for measurement and evaluation in human factors studies in control room modernization. *Nuclear Technology, 202*(2-3). <https://doi.org/10.1080/00295450.2018.1455461>
- Lewis, S., & Phillips, C. (2015). Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research, 44*(1), 27–46. <https://doi.org/10.1007/s10936-014-9329-z>
- Lobato, L. (2003, December). *O que o professor de ensino básico deve saber de Lingüística?*. Presented at 2a Reunião Regional da Sociedade Brasileira para o Progresso da Ciência (SBPC), Fortaleza, Brasil
- Mahdavi, M. (2014). An overview: Metacognition in education. *International Journal of Multidisciplinary Current Research, 2*, 529–535.
- Maia, M., Franchetto, B., Storto, L., & Sandalo, F. (2002). A construção do conhecimento lingüístico: Do saber do falante à pesquisa. *Cadernos de Educação Escolar Indígena, 1*(1), p. 47–78.
- Maia, M. (2006). *Manual de linguística: Subsídios para formação de professores indígenas na área de linguagem*. Brasília: UNESCO/MEC.
- Maia, M. (2014a). Teoria gramatical, processamento de frases e sintaxe experimental. *Revista da ABRALIN, 13*(2), 93–117.
- Maia, M. (2014b). Efeito da lacuna preenchida e plausibilidade semântica no processamento de frases em Português Brasileiro. *Cadernos de Letras da UFF, 49*, 23-46. <https://doi.org/10.22409/cadletrasuff.2014n49a5>
- Maia, M. (2016). A computational efficiency principle in action in the processing of recursively embedded PPs in Brazilian Portuguese and in Karajá. *Gragoatá, 21*(40), 157-174.
- Maia, M., Moura, A. R. A., & Souza, M. (2016). Ilhas sintáticas e plausibilidade semântica: Um estudo de rastreamento ocular de frases com lacunas preenchidas em português brasileiro. *Scripta, 20*, 293–311. <https://doi.org/10.5752/P.2358-3428.2016v20n38p287>
- Maia, M. (2019a). Pensando (Psico)linguisticamente, experimentalmente, educacionalmente. In E. Pilati, R. Naves & H. Salles. (Orgs.), *Novos olhares para a gramática na sala de aula: Questões para estudantes, professores e pesquisadores* (pp. 93–118). Campinas: Editora Pontes.
- Maia, M. (Org.). (2019b). *Psicolinguística e metacognição na escola*. Campinas: Mercado de Letras.
- Maia, M., & Nascimento, G. (2020). Anomalias de forma e de conteúdo em Português Brasileiro: Um estudo preliminar de rastreamento ocular da leitura e de avaliação de aceitabilidade. *Cadernos De Tradução, 40*, 45–73. <https://doi.org/10.5007/2175-7968.2020v40nesp2p45>
- Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Using eye tracking technology as an indirect instruction tool to improve text and picture processing and learning. *British Journal of Educational Technology, 47*(6), 1083–1095. <https://doi.org/10.1111/bjet.12271>
- Metcalfe, J. (2008). Evolution of metacognition. In J. Dunlosky & R. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 29–46). New York: Psychology Press.
- Morais, J. (2017). Literacy and democracy. *Language, Cognition and Neuroscience, 33*(3), 351–372, <https://doi.org/10.1080/23273798.2017.1305116>

- Oxford English Dictionary. (2019). WYSIWYG. In Oxford English Dictionary. Retrieved from <https://www.oxfordreference.com/view/10.1093/acref/9780191884276.001.0001/acref-9780191884276-e-3582?rskey=S5OPsH&result=1>
- Pilati, E. (2017). *Linguística, gramática e aprendizagem ativa* (1st ed.) Campinas: Pontes.
- Pilati, E., Vicente, H.G., Naves, R., & Lima-Salles, H. M. M. (Orgs.). (2014). *Linguística e ensino de línguas* (1st ed.) Brasília: UnB, v. 2.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3): 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Ribeiro, A. J. C. (2012). Late closure e good-enough no processamento de frases garden-path do português do Brasil: Evidências de eyetracking. *Revista Virtual de Estudos da Linguagem*, 10, 84–106.
- Roeper, T. & Oseki, Y. (2018). Recursion in the acquisition path for hierarchical syntactic structure. In L. Amaral, M. Maia, A. Nevins & T. Roeper (Eds.), *Recursion across domains* (pp. 267–278). Cambridge: Cambridge University Press.
- Roeper, T. Maia, M., & Pilati, E. (2020). *Experimentando linguística na escola: Conhecimento gramatical, leitura e escrita*. Campinas: Pontes.
- Santos, S. L. (2019). *Relações causais e o uso dos conectivos “porque” e “já que”: contribuições da psicolinguística para a educação* (Tese de Doutorado, UFRJ, Rio de Janeiro).
- Schroeder, S., Hyönä, J., & Liversedge, S. (2015). Developmental eye tracking research in reading: Introduction to the special issue. *Journal of Cognitive Psychology*, 27(5), 500–510. <https://doi.org/10.1080/20445911.2015.1046877>
- Scott, C. M. (2009). A case for the sentence in reading comprehension. *Language, Speech, and Hearing Services in Schools*, 40(2), 184–191. [https://doi.org/10.1044/0161-1461\(2008/08-0042\)](https://doi.org/10.1044/0161-1461(2008/08-0042))
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359–393. <https://doi.org/10.1111/j.1468-0017.2010.01394.x>
- Stowe, L. (1986). Parsing wh–constructions: Evidence for on–line gap location. *Language and Cognitive Processes*, 1(3), 227–245. <https://doi.org/10.1080/01690968608407062>
- Tankersley, K. (2003). *The threads of reading: Strategies for literacy development*. Alexandria, VA: Association for Supervision and Curriculum development.
- Tankersley, K. (2005). *Literacy strategies for grades 4-12: Reinforcing the threads of reading*. Alexandria, Va: Association for Supervision and Curriculum Development.
- Vyshedskiy, A. (2019). Language evolution to revolution: The leap from rich-vocabulary non-recursive communication system to recursive language 70,000 years ago was associated with acquisition of a novel component of imagination, called Prefrontal Synthesis, enabled by a mutation that slowed down the prefrontal cortex maturation simultaneously in two or more children – the Romulus and Remus hypothesis. *Research Ideas and Outcomes*, 5(4). <https://doi.org/10.3897/rio.5.e38546>
- Vyshedskiy, A., Mahapatra, S., & Dunn, R. (2017). Linguistically deprived children: Metaanalysis of published research underlines the importance of early syntactic language use for normal brain development. *Research Ideas and Outcomes*, 3. <https://doi.org/10.3897/rio.3.e20696>
- Wagers, M.W., & Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*, 45, 395–433. <https://doi.org/10.1017/S0022226709005726>
- WHO. (2020, June 25). *First WHO Infodemiology Conference*. World Health Organization. <https://www.who.int/news-room/events/detail/2020/06/30/default-calendar/1st-who-infodemiology-conference>.
- Wright, R.D., & Ward, L.M. (2008). *Orienting of attention*. Cambridge: Oxford University Press.

[recebido em 30 de outubro de 2021 e aceite para publicação em 28 de fevereiro de 2022]

## DIFFERENCES BETWEEN STEREOTYPICAL GENDER AND DEFINITIONAL GENDER IN PRONOMINAL ANTECEDENT RETRIEVAL IN BRAZILIAN PORTUGUESE

### AS DIFERENÇAS ENTRE GÊNERO DE ESTEREÓTIPO E GÊNERO DEFINIDO NA RECUPERAÇÃO DE ANTECEDENTES PRONOMINAIS EM PORTUGUÊS BRASILEIRO

Michele Calil dos Santos Alves\*  
michelecalil@gmail.com

Together with Alves (2019), this paper is part of a research work dedicated to investigating the role of gender cues in pronominal antecedent retrieval in Brazilian Portuguese. The main aim is to determine whether Principle B structural constraints work as an initial filter in pronominal antecedent retrieval in Brazilian Portuguese, or whether gender cues play a major role at early processing phases since this language has visible redundant morphological markings. It is also examined whether antecedent candidates carrying different gender cues have different weights in memory. In the present work, a comparison between antecedent candidates carrying stereotypical gender and definitional gender was provided. The results of two eye-tracking experiments conducted with native speakers of Brazilian Portuguese showed that gender morphological cues play a role at the very beginning of pronoun resolution processing, and only at a later processing phase, structural constraints of Principle B seemed to help the parser to select the correct antecedent. Like the previous work, the results indicated Principle B structural constraints do not work as an initial filter blocking structurally unacceptable candidates in Brazilian Portuguese; on the contrary, all candidates seemed to be taken into account despite of violating Principle B structural constraints. Similarly to the previous work, masculine antecedent candidates were preferable to be retrieved by memory in comparison to feminine candidates. Moreover, the results indicated that candidates with definitional gender rather than stereotypical gender are preferable to be retrieved by memory since the former is lexically specified, while the latter relies on probabilities based on world knowledge inferences.

**Palavras-chave:** Pronoun resolution. Gender cues. Memory retrieval. Brazilian Portuguese.

Juntamente com Alves (2019), este artigo é parte de uma pesquisa dedicada a investigar o papel das pistas de gênero na recuperação de antecedentes pronominais em português brasileiro. O objetivo principal é determinar se as restrições estruturais do Princípio B atuam como um filtro inicial na recuperação de antecedentes pronominais em português brasileiro, ou se as pistas de gênero são predominantes em fases iniciais do processamento uma vez que trata-se de uma língua com marcas morfológicas visíveis redundantes. Também examinou-se se candidatos que carregam diferentes marcas de gênero têm diferentes pesos na memória. Neste trabalho, serão

---

\* Centro de Educação e Letras, Universidade Federal do Acre, Cruzeiro do Sul, Brasil. ORCID: <https://orcid.org/0000-0002-9514-6795>

comparados candidatos a antecedentes com gênero de estereótipo e com gênero definido. Os resultados de dois experimentos de rastreamento ocular realizados com falantes nativos de português brasileiro mostraram que as pistas de gênero participam desde o início do processamento da resolução pronominal, e que somente em um estágio mais tardio, as restrições estruturais do Princípio B pareceram ajudar o processador sintático a selecionar o antecedente correto. Assim como no último trabalho, os resultados indicaram que as restrições estruturais do Princípio B não atuam como um filtro inicial bloqueando candidatos inaceitáveis sob o ponto de vista estrutural; ao contrário, todos os candidatos pareceram ser levados em conta apesar de violarem as restrições estruturais do Princípio B. Em semelhança com o último trabalho, nossa memória parece preferir recuperar candidatos a antecedentes masculinos em comparação com antecedentes femininos. Além disso, os resultados indicaram que nossa memória também parece preferir recuperar candidatos a antecedentes com gênero definido ao invés de candidatos com gênero de estereótipo já que o primeiro é especificado no léxico, enquanto o segundo é resultado de probabilidades baseadas em inferências de conhecimento de mundo.

**Keywords:** Resolução pronominal. Pistas de gênero. Recuperação na memória. Português brasileiro.

•

## 1. Introduction

One of psycholinguists concerns is about the mental processes that underlie language comprehension. What kinds of information are involved in language processing? Can one piece of linguistic information be more important than the other or do they all weigh the same? Can non-linguistic information play a role as well? How is the relation between language and other cognitive systems such as memory? This paper seeks to answer these questions by investigating pronouns and how they retrieve antecedents in discourse.

The syntactic constraints involved in pronoun resolution are known in the generative binding theory as Principle B, which states that pronouns must be free in their local (c-command) domains, roughly the clause where they are located (Chomsky 1993, 1995). In other words, pronouns and antecedents cannot be too close to each other. Besides that, Principle B also states that pronouns and antecedents must agree in gender and number. Finally, pronouns tend to retrieve discursive salient antecedents (Arnold *et al.* 2000; Gordon & Hendrick 1998; Foraker & McElree 2007; Greene, McKoon & Ratcliff 1992; Grosz, Weinstein & Joshi 1995; Kehler 2007; Rigalleau, Caplan & Baudiffier 2004). Given the fact that pronoun resolution processing involves syntactic, morphological, and discursive factors, it is an interesting way to investigate how mind deals with different types of information. Pronoun processing is also a good way to study the relation between language and memory since antecedents must be retrieved from a set of previous stored items. And one way to study the influence of non-linguistic factors in language processing is through gender stereotypes, as they are socially natured and depends on world knowledge. This way, this paper will study how gender-stereotyped antecedents are processed in pronoun resolution.

Previous research in Psycholinguistics have shown controversial results over the role of syntactic cues on how pronouns retrieve their antecedents. The traditional view

argues that the syntactic constraints work as an initial filter, blocking all candidates that violate Principle B structural constraints (Leitão, Peixoto & Santos 2008; Nicol & Swinney 1989; Sturt 2003, among others). However, a few works have shown that Principle B structural constraints do not work as an initial filter since structurally unacceptable antecedent candidates that agree with the pronouns can interfere in pronoun resolution at early processing phases (Alves 2019; Badecker & Straub 2002; Patil, Vasisht & Lewis 2016), even being occasionally misretrieved (Alves 2019; Engelmann, Jäger and Vasisht 2015). One may find a complete discussion over this question in Alves (2019).

The present paper aims at determining whether Principle B structural constraints work as initial filter in pronoun resolution in Brazilian Portuguese. The study reported here will investigate the role of gender morphology in pronoun resolution processing in order to find out whether structurally unacceptable candidates that agree in gender with the pronouns can interfere in pronoun resolution processing despite of violating Principle B structural constraints. In Alves (2019), it was shown that Principle B structural constraints do not work as an initial filter in Brazilian Portuguese. However, unlike our previous paper, which investigated semantic gender and grammatical gender, the experiment reported here will investigate stereotypical gender and definitional gender.

In Alves (2019), the results revealed that masculine antecedents are more prominent in memory than feminine antecedents in pronoun resolution processing in Brazilian Portuguese, because masculine works as a default gender in Romance languages (Alves 2019). Cacciari and Padovani (2007) explained that masculine spreads activation to conceptually masculine and feminine units because masculine is unmarked in Italian; on the other hand, feminine is marked, and it only activates the feminine unit. Similarly, according to Casado, Palma and Paolieri (2017), speakers include female and male representations when they process masculine nouns in Spanish. And Alves (2014) found out that masculine pronouns can equally retrieve either masculine or feminine antecedents in Brazilian Portuguese. One may find a complete discussion over masculine in Alves (2019). This way, the present paper will examine whether masculine and feminine definitional/stereotypical antecedent candidates weigh different in memory.

## 2. Stereotypical gender

According to Oakhill, Garnham and Reynolds (2005), background knowledge influences the construction of the information providing cognitive economy. The authors studied one specific type of background knowledge: stereotypical gender information. Psycholinguists often use the following riddle to exemplify stereotypical gender:

This morning a father and his son were driving along the highway to work, when they were involved in a horrible accident. The father was killed, and the son was quickly driven to the hospital, severely injured. When the boy was taken into the operating theatre, the surgeon exclaimed: “Oh, my God, this is my son!”

The riddle is based on the fact that when there is not any specific information about the gender associated with an occupation or role noun, as *surgeon*, prior knowledge, in

this case, gender stereotype, would be used by the reader to infer the more likely gender, which, in this case, is masculine. However, once that inference is proved incorrect by the end of the story, readers would have to reconstruct their character representation, which is more costly for processing.

In English, most terms for occupations are not gender marked, and many of them are gender stereotyped. So, if a reader supposes a surgeon is male and a secretary is female, it is so due to real statistics and world knowledge. However, a big question in the literature is whether stereotyped information is immediately activated by the time a stereotyped word is encountered, or only later, when required by discourse, as for example, when a pronoun needs to be resolved.

Oakhill *et al.* (2005) conducted a series of experiments in English to check whether gender stereotypical information is automatic and unable to be suppressed. In their experiments, the participants were instructed to decide whether two terms (an occupation and a kinship term) could refer to the same person (e.g.: nurse-aunt; nurse-uncle). The authors found out people were biased by gender stereotypes. For example, participants made more errors or had longer reaction times to accept pairs like *engineer-mother* than *engineer-father* or with gender-neutral terms. This way, both accuracy and reaction time suffered when stereotypical gender conflicted with the gender of the kinship term. Since the authors did not find facilitation effects of the congruent pairs, they affirmed stereotyped information is a case of interference. Moreover, according to Oakhill *et al.* (2005), gender stereotype might have a very strong and persistent effect because participants could not suppress it even after being strongly warned about it. The authors concluded gender stereotype associated with occupations and roles is incorporated into the information representation immediately as soon as an occupation or a role name is read. In other words, it is automatic.

Carreiras *et al.* (1996) argued that when reading, a representation of the situation is incrementally formed in our mental model through text input, context, and world knowledge. A particular type of world knowledge is gender stereotype. However, the authors alerted that, unlike English, in Spanish, stereotype information is not used to infer a person's gender because gender is explicitly given by the definite article. In order to compare how stereotypical gender influences comprehension in those two languages, self-paced reading was used to present short texts with a stereotypical role name followed by a pronoun (he/she). In (1), "futebolista" (footballer) is masculine stereotyped:

- (1) Examples of the materials in Carreiras *et al.* (1996)
  - a. The footballer wanted to play in the match.  
He/She had been training very hard during the week.
  - b. El futbolista quería jugar el partido.  
El había estado entrenando mucho durante la semana.
  - c. La futbolista quería jugar el partido.  
Ella había estado entrenando mucho durante la semana.

On the one hand, the results in Carreiras *et al.* (1996) showed reading times for the second sentences in English were longer when there was a mismatch between the gender of the pronoun and the stereotyped gender of the role noun. On the other hand, the results for Spanish showed reading times for the first sentences were longer when there was a mismatch between the gender of definite article and the gender of the stereotyped role noun. Carreiras *et al.* (1996) concluded stereotypes can influence the representation for a particular character in a text as soon as they become available. When they are the only source of gender information (as in the first sentences of the examples in English) they can be used to infer the gender of the characters in the mental model. However, if some upcoming information that mismatches the stereotyped inference appears in the text, that previous inference must be overridden, which might explain the difficulties the readers had in these cases. Although gender assignment in Spanish does not require a stereotyped inference as in English since gender information is already given in the articles, readers may also have to encode something new in their mental models: the clash between the gender of the articles and the stereotypical gender of the characters.

Kennison and Trofe (2003) also examined the role of gender-stereotyped nouns during language comprehension. First, they conducted a rating study with 405 nouns and compound nouns. Out of these nouns, 32 words were strongly stereotyped to refer to males and 32 to females. Thus these 64 nouns were used in a self-paced reading study on pronoun resolution, but unlike Carreiras *et al.* (1996), in which the participants saw the whole sentence, Kennison and Trofe (2003) used the phrase-by-phrase moving window paradigm. This way, they believed they would obtain more information about the time course of information than Carreiras *et al.* (1996). With respect to the self-paced reading study, reading times were longer at two regions after the pronouns when the gender stereotype of the antecedent mismatched the gender of the pronoun than when they matched. This result replicated Carreiras *et al.* (1996). Moreover, reading times at the pronoun *she* were longer than at the pronoun *he*. The authors believed it was because *he* is more frequent than *she*. Kennison and Trofe (2003) concluded word-specific gender stereotypes influence resolution of pronouns. They hypothesized comprehenders must have a representation for each word stored in memory, and this representation includes information about the word's gender stereotype, which is computed according to the relative frequency of usage in referring to males and females.

Cacciari and Padovani (2007) conducted two priming experiments aiming at investigating gender stereotypes in Italian. In Experiment 1, participants were shown pairs of stimuli formed by a prime word and a target gender-marked pronoun. The participants were asked to decide the grammatical gender of the pronoun ignoring the prime word. The primes were role nouns with and without gender stereotypes, and the stereotyped gender of the role nouns would match or mismatch the gender of the pronouns. No statistically significant priming results were found in Experiment 1. Thus, the authors decided to give the participants more time to process the prime in Experiment 2 so that the stimulus presentation time increased from 200 to 300ms and the interval between the prime and the target increased from 100 to 200ms. The results of Experiment 2 in Cacciari and Padovani (2007) showed reaction times were faster when the gender of the pronoun, matched the stereotyped gender of the role noun. Thus, although it is less

stable than semantic gender, the authors concluded stereotypical gender might be part of the nouns's mental representation. In addition, the results indicated when feminine stereotyped role nouns mismatched the pronouns (teacher-he), an inhibition effect was found in the response. However, no inhibition effect was found for masculine stereotyped nouns that mismatched the gender of the pronouns (engineer-she). The authors stated since masculine is unmarked in Italian, it might spread activation to conceptually masculine and feminine units, while the marked feminine would activate only the feminine unit.

Duff and Keir (2004) analyzed whether processing of individual words could be modified by discourse context. Their first goal was to show how gender stereotypes are accessed during comprehension and how they interfere in processing when stereotypes are violated. Their second goal was to investigate whether such interference effect could be reduced by discourse context. In Experiment 1, Duff and Keir (2004) monitored the participants' eye movements while reading sentences containing role nouns or occupation terms followed by reflexives, which could match or mismatch the stereotypical gender of the antecedents. In Experiment 2, the authors created a series of paragraphs focused on a particular character identified with a stereotyped role or occupation. In the disambiguating condition, the gender of the character was explicitly stated before the character was introduced informing the participants whether the character was a woman or a man. On the other hand, in the neutral condition, the gender of the character was not stated.

Duff and Keir (2004) found out in Experiment 1 and in the neutral condition of Experiment 2 reading processing suffered interference when text and the gender stereotyped information conflicted (mismatch-cost). This could possibly mean gender stereotypes are immediately activated by the time the role names were found, resulting in a conflict when the reflexives were subsequently encountered, or that gender information was not activated until the reflexives were encountered. Their study could not disentangle between these two possibilities. It should be highlighted there was no interference in the disambiguating condition in Experiment 2, that is, when prior context specified the gender of the role names. Duff and Keir (2004) concluded gender stereotypes are automatic, but they can be overridden, at least for a short time, by early explicit specification of gender in the discursive context.

### **3. Comparing stereotypical gender and definitional gender**

Kneiner, Sturt and Garrod (2008) were concerned with the nature of the stereotypical gender. According to the mental models approach, stereotypical gender is inferred from world knowledge, while according to the lexical view it is stored in the lexicon as part of the lexical representation. A good way to disentangle between these two proposals is comparing stereotypical gender with definitional gender. Both lexical and inferential approaches predict processing difficulty in mismatching conditions, that is, when stereotypical or definitional genders mismatch the gender of the reflexives. However, according to the lexical account, there may be some quantitative probabilistic differences between definitional and stereotypical nouns. Definitional gender is more polarized, or

even categorical, whereas stereotypical gender is more graded biased. This way, a larger mismatch-cost would be expected for definitional gender than for stereotypical gender. On the other hand, according to the inferential view, definitional gender and stereotypical gender are qualitative different, that is, the former might be specified in the lexicon, while the latter might be a result of probabilities based on world knowledge.

Kneiner *et al.* (2008) showed the results of two eye-tracking experiments differed in function of the order by which the gender information is given in discourse. In Experiment 1, the role nouns were presented earlier than the reflexives (anaphora), while in Experiment 2, the reflexives were presented earlier than the role nouns (cataphora). In Experiment 1, the mismatch-cost were similar for both stereotypical and definitional gender. However, in Experiment 2, the mismatch-cost was only found for definitional gender nouns. The authors concluded that, unlike definitional gender, stereotypical gender could be overridden by prior discourse. This way, there is a qualitative difference between definitional gender and stereotypical gender, which is evidence in favor of the inferential approach. The gender of definitional nouns might be lexically represented, while stereotypical gender might be pragmatically inferred. Finally, Kneiner *et al.* (2008) argued in favor of an interactive model in which lexical, pragmatic, and syntactic representations would efficiently communicate throughout processing allowing readers to incrementally integrate information from different linguistic levels.

Osterhout, Bersick and McLaughlin (1997) were concerned about the mental representation and processes underlying stereotype-violating reflexives compared to reflexives that are consistent with the stereotypes. In addition, they were concerned whether the stereotype violation in an ERP study would resemble to pragmatically implausible words (N400) or syntactically anomalous words (P600). In the second scenario, a P600 brain response would be similar to the brain response elicited by definitional male or female nouns whose grammatical features are associated with the word's lexical representation.

The materials used for reflexive antecedents in Osterhout *et al.* (1997) were nouns specifying occupations (as words like *actress*), states (as words like *bachelor*) and titles (as words like *duke*). The first type of nouns was gender stereotyped, while the latter two types were definitional male or female nouns. Half of the definitional nouns were male (as *king*) or female (as *queen*), and half of the stereotypical nouns were female stereotyped (as *babysitter*) or male stereotyped (as *pilot*).

Osterhout *et al.* (1997) reported violations of definitional gender evoked a larger P600 than violations of gender stereotypes. Moreover, the positive wave of P600 evoked for violations of gender stereotypes persisted even when the participants judged the sentences to be acceptable. The fact that both types of violations evoked P600 effects indicate the lexical representation of definitional and stereotypical nouns is similar, and they participate in the grammatical rules requiring agreement, that is, their gender information is encoded within grammar.

According to Osterhout *et al.* (1997), the different amplitudes of P600 between violations of definitional and stereotypical gender might suggest participants experienced more difficulties from recovering from a definitional gender violation than a stereotypical gender violation. In other words, violations with gender definitions result in an

unavoidable ungrammaticality, whereas violations with gender stereotypes result in reanalysis. It seems initially, readers assign the stereotypical reading, but by the time they encounter the reflexives and they realize their first reading was not adequate, they are ultimately forced to assign the less preferred gender feature.

Canal, Garnham and Oakhill (2015) aimed at replicating the results found by Osterhout *et al.* (1997). Unlike them, Canal *et al.* (2015) claimed the difference between definitional and stereotypical nouns is not quantitative. Instead of finding P600 effects in the mismatch conditions for both types of antecedents as Osterhout *et al.* (1997), Canal *et al.* (2015) found out that when the gender of the reflexives mismatched the stereotypical gender of the antecedents, the ERP results were biphasic showing negative effects in frontal left electrodes (Nref) and a positive effect in parietal electrodes (P600). It is worth mentioning that like Osterhout *et al.* (1997), they also found P600 effects when the gender of the reflexives mismatched the gender of the definitional antecedents. Based on other works, Canal *et al.* (2015) interpreted the Nref effect reflecting a search for additional information to link the reflexives to the stereotypical antecedents. In other words, as stereotypical gender does not determine gender categorically, but it is based on a probabilistic bias, comprehenders might need to look for additional information to realize antecedents and reflexives are coreferential. They explained the difference between their results and Osterhout *et al.* (1997)'s may be linked to a couple of factors: (a) Osterhout *et al.* (1997) used less electrodes, which might indicate they could have missed some effects; (b) the materials in Osterhout *et al.* (1997) were more stereotypically biased due to the use of modifiers; or (c) a social change - today's society is more liberal than 15 years ago.

#### 4. The present study

The present study aimed at investigating the differences between definitional gender and stereotypical gender in memory during antecedent retrieval in Brazilian Portuguese. By comparing these two types of genders, it would be possible to verify whether processing coreference dependencies with structurally unacceptable candidates carrying definitional gender would be more costly than processing sentences with structurally unacceptable candidates carrying stereotypical gender. It is hypothesized that antecedent candidates with definitional gender are preferable to be retrieved in memory than antecedent candidates with stereotypical gender. A reason for that would lie in the fact that definitional gender is lexically specified, while stereotypical gender is a result of probabilistic world knowledge inferences (Carreiras *et al.* 1996; Kneiner *et al.* 2008; Oakhill *et al.* 2005). Thus, if definitional gender were preferable than stereotypical gender, pronoun processing with structurally unacceptable antecedents carrying definitional gender would cause larger interference effects, because there would be more competition between them and structurally acceptable antecedents, and, consequently, antecedent retrievals would be slower.

Similarly, it is hypothesized that masculine antecedent candidates are more preferable than feminine antecedent candidates in memory. Masculine seems to weigh more than feminine because masculine is the default gender in languages like Portuguese

(Alves 2014, 2019; Cacciari & Padovani, 2007; Casado *et al.* 2017). Therefore, it is expected that masculine structurally unacceptable candidates would be responsible for slower antecedent retrievals as they would cause greater interference effects as a stronger competition between them and the structurally acceptable antecedents would occur.

Both Experiments 1a and 1b tested for structurally unacceptable antecedents with definitional and stereotypical genders; however, the former tested for those types of gender in the feminine and the latter in the masculine. It was necessary to divide the stimuli into two experiments due to the number of conditions.

#### 4.1. Participants

32 native speakers of Brazilian Portuguese (21 female and 11 male, mean age of 22 years) participated in Experiment 1a; and 36 (24 female and 12 male, mean age of 22 years) participated in Experiment 1b. The participants were undergraduate students with normal or corrected-to-normal vision. They were randomly selected to participate in the experiments as volunteers and they were not aware of the object of study of this research. They all signed a consent form giving permission to the experimenter to publish the results. The participants have not received any credits for their participation in the study; instead, they received 3 hours of Cultural-Scientific Activities (*Atividades-Científico-Culturais Discentes*, AACC) as compensation for their work.

#### 4.2. Materials and design

The experimental materials of each experiment consisted of 48 sentences distributed in 4 conditions. The experimental trials were arranged into 4 lists using a Latin Square. Each list was pseudo-randomized and contained 12 experimental items and 24 fillers. Each and every trial was followed by a comprehension question. The experimental trials were composed by embedded third person singular pronouns (*ele/ela*), whose antecedents in the main clause could be masculine or feminine common nouns. However, between the antecedents and the pronouns there were structurally unacceptable antecedents that could match or mismatch the third person pronouns in gender.

The materials followed the same structure [Subject DP1] [VP1 [Object DP2<sub>i</sub>]] [que [Subject t<sub>1</sub>] [VP2 [Object ele/ela] Adv PP]] and there was length control in the regions before and after the pronouns. The antecedent was in the subject position within DP1 and was composed by a definite article and a masculine/feminine common noun. Following DP1, there was a transitive verb within VP1, which also included the object DP2, composed by a definite article and a feminine/masculine definitional or stereotypical noun. After that, there was a subject relative clause introduced by the relative pronoun *que*. The relative pronoun was followed by VP2, which contained a transitive verb with 5-6 characters. The subject of the relative clause was the moved Object DP2 in the main clause and the object of relative clause was the personal pronoun *ele/ela*. After the pronoun there was a manner adverb ending in *-mente* from 9 to 11 characters followed by a PP. The personal pronouns *ele/ela* were used in the materials instead of the oblique pronouns *o/a*, because the latter are less salient and attached to the verb, which makes it

difficult to get them analyzed. Besides that, the oblique pronouns are very rare in colloquial Brazilian Portuguese nowadays.

It should be highlighted that the accessibility of the structurally acceptable antecedents and the distractors were quite balanced in the materials. On one hand, the structurally acceptable antecedents used in the materials would be the preferable antecedent candidates not only due to structural constraints related to Principle B, but also due to discursive factors. They were the subjects of the main clause, which means they were highly accessible, and they could be considered as the discursive topic. On the other hand, the structurally unacceptable antecedent candidates used in the materials were the subject of the embedded clause, and even though they were not as prominent as the structurally acceptable antecedents, they are also highly accessible in discourse due to its recentness.

The independent variables of the experiment were: a) *antecedent matching the gender of the pronoun*, which is a factor that is directly related to Principle B structural constraints; and b) *distractor matching the gender of the pronoun*, which is factor that relies purely on agreement cues. This way, the experimental design was 2x2, with four main conditions. There was also a control for the *distractor type of gender*; therefore, half of the experimental trials contained distractors with definitional gender and the other half contained distractors with stereotypical gender. It should be noticed that all distractors of Experiment 1a were feminine and all distractors of Experiment 1b were masculine. A sample of the materials of Experiment 1a can be found in Tables 2 and 3 and a sample of the materials of Experiment 1b can be found in Tables 4 and 5. Brackets delimit the region of interest, which was the pronoun. One may find a complete list of the materials in the Appendix.

**Table 1. Sample of materials for distractors with feminine definitional gender used in Experiment 1a**

	Antecedent mismatch	Antecedent match
Distractor mismatch	A <u>enfermeira</u> conhecia a <u>mulher</u> que matou [ <b>ele</b> ] brutalmente na frente da casa da família.  (The <sub>[fem]</sub> nurse <sub>[fem]</sub> knew the <u>woman</u> who brutally killed <b>him</b> in front of the family's house.)	O <u>enfermeiro</u> conhecia a <u>mulher</u> que matou [ <b>ele</b> ] brutalmente na frente da casa da família.  (The <sub>[masc]</sub> nurse <sub>[masc]</sub> knew the <u>woman</u> who brutally killed <b>him</b> in front of the family's house.)
Distractor match	O <u>enfermeiro</u> conhecia a <u>mulher</u> que matou [ <b>ela</b> ] brutalmente na frente da casa da família.  (The <sub>[masc]</sub> nurse <sub>[masc]</sub> knew the <u>woman</u> who brutally killed <b>her</b> in front of the family's house.)	A <u>enfermeira</u> conhecia a <u>mulher</u> que matou [ <b>ela</b> ] brutalmente na frente da casa da família.  (The <sub>[fem]</sub> nurse <sub>[fem]</sub> knew the <u>woman</u> who brutally killed <b>her</b> in front of the family's house.)

**Table 2. Sample of the materials for distractors with feminine stereotypical gender used in Experiment 1a**

	Antecedent mismatch	Antecedent match
Distractor mismatch	<p>A <u>bibliotecária</u> seguiu a <u>recepcionista</u> que guiou [<b>ele</b>] brevemente através do corredor do grande gabinete real.</p> <p><i>(The<sub>[fem]</sub> <u>librarian</u><sub>[fem]</sub> followed the<sub>[fem]</sub> <u>receptionist</u> who briefly guided <b>him</b> through the hallway of the large royal office.)</i></p>	<p>O <u>bibliotecário</u> seguiu a <u>recepcionista</u> que guiou [<b>ele</b>] brevemente através do corredor do grande gabinete real.</p> <p><i>(The<sub>[masc]</sub> <u>librarian</u><sub>[masc]</sub> followed the<sub>[fem]</sub> <u>receptionist</u> who briefly guided <b>him</b> through the hallway of the large royal office.)</i></p>
Distractor match	<p>O <u>bibliotecário</u> seguiu a <u>recepcionista</u> que guiou [<b>ela</b>] brevemente através do corredor do grande gabinete real.</p> <p><i>(The<sub>[masc]</sub> <u>librarian</u><sub>[masc]</sub> followed the<sub>[fem]</sub> <u>receptionist</u> who briefly guided <b>her</b> through the hallway of large royal office.)</i></p>	<p>A <u>bibliotecária</u> seguiu a <u>recepcionista</u> que guiou [<b>ela</b>] brevemente através do corredor do grande gabinete real.</p> <p><i>(The<sub>[fem]</sub> <u>librarian</u><sub>[fem]</sub> followed the<sub>[fem]</sub> <u>receptionist</u> who briefly guided <b>her</b> through the hallway of the large royal office.)</i></p>

**Table 3. Sample of the materials for distractors with masculine definitional gender used in Experiment 1b**

	Antecedent mismatch	Antecedent match
Distractor mismatch	<p>O <u>advogado</u> hostilizou o <u>rei</u> que tratou [<b>ela</b>] rudemente na frente de alguns convidados na festa.</p> <p><i>(The<sub>[masc]</sub> <u>lawyer</u><sub>[masc]</sub> antagonized the <u>king</u> who treated <b>her</b> rudely in front of some guests at the party.)</i></p>	<p>A <u>advogada</u> hostilizou o <u>rei</u> que tratou [<b>ela</b>] rudemente na frente de alguns convidados na festa.</p> <p><i>(The<sub>[fem]</sub> <u>lawyer</u><sub>[fem]</sub> antagonized the <u>king</u> who treated <b>her</b> rudely in front of some guests at the party.)</i></p>
Distractor match	<p>A <u>advogada</u> hostilizou o <u>rei</u> que tratou [<b>ele</b>] rudemente na frente de alguns convidados na festa.</p> <p><i>(The<sub>[fem]</sub> <u>lawyer</u><sub>[fem]</sub> antagonized the <u>king</u> who treated <b>him</b> rudely in front of some guests at the party.)</i></p>	<p>O <u>advogado</u> hostilizou o <u>rei</u> que tratou [<b>ele</b>] rudemente na frente de alguns convidados na festa.</p> <p><i>(The<sub>[masc]</sub> <u>lawyer</u><sub>[masc]</sub> antagonized the <u>king</u> who treated <b>him</b> rudely in front of some guests at the party.)</i></p>

**Table 4. Sample of the materials for distractors with masculine stereotypical gender used in Experiment 1b**

	Antecedent mismatch	Antecedent match
Distractor mismatch	<p>O <u>cabeleireiro</u> detestou o <u>piloto de corrida</u> que olhou [ela] lentamente dos pés a cabeça antes da entrevista no camarim.</p> <p>(The<sub>[masc]</sub> <u>hairdresser</u><sub>[mas]</sub> disliked the<sub>[masc]</sub> <u>race car driver</u> who slowly stared at <b>her</b> in the dressing room before the interview.)</p>	<p>A <u>cabeleireira</u> detestou o <u>piloto de corrida</u> que olhou [ela] lentamente dos pés a cabeça antes da entrevista no camarim.</p> <p>(The<sub>[fem]</sub> <u>hairdresser</u><sub>[fem]</sub> disliked the<sub>[masc]</sub> <u>race car driver</u> who slowly stared at <b>her</b> in the dressing room before the interview.)</p>
Distractor match	<p>A <u>cabeleireira</u> detestou o <u>piloto de corrida</u> que olhou [ele] lentamente dos pés a cabeça antes da entrevista no camarim.</p> <p>(The<sub>[fem]</sub> <u>hairdresser</u><sub>[fem]</sub> disliked the<sub>[masc]</sub> <u>race car driver</u> who slowly stared at <b>him</b> in the dressing room before the interview.)</p>	<p>O <u>cabeleireiro</u> detestou o <u>piloto de corrida</u> que olhou [ele] lentamente dos pés a cabeça antes da entrevista no camarim.</p> <p>(The<sub>[masc]</sub> <u>hairdresser</u><sub>[masc]</sub> disliked the<sub>[masc]</sub> <u>race car driver</u> who slowly stared at <b>him</b> in the dressing room before the interview.)</p>

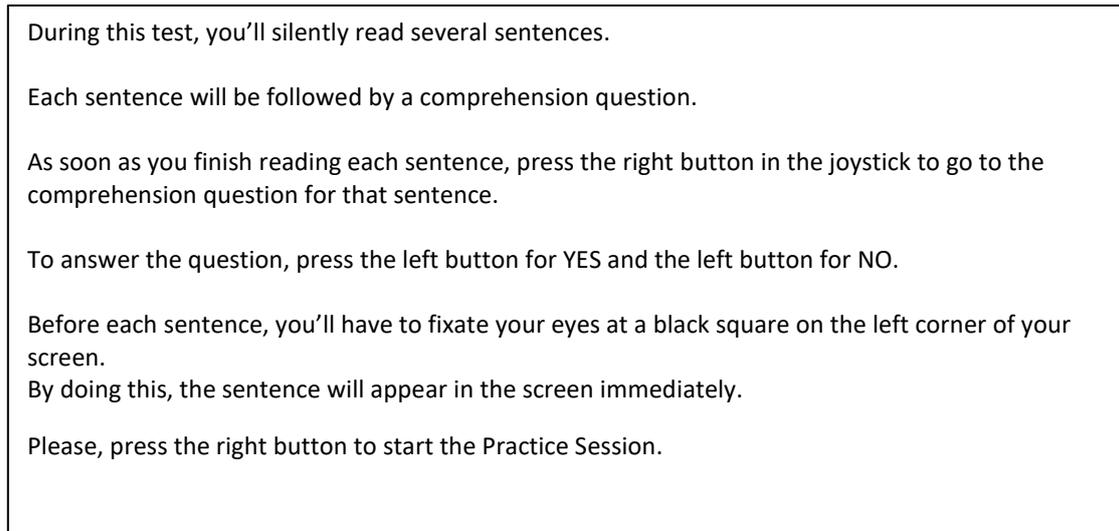
The dependent variables were First Fixation Duration, First Pass, and Regression Path at the pronoun region. First Fixation corresponds to the very beginning of reading processing, and it measures the duration of the first eye fixation at a certain region in the text. First Pass also captures early processing phases as it measures the sum of all eye fixations at a certain region of the text before our eyes move to the right or left. Finally, Regression Path, or Go-Past Time, captures late processing phases. It corresponds to the duration of all fixations from first fixation at a region to first moving to the right – including regressions back to earlier parts of the sentence before moving on.

### 4.3. Procedure

The experiment was conducted in a psycholinguistics laboratory (Laboratório de Psicolinguística Experimental – LAPEX) in Rio de Janeiro, Brazil. The eye-tracker used in this experiment was *Eye Link 1000* and the experiment was programmed and conducted on *Eye Track 7.10*<sup>1</sup> software. All trials were typed in font *Monaco* size 12. The participants were instructed to seat comfortable and were given written and oral instructions. The instructions screen is illustrated in Figure 1.<sup>2</sup>

<sup>1</sup> The primary developers of *Eye Track* were David Stracuzzi and Jeff Kinsey and it is conceptually based on software written by Saarbruken and provided to UMASS by Christoph Scheepers. *Eye Track* can be downloaded for free on <https://blogs.umass.edu/eyelab/software/>.

<sup>2</sup> The participants received the instructions in Portuguese; the translation to English is only for the purpose of this paper.



**Figure 1. Instructions screen of Experiments 1a and 1b**

After receiving the instructions, the calibration process would start followed by a short practice with 6 filler sentences so that the experimenter would check whether the participants understood the task and were performing it at a natural speed. The experiment duration was of 20 minutes approximately.

#### **4.4. Analysis**

The eye-tracking data was analyzed using the following tools: Visual EDF to ASC, to convert the .EDF files that Eye Link 1000 generates; *Robodoc.py*<sup>3</sup>, to clean eye blinks and long saccades (longer than 80ms); *Question\_acc.py*<sup>4</sup> to compute the comprehension questions accuracy; *EyeDry*<sup>5</sup> to compute the reading measures; and R for the data exploration and statistical analysis.

Some experimental trials had to be excluded due to eye blinks and long saccades at the regions of interest (22% in Experiment 1a, and 15% in Experiment 1b).

As Experiment 1a and Experiment 1b contained the same materials, except for the distractors, the data was analyzed together in a between-subjects analysis. The observations were normalized with log-transformations and the means were centered in order to minimize the effect of data variability and outliers. After that, linear mixed effects models (LMEs) were created with the help of *lmerTest*<sup>6</sup> package in order to analyze the role of each independent variable in the results. The fixed effects of the lmes were: a) antecedent matching the gender of the pronoun (match/mismatch); b) distractor matching the gender of the pronoun (match/mismatch); c) distractor type of gender

<sup>3</sup> *Rododoc.py* is a python script created by Adrian Staub and Chuck Clifton, and the 2016 version was revised by Jesse Harris. It can also be downloaded on <https://blogs.umass.edu/eyelab/software/>.

<sup>4</sup> *Question\_acc.py* is a python script that comes with *Robodoc.py* utils to check questions accuracy and their reaction times.

<sup>5</sup> *EyeDry* was created by Chuck Clifton and can be downloaded on <https://blogs.umass.edu/eyelab/software/>.

<sup>6</sup> Alexandra Kuznetsova, Per Bruun Brockhoff and Rune Haubo Bojesen Christensen (2015). *lmerTest: Tests in Linear Mixed Effects Models*. R package version 2.0-29 on <https://CRAN.R-project.org/package=lmerTest>.

(definitional/stereotypical); and d) distractor gender (masculine/feminine). On the other hand, the random effects were the participants and the items. LME analysis was chosen as it verifies the influence of the independent variables as well as any “noise” influence such as items peculiarities and participants differences.

Full random LME models (with random slopes for each of the random effects) were created to analyze the data. For space reasons, only the statistically significant results were reported in the text. One may check the Appendix for a complete list of the LME results. After running the LME models, post-hoc tests were conducted in order to figure out which pairwise comparisons among the conditions of the experiments were statistically significant. Differences between the least squared means (lsmeans), which are predicted marginal means of LME models, were estimated with the help of emmeans package.<sup>7</sup>

Aiming to clarify the statistically significant results, histograms with error bars were created using ggplot2 package.<sup>8</sup> In order to facilitate the comprehension of the graphs, all statistically significant pairwise comparisons were indicated in the figures with “\*\*”.

## 4.5. Results

It should be mentioned that the participants answered the comprehension questions with an average of accuracy of 90% in Experiment 1a and 89% in Experiment 1b, which means that the participants were paying attention to the task and reading the sentences properly.

Results for the eye-tracking measures, chosen as dependent variables – First Fixation, First Pass, and Regression Path – will be reported and discussed below.

### 4.5.1. First Fixation

The results of the LME for First Fixation at the pronoun region can be seen in Table 5.

**Table 5. LME results for First Fixation at the pronoun region.**

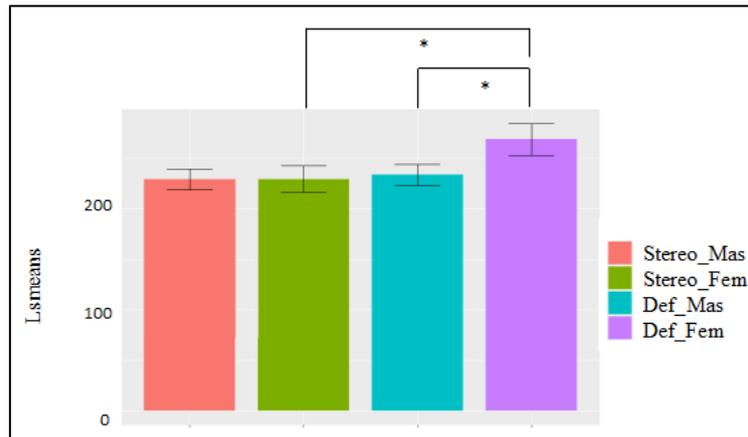
Parameter	Pronoun			
	$\beta_0$	SE	t-value	p-value
(Intercept)	0.11	0.23	0.48	0.63
Type of gender x Gender	1.20	0.41	2.90	0.004
Antecedent x Type of gender x Gender	-1.15	0.57	2.00	0.046

In post-hoc tests, the interaction *Type of gender x Gender* was found statistically significant in two pairwise comparisons. The differences between the lsmeans ( $\beta_0=39.26$ ,  $SE=16.53$ ,  $t\text{-value}=2.37$ ,  $p=0.019$ ) indicated that reading times at the pronoun region were longer in sentences containing distractors with feminine definitional gender ( $lsmean=269$ ,  $SE=15.5$ ) than in sentences containing distractors with feminine stereotypical gender

<sup>7</sup> Russel Lenth (2019). emmeans: estimated marginal means, aka Least-Squares Means. R package version 1.4. <https://CRAN.R-project.org/package=emmeans>.

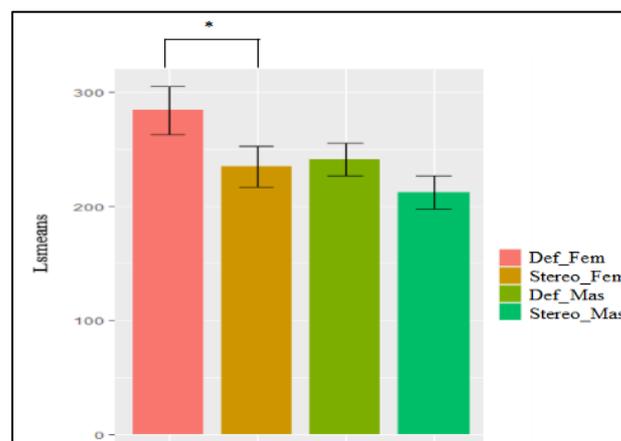
<sup>8</sup> H.Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

( $l_s$ mean=229, SE=13.3). Besides that, the differences between the  $l_s$ means ( $\beta_0=34.95$ , SE=16.81,  $t$ -value=2.07,  $p=0.043$ ) indicated that reading times at the pronoun region were longer in sentences containing distractors with feminine definitional gender ( $l_s$ mean=269, SE=15.5) than in sentences containing masculine definitional gender ( $l_s$ mean=234, SE=10.1). Figure 2 illustrates the interaction *Type of gender x Gender* for First Fixation Duration at the pronoun region.



**Figure 2.** First Fixation at the pronoun region by the type of gender and gender of the distractors.

The post-hoc tests of the interaction *Antecedent x Type of gender x Gender* was only statistically significant in one pairwise comparison. The differences between the  $l_s$ means ( $\beta_0=49.12$ , SE=23.83,  $t$ -value=2.06,  $p=0.044$ ) indicated that sentences in which the antecedents mismatched the pronouns, reading times at the pronoun region were longer in sentences containing distractors with feminine definitional gender ( $l_s$ mean=284, SE=20.9) than in sentences containing distractors with feminine stereotypical gender ( $l_s$ mean=235, SE=17.9). Figure 3 illustrates the interaction [*Mismatching*] *Antecedent x Type of gender x Gender* for First Fixation Duration at the pronoun region.



**Figure 3.** First Fixation at the pronoun region by mismatching antecedents, type of gender and gender.

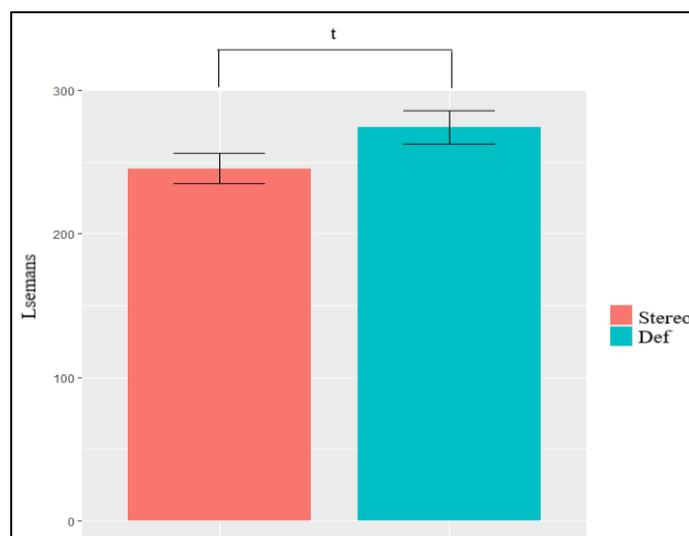
#### 4.5.2. First Pass

The results of the LME for First Pass at the pronoun region can be seen in Table 6.

**Table 6. LME results for First Pass at the pronoun region.**

Parameter	Pronoun			
	$\beta_0$	SE	t-value	p-value
(Intercept)	0.10	0.21	0.48	0.63
Type of gender	-0.69	0.30	-2.24	0.026
Type of gender x Gender	1.10	0.39	2.80	0.005

The post-hoc test of the main effect of *Type of Gender* showed a trend towards statistical significance regarding the difference between stereotypical gender and definitional gender ( $\beta_0=28.57$ ,  $SE=14.57$ ,  $t\text{-value}=1.96$ ,  $p=0.068$ ). Reading times at the pronoun were longer in sentences containing distractors with definitional gender ( $l\text{mean}=274$ ,  $SE=11.8$ ) than in sentences containing distractors with stereotypical gender ( $l\text{mean}=246$ ,  $SE=10.8$ ). Figure 4 illustrates the main effect of *Type of Gender* for First Pass Duration at the pronoun region.



**Figure 4. First Pass at the pronoun region by the type of gender of the distractors.**

The interaction *Type of gender x Gender* was found statistically significant in the two pairwise comparisons. The differences between the  $l\text{means}$  ( $\beta_0=52.39$ ,  $SE=21.14$ ,  $t\text{-value}=2.47$ ,  $p=0.02$ ) indicated that reading times at the pronoun region were longer in sentences containing distractors with definitional feminine gender ( $l\text{mean}=293$ ,  $SE=18.56$ ) than in sentences containing distractors with feminine stereotypical gender ( $l\text{mean}=241$ ,  $SE=15.6$ ). Besides that, the differences between the  $l\text{means}$  ( $\beta_0=37.96$ ,  $SE=18.74$ ,  $t\text{-value}=2.02$ ,  $p=0.047$ ) indicated that reading times at the pronoun region were longer in sentences containing distractors with feminine definitional gender ( $l\text{mean}=293$ ,  $SE=18.1$ ) than in sentences containing distractors with masculine definitional gender

(lsmean=255, SE=12.3). Figure 5 illustrates the interaction *Type of gender x Gender* for First Pass Duration at the pronoun region.

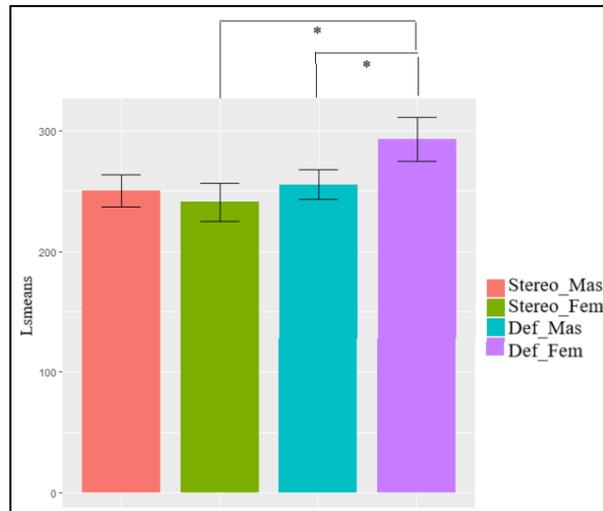


Figure 5. First Pass at the pronoun region by the type of gender and gender of the distractors.

#### 4.5.3. Regression Path

The results of the LME for Regression at the pronoun region can be seen in Table 7.

Table 7. LME results for Regression Path at the pronoun region.

Pronoun				
Parameter	$\beta_0$	SE	t-value	p-value
(Intercept)	-0.44	0.25	-1.73	0.101
Antecedent	0.89	0.33	2.65	0.009
Antecedent x Distractor	-0.99	0.42	-2.33	0.020
Antecedent x Gender	-0.97	0.39	-2.45	0.015
Distractor x Gender	-1.01	0.37	-2.70	0.007
Antecedent x Distractor x Gender	1.62	0.53	3.02	0.002

The post-hoc test of the main effect of *Antecedent* showed that the difference between the lsmeans ( $\beta_0=-146.66$ ,  $SE=57.49$ ,  $t\text{-value}=-2.55$ ,  $p=0.018$ ) indicated that reading times at the pronoun region were faster when the antecedents and pronouns matched in gender (lsmean=375,  $SE=38.6$ ) than when they mismatched (lsmean=521,  $SE=64.5$ ). Figure 6 illustrates the main effect of *Antecedent*.

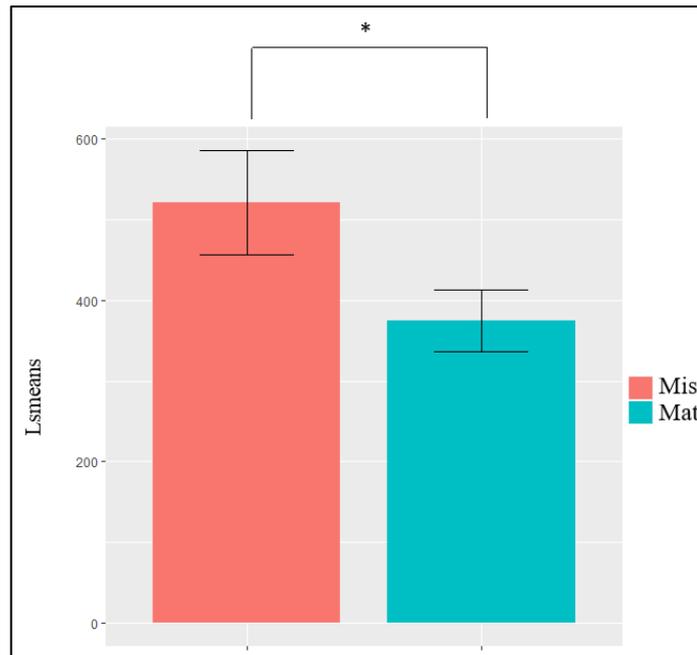


Figure 6. Regression Path at the pronoun region by antecedents.

The post-hoc tests of the interaction *Antecedent x Distractor* showed that the differences between the lsmeans ( $\beta_0 = -151.03$ ,  $SE = 72.65$ ,  $t\text{-value} = -2.07$ ,  $p = 0.042$ ) indicated that when distractors matched the pronouns in gender, reading times at the pronoun region in sentences in which the antecedents also matched the pronouns (lsmean=347,  $SE = 55.9$ ) were faster than in sentences in which the antecedents mismatched the pronouns in gender (lsmean=498,  $SE = 78.3$ ). Figure 7 illustrates the interaction *Antecedent x Distractor*.

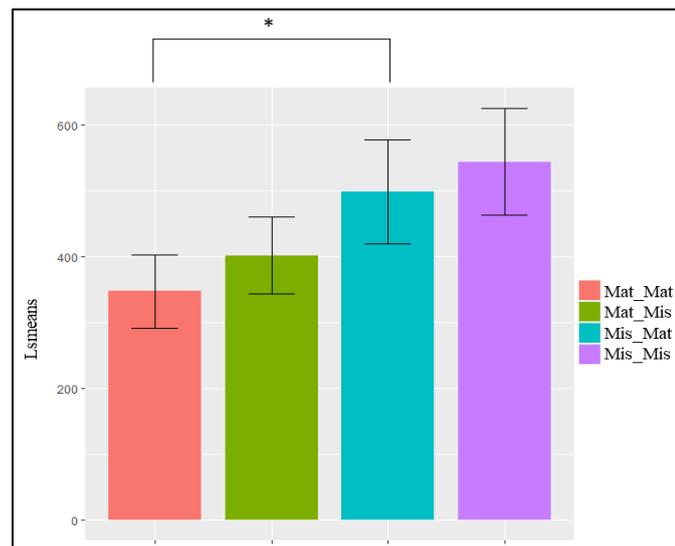
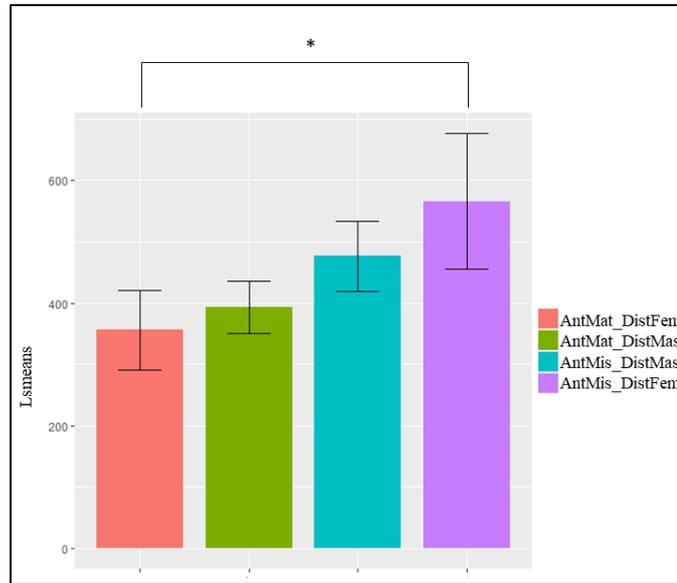


Figure 7. Regression Path at the pronoun region by antecedents and distractors.

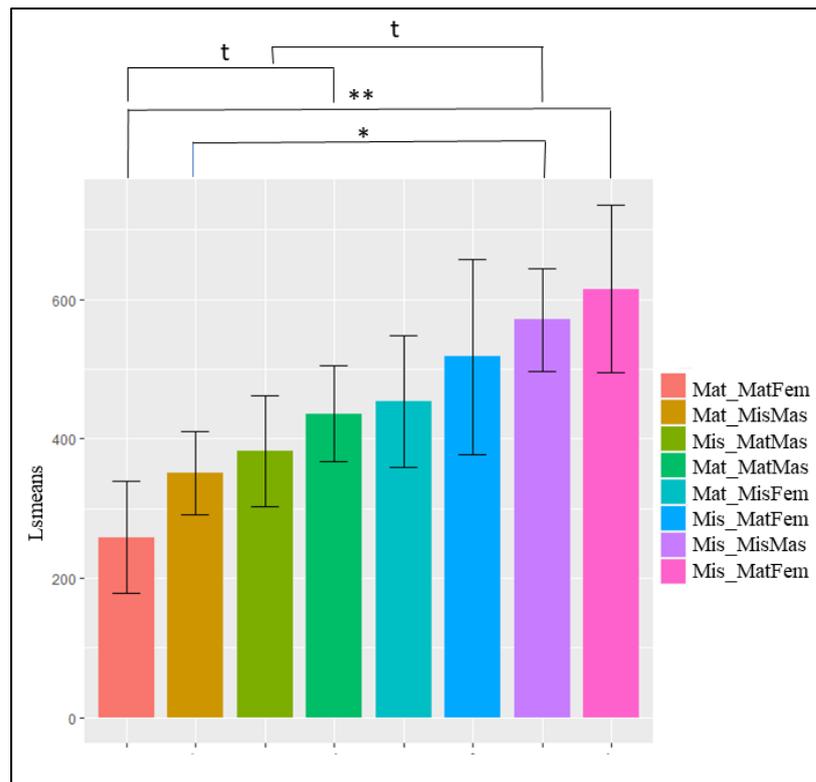
Post-hoc tests of the interaction *Antecedent x Gender* showed that the differences between the lsmeans ( $\beta_0 = -210.3$ ,  $SE = 91.74$ ,  $t\text{-value} = -2.29$ ,  $p = 0.030$ ) indicated that in sentences with feminine distractors, reading times at the pronoun region were faster when antecedents matched the pronoun in gender (lsmean=356,  $SE = 64.7$ ) than when they

mismatched (lsmean=566, SE=110.7). Figure 8 illustrates the interaction *Antecedent x Gender*.



**Figure 8.** Regression Path at the pronoun region by antecedents and distractors gender.

Post-hoc tests of the interaction *Antecedent x Distractor x Gender* indicated two statistically significant pairwise comparisons. The differences between the lsmeans ( $\beta_0 = -356.01$ ,  $SE = 113.9$ ,  $t\text{-value} = -3.12$ ,  $p = 0.0027$ ) indicated that when distractors matched the pronouns in feminine gender, reading times at the pronoun region were faster when antecedents also matched the pronouns in gender (lsmean=259, SE=79.6) than when they mismatched (lsmean=615, SE=120.4). The differences between the lsmeans ( $\beta_0 = -219.98$ ,  $SE = 85.7$ ,  $t\text{-value} = -2.56$ ,  $p = 0.011$ ) also indicated that when masculine distractors mismatched the pronouns, reading times at the pronouns region were faster when antecedents matched the pronouns in gender (lsmean=351, SE=60.2) than when they mismatched (lsmean=571, SE=73.5). Besides that, two pairwise comparisons showed a trend towards significance. The differences between the lsmeans ( $\beta_0 = -177.1$ ,  $SE = 92.61$ ,  $t\text{-value} = -1.91$ ,  $p = 0.059$ ) showed that when both antecedents and distractors matched the pronouns in feminine (lsmeans=259, SE=79.6), reading times at the pronoun region tended to be faster than when they matched in masculine (lsmeans=436, SE=68.5). The differences between the lsmeans ( $\beta_0 = -188.76$ ,  $SE = 98.88$ ,  $t\text{-value} = -1.90$ ,  $p = 0.063$ ) also showed that when antecedents mismatched the pronouns in gender, reading times at the pronoun region tended to be faster when masculine distractors matched the pronouns (lsmean=382, SE=79.7) than when they mismatched (lsmean=571, SE=73.5). Figure 9 illustrates the *Antecedent x Distractor x Gender* interaction.



**Figure 9. Regression Path at the pronoun region by antecedents and distractors and gender.**

## 5. Discussion

The results reported in this paper corroborates Alves (2019), that is, Principle B structural constraints do not work as an initial filter in Brazilian Portuguese since structurally unacceptable antecedent candidates interfered in pronoun resolution processing at very early processing phases as shown by First Fixation and First Pass eye-tracking measures. More than that, it seems that Principle B structural constraints only play a role at late processing phases as shown by Regression Path results. This way, it seems that in the beginning of processing all candidates are taken into account by memory as potential candidates; and at a later stage, Principle B structural constraints help the parser to select the correct candidate.

As predicted, definitional gender weighs more in memory than stereotypical gender since structurally unacceptable candidates with definitional gender caused more inhibitory effects than structurally unacceptable candidates with stereotypical gender. Candidates with definitional gender are more preferable to be retrieved, increasing the competition between them and the structurally acceptable candidates, resulting in longer processing.

With respect to masculine, the case is more complex since both facilitatory and inhibitory effects were found. At early processing stages, as shown by First Fixation and First Pass, the facilitatory effect of masculine distractors when compared to feminine distractors might indicate the former are being misretrieved by memory (Alves 2019; Engelmann *et al.* 2015). On the other hand, at late processing phases, as shown by Regression Path (although the results only show a trend towards statistical significance),

when the structurally acceptable antecedents agree with the pronouns, pronoun resolution tends to take longer with masculine distractors than with feminine distractors. This inhibitory effect happens because the competition between structurally acceptable antecedents and masculine distractors is greater compared to the competition between structurally acceptable antecedents and feminine distractors. This means that masculine structurally unacceptable antecedents tend to interfere in pronoun resolution even when there is already a structurally acceptable antecedent available in the sentence. In addition, when structurally unacceptable antecedents do not agree with the pronouns, masculine distractors that agree with the pronouns are responsible for faster pronoun resolutions than masculine distractors that do not agree with the pronouns. This seems to be another case of masculine structurally unacceptable antecedent misretrieval. Causing either facilitatory or inhibitory effects, masculine seems to be more preferable to be retrieved by memory than feminine, indicating that masculine weighs more in memory than feminine since it is the default gender in Brazilian Portuguese (Alves 2014, 2019; Cacciari & Padovani, 2007; Casado *et al.* 2017).

Finally, the results found in Regression Path indicate that Principle B is relevant to pronoun resolution in Brazilian Portuguese since there were longer reading times at the pronoun when structurally acceptable antecedents mismatched the pronouns in gender.

Our results show that morphological cues as well as Principle B structural constraints are very important in antecedent retrieval in Brazilian Portuguese. However, the results reported here indicate that gender cues play a role before Principle B structural constraints in the time course of pronoun resolution processing. This means that memory tends to retrieve more easily masculine candidates and candidates with definitional gender. Masculine weighs more in memory because it is the default gender, while definitional gender might be lexically specified unlike stereotypical gender. Stereotypical gender seems to depend on probabilities based on world knowledge inferences, which might be more cognitively costly.

The results reported here seemed to contradict Carreiras *et al.* (1996), Duff and Keir (2004), and Kneiner *et al.* (2008). The definite articles before the gender stereotyped distractors in our materials were not enough to override stereotyped inferences as it was clear in our results that sentences with distractors with stereotypical gender and definitional gender were processed differently. In other words, stereotypical gender inferences are strong and cannot be suppressed even in cases where they would not be needed to comprehend and process a sentence, as in our materials. This corroborates the hypothesis that stereotypical gender is automatic and computed as soon as the word is encountered, and not only when this information is needed (Oakhill *et al.* 2005).

Contradicting the lexical approach (Cacciari & Padovani, 2007; Kennison & Trofe, 2003; Osterhout *et al.* 1997) this paper corroborates the inferential approach suggesting that stereotypical gender seems to be qualitatively different from definitional gender (Canal *et al.* 2015; Carreiras *et al.* 1996; Kneiner *et al.* 2008; Oakhill *et al.* 2005). In other words, while definitional gender might be lexically specified, as an idiosyncrasy, stereotypical gender might be a result of probabilities based on world knowledge inferences. This way, stereotypical gender is more cognitively demanding than

definitional gender, which explains why items with definitional gender are more preferable to be retrieved by memory.

The conclusion is that this paper showed evidence that language processing relies on both morphological (gender) and syntactic information, for instance. Corroborating Alves (2019), it seems that Principle B structural constraints do not work as an initial filter in pronoun resolution in Brazilian Portuguese; and that stereotypical gender does not seem to be lexically specified, but it is a result of world knowledge inferences. This means that language processing makes use of both linguistic and non-linguistic information. However, it seems that not all information weighs the same; memory retrieves more easily items that carry default linguistic information as it seems to be the case of masculine and definitional gender in opposition to feminine and stereotypical gender.

**Acknowledgements:** I would like to thank the volunteers who participated in the experiments reported here. I also would like to thank the valuable comments and suggestions given by the anonymous reviewers. This research was financially supported by *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), 140615/2014-8.

## References

- Alves, M. C. dos S. (2014). *Processamento do traço de gênero na correferência pronominal com antecedentes sobrecomuns e comuns de dois gêneros no português do Brasil* (Dissertação de mestrado, UFRJ, Rio de Janeiro).
- Alves, M. C. dos S. (2019). As diferenças entre gênero gramatical e gênero semântico na recuperação de antecedentes pronominais em Português Brasileiro. *Diacrítica*, 33(2), 89–115. <https://doi.org/10.21814/diacritica.403>
- Arnold, J., Eisenband, J., Brown-Schmidt, S., & Trueswell, J. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76(1), 13–26. [https://doi.org/10.1016/s0010-0277\(00\)00073-1](https://doi.org/10.1016/s0010-0277(00)00073-1)
- Badecker, W. & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphora. *Journal of Experimental Psychology Learning Memory and Cognition*, 28, 748–769. <https://doi.org/10.1037/0278-7393.28.4.748>
- Cacciari, C., & Padovani, R. (2007). Further evidence of gender stereotype priming in Language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics*, 28, 277–293. <https://doi.org/10.1017/S0142716407070142>
- Canal, P., Garnham, A., & Oakhill, J. (2015). Beyond gender stereotypes in language comprehension: self sex-role descriptions affect the brain's potentials associated with agreement processing. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01953>
- Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology*, 49A(3), 639–663. <https://doi.org/10.1080/713755647>
- Casado, A., Palma, A., & Paolieri, D. (2017). The influence of sex information on gender word processing. *Journal of Psycholinguistic Research*, 47(3), 1–27. <https://doi.org/10.1007/s10936-017-9546-3>
- Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale, S. J. Keyser (Eds.). *The view from building 20* (pp. 1–52). Cambridge: The MIT Press. <https://doi.org/10.1007/s10936-017-9546-3>

- Chomsky, N. (1995) *The Minimalist Program*. Cambridge: The MIT Press.
- Duff, S., & Keir, J. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory and Cognition*, 32(4), 551–559. <https://doi.org/10.3758/BF03195846>
- Engelmann, F., Jäger, L., & Vasishth, S. (2015). *The determinants of retrieval interference in dependency resolution: Review and computational modeling*. Unpublished manuscript.
- Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: active versus passive representations. *Journal of Memory and Language*, 56(3), 357–383. <https://psycnet.apa.org/doi/10.1016/j.jml.2006.07.004>
- Gordon, P., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science*, 22(4), 389–424. [https://doi.org/10.1016/S0364-0213\(99\)80045-7](https://doi.org/10.1016/S0364-0213(99)80045-7)
- Greene, S. B., McKoon, G., & Ratclif, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18(2), 266–283. <https://doi.org/10.1037//0278-7393.18.2.266>
- Grosz, B., Weinstein, S., & Joshi, A. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2), 203–225.
- Kehler, A. (2007). Rethinking the SMASH approach to pronoun interpretation. In J. Gundel & N. Hedberg (Eds.), *Interdisciplinary perspectives on reference processing* (pp. 95–122). Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195331639.001.0001>
- Kneiner, H., Sturt, P., & Garrod, S. (2008). Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language*, 58(2), 239–261. <https://doi.org/10.1016/j.jml.2007.09.003>
- Leitão, M., Peixoto, P., & Santos, S. (2008). Processamento da co-referência intra-sentencial em português brasileiro. *Veredas*, 12(2), 50–61.
- Nicol, J., & Swinney, D.A. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistics Research*, 18(1), 5–19. <https://doi.org/10.1007/BF01069043>
- Oakhill, J., Garnham, A., & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory and Cognition*, 33(6), 972–983. <https://doi.org/10.3758/BF03193206>
- Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory and Cognition*, 25(3), 273–285. <https://doi.org/10.3758/BF03211283>
- Patil, U., Vasishth, S., & Lewis, R. (2016). Retrieval interference in syntactic processing: The case of reflexive binding in English. *Frontiers of Psychology*, 7, 329. <https://doi.org/10.3389/fpsyg.2016.00329>
- Rigalleau, F., Caplan, D, Baudiffier, V. (2004). New arguments in favor of an automatic gender pronominal process. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 57(5), 893–933. <https://doi.org/10.1080/02724980343000549>
- Sturt, P. (2003). The time-course of the application of binding constraints in reference. *Journal of Memory and Language*, 48(3), 542–562. [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3)

## Appendix

### List of the experimental sentences of Experiment 1a

#### *Distractors with definitional gender*

a) O veterinário/A veterinária reconheceu a nora que feriu ele/ela fortemente por trás da cabeça no momento do assalto.

O ferimento por trás da cabeça foi no momento do assalto?

*The<sub>[MASC]</sub>vet<sub>[MASC]</sub>/The<sub>[FEM]</sub>vet<sub>[FEM]</sub> recognized the daughter-in-law who forcefully hit him/her behind the head at the moment of the assault.*

*Was the wound behind the head at the moment of the assault?*

b) O comissário de bordo/A comissária de bordo processou a moça que xingou ele/ela ferozmente devido ao um mal entendido.

A moça xingou devido ao um mal entendido?

*The<sub>[MASC]</sub>flight attendant<sub>[MASC]</sub>/The<sub>[FEM]</sub>flight attendant<sub>[FEM]</sub> sued the young lady who fiercely cursed him/her due to a misunderstanding.*

*Did the young lady curse due to a misunderstanding?*

c) O pedagogo/A pedagoga elogiou a mãe que tocou ele/ela totalmente com uma das lições mais belas de coragem.

A mãe deu uma bela lição de coragem?

*The<sub>[MASC]</sub>pedagogue<sub>[MASC]</sub>/The<sub>[FEM]</sub>pedagogue<sub>[FEM]</sub> complimented the mother who completely touched him/her with one of the most beautiful lessons of courage.*

*Did the mother give a beautiful lesson of courage?*

d) O advogado/A advogada hostilizou a rainha que tratou ele/ela rudemente na frente de alguns convidados na festa.

A rainha foi rude no café-da-manhã?

*The<sub>[MASC]</sub>lawyer<sub>[MASC]</sub>/The<sub>[FEM]</sub>lawyer<sub>[FEM]</sub> antagonized the queen who treated him/her rudely in front of some of guests at the party.*

*Was the queen rude at the breakfast?*

e) O criminoso/A criminosa ameaçou a dama que notou ele/ ela de repente próximo ao prédio alguns dias depois do crime.

A dama foi assassinada?

*The<sub>[MASC]</sub>criminal<sub>[MASC]</sub>/The<sub>[FEM]</sub>criminal<sub>[FEM]</sub> threatened the lady who suddenly noticed him/her near the building some days after the crime.*

*Was the lady murdered?*

f) O enfermeiro/A enfermeira conhecia a mulher que matou ele/ela brutalmente na frente da casa da família.

O crime aconteceu na frente do shopping?

*The<sub>[MASC]</sub> nurse<sub>[MASC]</sub>/The<sub>[FEM]</sub> nurse<sub>[FEM]</sub> knew the woman who brutally killed him/her in front of the family's house.*

*Did the crime happen in front of the shopping mall?*

*Distractors with stereotypical gender*

a) O perito/A perita acatou a assistente social que chamou ele/ela novamente por volta do meio dia no escritório.

A assistente social chamou alguém por volta do meio-dia no escritório?

*The<sub>[MASC]</sub> investigator<sub>[MASC]</sub>/The<sub>[FEM]</sub> investigator<sub>[FEM]</sub> answered the<sub>[FEM]</sub> social worker who called him/her again around noon at the office.*

*Did the<sub>[FEM]</sub> social worker call someone around noon at the office?*

b) O cenógrafo/A cenógrafa despediu a modelo que ofendeu ele/ela puramente por causa de fofocas nos bastidores.

Houve fofocas nos bastidores?

*The<sub>[MASC]</sub> set designer<sub>[MASC]</sub>/The<sub>[FEM]</sub> set designer<sub>[FEM]</sub> fired the<sub>[FEM]</sub> model who solely offended him/her due to gossips in the side scenes.*

*Were there gossips in the side scenes?*

c) O bibliotecário/A bibliotecária seguiu a recepcionista que guiou ele/ela brevemente através do corredor do grande gabinete real.

Há um corredor no grande gabinete real?

*The<sub>[MASC]</sub> librarian<sub>[MASC]</sub>/The<sub>[FEM]</sub> librarian<sub>[FEM]</sub> followed the<sub>[FEM]</sub> receptionist who briefly guided him/her through the hallway of the large royal office.*

*Is there a hallway in the large royal office?*

d) O cozinheiro/A cozinheira agradeceu a nutricionista que ajudou ele/ela rapidamente no meio da confusão daquele restaurante.

A confusão foi na padaria?

*The<sub>[MASC]</sub> cook<sub>[MASC]</sub>/The<sub>[FEM]</sub> cook<sub>[FEM]</sub> thanked the<sub>[FEM]</sub> nutritionist who quickly helped him/her in the middle of the mess of that restaurant.*

*Was the mess at the bakery?*

e) O bailarino/A bailarina admira a ginecologista que ajudou ele/ela gentilmente depois de uma das fases mais difíceis de sua vida.

A ginecologista é cruel com seus pacientes?

*The<sub>[MASC]</sub> dancer<sub>[MASC]</sub>/The<sub>[FEM]</sub> dancer<sub>[FEM]</sub> admires the<sub>[FEM]</sub> gynecologist who gently helped him/her after one of the most difficult times in life.*

*Is the<sub>[FEM]</sub> gynecologist cruel with his patients?*

f) O cabeleireiro/A cabeleireira detestou a esteticista que olhou ele/ela lentamente dos pés a cabeça durante a entrevista.

A esteticista foi quem conduziu a entrevista?

*The<sub>[MASC]</sub> hairdresser<sub>[MASC]</sub>/The<sub>[FEM]</sub> hairdresser<sub>[FEM]</sub> disliked the<sub>[FEM]</sub> beautician who slowly stared at him/her before the interview.*

*Was the<sub>[FEM]</sub> beautician who conducted the interview?*

### List of the experimental sentences of Experiment 1b

*Distractors with definitional gender*

a) O veterinário/A veterinária reconheceu o genro que feriu ele/ela fortemente por trás da cabeça no momento do assalto.

O ferimento por trás da cabeça foi no momento do assalto?

*The<sub>[MASC]</sub> vet<sub>[MASC]</sub>/The<sub>[FEM]</sub> vet<sub>[FEM]</sub> recognized the son-in-law who forcefully hit him/her behind the head at the moment of the assault.*

*Was the wound behind the head at the moment of the assault?*

b) O comissário de bordo/A comissária de bordo processou o rapaz que xingou ele/ela ferozmente devido ao um mal entendido.

O rapaz xingou devido ao um mal entendido?

*The<sub>[MASC]</sub> flight attendant<sub>[MASC]</sub>/The<sub>[FEM]</sub> flight attendant<sub>[FEM]</sub> sued the young man who fiercely cursed him/her due to a misunderstanding.*

*Did the young man curse due to a misunderstanding?*

c) O pedagogo/A pedagoga elogiou o pai que tocou ele/ela totalmente com uma das lições mais belas de coragem.

O pai deu uma bela lição de coragem?

*The<sub>[MASC]</sub> pedagogue<sub>[MASC]</sub>/The<sub>[FEM]</sub> pedagogue<sub>[FEM]</sub> complimented the father who completely touched him/her with one of the most beautiful lessons of courage.*

*Did the father give a beautiful lesson of courage?*

d) O advogado/A advogada hostilizou o rei que tratou ele/ela rudemente na frente de alguns convidados na festa.

O rei foi rude no café-da-manhã?

*The<sub>[MASC]</sub> lawyer<sub>[MASC]</sub>/The<sub>[FEM]</sub> lawyer<sub>[FEM]</sub> antagonized the king who treated him/her rudely in front of some of guests at the party.*

*Was the king rude at the breakfast?*

e) O criminoso/A criminosa ameaçou o cavalheiro que notou ele/ela de repente próximo ao prédio alguns dias depois do crime.

O cavalheiro foi assassinado?

*The<sub>[MASC]</sub> criminal<sub>[MASC]</sub>/The<sub>[FEM]</sub> criminal<sub>[FEM]</sub> threatened the gentleman who suddenly noticed him/her near the building some days after the crime.*

*Was the gentleman murdered?*

f) O enfermeiro/A enfermeira conhecia o homem que matou ele/ela brutalmente na frente da casa da família.

O crime aconteceu na frente do shopping?

*The<sub>[MASC]</sub> nurse<sub>[MASC]</sub>/The<sub>[FEM]</sub> nurse<sub>[FEM]</sub> knew the man who brutally killed him/her in front of the family's house.*

*Did the crime happen in front of the shopping mall?*

*Distractors with stereotypical gender*

a) O perito/A perita acatou o detetive que chamou ele/ela novamente por volta do meio-dia no escritório.

O detetive chamou alguém por volta do meio-dia no escritório?

*The<sub>[MASC]</sub> investigator<sub>[MASC]</sub>/The<sub>[FEM]</sub> investigator<sub>[FEM]</sub> answered the<sub>[MASC]</sub> detective who called him/her again around noon at the office.*

*Did the<sub>[MASC]</sub> detective call someone around noon at the office?*

b) O cenógrafo/A cenógrafa despediu o electricista que ofendeu ele/ela puramente por causa de fofocas nos bastidores.

Houve fofocas nos bastidores?

*The<sub>[MASC]</sub> set designer<sub>[MASC]</sub>/The<sub>[FEM]</sub> set designer<sub>[FEM]</sub> fired the<sub>[MASC]</sub> electrician who solely offended him/her due to gossips in the side scenes.*

*Were there gossips in the side scenes?*

c) O bibliotecário/O bibliotecário seguiu o guarda que guiou ele/ela brevemente através do corredor do grande gabinete real.

Há um corredor grande no gabinete real?

*The<sub>[MASC]</sub> librarian<sub>[MASC]</sub>/The<sub>[FEM]</sub> librarian<sub>[FEM]</sub> followed the<sub>[MASC]</sub> guard who briefly guided him/her through the hallway of the large royal office.*

*Is there a hallway in the large royal office?*

d) O cozinheiro/A cozinheira agradeceu o taxista que ajudou ele/ela rapidamente no meio da confusão daquele restaurante.

A confusão foi na padaria?

*The<sub>[MASC]</sub> cook<sub>[MASC]</sub>/The<sub>[FEM]</sub> cook<sub>[FEM]</sub> thanked the<sub>[MASC]</sub> taxi driver who quickly helped him/her in the middle of the mess of that restaurant.*

*Was the mess at the bakery?*

e) O bailarino/A bailarina admira o surfista que ajudou ele/ela gentilmente depois de uma das fases mais difíceis de sua vida.

O surfista é cruel com seus amigos?

*The<sub>[MASC]</sub> dancer<sub>[MASC]</sub>/The<sub>[FEM]</sub> dancer<sub>[FEM]</sub> admires the<sub>[MASC]</sub> surfer who gently helped him/her after one of the most difficult times in life.*

*Is the<sub>[MASC]</sub> surfer cruel with his friends?*

f) O cabeleireiro/A cabeleireira detestou o piloto de corrida que olhou ele/ela lentamente dos pés a cabeça antes da entrevista no camarim.

O piloto de corrida foi quem conduziu a entrevista?

*The<sub>[MASC]</sub> hair dresser<sub>[MASC]</sub>/The<sub>[FEM]</sub> hairdresser<sub>[FEM]</sub> disliked the<sub>[MASC]</sub> race car driver who slowly stared at him/her before the interview.*

*Was the<sub>[MASC]</sub> race car driver who conducted the interview?*

**LME results for First Fixation at the pronoun region.**

Pronoun				
Parameter	$\beta_0$	SE	t-value	p-value
<i>(Intercept)</i>	0.11	0.23	0.48	0.63
<i>Antecedent</i>	0.44	0.32	1.393	0.16
<i>Distractor</i>	0.16	0.32	0.49	0.62
<i>Type of gender</i>	-0.61	0.33	-1.84	0.070
<i>Gender</i>	-0.51	0.29	-1.74	0.084
<i>Antecedent x Distractor</i>	-0.44	0.45	-0.98	0.32
<i>Antecedent x Type of gender</i>	0.18	0.47	0.39	0.69
<i>Distractor x Type of gender</i>	0.35	0.44	0.79	0.42
<i>Antecedent x Gender</i>	-0.11	0.39	-0.29	0.77
<i>Distractor x Gender</i>	0.30	0.39	0.75	0.44
<i>Type of gender x Gender</i>	1.20	0.41	2.90	0.004
<i>Antecedent x Distractor x Type of gender</i>	-0.24	0.64	-0.37	0.70
<i>Antecedent x Distractor x Gender</i>	1.19	0.55	0.34	0.73
<i>Antecedent x Type of gender x Gender</i>	-1.15	0.57	2.00	0.046
<i>Distractor x Type of gender x Gender</i>	-1.01	0.56	-1.80	0.072
<i>Antecedent x Distractor x Type of gender x Gender</i>	0.97	0.81	1.18	0.23

**LME results for First Pass at the pronoun region.**

Pronoun				
Parameter	$\beta_0$	SE	t-value	p-value
<i>(Intercept)</i>	0.10	0.21	0.48	0.63
<i>Antecedent</i>	0.38	0.32	1.17	0.24
<i>Distractor</i>	0.005	0.31	0.016	0.98
<i>Type of gender</i>	-0.69	0.30	-2.24	0.026
<i>Gender</i>	-0.37	0.27	-1.39	0.16
<i>Antecedent x Distractor</i>	-0.05	0.44	-0.12	0.89
<i>Antecedent x Type of gender</i>	0.40	0.47	0.85	0.39
<i>Distractor x Type of gender</i>	0.63	0.43	1.46	0.14
<i>Antecedent x Gender</i>	-0.21	0.38	-0.54	0.58
<i>Distractor x Gender</i>	0.18	0.39	0.48	0.62
<i>Type of gender x Gender</i>	1.10	0.39	2.80	0.0054
<i>Antecedent x Distractor x Type of gender</i>	-0.97	0.64	-1.51	0.13
<i>Antecedent x Distractor x Gender</i>	0.23	0.55	0.41	0.67
<i>Antecedent x Type of gender x Gender</i>	-1.02	0.57	-1.78	0.074
<i>Distractor x Type of gender x Gender</i>	-0.93	0.55	-1.69	0.090
<i>Antecedent x Distractor x Type of gender x Gender</i>	1.06	0.81	1.31	0.19

## LME results for Regression Path at the pronoun region.

Pronoun				
Parameter	$\beta_0$	SE	t-value	p-value
<i>(Intercept)</i>	-0.44	0.25	-1.73	0.10
<i>Antecedent</i>	0.89	0.33	2.65	0.0098
<i>Distractor</i>	0.68	0.34	1.98	0.051
<i>Type of gender</i>	-0.02	0.34	-0.08	0.93
<i>Gender</i>	0.39	0.26	1.49	0.13
<i>Antecedent x Distractor</i>	-0.99	0.42	-2.33	0.020
<i>Antecedent x Type of gender</i>	-0.04	0.45	-0.09	0.92
<i>Distractor x Type of gender</i>	-0.40	0.46	-0.86	0.38
<i>Antecedent x Gender</i>	-0.97	0.39	-2.45	0.015
<i>Distractor x Gender</i>	-1.01	0.37	-2.70	0.0072
<i>Type of gender x Gender</i>	0.12	0.37	0.32	0.74
<i>Antecedent x Distractor x Type of gender</i>	0.36	0.60	0.60	0.54
<i>Antecedent x Distractor x Gender</i>	1.62	0.53	3.02	0.0026
<i>Antecedent x Type of gender x Gender</i>	-0.07	0.54	-0.14	0.088
<i>Distractor x Type of gender x Gender</i>	0.87	0.52	1.66	0.096
<i>Antecedent x Distractor x Type of gender x Gender</i>	-0.26	0.77	-0.33	0.73

[recebido em 30 de outubro de 2021 e aceite para publicação em 08 de fevereiro de 2022]

## LEITURA DE SENTENÇAS EM ALEMÃO POR FALANTES DE HUNSRIQUEANO: ACESSO LEXICAL NÃO SELETIVO E COATIVAÇÃO SINTÁTICA

### GERMAN SENTENCE READING BY HUNSRÜCKISCH SPEAKERS: NON-SELECTIVE LEXICAL ACCESS AND SYNTACTIC COACTIVATION

Bernardo Kolling Limberger\*  
bernardo.limberger@ufpel.edu.br

Estudos sobre o processamento da leitura bilíngue/multilíngue têm evidenciado de forma robusta que o acesso lexical é não seletivo, havendo uma busca paralela em todos os subsistemas linguísticos. Esse efeito tem sido replicado nas pesquisas mais recentes nas quais são apresentadas sentenças aos participantes. O contexto de sentença também mostra que pode haver coativação sintática, de modo que as representações sintáticas de uma língua possam facilitar o processamento de estruturas semelhantes na outra língua. De forma geral, tais as investigações incluem línguas hegemônicas e majoritárias. Diante disso, o objetivo deste estudo foi investigar o acesso lexical durante o processamento de sentenças em alemão padrão por falantes de hunsriqueano, língua de imigração alemã falada no Brasil. Foi aplicada uma Tarefa de compreensão de sentenças em alemão. Dois grupos participaram da pesquisa: um composto por falantes de hunsriqueano, o outro por não falantes de qualquer língua minoritária de origem alemã. Todos os participantes estudam alemão como língua estrangeira. Os resultados mostram um efeito do compartilhamento de representações semânticas, fonológicas (acesso lexical não seletivo) e sintáticas (coativação sintática) entre hunsriqueano e alemão no processamento de sentenças em alemão. Por meio dos resultados, visamos contribuir com a pesquisa em línguas minoritárias e processamento da leitura, uma relação nem sempre trivial.

**Palavras-chave:** Acesso lexical. Coativação sintática. Leitura. Hunsriqueano. Alemão padrão.

Studies on bilingual/multilingual reading processing have robustly shown that lexical access is non-selective, with parallel search across all language subsystems. This effect has been replicated in more recent research, in which participants read sentences. The sentence context also shows that there can be syntactic coactivation, so that syntactic representations in one language can facilitate processing of similar structures in the other language. In general, such investigations include both hegemonic and majority languages. The aim of this study was to investigate lexical access during sentence processing in German by speakers of Hunsrückisch, an German immigrant language spoken in Brazil. A German sentence comprehension task was applied. Two groups participated in the research: one composed of Hunsrückisch speakers, the other one composed of non-speakers of any minority language with German origin. All participants study German as a foreign language. The results show an effect of sharing semantic, phonological (non-selective

---

\* Professor do Centro de Letras e Comunicação - Área de Alemão e Programa de Pós-Graduação em Letras. Universidade Federal de Pelotas (UFPel). Pelotas, RS - Brasil. ORCID: 0000-0001-5504-2361

lexical access) and syntactic (syntactic coactivation) representations between Hunsrückisch and German on sentence processing. Through the results, we aim to contribute to research on minority languages and reading processing, a not always trivial relationship.

**Keywords:** Lexical Access. Syntactic coactivation. Reading. Hunsrückisch. Standard German.

•

## 1. Introdução

Estudos sobre acesso lexical em falantes bilíngues e multilíngues têm mostrado efeitos de coativação linguística na leitura, como já foi amplamente demonstrado na literatura (Dijkstra 2005; Lauro & Schwartz 2017; Palma & Titone 2020; Toassi & Mota 2015). Entende-se por acesso lexical o processo de recuperar informações semânticas e gramaticais por meio da forma ortográfica e/ou fonológica da palavra (Dijkstra 2005). As investigações com bilíngues e multilíngues têm sugerido que o acesso lexical é não seletivo. Então, a tendência é de que as representações das duas ou mais línguas sejam ativadas simultaneamente e que compitam durante a seleção lexical, de forma a levar a uma busca paralela em todos os subsistemas linguísticos.

No Brasil, as pesquisas sobre acesso lexical na leitura ainda são escassas. Os estudos brasileiros sobre essa temática contam com a participação de falantes de um grupo crescente de pessoas que buscam cursos de línguas adicionais, como inglês, francês ou alemão (Barcelos 2016; Blank 2013; Duarte & Blank 2019; Pickbrenner 2017; Toassi, Mota, & Teixeira 2020). Essas investigações têm incluído configurações de línguas hegemônicas (como, por exemplo, português, francês, inglês e alemão), em detrimento de línguas minoritárias.

O presente estudo, fundamentado no trabalho de doutorado de Limberger (2018), busca preencher essa lacuna, ao incluir a língua minoritária *Hunsrückisch*/hunsriqueano, uma variedade linguística suprarregional do alemão usada principalmente no sul do Brasil (Altenhofen 1996). Essa língua se relaciona ao alemão *standard*/padrão (referido aqui como alemão) pela etimologia e pela tipologia linguística, de modo que a língua minoritária pode subsidiar a aprendizagem de alemão como língua estrangeira (Messa 2009; Spinassé 2005; Steffen 2008). Esses efeitos são motivados pelo léxico, pois há diversas palavras de origem germânica que são cognatas entre essas duas línguas e pela morfossintaxe. A inclusão de uma língua minoritária na investigação sobre esses fenômenos pode contribuir para a compreensão de diferentes configurações multilíngues e exibir facetas da interação entre a oralidade e a escrita durante o processamento multilíngue. Pesquisas com falantes de línguas e variedades linguísticas predominantemente orais podem evidenciar os efeitos do conhecimento de uma língua oral no uso e no processamento de outra(s) língua(s) na modalidade escrita.

Além dos efeitos cognatos, parte-se do pressuposto de que o compartilhamento de uma estrutura entre duas línguas possa ocasionar efeitos de facilitação (Cai, Pickering, Yan, & Branigan 2011; Hopp 2017; Westergaard, Mitrofanova, Mykhaylyk, & Rodina 2016). No caso do hunsriqueano e alemão, a *Satzklammer*, estrutura analítica com dois

verbos, é compartilhada entre as línguas. Essa estrutura é utilizada, por exemplo, no *Perfekt*, pretérito perfeito, estrutura-alvo deste estudo. Nos exemplos abaixo, ou seja, traduções para a frase ‘Ele comeu um pedaço de bolo’ em alemão (Exemplo 1) e em hunsriqueano (Exemplo 2), pode-se encontrar a forma finita do verbo auxiliar (*hat/hot*) na segunda posição (V2) e o verbo principal em posição final.

(1) *Er hat ein Stück Kuchen gegessen.*

(2) *Er hot en Stick Bolo gess.*  
‘Ele comeu um pedaço de bolo’.

Em hunsriqueano, essa estrutura analítica ocorre também no presente (Exemplo 3), pois são usadas perífrases com o verbo auxiliar *tun*. No caso dessa estrutura analítica, o verbo principal se localiza na *Satzklammer* direita (posição final) e, por isso, é possível introduzir constituintes no campo intermediário, e o verbo auxiliar permanece na posição V2. Em alemão padrão, somente a posição V2 é preenchida.

(3) *Ich tun viel esse.*

(4) *Ich esse viel.*  
‘Eu como muito’.

Essa coexistência de representações gramaticais entre hunsriqueano, a língua aprendida na infância, e alemão, a língua aprendida tardiamente em contexto escolar, pode resultar em efeitos de coativação sintática. Como pode-se perceber no Exemplo 3, em hunsriqueano, a estrutura de *Satzklammer* aparece inclusive no *Präsens*, devido ao uso de perífrases com o verbo *tun*. Como mostram Altenhofen *et al.* (2018), a perífrase é a estrutura dominante na região *Deutsch*, onde se fala uma variedade de hunsriqueano menos próxima ao alemão padrão, em comparação ao tipo *Deutsch* (Altenhofen 2010).

Então, diante dessas características das línguas e dos estudos que encontram efeitos de facilitação na aprendizagem, partiu-se, neste estudo, do pressuposto de que pode haver influências do hunsriqueano no processamento de sentenças em alemão. O objetivo deste estudo foi investigar o acesso lexical durante o processamento de sentenças em alemão por falantes multilíngues de hunsriqueano, português brasileiro e alemão, escritas com a mesma estrutura (*Satzklammer*). Investigou-se sob que condições (verbo cognato *x* verbo não cognato) e tempos verbais (*Perfekt x Präsens*) o compartilhamento de representações ortográficas/fonológicas dos cognatos e sintáticas da *Satzklammer* entre as línguas pode influenciar o processamento da leitura. As variáveis analisadas foram tempo de resposta e acurácia, e foram comparados falantes de hunsriqueano com não falantes de língua minoritária germânica, que também são multilíngues (falantes de português, alemão e inglês). A hipótese era que os falantes de hunsriqueano (grupo GHA) teriam tempos de resposta mais baixos e níveis de acurácia mais elevados no processamento de sentenças em alemão com verbo cognato, em comparação ao processamento de sentenças com verbo não cognato e em comparação com o grupo de pessoas que não fala hunsriqueano.

O presente texto começa por uma revisão de literatura que contempla estudos sobre acesso lexical, focalizando o processamento de sentenças e a coativação sintática. Em seguida, relata-se sobre a situação sociolinguística dos falantes de hunsriqueano. Essa seção teórica é seguida pelo método do estudo, no qual são apresentadas informações sobre os participantes, a Tarefa de compreensão de sentenças em alemão, o paradigma experimental e os procedimentos de coleta de dados. Por fim, são apresentados os resultados do estudo seguidos da discussão e da conclusão.

## 2. Revisão de literatura

### 2.1. Acesso lexical

No bilinguismo/multilinguismo, investigar acesso lexical é algo complexo, porque dois ou mais subsistemas linguísticos estão em jogo, interagindo entre si. Os bilíngues e multilíngues não reconhecem palavras da mesma forma que monolíngues (Assche, Duyck, & Hartsuiker 2012; Kroll & Tokowicz 2005). Há duas hipóteses sobre o acesso lexical bilíngue/multilíngue: a primeira postula que o acesso às informações armazenadas no léxico mental é seletivo com relação à língua, ao passo que a segunda postula que o acesso lexical não é seletivo (de Bot 2004; Dijkstra, Grainger, & van Heuven 1999). Isso significa que, na primeira hipótese, por exemplo, na leitura de um texto em alemão, somente a palavra em alemão, como *Fleisch* ‘carne’, poderia ser acessada quando um falante de hunsriqueano e alemão lesse essa palavra em um texto escrito em alemão. Em hunsriqueano, a palavra seria grafada de forma idêntica. Dessa forma, as línguas não competiriam por ativação. Entretanto, na hipótese da não seletividade, influências de uma língua na outra deveriam ocorrer, porque ambas as representações são ativadas e competem durante a seleção lexical, ocorrendo uma busca paralela em todos os subsistemas (Linck, Hoshino, & Kroll 2008; Rodriguez-Fornells, Rotte, Heinze, Nösselt, & Münte 2002). Ambas as hipóteses relacionadas à ativação podem estar ou não pautadas em modelos que visam a simular o processamento bilíngue/multilíngue e fazer previsões sobre os fatores que influenciam o acesso lexical.

Um dos modelos mais utilizados na literatura é o *Bilingual Interactive Activation Model* – BIA (van Heuven, Dijkstra, & Grainger 1998), que foi desenvolvido com base no modelo interativo de processamento de palavras (Rumelhart & McClelland 1981). Esse modelo assume um léxico compartilhado para as duas línguas, que se estende à reformulação do modelo, *Bilingual Interactive Activation Plus Model* – BIA+ (Dijkstra & van Heuven 2002). No modelo atualizado, a apresentação visual de uma palavra conduz à coativação das representações ortográficas e fonológicas nas duas línguas, que ativam representações semânticas e, depois, os nós linguísticos. A ativação pode fluir também no sentido contrário: as representações semânticas ativam as diferentes ortografias. Nessa visão, cognatos possuem duas entradas ortográficas e/ou fonológicas que compartilham uma mesma representação semântica. Na ativação simultânea dessas entradas ortográficas, a representação semântica, em parte, envia *feedback* para o nível ortográfico, que, por sua vez, amplifica a ativação. Em uma das duas línguas, então, a entrada do cognato alcançará o limiar de ativação mais rapidamente que a entrada de um não cognato. O modelo fornece o melhor quadro teórico para interpretar o reconhecimento

visual de palavras no bilinguismo/multilinguismo, mas não se consideram aspectos morfológicos e traços semânticos (Van Assche, Duyck, & Brysbaert 2013).

O modelo BIA+ tem sido utilizado para explicar o efeito de facilitação no processamento de cognatos, aqui referido como “efeito cognato” (Dijkstra 2005), robustamente encontrado em estudos empíricos sobre acesso lexical com tarefas de decisão lexical (Allen, Conklin, & Miwa 2021; Barbosa Pinto & Arêas da Luz Fontes 2020; Caramazza & Brones, 1979; de Groot & Nas, 1991; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen 2010; Lemhöfer & Dijkstra 2004; Szubko-Sitarek 2012; Van Hell & Dijkstra 2002). Tais estudos mostram que o efeito parece ser maior no processamento da língua adicional em comparação ao processamento da língua materna e foi encontrado em crianças e adultos em fase de aprendizagem da leitura (Abutalebi *et al.* 2007; Arêas da Luz Fontes, de Souza Brentano, Toassi, Sittig, & Finger 2021; Limberger 2021; Schröter & Schroeder 2016). O efeito parece ser modulado pelo grau de similaridade ortográfica e pela frequência de palavras (Dijkstra *et al.* 2010; Sherkina-Lieber 2004), pelo compartilhamento de representações ortográficas e fonológicas na leitura (Kirk, Kempe, Scott-Brown, Philipp, & Declerck 2018; Lemhöfer & Dijkstra 2004) e pela tarefa aplicada (Allen *et al.* 2021; Bultena, Dijkstra, & van Hell 2014). Outro fator que exerce influência sobre o efeito cognato é o nível de proficiência (Barcelos 2016; Bultena *et al.* 2014; Pickbrenner 2017; Van Hell & Dijkstra 2002). Nesses estudos, o efeito cognato nem sempre foi alcançado em falantes com baixo nível de proficiência.

## 2.2. Acesso lexical em contexto de sentença

Nos estudos reportados acima, os experimentos são compostos quase exclusivamente por palavras. Porém, atualmente, as investigações buscam verificar se efeitos cognatos permanecem na leitura em contexto de sentença ou de texto, uma vez que é evidente que o contexto linguístico no qual a palavra-alvo está inserida exerce influência sobre a ativação da palavra.

O conjunto de estudos tem mostrado que os efeitos cognatos se mantêm, de fato, no processamento de sentenças. No entanto, as sentenças que apresentam tais resultados são de baixa restrição semântica, ou seja, compostas por combinações menos previsíveis de argumentos e verbos (Duyck, Assche, Drieghe, & Hartsuiker 2007; Hopp 2017; Lauro & Schwartz 2017; Libben & Titone 2009; Schwartz, Kroll, & Diaz 2007; Schwartz & Kroll 2006). Dessa forma, efeitos de facilitação cognata são encontrados mais frequentemente quando o leitor não tem a possibilidade de inferir a palavra-alvo por meio do contexto da sentença. Pesquisas mais atuais têm controlado esse fator, apresentando aos participantes somente sentenças com baixa restrição semântica (Bosma & Nota 2020; Toassi *et al.* 2020) e, com esse conjunto de estímulos, encontrado efeitos de facilitação. Os efeitos a nível da sentença também têm sido explicados à luz do modelo BIA+ (Dijkstra & van Heuven 2002), que preconiza que o reconhecimento de palavras inseridas em contexto de sentença é sensível a informações sintáticas e semânticas e que o contexto interage com o reconhecimento da palavra-alvo da sentença. Isso indica que processos descendentes na compreensão interagem diretamente com processos ascendentes de acesso lexical.

### 2.3. Coativação sintática

Além da baixa restrição semântica, outro fator que pode contribuir para efeitos de facilitação é o compartilhamento da estrutura sintática. Quando há uma estrutura comum em duas línguas (por exemplo, a *Satzklammer*, no caso das línguas envolvidas neste estudo), pode haver coativação da estrutura, o que pode ter um efeito de facilitação: os aprendizes de L3 integram os estímulos (*parsing*) e podem fazer previsões por meio da consulta das gramáticas previamente adquiridas (Westergaard *et al.* 2016).

O compartilhamento da estrutura pode resultar no chamado *cognate boost* (Cai *et al.* 2011; Hopp 2017) quando combinado com o compartilhamento lexical (cognatos). Esse efeito é provocado quando há coexistência de representações lexicais em duas línguas, o que pode acarretar uma aceleração no processamento geral da sentença. Quando verbos não cognatos são processados, os efeitos de *boost* seriam reduzidos ou ausentes. Ou seja, a coexistência da representação da palavra-alvo em duas línguas pode levar a uma aceleração no processamento da sentença.

### 2.4. Hunsriqueano

A língua minoritária hunsriqueano/*Hunsrückisch* (Altenhofen 1996), falada no contexto brasileiro onde a pesquisa foi realizada, é considerada atualmente uma língua fora da matriz, porque é uma língua de imigração. Em contato com o português brasileiro (PB), o hunsriqueano possui diversos empréstimos dessa língua (Altenhofen *et al.* 2007). Segundo esses autores, a língua minoritária brasileira é essencialmente falada e não dispõe de uma prática de escrita uniforme e amplamente utilizada. Além disso, a língua possui propostas de convenções de conversão grafema-fonema (Altenhofen *et al.* 2007; Wiesemann 2008) que não são padronizadas. Essa língua se fundamenta, segundo Altenhofen (1996), num contínuo dialetal formado essencialmente pelos dialetos trazidos pelos imigrantes alemães, principalmente a partir de 1824, cuja grande parte era oriunda da região do Hunsrück, no oeste da Alemanha. No Brasil, o hunsriqueano assumiu, devido à maior semelhança com o alemão padrão e ao maior número de falantes<sup>1</sup>, o papel de variedade comum entre as comunidades.

Diferentemente das variedades linguísticas diatópicas da Alemanha, o hunsriqueano do Brasil pode ser caracterizado como um arquipélago linguístico, ao invés de configurar ilhas isoladas (Steffen & Altenhofen 2014). Os falantes da língua não dispõem de uma região específica para o seu uso, mas sim de colônias que se localizam espalhadas e compõem o contínuo variacional da língua (Altenhofen 2010; Bellmann 1983; Machado 2016). No estado do Rio Grande do Sul, por exemplo, os descendentes de alemães fundaram outras colônias, as denominadas Colônias Novas (Verband Deutscher Vereine 1999). Essas diferentes comunidades compõem duas principais variedades do hunsriqueano - o tipo *Deitsch* e o tipo *Deutsch*. A oposição *Deitsch/Deutsch* é baseada na denominação dos falantes de hunsriqueano de ambas as

---

<sup>1</sup> Estima-se que hunsriqueano seja a língua de referência de mais de um milhão de brasileiros (Altenhofen *et al.* 2018).

variedades (Altenhofen 2010). O tipo *Deutsch* [+ alto alemão] é uma variedade mais próxima ao alemão *standard*, em comparação ao tipo *Deitsch* [+ dialetal].

Hunsriqueano (ambas as variedades) e alemão compartilham aspectos lexicais e estruturas gramaticais. Privilegiamos, neste estudo, uma breve descrição das semelhanças nos substantivos e nos verbos e na estrutura da *Satzklammer* - classes de palavras-alvo e estrutura contempladas nos experimentos. No âmbito lexical, hunsriqueano e alemão possuem uma série de cognatos (Altenhofen 2010), por exemplo, *Hos* e *Hose* ‘calça’, *Schul* e *Schule* ‘escola’, *uffmache* e *aufmachen* ‘abrir’. No nível sintático, em ambas as línguas, a posição do verbo finito é a segunda (V2) e, quando este verbo é auxiliar, no caso do *Perfekt* (pretérito perfeito), o verbo principal se localiza em posição final (Duden, 2005). Essa estrutura, denominada *Satzklammer*, aparece, em hunsriqueano, inclusive no presente (perífrase com o verbo auxiliar *tun*). Isso significa que, na língua minoritária, a *Satzklammer* é ainda mais presente (por exemplo, *Ich tun viel treeme*. ‘Eu sonho muito’).

A língua tem, portanto, as seguintes características principais: a) possui status de língua minoritária do Brasil (Altenhofen 2013), cooficializada em três municípios (Morello, 2021); b) localiza-se fora da sua matriz de origem germânica (Altenhofen 1996); c) não possui um sistema padronizado de escrita (Altenhofen *et al.* 2007; Limberger 2018); d) é adquirida na infância (Altenhofen, 1996; Flores, Rinke, & Wagner 2020) e, por fim, e) possui um repertório linguístico dinâmico, variável e dependente das práticas da comunidade de fala (Altenhofen 2010; Bellmann 1983; Flores *et al.* 2020; Machado 2016).

### 3. Método

#### 3.1. Participantes

A amostra para este estudo foi recrutada por conveniência, inicialmente na região metropolitana de Porto Alegre (Rio Grande do Sul, Brasil), onde a maioria dos participantes residia. Para integrar a amostra da pesquisa, os participantes deviam, antes de tudo, assinar e estar em conformidade com o termo de consentimento livre e esclarecido (aprovado pelo Comitê de Ética da Pontifícia Universidade Católica do Rio Grande do Sul) e apresentar condições de saúde apropriadas (visão normal ou corrigida e ausência de transtornos de linguagem e aprendizagem) para realizar as tarefas. Ainda, era pré-requisito que eles usassem português e alemão (e, para o grupo GHA, hunsriqueano) no cotidiano. Quanto à idade, esperávamos que os participantes tivessem entre 18 e 45 anos. Essa restrição da idade se relaciona com a intenção de conseguir parear os grupos e controlar a variabilidade nos resultados. Para o fator escolaridade, esperávamos que os participantes estivessem cursando ou tivessem concluído o Ensino Superior. No julgamento dos estímulos, participaram do estudo piloto alemães de 19 a 35 anos (média de idade = 25,1 anos; DP = 4,2), todos estudantes da Universidade de Freiburg (Alemanha).

Como critérios de inclusão específicos para os participantes da amostra-alvo, os falantes de hunsriqueano deveriam ser bilíngues precoces e falar a variedade do tipo *Deitsch* (Altenhofen, 2010; Machado, 2016). Para conseguirem fazer as tarefas, o nível

de proficiência em alemão, aprendido em educação formal, precisaria ser alto (desejável: nível B2/C1 (Conselho da Europa 2005). O nível de proficiência foi verificado a partir dos resultados dos testes de proficiência que os participantes já tinham: *Deutsches Sprachdiplom I e II*, do Conselho de Ministros da Educação e Cultura da Alemanha, *Kleines Sprachdiplom*, *Großes Sprachdiplom* e *TestDaF*, do Instituto Goethe, e *Deutsche Sprachprüfung für den Hochschulzugang*, vinculada a universidades alemãs. Os resultados dos testes de proficiência que eles já tinham foram correlacionados ao resultado no desempenho da tarefa de vocabulário receptivo do *Institut für Test-forschung und Testentwicklung*, que fornece uma estimativa de nível de proficiência, e às medidas de proficiência autoavaliada. A partir desses dados, foi criado um escore que corresponde a um dos seis níveis do Quadro Comum Europeu de Referências para Línguas (Conselho da Europa 2005).

Depois da aplicação dos critérios de inclusão específicos para cada grupo, a amostra final abrangeu 56 participantes. Os dois grupos foram compostos por pessoas que aprenderam PB durante a infância e alemão como língua estrangeira. A diferença entre ambos os grupos foi o conhecimento em hunsriqueano. Os integrantes do denominado GHA dispõem de conhecimentos de hunsriqueano adquiridos na infância. O grupo GA contou com a participação de falantes que não possuem conhecimento de qualquer língua minoritária de origem alemã (por exemplo, hunsriqueano, pomerano, westfaliano ou outras). Os participantes desse grupo falam alemão padrão. Para ambos os grupos, conhecimentos de outras línguas, sobretudo de inglês e espanhol, não impossibilitaram a participação, configurando o multilinguismo. No entanto, os estímulos do experimento foram controlados para que a influência dessas línguas fosse a menor possível.

As características gerais sobre idade, sexo/gênero, anos de escolaridade foram pareadas entre ambos os grupos (Tabela 1): Grupo de falantes de português brasileiro, hunsriqueano, alemão e outras línguas (GHA) e Grupo de falantes de PB, alemão e outras línguas (GA). Os grupos também foram equiparados com relação às médias de idade de aquisição de alemão, nível de proficiência, alcance de memória de trabalho e efeito Stroop. As características demográficas e linguísticas foram levantadas por meio do questionário (Scholl & Finger 2013) adaptado para este estudo. As características cognitivas foram mensuradas por meio da Tarefa de alcance de computação (Wood, Carvalho, Rothe-Neves, & Haase 2001) e pela Tarefa Stroop (Stroop 1935). Essas tarefas foram aplicadas para parear os grupos com relação a construtos cognitivos que influenciam o bilinguismo e a leitura (Bialystok 2009; Daneman & Merikle 1996; Green 1998; Kramer 2011; Limberger & Buchweitz 2014). Testes-T para amostras independentes mostraram que não houve diferenças significativas entre os grupos no que tange a essas características. Para informações detalhadas sobre os participantes, ver Limberger (2018).

**Tabela 1. Dados gerais e linguísticos dos grupos participantes do estudo**

	GHA (n = 28)	GA (n = 28)
Idade (DP)	27,8 (6,9)	24,5 (6,0)
Faixa etária	18 – 43	18 – 39
Sexo/gênero (F/M)	23/5	22/6
Anos de escolaridade (DP)	15,8 (2,4)	15,9 (3,7)
Idade de aquisição de alemão (DP)	12,8 (6,1)	10,7 (4,6)
Nível de proficiência* (DP)	4,9 (0,9)	4,6 (0,8)
Alcance de memória de trabalho (DP)	2,96 (0,87)	3,29 (1,06)
Efeito Stroop (DP)	59 (254)	66 (346)

*Nota.* DP: desvio padrão; GHA: Grupo de falantes de português brasileiro, hunsriqueano, alemão e outras línguas; GA: Grupo de falantes de português brasileiro, alemão e outras línguas; F: sexo feminino/ M: sexo masculino.

\* Escore equivalente ao nível de proficiência B2 (Conselho da Europa 2005)

Todos os participantes eram voluntários, sendo a maioria do sexo feminino. Todos possuíam nível de escolaridade superior em curso ou completo. Compõem a amostra participantes com ocupações/profissões variadas. Pode-se dizer que os perfis dos grupos são similares entre si, com pouca variabilidade no que concerne às médias dos grupos. Boa parte dos participantes do grupo GHA (9/28, 35%) já fez uso do hunsriqueano também na modalidade escrita. Contudo, quase todos mencionaram que não usam a escrita do hunsriqueano com frequência. Ambos os grupos têm, em média, níveis de proficiência em alemão semelhantes.

### 3.2. Tarefa de compreensão de sentenças em alemão

#### 3.2.1. Estímulos

Para a tarefa de compreensão de sentenças, foram selecionados primeiramente verbos cognatos e verbos não cognatos entre alemão e hunsriqueano. Considerando o objetivo da pesquisa, o experimento foi composto somente por sentenças em alemão, além de sentenças *fillers* (sentenças com estruturas variadas). Os 60 verbos selecionados nas condições de cognatos e de não cognatos (Tabela 2) foram pareados com relação à extensão, ao número de sílabas e à frequência segundo o corpus SUBTLEX-DE (Brysbaert *et al.* 2011). Para as palavras cognatas, foi mensurada também a distância Levenshtein (Heeringa 2004; Möller 2011), considerando fonemas e grafemas (estes de forma estimada, pois a língua hunsriqueana não possui um padrão de escrita). O status dos verbos (cognatos e não cognatos) também foi julgado por cinco falantes de hunsriqueano da região *Deutsch*, que não participaram do experimento. Selecionamos verbos transitivos, verificados na gramática de valência de verbos (Helbig & Schenkel 1991) e no dicionário de valência E-VALBU do *Institut für Deutsche Sprache* (IDS).<sup>2</sup> Todos os verbos em alemão são passíveis de serem preenchidos com um sujeito animado (nominativo), com papel temático de agente e um complemento que desempenha um papel de paciente (acusativo). Depois de selecionados, os verbos foram escritos também no particípio II, para formular sentenças com *Satzklammer* no pretérito perfeito (*Perfekt*). Devido à dificuldade de selecionar estímulos cognatos e não cognatos pareados de acordo

<sup>2</sup> Disponível em: <https://grammis.ids-mannheim.de/verbvalenz> Último acesso em 21 agosto 2021.

com os critérios psicolinguísticos, não foi possível controlar os verbos de acordo com a complexidade morfológica, ou seja, verbos regulares e irregulares. Também foram selecionados 60 verbos *fillers*, para garantir a variabilidade de sentenças.

**Tabela 2. Verbos selecionados para a Tarefa de compreensão de sentenças: médias e desvios padrão (entre parênteses)**

	<i>Präsens</i>		<i>Perfekt</i>	
	Cognatos	Não cognatos	Cognatos	Não cognatos
Número de letras (DP)	7,30 (1,85)	7,70 (1,21)	8,51 (1,52)	7,97 (1,22)
Número de sílabas (DP)	2,26 (0,52)	2,63 (0,49)	2,72 (0,59)	2,70 (0,6)
Frequência/milhão (DP)	39,43 (40,44)	39,62 (28,96)	28,67 (46,03)	30,72 (39,44)
Distância LV fonológica (DP)	0,48 (0,14)	NA	0,55 (0,10)	NA
Distância LV ortográfica (DP)	0,25 (0,10)	NA	0,22 (0,14)	NA

Nota. LV: Levenshtein; DP: Desvio padrão; NA: não se aplica.

Com esse conjunto de verbos, foram formuladas sentenças sempre com a mesma complexidade sintática. Ao todo, foram formuladas 60 sentenças no *Präsens* e 60 sentenças no *Perfekt*, resultando em 120 sentenças. As sentenças no *Präsens* foram formuladas sem verbo modal, de modo que somente a posição V2 fosse preenchida e não tivessem a estrutura de *Satzklammer*. Dos cruzamentos do status dos verbos com a estrutura selecionada, surgiram quatro condições experimentais: sentenças no *Präsens* com verbos cognatos, *Präsens* com verbos não cognatos, *Perfekt* com verbos cognatos e *Perfekt* com verbos não cognatos. Para formular sentenças com conteúdo não previsíveis, ou seja, de baixa restrição semântica, mas plausíveis, utilizamos os corpora *Wortschatz* (Uni Leipzig)<sup>3</sup> e o E-VALBU. Formulamos sentenças com cada verbo nos dois tempos. Cada verbo resultou, então, em duas sentenças, uma escrita em cada tempo verbal.

As sentenças no *Präsens* possuem 8 ou 9 palavras, e as sentenças no *Perfekt*, 9 ou 10 palavras. A primeira posição da sentença foi preenchida por sintagmas nominais compostos por sujeitos animados, podendo ter adjetivo ou não. Para que o conjunto de sentenças parecesse mais natural, em algumas sentenças, foram feitas inversões entre sujeito e verbos, resultantes do deslocamento do adjunto adverbial (*Seine Tante backt jeden Morgen gerne Brötchen. Jeden Morgen backt seine Tante gerne Brötchen.* ‘Todas as manhãs, a sua tia gosta de fazer pãezinhos.’). O número de inversões em cada uma das condições é o mesmo.

A palavra precedente ao alvo também foi controlada, pois poderia ter um efeito de *priming* sobre o alvo (Van Assche *et al.* 2013). Em cada condição, metade das palavras que precedem o verbo-alvo é cognata, e a outra metade não. Foi possível controlar essas palavras de acordo com a extensão e a frequência (Tabela 3).

**Tabela 3. Critérios psicolinguísticos controlados para as palavras que precedem o verbo-alvo em cada condição: médias e desvio padrão (entre parênteses)**

	<i>Präsens</i>		<i>Perfekt</i>	
	Cognatos	Não cognatos	Cognatos	Não cognatos
Número de letras (DP)	7 (2,72)	6,97 (2,34)	6,67 (2,91)	6,56 (2,16)
Frequência/milhão (DP)	149,98 (308,86)	148,20 (231,1)	86,35 (157,24)	84,61 (139,77)

<sup>3</sup> Corpora disponíveis em: <http://wortschatz.uni-leipzig.de/>. Último acesso em 22 setembro 2021.

Depois de formuladas as 120 sentenças, aplicamos um teste *cloze* com 26 falantes de alemão com alto nível de proficiência (nativos ou com nível C2), por meio de um formulário *Google*. Esses leitores não participaram do estudo piloto na Alemanha nem dos experimentos no Brasil. O objetivo do teste foi verificar a predizibilidade do contexto nas sentenças, de modo a manter somente sentenças de baixa restrição semântica (Hopp 2017; Van Assche *et al.* 2013). Foram formuladas duas versões do teste *cloze*, a fim de que cada participante lesse cada verbo somente uma vez. Dessa forma, cada versão do teste foi composta por 60 sentenças, e cada grupo de participantes preencheu somente um conjunto de sentenças. Foram removidas palavras cruciais para a sentença, como o núcleo do sintagma nominal na posição do sujeito ou do objeto, constituintes obrigatórios na estrutura selecionada e determinados semanticamente pelo verbo escolhido (Duden 2005). Os participantes receberam a instrução para preencherem as 60 sentenças com a primeira palavra que conseguiram acessar espontaneamente. Uma vez que foi controlada a palavra precedente ao alvo, nas sentenças no *Präsens*, como descrito anteriormente, deixamos como lacuna o substantivo localizado na posição de objeto, e nas sentenças no *Perfekt*, foi suprimido o substantivo principal localizado na posição de sujeito. Mantivemos os verbos no teste *cloze*, devido à dificuldade de parear os verbos entre as condições. Como exemplo, apresentamos a seguinte sentença no *Präsens*: *Der Bürgermeister ruft sein... \_\_\_\_\_ in sein Büro* ‘O prefeito chama seu(s)/sua(s) \_\_\_\_\_ no seu escritório’. A palavra com a qual a sentença tinha sido originalmente formulada foi *Freunde* (amigos). Nenhum dos participantes preencheu a sentença com essa palavra, mas com palavras variadas. Uma sentença de baixa restrição semântica deve ser completada com uma variedade de palavras, e os resultados do teste *cloze* permitem verificar a manipulação semântica para otimizar as sentenças. Dessa forma, pudemos manter essa sentença como originalmente formulada e caracterizá-la como sentença de baixa restrição semântica. Em geral, as sentenças formuladas inicialmente foram de baixa restrição (média = 9,6%; DP = 14,6%). Como indica o alto desvio padrão, houve variabilidade na média de porcentagem de acertos; por isso, algumas sentenças tiveram um número considerável de acertos. Então, foi necessário alterar essas sentenças, utilizando uma palavra que não foi preenchida no teste *cloze* e aprimorando levemente algumas características gramaticais. Como tinha sido previsto no enunciado do teste, alguns leitores sugeriram possíveis melhorias nas sentenças, principalmente com relação a traços semânticos e sintáticos.

As sentenças puderam ser pareadas dentro de cada um dos tempos verbais de acordo com o número de palavras e letras, a quantidade aproximada de cognatos na sentença e o número de palavras que precedem o alvo (Tabela 4). Foram aplicados Testes-T de amostras independentes, para verificar se os conjuntos de estímulos foram uniformes em cada tempo verbal, e as diferenças significativas foram nulas.

**Tabela 4. Médias gerais e desvios padrão (entre parênteses) das sentenças nas quatro condições da Tarefa de compreensão de sentenças**

	<i>Präsens</i>		<i>Perfekt</i>	
	Cognatos	Não cognatos	Cognatos	Não cognatos
Número de palavras (DP)	7,67 (0,48)	7,7 (0,53)	8,6 (0,5)	8,6 (0,50)
Número de letras (DP)	43,20 (4,92)	43,23 (3,99)	47,77 (3,26)	47,90 (2,76)
Número de cognatos (DP)	4,6 (1,54)	4,65 (1,2)	5,53 (0,78)	5,43 (1,36)
Palavras que precedem o alvo (DP)	2,4 (0,49)	2,6 (0,81)	7,6 (0,49)	7,6 (0,50)

Para todas as sentenças, foram formuladas perguntas fechadas. Para escrever as perguntas, foram controlados os seguintes fatores: número de palavras, número de letras, quantidade de cognatos e foco da pergunta (sujeito ou objeto). Além disso, metade das perguntas teve resposta positiva e a outra metade teve resposta negativa (para estes resultados, ver Limberger 2018).

Depois de todo o processo da formulação das sentenças, julgamento por falantes e (re)formulação, as 180 sentenças foram lidas pela professora Dra. Evelyn Ferstl. Por fim, por sugestão da professora, todas as sentenças e as respectivas perguntas formuladas no *Präsens* foram transformadas para o *Perfekt* e vice-versa. Com esse procedimento, asseguramos que os eventuais efeitos adviriam do compartilhamento da estrutura sintática e não de algum tipo de facilitação no nível semântico, ocasionado pela repetição das sentenças. Desse modo, temos dois conjuntos de sentenças, sendo cada um deles aplicado de modo contrabalanceado entre os participantes, de acordo com o procedimento explicado a seguir. Depois disso, foi aplicado um estudo piloto na Universidade de Freiburg, de modo a assegurar que eventuais efeitos na leitura de sentenças e de perguntas não ocorreriam devido a fatores não controlados. Como esperado, foi possível verificar que não houve diferenças no processamento de cognatos e não cognatos. As sentenças e as perguntas nas quais os participantes tiveram médias de acurácia mais baixas que 80% foram reanalisadas, e algumas nuances gramaticais e semânticas foram reformuladas com a ajuda da professora Evelyn Ferstl. Trata-se de oito sentenças, que tinham sutis inadequações (por exemplo, emprego inapropriado de pronome possessivo). O Quadro 1 mostra exemplos das sentenças nos dois tempos verbais e as respectivas perguntas.

**Quadro 1. Exemplos das sentenças e perguntas**

<i>Präsens</i>	Sentenças com verbos Cognatos	Sentenças com verbos não cognatos	Extensão
Sentença	<i>Die Maurer <u>bauen</u> eine Brücke in der Stadt.</i> Os pedreiros constroem uma ponte na cidade.	<i>Die Bauer <u>liefern</u> den Rahm in großen Eimern.</i> Os agricultores fornecem nata em baldes grandes.	8 ou 9 palavras
Pergunta	<i>Bauen sie ein Haus?</i> Eles constroem uma casa?	<i>Liefern sie Rahm?</i> Eles fornecem nata?	3 ou 4 palavras
Resposta	<i>NEIN</i> Não	<i>JA</i> Sim	
<i>Perfekt</i>	Sentenças com verbos Cognatos	Sentenças com verbos não cognatos	Extensão
Sentença	<i>Mein Mann <u>hat</u> unsere große Mauer selbst <u>gebaut</u>.</i> O meu marido construiu o nosso muro alto.	<i>Sie <u>haben</u> die Produkte sogar ins Ausland <u>geliefert</u>.</i> Eles fornecerem os produtos até mesmo para o exterior.	9 ou 10 palavras
Pergunta	<i>Hat er eine Mauer gebaut?</i> Ele construiu um muro?	<i>Haben sie Milch geliefert?</i> Eles forneceram leite?	4 ou 5 palavras
Resposta	<i>JA</i> Sim	<i>NEIN</i> Não	

O conjunto de estímulos e dados está disponível em <https://osf.io/f4dg7/>.

### 3.2.2. Paradigma e procedimento de coleta de dados

Antes de iniciar o experimento, os participantes foram informados que leriam sentenças na tela do computador e responderiam a perguntas. As sentenças e as perguntas foram apresentadas por inteiro no centro da tela do computador na fonte Arial (tamanho 18), com a fonte branca e a tela preta. Depois de cada sentença, foi apresentado um ponto de fixação (400 ms), seguido da respectiva pergunta (apresentada por 200 ms a 3 segundos).

As sentenças foram apresentadas em ordem aleatória e desapareciam quando o participante pressionava a tecla de espaço ou após 10 segundos. Os participantes foram instruídos a lerem as sentenças com base no treino que tiveram antes da tarefa (oito estímulos, dois de cada condição). Depois de cada sentença, os participantes pressionavam a tecla de espaço, para sinalizarem assim que terminassem de ler a sentença. Dessa forma, foi contabilizado o tempo de leitura das sentenças. Então, para cada sentença, os participantes responderam a uma pergunta de compreensão. Se a resposta era SIM, eles apertavam a tecla S do computador; se a resposta era NÃO, a tecla L, sinalizadas com as cores verde e vermelho, respectivamente.

A tarefa foi programada e aplicada com o *software PsychoPy 2* (Peirce 2008). Foi registrada a latência da resposta e avaliada a acurácia de leitura. O conjunto de 180 sentenças foi distribuído em cinco blocos de 36 sentenças cada. Houve duas versões da tarefa. As estruturas e as condições experimentais foram pareadas nos blocos e os verbos não se repetiam dentro de cada bloco. Entre os blocos, havia um intervalo com o tempo gerenciável pelo participante.

Todo o procedimento foi explicado por meio do termo de consentimento livre e esclarecido. Em seguida, os participantes foram solicitados a preencherem o questionário. O experimento foi aplicado em um *notebook* HP Pavilion 14'. O local da coleta de dados variou de acordo com o desejo e a disponibilidade dos participantes, contanto que a coleta pudesse acontecer em alguma sala iluminada e silenciosa. Na maioria dos casos, era o local de escolha do participante.

### 3.3. Análise

O desempenho dos participantes foi analisado estatisticamente com software SPSS® 19.0 (IBM Corp. 2010). As variáveis dependentes analisadas foram o tempo de resposta (TR) e a acurácia. Além disso, foi calculada uma estimativa da leitura de palavras por minuto. Respostas omissas ou erradas não foram computadas na análise do TR, porque podem ocasionar falsas interpretações dos resultados.

Para avaliar múltiplas interações entre as variáveis, foi aplicado um modelo de equações de estimações generalizadas (GEE, do inglês *Generalized Estimating Equations*). Esse modelo focaliza as mudanças médias nas respostas em diferentes momentos e o impacto das covariáveis sobre essas mudanças também entre os grupos (Ma, Mazumdar, & Memtsoudis 2012). Ele pode acomodar vários tipos de dados que não possuem necessariamente distribuição normal, uma vez que diferentes funções podem ser usadas para modelar as múltiplas relações entre as variáveis. Foi comparado o desempenho dos dois grupos nas condições: tempo verbal (*Perfekt x Präsens*) e status do verbo (cognato *x* não cognato). Em todos os testes estatísticos, foi adotado o nível de significância de  $p \leq 0,05$ .

## 4. Resultados

Os resultados do experimento foram analisados por um modelo de GEE. O objetivo dessa análise foi investigar as interações entre os fatores da compreensão de sentenças (tempo verbal e *status* do verbo) e o desempenho de cada grupo. Na tabela 5, são disponibilizadas as médias do TR, da acurácia e da estimativa de palavras por minuto.

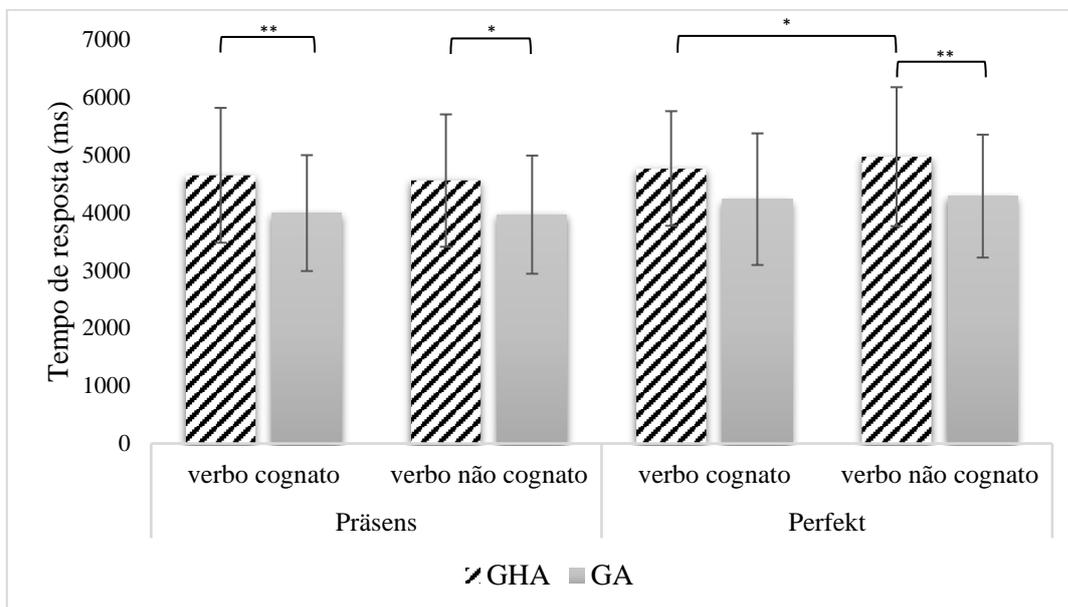
**Tabela 5. Médias dos tempos de resposta (em ms), índices de acurácia (em porcentagem) e estimativas de palavras por minuto com os desvios padrão (entre parênteses) para os grupos GHA e GA na leitura de sentenças em alemão na Tarefa de compreensão de sentenças**

Tempo verbal	Status do verbo	GHA			GA		
		TR	P/Min.	Acurácia	TR	P/Min.	Acurácia
<i>Präsens</i>	Cognato (A)	4642 (1167)	103,8 (20,7)	96 (6)	3988 (1003)	120,2 (23,3)	99 (2)
	Não cognato (B)	4551 (1146)	103,7 (16,3)	96 (8)	3961 (1023)	119,8 (24,8)	98 (3)
	Efeito A sobre B	-91	0,1	0	-27	0,4	0
	Cognato (A)	4760 (992)	112,1 (16,4)	95 (9)	4229 (1139)	124,3 (16,9)	98 (4)
<i>Perfekt</i>	Não cognato (B)	4965 (1202)	109,6 (17,2)	94 (12)	4282 (1064)	123,2 (17,9)	98 (5)
	Efeito A sobre B	205*	2,5	-3	53	1,1	-1

Nota. Diferenças significativas são indicadas por: \*  $p < 0,05$ . P/min: palavras por minuto (estimativa)

Nas médias de TR da leitura sentença, houve efeitos significativos para condição [ $W(2) = 5,42, p < 0,020$ ] e para grupo [ $W(2) = 4,41, p < 0,036$ ], mas não para as interações condição\*grupo. Também foi calculada uma estimativa do número de palavras lidas por minuto. Houve efeito de grupo também para essa variável [ $W(1) = 38,01, p < 0,001$ ]. Para a variável acurácia, as diferenças foram nulas tanto dentro dos grupos quanto entre os grupos.

Nas comparações dentro dos grupos, testes *post hoc* de Bonferroni revelaram que somente os participantes do GHA foram significativamente mais rápidos para lerem as sentenças com verbos cognatos no *Perfekt* em comparação a sentenças com verbos não cognatos nesse tempo verbal ( $p = 0,016$ ). Nas comparações entre os grupos para a leitura de sentenças no *Präsens*, houve diferença significativa favorecendo o GA tanto nas sentenças com verbos cognatos ( $p = 0,023$ ) quanto nas com verbos não cognatos ( $p = 0,040$ ). Essas diferenças permaneceram na estimativa do número de palavras lidas por minuto no *Präsens* e no *Perfekt*. Para ambas as condições, as diferenças foram altamente significativas ( $p < 0,001$ ). Nas sentenças no *Perfekt*, apesar de terem lido mais palavras por minuto, os participantes do GA foram significativamente mais rápidos nas sentenças com os verbos não cognatos ( $p = 0,025$ ). Nas sentenças com verbos cognatos, a diferença entre os grupos foi marginalmente significativa ( $p = 0,061$ ). No gráfico abaixo, podemos visualizar melhor a diferença dentro do GHA e entre os grupos na leitura das sentenças.



**Gráfico 1. Médias dos tempos de resposta (em ms) e dos desvios padrão dos dois grupos na leitura de sentenças em alemão**

Nota. Diferenças significativas são indicadas por: \*  $p < 0,05$ ; \*\*  $p < 0,01$ .

Para as perguntas, o modelo não indicou quaisquer diferenças significativas no TR. A ausência de diferenças significativas se repetiu nas médias de acurácia (média 91,5 para o grupo GHA e média 91 para o grupo GA). Os resultados na compreensão das sentenças sugerem que ambos os grupos conseguiram responder às perguntas sobre as sentenças de

forma comparável. Na resposta às perguntas, portanto, não houve efeitos dos cognatos com relação aos não cognatos.

## 5. Discussão

Os resultados da Tarefa de compreensão de sentenças indicam um efeito da língua minoritária no processamento da leitura em alemão pelos multilíngues do grupo GHA, grupo de falantes de hunsriqueano. Os participantes desse grupo foram mais rápidos na leitura de sentenças com verbos cognatos no tempo verbal *Perfekt* em comparação a sentenças com verbos não cognatos. Por outro lado, o efeito de facilitação para os falantes de hunsriqueano não ocorreu na leitura de sentenças com verbos cognatos no *Präsens* nem na leitura das perguntas.

Os resultados do experimento aplicado neste estudo indicam que a hipótese do acesso lexical não seletivo (de Bot 2004; Dijkstra *et al.* 1999) pode ser mantida para o processamento de sentenças, pois estas foram moduladas para serem interpretadas como de baixa restrição semântica, ou seja, com combinações menos previsíveis de argumentos e verbos. Essa escolha metodológica está em consonância com uma série de estudos que encontraram efeito cognato em sentenças (Bosma & Nota 2020; Duyck *et al.* 2007; Hopp 2017; Lauro & Schwartz 2017; Schwartz *et al.* 2007; Toassi *et al.* 2020).

Além da baixa restrição semântica, outro fator que pode ter contribuído para o efeito de facilitação foi o compartilhamento da estrutura sintática (*Satzklammer*), uma vez que essa estrutura é bastante recorrente no hunsriqueano (Altenhofen *et al.* 2018), inclusive no *Präsens* (perífrases com o verbo auxiliar *tun*). Dessa forma, os verbos no presente preenchem a posição V2 e a posição final da sentença. Em alemão padrão, essa estrutura só ocorre no *Präsens* quando há um verbo modal. Entretanto, o efeito foi encontrado apenas combinado com os verbos cognatos. A facilitação pode resultar de uma espécie de *cognate boost* nas sentenças (Cai *et al.* 2011; Hopp 2017), provocado quando há coexistência das representações lexicais em duas línguas, que pode acarretar uma aceleração no processamento geral da sentença. O compartilhamento da estrutura e da palavra-alvo atuou como uma influência facilitatória (Westergaard *et al.* 2016). A ausência desse efeito nas sentenças escritas no *Präsens* também pode explicar que, para haver um efeito de facilitação, seria necessário o compartilhamento de representações sintáticas e lexicais. No caso de ambas as línguas focalizadas no estudo, nas sentenças no presente, há somente o compartilhamento das representações lexicais, pois a estrutura do presente em alemão (posição V2 preenchida) é diferente da estrutura em hunsriqueano (perífrases com o verbo auxiliar *tun*, com a posição V2 e a posição final preenchidas).

No que concerne às perguntas, todas fechadas (perguntas *ja/nein*), a baixa restrição semântica pode ter sido anulada na compreensão, porque o verbo e as informações argumentais já eram conhecidos pelos participantes. Embora a estrutura das perguntas também seja compartilhada entre as línguas (verbo na primeira posição), a ausência de efeitos cognatos pode ser explicada por um efeito do contexto. Os participantes já tinham lido o verbo na sentença, tanto cognato quanto não cognato, e esse contato pode ter influenciado o efeito de coativação das palavras lidas nas perguntas. O modelo BIA+ (Dijkstra & van Heuven 2002) pode explicar resultados como esse: o contexto da sentença

interage com o reconhecimento da palavra-alvo. Então, informações contextuais podem facilitar o acesso lexical inclusive de palavras não cognatas.

Por fim, cabe discutir os efeitos da tarefa selecionada para este estudo. Da mesma forma que nos estudos com palavras isoladas, tarefas e técnicas de investigação diferentes podem provocar resultados diferentes (Allen *et al.* 2021; Bultena *et al.* 2014). Então, a aplicação da Tarefa de compreensão de sentenças com o monitoramento dos movimentos oculares poderia resultar em outros efeitos, como a facilitação também na leitura de sentenças no *Präsens* e/ou nas perguntas. No entanto, não dispúnhamos dessa ferramenta, e permanece uma lacuna para pesquisas futuras.

A hipótese para este experimento previa que falantes de hunsriqueano (grupo GHA) teriam médias de TR mais baixas e níveis de acurácia mais elevados no processamento de sentenças com verbo cognato em alemão, em comparação ao processamento de sentenças com verbos não cognatos e em comparação ao desempenho de não falantes de hunsriqueano (grupo GA). Esse efeito seria motivado pela presença dos verbos cognatos nas sentenças e pelo compartilhamento da estrutura sintática pelas duas línguas. Os resultados do experimento confirmaram parcialmente a hipótese. O grupo GHA apresentou efeito cognato, previsto na hipótese, para a leitura de sentenças no *Perfekt*, motivado pelo compartilhamento da estrutura e pela presença de verbos cognatos nas sentenças. A hipótese também versa sobre diferenças entre os grupos, que apareceram somente no tempo de processamento das sentenças, sendo o grupo GA mais rápido nas sentenças no *Präsens* e no *Perfekt* sem verbo cognato. Esses resultados indicam níveis de competição linguística maiores para os falantes do grupo GHA, que postergaram, um pouco, o tempo de processamento de sentenças. Além disso, tais resultados podem estar relacionados com variação individual dentro dos grupos. Embora as médias dos grupos estejam pareadas com relação às principais características linguísticas e cognitivas, a sua mensuração tem limitações, e há variação em todas as características (por exemplo, na idade e na escolaridade). Foi usado apenas um instrumento para avaliar a memória de trabalho e o controle inibitório, e o conjunto de instrumentos para caracterizar o nível de proficiência na língua estrangeira não é tão preciso, porque o teste de proficiência foi realizado antes da realização da presente pesquisa. Nesse sentido, fatores individuais (linguísticos e cognitivos) também podem ter influenciado o desempenho.

Os resultados da Tarefa de compreensão de sentenças demonstraram diferenças consistentes entre os grupos na leitura de sentenças. O grupo GA foi mais rápido em comparação ao grupo GHA para ler as sentenças. As diferenças entre os grupos foram verificadas mesmo com o controle dos estímulos e o pareamento dos grupos quanto a critérios linguísticos. São raros os estudos que comparam dois ou três grupos de bilíngues/multilíngues quanto a níveis de proficiência, idade de aquisição ou *background* linguístico; os únicos que encontramos sobre o acesso lexical com sentenças foram os estudos conduzidos por (Hopp 2017; Toassi *et al.* 2020). No primeiro estudo, o autor não encontrou diferenças entre os grupos de falantes inglês como L2 ou L3 nas medidas de rastreamento ocular. As diferenças ocorreram somente entre os falantes tardios de inglês em comparação a falantes nativos. No segundo estudo, as autoras encontraram diferenças

entre os grupos relacionadas ao nível de proficiência na L2: os níveis de coativação foram maiores para o grupo de bilíngues menos proficientes.

No presente estudo, acreditamos que níveis mais altos de competição linguística pelo GHA em comparação ao grupo GA possam explicar a diferença entre os grupos. Os falantes de hunsriqueano possuem dois subsistemas linguísticos semelhantes, que competem entre si, uma vez que as alternativas lexicais permanecem disponíveis para a seleção (Linck *et al.* 2008; Rodriguez-Fornells *et al.* 2002). Segundo a hipótese de Hopp (2017), quando os verbos são os alvos da sentença, e por serem de mais difícil processamento (Van Assche *et al.* 2013), ainda mais em sentenças de baixa restrição semântica, é possível que haja uma capacidade menor de inibir a sintaxe da L1 durante o *parsing* da L2/L3. Como consequência, o fluxo de ativação da sintaxe na L1 se tornaria visível por meio de diferenças no tempo de processamento. A completa supressão da L1 tipologicamente semelhante seria, segundo os modelos de ativação (de Bot 2004; Dijkstra & van Heuven 2002), impossível, o que pode explicar também a influência da L1 e as diferenças entre os grupos.

Os resultados da comparação entre os grupos podem ser explicados, ainda, à luz do Modelo de controle inibitório (Green 1998). Construtos cognitivos, especificamente, o controle inibitório, operam para controlar a língua irrelevante na Tarefa de compreensão de sentenças. Uma hipótese de explicação é que, devido à semelhança entre as línguas, esses mecanismos atuam um pouco mais tarde nos falantes de hunsriqueano e alemão, em comparação a falantes de um par linguístico com línguas não tão semelhantes (português e alemão, por exemplo). Então, as similaridades entre as línguas podem ter prolongado, ligeiramente, o tempo de processamento das perguntas e sentenças pelo grupo falante de hunsriqueano.

Da mesma forma que em outros experimentos que envolvem línguas minoritárias (Bosma & Nota 2020; Limberger 2021), a escassez de estudos que comparam diferentes perfis de multilíngues dificulta a discussão, tornando-se imprescindível considerar várias possibilidades.

## 8. Conclusão

Este estudo mostrou um efeito do compartilhamento de representações semânticas, fonológicas (acesso lexical não seletivo) e sintáticas (coativação sintática) entre hunsriqueano e alemão no processamento de sentenças em alemão. Os tempos de leitura de sentenças foram modulados pelo conhecimento em hunsriqueano, ou seja, o conhecimento da língua minoritária influenciou o tempo de leitura de sentenças com *Satzklammer*.

Por meio desses resultados, visamos contribuir com a pesquisa em línguas minoritárias/ágrafas e processamento da leitura. Via de regra, as investigações sobre efeitos do compartilhamento de representações no processamento privilegiam as línguas majoritárias que possuem ortografia padronizada. O estudo de cunho psicolinguístico sobre línguas minoritárias pode oferecer uma contribuição relevante para entendimento do multilinguismo em suas diversas formas e exibir outras facetas da interação entre a

oralidade e a escrita, entre as representações fonológicas e ortográficas durante o processamento linguístico.

Lacunas no presente estudo estão relacionadas com a falta de *corpora* e materiais escritos em hunsriqueano, o que dificulta a seleção de estímulos e a torna menos precisa. Além disso, obstáculos na coleta de dados (falta de um laboratório para a coleta), na descrição dos participantes e na falta de equipamento de rastreamento ocular devem ser considerados na generalização dos resultados.

Por fim, almejamos que este estudo possa ter implicações para o ensino de alemão como língua estrangeira. As representações semânticas e fonológicas compartilhadas (cognatos) e estruturas sintáticas coexistentes em hunsriqueano e alemão, que acarretam efeitos de facilitação no processamento, podem ser aproveitadas pelo professor que possui alunos falantes de uma variedade alemã.

**Financiamento:** Esta pesquisa recebeu financiamento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasil), pela bolsa de doutorado no Brasil. Também agradeço ao Serviço Alemão de Intercâmbio Acadêmico (DAAD) e à CAPES, pela bolsa de doutorado no exterior. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

**Agradecimentos:** Agradeço a todos os voluntários que participaram deste estudo, por terem compartilhado as suas histórias e os seus desempenhos. Também agradeço aos meus orientadores, Prof. Dr. Augusto Buchweitz e Profa. Dra. Evelyn Ferstl, e à Aline Behling Duarte pela leitura crítica deste artigo. Agradeço, ainda, aos avaliadores anônimos cujas contribuições aprimoraram o texto.

## Referências

- Abutalebi, J., Keim, R., Brambati, S. M., Tettamanti, M., Cappa, S. F., De Bleser, R., & Perani, D. (2007). Late acquisition of literacy in a native language. *Human Brain Mapping*, 28(1), 19–33. <https://doi.org/10.1002/hbm.20240>
- Allen, D., Conklin, K., & Miwa, K. (2021). Cross-linguistic lexical effects in different-script bilingual reading are modulated by task. *International Journal of Bilingualism*, 25(1), 168–188. <https://doi.org/10.1177/1367006920943974>
- Altenhofen, C. V., Frey, J., Käfer, M. L., Klassmann, M. S., Neumann, G. R., & Spinassé, K. P. (2007). Fundamentos para a escrita do Hunsrückisch falado no Brasil. *Revista Contingentia*, 2(1), 73–87.
- Altenhofen, C. V. (1996). *Hunsrückisch in Rio Grande do Sul: Ein Beitrag zur Beschreibung einer deutschbrasilianischen Dialektvarietät im Kontakt mit dem Portugiesischen*. Stuttgart: Steiner.
- Altenhofen, C. V. (2010). Dachsprachenwechsel und Varietätenabgrenzung im Kontakt zwischen Hunsrückisch und Portugiesisch in Brasilien. In F. Boller (Ed.), *Festschrift für Harald Thun zum 60. Geburtstag* (pp. 1–26). Kiel: Westensee-Verlag.
- Altenhofen, C. V. (2013). Bases para uma política linguística das línguas minoritárias no Brasil. In C. Nicolaidis, K. A. da Silva, R. Tilio, & C. H. Rocha (Eds.), *Política e Políticas Linguísticas* (pp. 93–116). Campinas: Pontes Editores.
- Altenhofen, C. V., Prediger, A., & Habel, J. M. (2018). A escrita do Hunsrückisch. In: C. V. Altenhofen, G. R. Neumann, J. M. Habel, & A. Prediger (Eds.), *Hunsrückisch em prosa &*

- verso (pp. 23–34). Rio Grande do Sul: Editora do Instituto de Letras - UFRGS.
- Arêas da Luz Fontes, A. B., de Souza Brentano, L., Toassi, P. F., Sittig, C., & Finger, I. (2021). Evidence of non-selective lexical access in children from a Portuguese-English bilingual school. *PROLÍNGUA*, 15(2), 183–197. <https://doi.org/10.22478/ufpb.1983-9979.2020v15n2.54901>
- Assche, E. V., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual word recognition in a sentence context. *Frontiers in Psychology*, 3, 174. <https://doi.org/10.3389/fpsyg.2012.00174>
- Barbosa Pinto, N., & Arêas da Luz Fontes, A. B. (2020). O acesso lexical em falantes multilíngues português-inglês-italiano. *Veredas - Revista de Estudos Linguísticos*, 24(1), 291–316. <https://doi.org/10.34019/1982-2243.2020.v24.30512>
- Barcelos, L. (2016). *O acesso lexical em trilíngues brasileiros falantes de português, inglês e francês* (Dissertação de mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil).
- Bellmann, G. (1983). Probleme des Substandards im Deutschen. In K. Mattheier (Ed.), *Aspekte der Dialekttheorie* (pp. 105–130). Tübingen: Niemeyer.
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(01), 3–11. <https://doi.org/10.1017/S1366728908003477>
- Blank, C. A. (2013). *A influência grafo-fônico-fonológica na produção oral e no processamento de priming em multilíngues: uma perspectiva dinâmica* (Tese de doutorado, Universidade Católica de Pelotas, Porto Alegre, Brasil).
- Bosma, E., & Nota, N. (2020). Cognate facilitation in Frisian-Dutch bilingual children's sentence reading: An eye-tracking study. *Journal of Experimental Child Psychology*, 189, 104699. <https://doi.org/10.1016/j.jecp.2019.104699>
- Brybaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, 58(5), 412–424. <https://doi.org/10.1027/1618-3169/a000123>
- Bultena, S., Dijkstra, T., & van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *Quarterly Journal of Experimental Psychology*, 67(6), 1214–1241. <https://doi.org/10.1080/17470218.2013.853090>
- Cai, Z. G., Pickering, M. J., Yan, H., & Branigan, H. P. (2011). Lexical and syntactic representations in closely related languages: Evidence from Cantonese–Mandarin bilinguals. *Journal of Memory and Language*, 65(4), 431–445. <https://doi.org/10.1016/j.jml.2011.05.003>
- Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, 13(4), 212–214. <https://doi.org/10.3758/BF03335062>
- Conselho da Europa. (2005). *Quadro Europeu Comum de Referência para Línguas*. Porto: Asa Editores.
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3(4), 422–433. <https://doi.org/10.3758/BF03214546>
- de Bot, K. (2004). The multilingual lexicon: Modelling selection and control. *International Journal of Multilingualism*, 1(1), 17–32. <https://doi.org/10.1080/14790710408668176>
- de Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, 30(1), 90–123. [https://doi.org/10.1016/0749-596X\(91\)90012-9](https://doi.org/10.1016/0749-596X(91)90012-9)
- Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41(4), 496–518. <https://doi.org/10.1006/jmla.1999.2654>
- Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, 62(3), 284–301. <https://doi.org/10.1016/j.jml.2009.12.003>
- Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5(3), <https://doi.org/10.1017/S1366728902003012>

- Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford: Oxford University Press.
- Duarte, A. B., & Blank, C. A. (2019). A influência do priming grafo-fônico-fonológico em uma tarefa de decisão lexical em multilíngues falantes de português (L1), inglês (L2) e francês (L3). *Caderno de Letras*, 35, 13–25. <https://doi.org/10.15210/cdl.v0i35.17189>
- Duden. (2005). *Die Grammatik: Unentbehrlich für richtiges Deutsch*. Mannheim: Dudenverlag.
- Duyck, W., Assche, E. V., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33(4), 663–679. <https://doi.org/10.1037/0278-7393.33.4.663>
- Flores, C., Rinke, E., & Wagner, C. (2020). To hón ich imma insistieat. *Linguistic Approaches to Bilingualism*. <https://doi.org/10.1075/lab.20041.flo>
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67. <https://doi.org/10.1017/S1366728998000133>
- Heeringa, W. (2004). *Measuring dialect pronunciation differences using Levenshtein distance* (Ph.D. Dissertation, University of Groningen, Groningen, The Netherlands).
- Helbig, G., & Schenkel, W. (1991). *Wörterbuch Valenz und Distribution deutscher Verben* (8.). Tübingen: Max Niemeyer Verlag.
- Hopp, H. (2017). Cross-linguistic lexical and syntactic co-activation in L2 sentence processing. *Linguistic Approaches to Bilingualism*, 7(1), 96–130. <https://doi.org/10.1075/lab.14027.hop>
- IBM Corp. (2010). Released 2010. IBM SPSS Statistics for Windows (Version 19.0.) [Computer software]. New York: IBM Corp.
- Kirk, N. W., Kempe, V., Scott-Brown, K. C., Philipp, A., & Declerck, M. (2018). Can monolinguals be like bilinguals? Evidence from dialect switching. *Cognition*, 170, 164–178. <https://doi.org/10.1016/j.cognition.2017.10.001>
- Kramer, R. (2011). *Effects of bilingualism on inhibitory control and working memory: a study with early and late bilinguals* (Dissertação de mestrado, Universidade Federal de Santa Catarina, Florianópolis, Brasil).
- Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism Psycholinguistic Approaches* (pp. 531–553). Oxford: Oxford University Press.
- Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language*, 92, 217–233. <https://doi.org/10.1016/j.jml.2016.06.010>
- Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition*, 32(4), 533–550. <https://doi.org/10.3758/bf03195845>
- Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: evidence from eye movements during reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 35(2), 381–390. <https://doi.org/10.1037/a0014875>
- Limberger, B. K. (2018). *Processamento da leitura e suas bases neurais: um estudo sobre o hunsriqueano*. (Tese de doutorado, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Brasil). <https://doi.org/10.13140/RG.2.2.28790.42563>
- Limberger, B. K. (2021). Leitura de palavras em língua minoritária: a construção do léxico ortográfico em hunsriqueano. *DELTA*, 37(2), 1–26. <https://doi.org/10.1590/1678-460X202148072>
- Limberger, B. K., & Buchweitz, A. (2014). The effects of bilingualism and multilingualism on executive functions. *Fórum Linguístico*, 11(3), 261–277. <https://doi.org/10.5007/1984-8412.2014v11n3p261>
- Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The Mental Lexicon*, 3(3), 349–374. <https://doi.org/10.1075/ml.3.3.06lin>
- Machado, L. L. (2016). *Standard e substandard em contato com o português: variação na*

- competência de fala em Hochdeutsch de falantes de Hunsrückisch* (Dissertação de mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil)
- Ma, Y., Mazumdar, M., & Memtsoudis, S. G. (2012). Beyond repeated-measures analysis of variance: advanced statistical methods for the analysis of longitudinal data in anesthesia research. *Regional Anesthesia and Pain Medicine*, 37(1), 99–105. <https://doi.org/10.1097/AAP.0b013e31823ebc74>
- Messa, R. M. (2009). *O papel do dialeto no aprendizado do alemão padrão* (Dissertação de mestrado, Universidade do Vale do Rio dos Sinos, São Leopoldo, Brasil).
- Möller, R. (2011). Wann sind Kognaten erkennbar? Ähnlichkeit und synchrone Transparenz von Kognatenbeziehungen in der germanischen Interkomprehension. *Linguistik Online*, 46(2), 79–101.
- Morello, R. (2021). Lista de línguas cooficiais em municípios brasileiros. Consultado em <http://ipol.org.br/lista-de-linguas-cooficiais-em-municipios-brasileiros/>
- Palma, P., & Titone, D. (2020). Bilingual lexical access and reading. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual lexical ambiguity resolution* (pp. 159–183). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316535967.008>
- Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2, 10. <https://doi.org/10.3389/neuro.11.010.2008>
- Pickbrenner, M. (2017). *O acesso lexical em multilíngues inglês (L2) e alemão (L3) no reconhecimento de palavras tipologicamente similares* (Tese de Doutorado, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil).
- Rodriguez-Fornells, A., Rotte, M., Heinze, H.-J., Nösselt, T., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, 415(6875), 1026–1029. <https://doi.org/10.1038/4151026a>
- Rumelhart, D., & McClelland, J. L. (1981). Interactive Processing Through Spreading Activation. In A. M. Lesgold & C. Perfetti (Eds.), *Interactive processes in reading* (pp. 37–60). Hillsdale, NJ: Lawrence Erlbaum.
- Scholl, A. P., & Finger, I. (2013). Elaboração de um questionário de histórico de linguagem para pesquisas com bilíngues. *Nonada Letras Em Revista*, 2(21), 1–17.
- Schröter, P., & Schroeder, S. (2016). Orthographic processing in balanced bilingual children: Cross-language evidence from cognates and false friends. *Journal of Experimental Child Psychology*, 141, 239–246. <https://doi.org/10.1016/j.jecp.2015.09.005>
- Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, 22(1), 106–129. <https://doi.org/10.1080/01690960500463920>
- Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, 55(2), 197–212. <https://doi.org/10.1016/j.jml.2006.03.004>
- Sherkina-Lieber, M. (2004). The Cognate Facilitation Effect in Bilingual Speech Processing: The Case of Russian-English Bilingualism. *Cahiers Linguistiques d'Ottawa*, 32, 108–121.
- Spinassé, K. P. (2005). *Deutsch als Fremdsprache in Brasilien: Eine Studie über kontextabhängige unterschiedliche Lernersprachen und muttersprachliche Interferenzen*. Frankfurt: Peter Lang.
- Steffen, J., & Altenhofen, C. V. (2014). Spracharchipele des Deutschen in Lateinamerika: Dynamik der Sprachvernetzungen im mehrsprachigen Raum. *Zeitschrift Für Dialektologie Und Linguistik*, 81(1), 34–60.
- Steffen, J. (2008). A vantagem de falar dialeto: Aproveitar as variedades não-padrão para a construção de comunidades multilíngües. *Revista Contingentia*, 3(2), 67–76.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. <https://doi.org/10.1037/h0054651>
- Szubko-Sitarek, W. (2012). Cognate facilitation effects in trilingual word recognition. *English*, 1(2), 189–208.
- Toassi, P. F. P., Mota, M. B., & Teixeira, E. N. (2020). O efeito de palavras cognatas no acesso lexical do inglês como terceira língua. *Cadernos de Tradução*, 40(esp2), 74–96. <https://doi.org/10.5007/2175-7968.2020v40nesp2p74>
- Toassi, P. F. P., & Mota, M. B. (2015). Acesso lexical de bilíngues e multilíngues. *Acta*

- Scientiarum. Language and Culture*, 37(4), 393.  
<https://doi.org/10.4025/actascilangcult.v37i4.25234>
- Van Assche, E., Duyck, W., & Brysbaert, M. (2013). Verb processing by bilinguals in sentence contexts. *Studies in Second Language Acquisition*, 35(2), 237–259.  
<https://doi.org/10.1017/S0272263112000873>
- Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9(4), 780–789. <https://doi.org/10.3758/bf03196335>
- van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, 39(3), 458–483.  
<https://doi.org/10.1006/jmla.1998.2584>
- Verband Deutscher Vereine. (1999). *Cem anos de germanidade no Rio Grande do Sul*. São Leopoldo: Editora Unisinos.
- Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2016). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism*, 21(6), 666–682.  
<https://doi.org/10.1177/1367006916648859>
- Wood, G. M. de O., Carvalho, M. R. S., Rothe-Neves, R., & Haase, V. G. (2001). Validação da Bateria de Avaliação da Memória de Trabalho (BAMT-UFMG). *Psicologia: Reflexão e Crítica*, 14(2), 325–341. <https://doi.org/10.1590/S0102-79722001000200008>

[recebido em 24 de outubro de 2021 e aceite para publicação em 18 de março de 2022]

## APRENDIZAGEM DE DEPENDÊNCIAS NÃO ADJACENTES POR ADULTOS EM UMA LÍNGUA ARTIFICIAL: EVIDÊNCIAS EXPERIMENTAIS

### LEARNING AND PROCESSING NON-ADJACENT DEPENDENCIES BY ADULTS IN AN ARTIFICIAL LANGUAGE: EXPERIMENTAL EVIDENCE

Cristina Name\*  
cristina.name@ufjf.br

Késsia Henrique\*\*  
kessiasilvahenrique@gmail.com

Investigamos a aprendizagem de dependências não adjacentes por adultos (DNAs) em uma língua artificial. DNAs é a coocorrência de marcadores morfossintáticos e/ou fonológicos com material interveniente, como em *These cat-s*, em que observamos a dependência entre o determinante e a marcação de plural no nome e a base *cat-* entre eles. A aprendizagem de DNAs por crianças tem sido amplamente estudada, mas pouco se sabe sobre como essa aprendizagem ocorre em adultos com uma L1 consolidada. Usando uma combinação de técnicas experimentais, foi conduzido um experimento de aprendizagem de língua artificial. Trinta e quatro adultos participaram divididos entre grupo experimental e grupo controle. Na fase de treinamento, que durou 3 minutos, os participantes do grupo experimental ouviram combinações de DPs inventados, enquanto participantes do grupo controle ouviram estímulos musicais. Na fase teste, participantes de ambos os grupos tiveram que escolher, entre dois novos nomes, o que combinava melhor com o determinante. Houve diferença estatisticamente significativa entre os grupos experimental e controle, com maior número de acertos no primeiro, o que é sugestivo de maior precisão nas respostas dadas nesse grupo. Esses resultados sugerem que, mesmo após um curto período de exposição aos estímulos, os participantes foram capazes de abstrair os padrões a partir dos estímulos do treinamento e aplicá-los a estímulos inéditos. Esses achados foram interpretados como evidência de aprendizagem de DNAs por adultos.

**Palavras-chave:** Dependências não adjacentes. Aprendizagem de língua artificial. Processamento adulto.

We investigate non-adjacent dependency (NAD) adult learning in an artificial language. NAD is the co-occurrence of morphosyntactic and/or phonological markers with intervening material, as in *These cat-s*, where we observe the dependency between the determiner and the noun plural

---

\* Departamento de Letras, Faculdade de Letras, Universidade Federal de Juiz de Fora, Juiz de Fora, Brasil. Bolsista de Produtividade em Pesquisa do CNPq. ORCID: 0000-0003-3879-3766

\*\* Departamento de Letras, Faculdade de Letras, Universidade Federal de Juiz de Fora, Juiz de Fora, Brasil. Bolsista CAPES. ORCID: 0000-0001-5625-9503

mark, and the root cat- between them. NAD learning by children has been widely studied, but little is known about how it occurs for adults having consolidated their L1 knowledge. Using a combination of experimental techniques, an artificial language learning experiment was conducted. Thirty-four adults were assigned to either the experimental group or the control group. In the training phase, which lasted 3 minutes, participants in the experimental group heard invented DP, whereas the participants in the control group heard music sound stimuli. In the test phase, both groups had to choose, between two new nouns, the one that matched the given determiner. There was a significant difference between the experimental and control groups, with the highest number of correct answers in the former, indicating greater precision in the answers given by this group. These results suggest that even after a short period of exposure to the stimuli, participants were able to abstract patterns from the stimuli and apply them to new stimuli. These findings were interpreted as evidence of NAD learning by adults.

**Keywords:** Non-adjacent dependencies. Artificial language learning. Adult processing.

•

## 1. Introdução

A complexidade e variedade dos mecanismos de concordância existentes na língua vêm sendo amplamente debatidas na literatura, tanto na proposição de modelos de processamento desse mecanismo (Henrique & Name 2020; Rodrigues 2005, 2006; Vigliocco 1996; Violti 2005), quanto no entendimento desse fenômeno em uma concepção gerativista (Chomsky 2000, 2001). A ideia de que a concordância pode ser entendida como “o fenômeno gramatical no qual a forma de uma palavra numa sentença ou sintagma é determinada pela forma de uma palavra com a qual tem alguma ligação gramatical” (Trask 2004, p. 61) é bastante consensual na literatura. Do ponto de vista do mecanismo em si, e assumindo-se uma abordagem teórica minimalista, entende-se concordância como um processo geral, baseado na distinção entre traços (interpretáveis e não-interpretáveis) e na relação entre sonda e alvo (Chomsky 2000, 2001). Considerando-se suas manifestações, a concordância pode ser descrita como uma covariação sistemática entre propriedades formais e/ou semânticas entre dois elementos (Corbett 2003). As duas perspectivas – de um lado, o caráter abstrato do fenômeno, de outro, a manifestação morfofonológica sistemática decorrente dessa operação – são igualmente fundamentais para o entendimento da aprendizagem de dependências não adjacentes, foco deste artigo.

Dependências não adjacentes (doravante, DNA) são padrões de coocorrência de marcadores morfosintáticos e/ou morfofonológicos<sup>1</sup> entre elementos, apresentando material interveniente (Laguardia 2016; Name & Laguardia 2013; Sandoval & Gómez 2013;) e expressando regras morfosintáticas abstratas de concordância. Nos exemplos

---

<sup>1</sup> Agradecemos a contribuição de um dos pareceristas ao observar que dependências não adjacentes vão além das manifestações de concordância, ocorrendo também, por exemplo, nas passivas verbais no português, entre o auxiliar e particípio do verbo. Ver, por exemplo, Lima Júnior & Corrêa (2018).

em 1, temos a concordância nominal de gênero (1a) e de número (1b) entre determinante e nome, e a concordância verbal de número e pessoa entre pronome e verbo (1c).

(1)

- a. **La** *niñ-a*
- b. **These** *book-s*
- c. **Nós** *compra-mos*

Nesses exemplos, podemos observar elementos codependentes (em negrito) separados por material interveniente (em itálico), em uma configuração AXB, sendo A e B variáveis correlacionadas. Em outros termos, tal codependência se dá entre categorias, indo além dos elementos concretos que a manifestam. Nesse sentido, a aquisição de DNAs pressupõe tanto o reconhecimento dos marcadores morfossintáticos, quanto a identificação do padrão da relação subjacente entre as categorias ou classes que eles compõem. A aprendizagem de DNAs ganhou especial atenção a partir de evidência empírica de que mecanismos de aprendizagem estatística estariam disponíveis desde o início da vida dos bebês e seriam usados para detecção de regularidades que constituem as línguas naturais, tais como padrões fonotáticos e relações preditivas, tanto por bebês quanto por adultos (Saffran 2001; Thiessen, Girard & Erickson 2016). Focalizando inicialmente o tratamento probabilístico (Saffran, Aslin & Newport 1996) e a abstração de regras algébricas subjacentes a dependências adjacentes (Marcus, Vijayan, Bandi Rao & Vishton 1999; Peña, Bonatti, Nespor & Mehler 2002) por bebês e adultos, as pesquisas logo se voltaram para as DNAs, que são amplamente encontradas nas línguas naturais, refletem dependências morfossintáticas e, portanto, seriam cruciais para a aquisição e o processamento sintático.

Este artigo focaliza a aprendizagem de DNAs por adultos, buscando compreender de que modo, após terem a aquisição de uma língua estabilizada, adultos processam as DNAs em uma língua totalmente desconhecida. Ainda, investiga se eles são capazes de aprender um padrão de combinação e aplicar a regra aprendida a estímulos inéditos. Nas seções a seguir, faremos um breve panorama de estudos sobre aquisição de DNAs, particularmente, aqueles com foco na aprendizagem de DNAs por adultos. Como será visto, de modo geral, os experimentos com adultos fazem uso de tarefas de aprendizagem explícita de DNAs. Em seguida, apresentaremos uma atividade experimental utilizando o paradigma de aprendizagem de língua artificial com medidas *online* e *offline*, desenvolvida com o objetivo de investigar a sensibilidade de adultos às DNAs, bem como sua aprendizagem. A atividade buscou promover a aprendizagem implícita de regras, de forma a se aproximar mais de uma situação natural de aquisição de língua. Por fim, concluímos retomando contribuições que o estudo pode oferecer para a pesquisa sobre esse tópico e apontamos alguns pontos para investigação futura.

## 2. Aprendizagem de DNAs – uma breve revisão

Estudos relativos à aprendizagem de DNAs por bebês apontam para uma sensibilidade aos elementos codependentes da língua sendo adquirida em meados do segundo ano de

vida. Santelmann e Jusczyk (1998) observaram que crianças de 18 meses adquirindo inglês americano reconhecem a dependência morfosintática entre *is* e *-ing*, em uma janela de até três sílabas intervenientes. Van Heugten & Shi (2010) verificaram que, aos 17 meses, bebês adquirindo francês canadense são capazes de identificar DNAs dispostas em sintagmas fonológicos distintos, i.e., separados por uma fronteira prosódica.

A partir de pesquisas acerca de mecanismos de aprendizagem estatística desenvolvidas nos anos 1990 (Gómez & Gerken 1999; Marcus, Vijayan, Bandi Rao & Vishton 1999; Saffran, Aslin, & Newport 1996; entre outros), uma série de estudos foi conduzida buscando-se investigar mais especificamente a aprendizagem de novas DNAs, usando-se, para isso, línguas artificiais. O trabalho pioneiro de Gómez (2002) se voltou para o reconhecimento de DNAs, assim como para a generalização de diferentes padrões de DNAs em outros estímulos, por bebês de 18 meses e adultos. Segundo a autora, DNAs são mais dificilmente detectáveis do que dependências adjacentes, de modo que uma variação maior dos elementos intervenientes poderia facilitar o reconhecimento das DNAs. Para investigar essa assunção, foram criadas duas línguas artificiais, com as seguintes configurações: língua 1 –  $aXd, bXe, cXf$ ; língua 2 –  $aXe, bXf, cXd^2$ . Foram criados três grupos experimentais, em função da quantidade de possibilidades para o material interveniente (X), que variava 3, 12 ou 24 vezes. Os bebês foram testados com a técnica de Escuta Preferencial (*Head-Turn Preference Procedure*)<sup>3</sup> e apenas o grupo exposto a 24 elementos intervenientes apresentou efeito de novidade, com tempo médio de escuta significativamente maior para a condição agramatical (i.e., cujos estímulos feriam os padrões da língua apresentada na etapa de familiarização, língua 1 ou 2). Esses resultados sugerem que maior variação do conjunto de elementos intervenientes (X) levaria ao destaque dos elementos codependentes não adjacentes previamente familiarizados (*a-d, b-e etc.*), o que permitiria seu reconhecimento por bebês aos 18 meses de idade. Em seguida, Gómez e Maye (2005) testaram bebês de 12, 15 e 17 meses, usando os mesmos estímulos do experimento anterior (Gómez 2002) e variando o número de elementos intervenientes em 12, 18 e 24. Os resultados indicaram que bebês aos 15 meses são capazes de identificar as DNAs em contexto de 24 variações do elemento X e, aos 17 meses, em contexto de apenas 18 variações de X. Os resultados dos bebês de 12 meses não foram significativos em nenhum dos contextos. Em conjunto, os resultados apontam para uma capacidade de reconhecimento de diferentes pares de elementos codependentes a partir de 15 meses, sendo a variação do elemento interveniente um fator de facilitação.

Esses estudos apresentam evidência de que bebês no segundo ano de vida aprendem rapidamente combinações de coocorrência de elementos não adjacentes, mas não fica claro se são capazes de aprender quando há mais de uma possibilidade de elementos codependentes, como acontece nas línguas naturais. Em outros termos, não há evidência de que o bebê seja capaz de apreender a regra subjacente manifesta pelas DNAs, abstraí-la e generalizá-la em novos estímulos, o que caracterizaria, efetivamente, a aprendizagem da regra. Para investigar essa questão, Name e colaboradoras (Laguardia, 2016; Laguardia

---

<sup>2</sup> No experimento de Gómez (2002), as letras correspondem aos seguintes elementos: a=pel; b=vot; c=dak; d=rud; e=jic; f=tood. O X corresponde ao material interveniente entre as dependências.

<sup>3</sup> Em português europeu, denominada Movimento Preferencial da Cabeça (cf. Frota & Name, 2017).

*et al.* 2015a, 2015b; Laguardia, Name & Shi 2013, 2015; Name, Shi & Kouluaguina 2011), conduziram uma série de experimentos com bebês de 11 meses, que buscaram verificar (i) a identificação e a generalização de DNAs entre determinantes e nomes no DP; (ii) o papel de fronteiras prosódicas na aquisição de DNAs; e (iii) a capacidade de categorização de novos nomes com base na aquisição de DNAs. As autoras criaram estímulos usando determinantes masculinos do português brasileiro combinados a pseudopalavras terminadas em  $-e$  e  $-a^4$ . Da combinação desses elementos, surgiram duas classes de nomes: *esse* e *meu* foram combinados a nomes terminados em  $-a$  e *nosso* e *seu* foram combinados a nomes terminados em  $-e^5$ , totalizando 16 sintagmas, apresentados continuamente durante 2 minutos na fase de familiarização. No teste, os bebês escutaram sintagmas formados por quatro novos nomes, não previamente familiarizados, em configurações de DNA congruentes ou incongruentes às regras de formação dos estímulos da fase de familiarização. Em conjunto, os resultados dos experimentos sugerem que bebês aos 11 meses, após breve exposição a estímulos de uma (pseudo)língua, foram sensíveis a DNAs entre determinantes e terminações de nomes, fizeram uso de fronteiras de sintagma fonológico para adquirir regularidades morfossintáticas e foram capazes de classificar novos nomes em subcategorias gramaticais, com base nos padrões de DNAs.

No que se refere à aprendizagem de DNAs por adultos, Gómez (2002) apresentou os mesmos estímulos usados com os bebês, com a diferença de que o material interveniente poderia variar 2, 6, 12 ou 24 vezes. Os participantes foram treinados com estímulos de uma das duas línguas, nos quatro padrões criados (ver nota 1) durante 18 minutos. Ao final, os participantes eram informados de que os estímulos haviam sido criados com base em regras de ordem de palavras. Na fase de teste, foram apresentados aleatoriamente doze DNAs, sendo metade já conhecidos do treinamento, e metade com padrão da outra língua não familiarizada, e os participantes deveriam identificar se o estímulo seguia ou não as regras da língua familiarizada, apertando teclas distintas. Os participantes acertaram mais vezes quando o estímulo era congruente com o que foi exposto no treinamento, e os participantes expostos ao grupo de 24 possibilidades de variação do material interveniente obtiveram melhores resultados em comparação aos demais grupos. De acordo com a autora, esses resultados sugerem que o aumento da variedade do material interveniente foi um facilitador da aprendizagem de DNAs pelos adultos. No entanto, os estímulos-alvo já tinham sido apresentados durante o treinamento e a outra metade era de novos estímulos, de modo que esse desequilíbrio de familiaridade pode ter interferido na escolha, que não refletiria necessariamente uma aprendizagem.

Onnis, Monaghan, Christiansen & Chater (2004) desenvolveram um experimento de aprendizagem de língua artificial utilizando os mesmos estímulos e mesma técnica experimental de Gómez (2002). A diferença é que os autores utilizaram estímulos inéditos

---

<sup>4</sup> Dessa forma, evitou-se identidade fonológica entre as terminações de determinantes e nomes e criou-se uma pseudolíngua. Ainda, os padrões prosódicos do PB foram preservados na produção dos estímulos.

<sup>5</sup> Um grupo de crianças ouviu esses padrões de combinação, e outro grupo de crianças ouviu padrões inversos. Os estímulos e os padrões sofreram algumas modificações nos diferentes experimentos, a fim de se adequarem a objetivos específicos. Para detalhes, ver as referências apresentadas.

na fase teste, de forma que os novos estímulos seguiam as mesmas regras dos estímulos do treinamento, mas eram estímulos que não foram ouvidos antes. Desse modo, neutralizou-se um possível efeito de desequilíbrio entre os estímulos do teste. A intenção dos autores foi investigar se a maior variabilidade do material interveniente também auxiliaria na abstração e generalização das regras de combinação das DNAs. Os resultados encontrados são compatíveis com os resultados de Gómez (2002). Segundo os autores, houve efeito de variação do tamanho do conjunto do elemento interveniente, uma vez que quando havia maior possibilidade de variação do material interveniente, os participantes apresentaram melhor desempenho.

Newport & Aslin (2004) também investigaram a aprendizagem de DNAs por adultos tendo como foco a aprendizagem estatística e as dependências não adjacentes dentro de palavras. Na atividade experimental desenvolvida pelos autores foi criada uma língua artificial contendo 20 palavras trissílabas. Essas palavras eram formadas por um elemento A + elemento B + elemento C. As primeiras e terceiras sílabas eram previsíveis e a segunda era a menos previsível. Os estímulos foram apresentados por áudio e os participantes foram informados de que a língua que ouviriam tinha uma gramática, mas não possuía um significado e obedeciam a um determinado padrão. Na fase de teste foi utilizada uma tarefa de escolha forçada com medida *offline*. Os participantes foram apresentados a pares de palavras e deveriam indicar em uma folha de papel qual palavra se parecia mais com os estímulos do treinamento. Os participantes circulavam na folha se era a palavra 1 ou a palavra 2 e nas análises foi utilizada como variável dependente a taxa de acertos. Os resultados não foram estatisticamente significativos. Segundo os autores, o tamanho da língua inventada pode ter sido muito grande ou de fato haveria uma dificuldade legítima na aprendizagem de DNAs. Em um segundo experimento, Newport & Aslin (2004) utilizaram a mesma metodologia experimental, mas testando DNAs entre segmentos fonéticos, uma vez que o experimento com DNAs entre sílabas não apresentou resultados significativos estatisticamente. No segundo experimento, os resultados foram significativos, o que, de acordo com os autores, sugere que DNAs entre segmentos fonéticos seriam mais facilmente aprendidas em comparação com as sílabas.

Com o passar dos anos e com o maior entendimento sobre o papel das DNAs, surgiu a necessidade de se investigarem outros aspectos que envolviam as DNAs e sua aprendizagem. Van Den Bos, Christiansen e Misyak (2012) investigaram a aprendizagem estatística de DNAs comparando DNAs determinísticas, em que um elemento é sempre precedido de um mesmo elemento, e probabilísticas, em que vários elementos têm uma certa probabilidade de eventualmente seguir outro elemento distinto, como acontece em diversas línguas naturais. Foram criadas duas línguas artificiais, uma contendo DNAs determinísticas e outra contendo DNAs probabilísticas. Antes da fase de treinamento, que durou 19 minutos, os participantes foram instruídos a ouvir atentamente os estímulos, porque o conhecimento deles sobre as sequências apresentadas seria testado. No começo da fase de teste, os participantes foram informados que as sequências que ouviram foram formadas seguindo determinadas regras, que seriam ouvidas oito sequências que seguiam as mesmas regras e oito que violavam essas regras e que eles deveriam julgar se cada sequência seguia as regras do treinamento ou não. O teste utilizou como medida as taxas

de acerto e os resultados mostraram que houve maior taxa de acertos nas dependências determinísticas em comparação com as dependências probabilísticas, o que, segundo os autores, vai ao encontro de resultados anteriores, que sugerem que dependências probabilísticas teriam a aquisição mais demorada e seriam mais difíceis de serem adquiridas.

Uddén, Ingvar, Hagoort e Petersson (2012) desenvolveram um estudo a fim de testar a hipótese segundo a qual a aquisição e o processamento de dependências não adjacentes e de dependências adjacentes seriam de natureza qualitativa ou quantitativa diferentes. Uma diferença qualitativa seria compatível com a ideia de que diferentes arquiteturas de memória estariam envolvidas na aquisição/no processamento de dependências não adjacentes e de dependências adjacentes. A atividade experimental desenvolvida pelos autores buscou superar as limitações apresentadas por estudos anteriores. Nos estímulos foi utilizada uma língua artificial gerando dependências não adjacentes em uma configuração cruzada e dependências adjacentes. Para testar a hipótese proposta, os autores utilizaram uma combinação de tarefas experimentais em uma atividade experimental que teve duração de 9 dias. Primeiramente foi feito um teste de base para servir de comparação depois de todas as sessões de aprendizagem. A atividade experimental teve outras 4 etapas. A primeira foi formada por seções de aquisição implícita, em que cada sequência era apresentada por 4 segundos de forma escrita e, em seguida, os participantes deveriam digitar no teclado qual era essa sequência. Essa etapa durava cerca de 30 min em cada dia e se estendeu por 7 dias. Depois das seções de aprendizagem implícita, foi realizado o que os autores chamaram de “teste de preferência de gramaticalidade”. Os participantes deveriam indicar o mais rápido possível se gostavam ou não de uma sequência baseados apenas na intuição. No último dia do teste essa tarefa foi refeita e os participantes foram informados de que os estímulos seguiam determinadas regras. Na finalização do nono e último dia, os participantes também responderam um questionário após o experimento, que foi formado por perguntas abertas sobre o que acharam do teste e quais seriam as sequências gramaticais. Foi ainda pedido que, se possível, tentassem reproduzir alguma sequência gramatical. Após tentarem se lembrar de alguma combinação gramatical, 10 sentenças geradas pela língua artificial foram apresentadas e os participantes deveriam informar se pertenciam ou não à gramática da língua. Segundo os autores, os resultados mostraram que a aquisição de dependências não adjacentes não teve diferença qualitativa em comparação a dependências adjacentes. Houve diferença quantitativa no que diz respeito ao tempo necessário para aquisição de DNAs, de forma que dependências não adjacentes levaram mais tempo para serem adquiridas.

Além dos resultados do trabalho de Uddén, Ingvar, Hagoort e Petersson (2012) serem muito interessantes, o que mais chama atenção é a combinação de tarefas experimentais. Apesar de ter sido um teste bastante longo, que necessitava de um grande engajamento dos participantes, a atividade experimental foi desenvolvida de modo a minimizar os efeitos de uma possível aquisição explícita das regras. Além disso, com o questionário pós-experimento, os autores conseguiram investigar a consciência dos participantes a respeito dos estímulos e das tarefas experimentais, o que minimiza

possibilidade de uma interpretação alternativa para os resultados ancorada na memorização dos estímulos e não na aprendizagem implícita das regras subjacentes.

Nessa breve seleção de estudos, observamos que a aprendizagem de DNAs parece ocorrer cedo no desenvolvimento linguístico. Ainda, a criança não precisaria de muito tempo de exposição aos estímulos para apreender, ainda que temporariamente, os padrões apresentados. Observamos, também, desafios metodológicos nas pesquisas voltadas para a população adulta. Por exemplo, o excesso de informações fornecidas aos participantes torna a tarefa mais consciente e afasta a possibilidade de investigação de aprendizagem implícita (Gómez 2002; Newport & Aslin 2004). Outra fragilidade adviria de questões relacionadas aos estímulos. Em Gómez (2002), os estímulos usados no teste foram os mesmos usados na familiarização, o que pode ter levado a sua recuperação na memória, e não ao reconhecimento das regras; os estímulos usados por Newport & Aslin (2004) se assemelham pouco a línguas naturais, assim como os de Van Den Bos, Christiansen e Misyak (2012), visto que dependências determinísticas não parecem ocorrer nas línguas naturais. Assim, os resultados podem ter uma interpretação alternativa de que, na verdade, a reação dos participantes se deve a um efeito de recenticidade da memória daquela combinação particular e não da identificação de uma regra subjacente. Ademais, durante o experimento desenvolvido pelos autores, assim como ocorreu em outros trabalhos, foram fornecidas muitas informações, tornando os participantes muito conscientes da atividade e do fenômeno investigado.

A partir de estudos prévios e da necessidade da ampliação das investigações a respeito da aquisição de DNAs, foi desenvolvida uma atividade experimental com o objetivo de investigar em que medida a capacidade de detectar a estrutura/regra subjacente às DNAs, observada em bebês na fase de aquisição (Gómez 2002; Laguardia 2016; Laguardia *et al.* 2015a, 2015b; Laguardia, Name & Shi 2013; Name & Shi 2015; Name, Shi & Koulaguina 2011; Sandoval & Gómez 2013) permanece em adultos que já têm a aquisição de uma L1 consolidada, na aquisição de uma L2.

### 3. Atividade experimental

A presente seção tem por objetivo apresentar a investigação realizada experimentalmente<sup>6</sup>. Conforme observamos na revisão da literatura, há um número expressivo de pesquisas a respeito de DNAs com adultos utilizando metodologia *offline* e com tarefas que envolvem aprendizagem de gramática artificial. Alguns autores criticam esse paradigma; no entanto, Folia *et al.* (2008) e Folia *et al.* (2010) identificaram que a área de Broca está envolvida no processamento sintático também de gramáticas artificiais, fazendo sentido, portanto, investigar-se a aprendizagem em adultos com uma L1 consolidada dessa forma e minimizando-se possíveis interferências semânticas nos estímulos.

A atividade experimental aqui realizada tem como objetivo investigar a aprendizagem de DNAs por adultos em uma língua artificial, assemelhando-se à

---

<sup>6</sup> Esta pesquisa foi aprovada no Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal de Juiz de Fora (UFJF) sob o número: CAAE 44529721.0.0000.5147.

aquisição de L2, de modo a explorar a habilidade de adultos de aprender/processar DNAs após a aquisição de sua L1 consolidada.

Apesar da existência de uma extensa gama de estudos investigando a aquisição de DNAs, o presente trabalho se faz relevante ao buscar superar algumas das limitações encontradas em trabalhos anteriores. Ademais, até onde vai nosso conhecimento, trata-se do primeiro estudo realizado com adultos falantes nativos do PB. Desse modo, buscamos ampliar e contribuir para a discussão acerca da aprendizagem de DNAs, em particular, por adultos.

Como variável independente foi considerada a congruência com a gramática do treinamento, com 2 níveis, sendo eles congruente e incongruente. As variáveis dependentes foram o tempo de escolha em cada *trial* no teste de escolha forçada e a taxa de acertos da palavra alvo.

Na atividade experimental descrita, na fase de aprendizagem implícita para o grupo experimental, os participantes foram expostos por via auditiva a um conjunto de 128 combinações de determinantes e nomes inventados que seguiam um determinado padrão por aproximadamente 3 minutos. Ao mesmo tempo em que ouviam os estímulos, apareciam na tela duas imagens abstratas para que o participante escolhesse a que mais lhe agradava. Na fase do teste, os participantes foram expostos a 16 pares de nomes escritos na tela, inéditos, que não foram ouvidos na fase de treinamento, precedidos pela apresentação sonora de um determinante. Os participantes deveriam escolher qual dos dois nomes se combinava com o determinante escutado seguindo a sua intuição. Tanto os estímulos da fase de aprendizagem quanto os estímulos da fase teste foram aleatorizados pelo próprio site utilizado na programação do teste. Já os participantes do grupo controle, no lugar da fase de aprendizagem implícita, ouviram um conjunto de 128 combinações de sons aleatórios produzidos por um violão, totalizando a mesma duração da fase de aprendizagem do grupo experimental. A fase de teste foi igual para ambos os grupos.

Na Tabela 1, apresentamos exemplos dos estímulos experimentais e a regra de combinação entre os determinantes e nomes inventados.

**Tabela 1. Estímulos experimentais**

DETERMINANTE	KOI / ZIS	DAI / NUR
<b>TERMINAÇÃO FINAL DO NOME</b>	[e] [u]	[a] [o]
<b>Exemplos de estímulos</b>	<b>Koi</b> [3e'tu] <b>Koi</b> [la'ne] <b>Zis</b> [mu'ze] <b>Zis</b> [xo'vu]	<b>Dai</b> [pi'fo] <b>Dai</b> [te'ba] <b>Nur</b> [nu'bo] <b>Nur</b> [zu'va]

Podemos observar pela tabela que a relação de combinação entre determinante e nome inventado não era de um para um, de forma que havia dois determinantes que poderiam se combinar com duas terminações de nomes diferentes e outros dois determinantes que poderiam se combinar com outras duas terminações.

Com a intenção de não adicionar mais uma variável ao *design experimental*, controlamos o número de sílabas nos nomes inventados e a tonicidade, de forma que todas as palavras criadas eram dissílabas e oxítonas. Também foram balanceados os fonemas utilizados nos estímulos em cada sílaba, de forma que todos os fonemas possuíam o mesmo número de ocorrências tanto na primeira quanto na segunda sílaba de todos os nomes utilizados, tanto na fase de aprendizagem implícita quanto na fase teste.

**Previsões:** Nossas previsões foram as seguintes:

- I. Os participantes, após um curto período de tempo, deverão ser capazes de aprender as regras de combinação das DNAs, refletindo em uma porcentagem maior de acertos na tarefa de escolha forçada;
- II. As respostas corretas na tarefa de escolha forçada apresentarão tempos de reação menores, sugerindo que uma reação mais rápida pode indicar a aprendizagem implícita dessas regras;
- III. Os participantes não serão capazes de identificar conscientemente as regras de combinação das DNAs, visto que a primeira etapa do teste envolveu uma tarefa que exigia consciência dos participantes (a escolha de imagens abstratas), minimizando a atenção consciente aos estímulos verbais.

### 3.1. Método

Nesse experimento, inspirado pela atividade desenvolvida por Uddén, Ingvar, Hagoort e Petersson (2012), foi utilizada uma combinação de técnicas experimentais que será descrita adiante. Foi utilizada uma tarefa de aprendizagem de gramática artificial com medidas *offline* e *online* realizado de forma remota com 3 etapas, sendo elas: uma etapa de aprendizagem implícita, um teste de escolha forçada e, finalmente, um questionário *offline* desenvolvido no *Google Forms*.

Para o desenvolvimento da tarefa experimental foi utilizado o site Ibex<sup>7</sup>, desenvolvido por Zehr e Schwarz (2018) na Universidade da Pensilvânia.

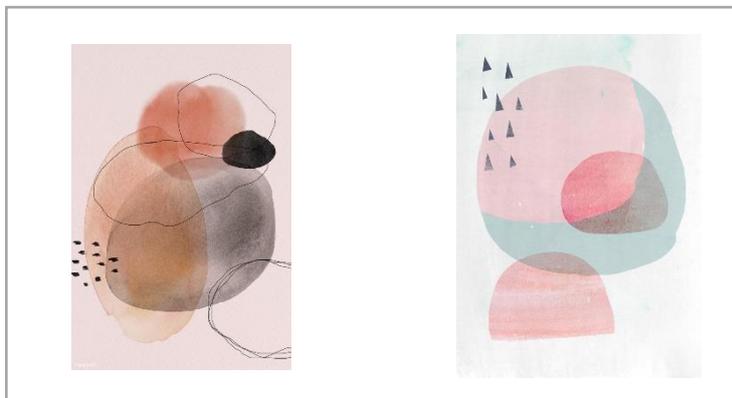
Participaram da atividade experimental 34 participantes adultos, divididos em dois grupos, sendo eles grupo experimental (média de idade 20 anos) e grupo controle (média de idade 25 anos). Todos os participantes fizeram a atividade experimental como voluntários e confirmaram a sua participação através de um termo de consentimento. Os participantes foram recrutados através das redes sociais e as únicas exigências eram de que tivessem o ensino superior concluído ou em curso e tivessem no mínimo 18 anos.

---

<sup>7</sup> O site é gratuito, de fácil acesso e fornece ferramentas para a criação de diversos experimentos psicolinguísticos que podem ser aplicados remotamente. Dessa forma foi possível a realização da atividade experimental de forma remota, visto que os experimentos presenciais não poderiam ser realizados por restrições impostas pela pandemia.

**Procedimento:** Os participantes foram treinados individualmente, cada um em sua residência. Devido a medidas de segurança sanitária, os participantes acessaram o experimento de maneira remota, através de um *link* a que cada um teve acesso. Todos foram instruídos a participarem da atividade em um local silencioso e utilizando fones de ouvido; além disso, também foram informados que só seria possível a realização da atividade em um computador. Depois de acomodados, ao clicar no *link*, uma primeira tela aparecia informando que antes do início da tarefa seria necessário que o participante lesse o termo de consentimento livre e esclarecido e clicasse na opção de aceitar o termo para prosseguir. Após informar seus dados, uma tela com as instruções para a realização da primeira fase do experimento era apresentada. As instruções informavam que na primeira etapa os participantes deveriam utilizar a tecla **C** do teclado para escolher a imagem da esquerda e a tecla **M** para escolher a imagem da direita, entre duas imagens que apareceriam na tela, de acordo com sua preferência. Enquanto faziam a escolha da imagem, os participantes ouviam combinações dos determinantes inventados com os nomes inventados seguindo o padrão da Tabela 1 para os participantes do grupo experimental. Os participantes do grupo controle ouviram sons de violão enquanto escolhiam as imagens. Essa fase durou aproximadamente 3 minutos e os participantes do grupo experimental foram expostos a 128 combinações de determinantes e palavras. As imagens utilizadas para a escolha foram imagens abstratas e não seguiam nenhum padrão.

Foram utilizadas imagens abstratas na fase do treinamento (Figura 1):



**Figura 1. Exemplo de imagens utilizadas na fase de treinamento, pré teste.**

Terminada a primeira etapa, surgia uma tela informando que a segunda etapa da atividade experimental começaria e, em seguida, uma tela de instruções. As instruções informavam que naquela etapa o participante ouviria um determinante e na tela apareceriam duas palavras por escrito e que ele/ela deveria selecionar, utilizando as mesmas teclas da primeira etapa, qual das duas palavras na tela combinava melhor com o determinante escutado. Essa etapa teve duração aproximada de 2 minutos.

A imagem ilustra um exemplo de tela da fase da tarefa de escolha forçada (Figura 2):

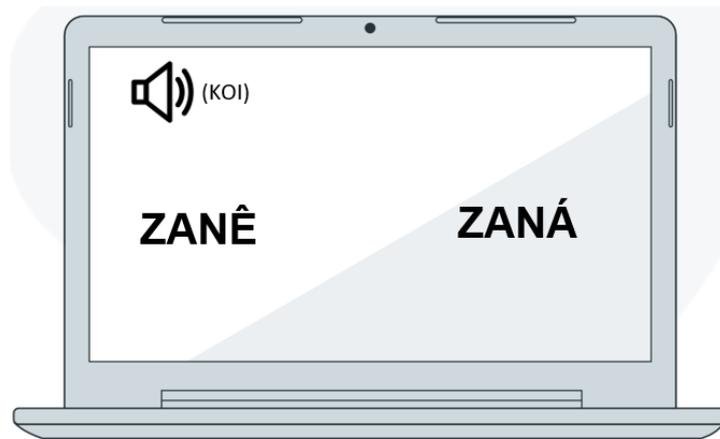


Figura 2. Exemplo de tela da fase teste.

Ao final do experimento *online*, os participantes receberam um *link* com acesso a um formulário desenvolvido no *Google Forms* com quatro perguntas abertas, sendo elas: “1) Nome”; “2) Você notou algum padrão nos estímulos?”; “3) Você consegue se lembrar de algo que ouviu durante o experimento?” e “4) O que você achou do experimento?”. Essas perguntas tiveram como objetivo avaliar se um possível resultado no experimento *online* poderia se justificar apenas por memorização dos estímulos e reconhecimento consciente das regras combinatórias ou se poderiam sugerir aquisição de conhecimento implícito da regra.

### 3.2. Resultados e discussão

Na análise dos resultados foram consideradas as variáveis **taxa de acertos** e **tempo de resposta** em cada *trial*. No que diz respeito à variável **taxa de acertos**, o grupo experimental apresentou taxas de acertos maiores em comparação com o grupo controle. O Gráfico 1 ilustra as taxas de erros e acertos em ambos grupos.

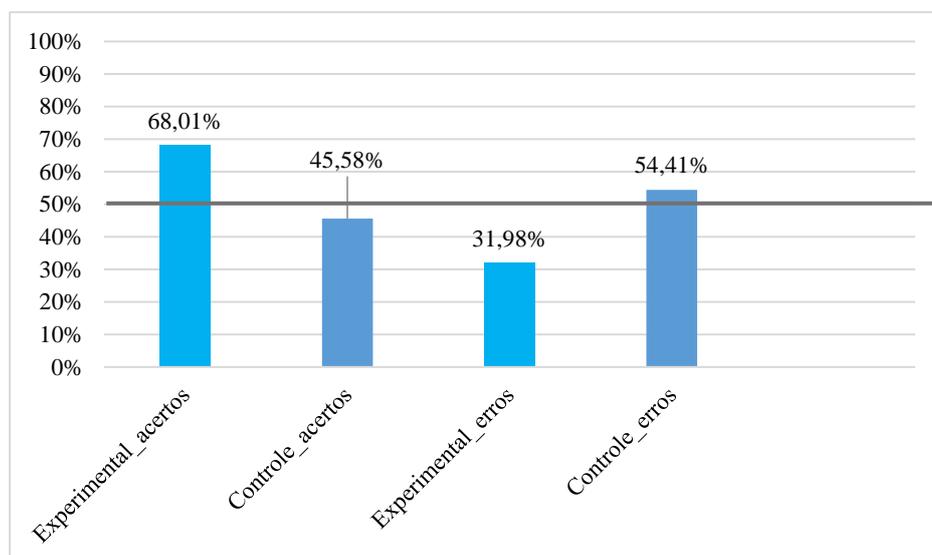


Gráfico 1. Taxa de acertos e erros dos grupos experimental e controle.

Inicialmente foi realizada análise estatística dos grupos separadamente, para verificar se haveria diferença entre taxas de acertos e erros dentro do grupo experimental e dentro do grupo controle. Os dados foram submetidos ao teste binomial no *software Statistical Package for the Social Sciences* (SPSS). No grupo controle, o teste mostrou que a taxa de erros e acertos está em proporção semelhante (acertos 54,41%, erros 45,58%,  $p=0.163$ ). Já no grupo experimental, o teste estatístico binomial revelou diferença estatisticamente significativa entre taxas de erros e acertos (acertos 68,01%, erros 31,98%,  $p=0.000$ ).

Também foi feita a análise dos dados comparando as diferenças entre taxas de erros e acertos entre os grupos experimental e controle. Foi realizado o teste McNemar, que mostrou que as proporções de erros e acertos nos dois grupos são diferentes ( $X^2_{(1)} = 25,175$ ;  $p=.000$ ). Como podemos ver no Gráfico 1, o grupo experimental apresentou taxas de acertos maiores, ao passo que o grupo controle permaneceu com as taxas de acerto e erro no nível da chance.

Em relação à variável tempo de respostas, primeiramente foi realizada uma análise exploratória dos dados para identificar possíveis *outliers*. O critério utilizado foi a seguinte equação: Limite inferior =  $Quartil1 - 1.5 * (Quartil3 - Quartil1)$  e Limite superior =  $Quartil3 + 1.5 * (Quartil3 - Quartil1)$ , e valores fora desses limites, foram considerados *OUTLIERS* (Crawley 2007, p. 155). Após a retirada de *outliers*, os dados foram organizados novamente, eliminando os valores classificados como *outliers*. Então foi realizado o teste de normalidade dos dados pelo teste de Shapiro-Wilk, o qual indicou que os dados não seguem uma distribuição normal, nesse sentido, foi utilizado o teste de Wilcoxon (Conover 1999, p. 272) para comparação das amostras.

Como pode ser visto na Tabela 2, não foi identificada uma diferença significativa entre os grupos experimental e controle, tanto para acertos (valor- $p=0,2802$ ), quanto para erros (valor- $p=0,9841$ ), portanto, o tempo de resposta entre os grupos controle e experimental não diferiu significativamente<sup>8</sup>:

**Tabela 2. Medianas do tempo (MS) para resposta do grupo experimental e grupo controle para acertos e erros e teste de comparação de medianas entre grupos controle e experimental.**

<b>Experimental Acerto</b>	<b>Controle Acerto</b>	<b>valor-p</b>
2344,5	2662,0	0,2802
<b>Experimental Erro</b>	<b>Controle Erro</b>	<b>valor-p</b>
2707,0	2500,0	0,9841

Também foi feita a análise separando os dados entre erros e acertos e também não foi identificada diferença significativa entre os tempos de resposta para acerto e erro, tanto no grupo experimental ( $p = 0,1679$ ) como no grupo controle ( $p = 0,7617$ ). Portanto,

<sup>8</sup> Valor de probabilidade de erro do teste de Wilcoxon.

respeitadas as condições do presente estudo, o tempo de resposta entre acertos e erros não diferiu significativamente, como mostra a Tabela 3:

**Tabela 3. Medianas do tempo (MS) para resposta do grupo experimental e grupo controle, para acertos e erros, e teste de comparação de medianas entre erros e acertos.**

<b>Experimental Acerto</b>	<b>Experimental Erro</b>	<b>valor-p1</b>
2344,5	2707,0	0,1679
<b>Controle Erro</b>	<b>Controle Erro</b>	<b>valor-p</b>
2662,0	2500,0	0,7617

Além disso, também foi realizada análise de correlação ponto bisserial a fim de comparar se há correlação entre tempos de respostas e taxas de acertos. Como pode ser visto na Tabela 4, o coeficiente de correlação para o grupo controle e experimental não foi significativo. No entanto, quando desconsiderados os grupos, considerando a totalidade das amostras, o coeficiente de correlação entre taxas de resposta e tempo de reação foi significativo.

O coeficiente de correlação foi negativo, ou seja, quanto maior o tempo de resposta, menor a probabilidade de acerto. No entanto, embora significativo, o coeficiente foi pequeno, indicando uma relação fraca entre as duas variáveis, como pode ser observado na Tabela 4.

**Tabela 4. Coeficientes de correlação ponto bisserial, entre a variável taxa de acertos (acerto = 1 ou erro = 0) e tempo de reação.**

	<b>Coeficiente de correlação ponto bisserial (rpbi)</b>	<b>valor-p</b>
<b>Grupo Controle</b>	-0,0752	0,2301
<b>Grupo Experimental</b>	-0,0871	0,1666
<b>Total</b>	-0,0957	0,03079

Apesar da existência de correlação entre taxas de acertos e tempos de resposta, esses dados não são relevantes, uma vez que só existe diferença significativa quando a divisão entre grupos controle e experimental não é levada em consideração.

Tomados em conjunto, os resultados das taxas de acerto sugerem efeito de treinamento, de forma que os participantes expostos ao conjunto de estímulos verbais na fase de treinamento, que durou aproximadamente 3 minutos, mostraram melhores desempenhos na fase teste em comparação com o grupo controle, que não teve acesso aos estímulos verbais na fase de treinamento, ficando impossibilitados de perceber algum padrão nos estímulos da fase de teste (*cf.* Gráfico 1).

<sup>9</sup> Correlação ponto bisserial significativa, a 5% de probabilidade.

Uddén, Ingvar, Hagoort e Petersson (2012), comparando dependências não adjacentes e de dependências adjacentes, defendem que dependências não adjacentes apresentam maior custo de processamento, uma vez que seus resultados apontaram tempo maior para a aprendizagem de dependências não adjacentes por adultos. No entanto, nossos dados sugerem que, após exposição a treinamento muito curto, de aproximadamente 3 minutos, os participantes foram capazes de identificar o padrão de combinação das dependências não adjacentes, ainda que inconscientemente. Apesar de deixar ainda em aberto a questão de maior ou menor dificuldade de aprendizagem desses dois tipos de dependências, nosso experimento apresenta evidência de relativa facilidade de aprendizagem de DNAs. Esses resultados são bastante surpreendentes, visto que havia mais de uma possibilidade de pareamento para cada determinante inventado, o que se assemelha a línguas naturais, como o português, em que um mesmo determinante pode ser combinado com diferentes terminações de nomes, e mais de um determinante pode ser combinado com a mesma terminação de nomes (*e.g.*, **Este/um menino, este/um pente, este/um gibi**), ou seja, a relação combinatória de DNAs em uma língua natural, raramente é de um para um.

Quanto ao questionário *offline* aplicado após o teste, a grande maioria dos participantes responderam que não se lembravam de nenhum estímulo do teste. Apenas dois participantes responderam que se lembravam, mas os estímulos que eles disseram se lembrar não existiam de fato. No que tange ao questionamento se haviam notado algum padrão nos estímulos, uma participante respondeu que não notou e os outros criaram hipóteses a respeito dos padrões combinatórios de cores das imagens na etapa da aprendizagem implícita, mas nenhum participante notou padrão nos estímulos verbais. Assim, as respostas ao questionário indicam que os resultados na taxa de acertos refletem a aprendizagem inconsciente das DNAs. Nesse sentido, os resultados sugerem que os participantes foram capazes de abstrair os padrões combinatórios entre determinantes e nomes inventados na primeira etapa e aplicaram esse padrão a novos estímulos, na fase teste, o que se assemelha ao que acontece nas línguas naturais, em que as dependências são probabilísticas (*cf.* Van Den Bos, Christiansen & Misyak 2012). Esse resultado é compatível com a nossa primeira previsão, de que os participantes apresentariam taxas de acertos acima do nível da chance no grupo experimental. Além disso, também são compatíveis com a previsão de que os participantes não seriam capazes de reconhecer conscientemente a regra combinatória subjacente aos estímulos, uma vez que o tempo de exposição foi curto e a tarefa de escolha de imagens realizada concomitantemente exigiu atenção consciente dos participantes para executá-la, ficando a escuta dos estímulos em segundo plano.

Em relação à nossa previsão sobre os tempos de reação (previsão II), nossos dados não foram estatisticamente significativos, pois as taxas de erros e acertos tiveram tempos de reação muito próximos, sugerindo que os estímulos gramaticais e agramaticais foram processados da mesma forma. Uma interpretação alternativa pode ser levantada devido à aplicação remota da tarefa, uma vez que os participantes realizaram a atividade experimental em casa e, apesar das instruções, distrações podem ter acontecido. Assim, a questão da aprendizagem de novas dependências não adjacentes por falantes adultos precisa ser investigada mais profundamente no futuro, em condições/ambientes mais bem

controlados que possam fornecer dados mais robustos, visto que a pandemia causou limitações na aplicação de experimentos com medidas *online*.

## 5. Conclusão

Os resultados do experimento vão ao encontro das previsões iniciais de que falantes adultos, apesar de expostos a novos padrões de DNAs em um curto período de tempo, são capazes de abstrair esses padrões de combinações de DNAs e generalizar o padrão aprendido a estímulos inéditos. Em trabalhos anteriores que investigaram a aprendizagem de DNAs por adultos em línguas artificiais, os participantes foram expostos durante um período longo de tempo aos estímulos ou recebiam informações que de alguma forma poderiam torná-los conscientes do fenômeno investigado. Nossa intenção foi investigar a aprendizagem implícita, em que o participante não aplica uma regra de forma racional, assemelhando-se ao que os bebês fazem ao adquirir uma L1.

Pode-se pressupor que a aprendizagem de DNAs demandaria duas etapas – uma de reconhecimento de padrões morfofonológicos e de variação probabilística entre os elementos, e uma etapa de caráter morfossintático, com o tratamento de tais padrões por mecanismo(s) de concordância. Nesse sentido, nosso experimento, assim como outros que usam línguas artificiais, trata claramente para o reconhecimento dos padrões morfofonológicos e de variação probabilística entre os elementos. Ainda é uma questão em aberto se esses experimentos tocam no nível morfossintático, à medida que as tarefas conduzidas não permitem que se observe claramente a implementação da concordância durante o processamento dos estímulos.

Por fim, vale destacar que os resultados do experimento conduzido apresentam evidência de capacidade de aprendizagem de DNAs disponível em adultos com uma L1 consolidada, após pouca exposição aos estímulos (comparado a situação natural de aquisição de L1 por bebês e crianças pequenas), e sem instrução consciente. Tal capacidade, portanto, poderia ser usada na aprendizagem de L2. No entanto, as dificuldades comumente enfrentadas por adultos aprendendo uma segunda língua, sobretudo no que se refere à identificação a seus processos de concordância, parecem sugerir o contrário. Por não se tratar do foco deste artigo, não avançaremos nessa discussão, mas entendemos que o processo de aprendizagem de L2 por adultos abrange uma série de questões que o distinguem da aquisição espontânea de L1, como a interferência da L1, a qualidade e a quantidade de interações linguísticas e a motivação do aprendiz, apenas para citar alguns pontos que podem interferir positivamente ou não nesse processo. De toda forma, ser capaz de identificar e abstrair padrões de DNAs e de aplicá-los a novos dados é, sem dúvida, condição relevante.

Assim, com esse trabalho, buscamos contribuir para o melhor entendimento da aprendizagem de dependências não adjacentes por adultos e de que forma mecanismos de aprendizagem presentes em bebês e crianças na fase de aquisição ainda permanecem disponíveis a adultos que já têm a aquisição da primeira língua consolidada. Ademais, buscamos entender em que medida essa capacidade pode colaborar na aprendizagem de uma L2. Além disso, buscamos contribuir com o desenvolvimento de uma nova metodologia experimental, que buscou superar algumas limitações de experimentos

anteriores feitos com adultos no paradigma de aprendizagem de língua artificial, de forma que fosse possível testar o conhecimento inconsciente dos participantes.

Mais pesquisas são necessárias para a ampliação das investigações a respeito da aprendizagem de DNAs por adultos. Por exemplo, ainda é preciso investigar o efeito do tempo na aprendizagem de novos DNAs e também o processamento de novas estruturas por adultos, além dos mecanismos usados nessa aprendizagem. Esperamos que nosso estudo fomente novas pesquisas, ampliando nosso conhecimento sobre esse tópico.

**Financiamento:** Esta pesquisa foi financiada pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## Referências

- Chomsky, N. (2000). Minimalist inquiries: The framework (MITOPL 15). In R. Martin, D. Michaels & J. Uriagereka (Eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik* (pp. 89–155). Cambridge: MIT Press. Consultado em <https://ci.nii.ac.jp/naid/10009706734/>
- Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: a life in language*. Cambridge (pp. 1–52). Cambridge: MIT Press.
- Folia, V., Uddén, J., De Vries, M., Forkstam, C., & Petersson, K. M. (2010). Artificial language learning in adults and children. *Language Learning*, 60(2), 188–220. <https://doi.org/10.1111/j.1467-9922.2010.00606.x>
- Folia, V., Uddén, J., Forkstam, C., Ingvar, M., Hagoort, P., & Petersson, K. M. (2008). Implicit learning and dyslexia. *Annals of the New York Academy of Sciences*, 1145(1), 132–150. <https://doi.org/10.1196/annals.1416.012>
- Frost, R. L. A., Jessop, A., Durrant, S., Peter, M. S., Bidgood, A., Pine, J. M., Rowland, C., & Monaghan, P. (2020). Nonadjacent dependency learning in infancy, and its link to language development. *Cognitive Psychology*, 120, 101291. <https://doi.org/10.1016/j.cogpsych.2020.101291>
- Frota, S., & Name, C. (2017). Questões de percepção em língua materna. In M. J. Freitas & A. L. Santos (Orgs.), *A aquisição de língua materna e não materna. Questões gerais e dados do português* (pp. 35–50). Berlin: Language Science Press.
- Gerken, L. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition*, 98(3), B67–B74. <https://doi.org/10.1016/j.cognition.2005.03.003>
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431–436. <https://doi.org/10.1111/1467-9280.00476>
- Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–135. [https://doi.org/10.1016/S0010-0277\(99\)00003-7](https://doi.org/10.1016/S0010-0277(99)00003-7)
- Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*, 7(5), 567–580. <https://doi.org/10.1111/j.1467-7687.2004.00381.x>
- Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7(2), 183–206. [https://doi.org/10.1207/s15327078in0702\\_4](https://doi.org/10.1207/s15327078in0702_4)
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579. <https://doi.org/10.1126/science.298.5598.1569>
- Henrique, K., & Name, C. (2020). O processamento de múltiplas dependências não adjacentes em orações relativas no PB. *Revista Linguística*, 16(1), 278–302. <https://doi.org/10.31513/linguistica.2020.v16n1a31631>

- Laguardia, M. (2016). *Etapas iniciais da aquisição lexical - Habilidades estatísticas e simbólicas no tratamento de dependências não adjacentes*. (Tese de doutorado, Universidade Federal de Juiz de Fora, Brasil).
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398), 77–80. <https://doi.org/10.1126/science.283.5398.77>
- Name, C., & Laguardia, M. (2013). Etapas iniciais da aquisição lexical. *PROLÍNGUA*, 8(2). Disponível em <https://periodicos.ufpb.br/ojs/index.php/prolingua/article/view/19347>
- Name, C., & Shi, R. (2015). Preverbal infants track and represent non-adjacent dependencies at an abstract level. In J. Scott & D. Waughtal (Eds.), *Proceedings of the 40<sup>th</sup> Boston University Conference on Language Development* (Vol. I, pp. 30–39). Boston: Cascadilla Press.
- Name, C., Shi, R., & Koulaguina, E. (2011, July). *Non-adjacent dependency learning and grammatical categorization in 11-month-old infants*. Presented at the 12th International Congress for the Study of Child Language. Montreal, Canada.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48(2), 127–162. [https://doi.org/10.1016/s0010-0285\(03\)00128-2](https://doi.org/10.1016/s0010-0285(03)00128-2)
- Onnis, L., Monaghan, P., & Christiansen, M. (2004). Variability is the spice of learning, and a crucial ingredient for detecting and generalizing in nonadjacent dependencies. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26(26), 1047–1052. Disponível em <https://escholarship.org/content/qt4558716n/qt4558716n.pdf>
- Rodrigues, E. (2005). O processamento da concordância verbal com construções partitivas no português brasileiro. Consultado em <https://revistas.ufrj.br/index.php/rl/article/viewFile/4371/3143>
- Rodrigues, E. (2006). *O processamento da concordância de número entre sujeito e verbo na produção de sentenças*. (Tese de doutorado, PUC Rio, Rio de Janeiro, Brasil).
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Sandoval, M., & Gómez, R. L. (2013). The development of nonadjacent dependency learning in natural and artificial languages. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(5), 511–522. <https://doi.org/10.1002/wcs.1244>
- Santelmann, L. (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition*, 69(2), 105–134. [https://doi.org/10.1016/s0010-0277\(98\)00060-2](https://doi.org/10.1016/s0010-0277(98)00060-2)
- Teixeira, M. (2012). *A contribuição das informações distribucionais de DNA nas etapas iniciais da aquisição lexical*. (Dissertação de mestrado, UFJF, Juiz de Fora, Brasil). Disponível em <https://repositorio.ufjf.br/jspui/handle/ufjf/1792>
- Thiessen, E., Girard, S., & Erickson, L. (2016). Statistical learning and the critical period: how a continuous learning mechanism can give rise to discontinuous learning. *WIREs Cognitive Science*, 7(4), 276–288. <https://doi.org/10.1002/wcs.1394>
- Trask, R. L. (2004). *Dicionário de linguagem e lingüística*. São Paulo: Contexto.
- Uddén, J., Ingvar, M., Hagoort, P., & Petersson, K. M. (2012). Implicit acquisition of grammars with crossed and nested non-adjacent dependencies: Investigating the push-down stack model. *Cognitive Science*, 36(6), 1078–1101. <https://doi.org/10.1111/j.1551-6709.2012.01235.x>
- Van den Bos, E., Christiansen, M. H., & Misyak, J. B. (2012). Statistical learning of probabilistic nonadjacent dependencies by multiple-cue integration. *Journal of Memory and Language*, 67(4), 507–520. <https://doi.org/10.1016/j.jml.2012.07.008>
- van Heugten, M., & Shi, R. (2010). Infants' sensitivity to non-adjacent dependencies across phonological phrase boundaries. *The Journal of the Acoustical Society of America*, 128(5), EL223–EL228. <https://doi.org/10.1121/1.3486197>
- Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, 61(3), 261–298.

[https://doi.org/10.1016/s0010-0277\(96\)00713-5](https://doi.org/10.1016/s0010-0277(96)00713-5)

Viotti, E. (2005). O Caso default no português do Brasil: revisitando o Caso dos inacusativos. *Revista de Estudos Da Linguagem*, 13(2), 53–71. <https://doi.org/10.17851/2237-2083.13.2.53-71>

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). *OSF*. <https://doi.org/10.17605/OSF.IO/MD832>

[recebido em 30 de outubro de 2021 e aceite para publicação em 18 de março de 2022]

## A AQUISIÇÃO DE CLÍTICOS EM PORTUGUÊS EUROPEU L2 E A HIPÓTESE DE RECONFIGURAÇÃO DOS TRAÇOS

### EUROPEAN PORTUGUESE L2 CLITICS ACQUISITION AND THE FEATURE REASSEMBLY HYPOTHESIS

Ronan Pereira\*  
a57730@campus.fcsh.unl.pt

A Hipótese de Reconfiguração de Traços (HRT) sugere que as dificuldades na aquisição de uma língua segunda (L2) ocorrem quando os aprendentes têm de reconfigurar os traços existentes na sua língua materna para apresentarem a configuração que a L2 possui. Este estudo objetivou testá-la relativamente à aquisição dos padrões de colocação dos clíticos do português europeu (PE) L2 e à aceitabilidade de respostas verbais, seguindo Martins (1994), que correlaciona os dois fenómenos com os traços-V fortes do PE. Divididos em dois grupos de proficiência distintos, 21 falantes nativos de espanhol executaram duas tarefas: uma de completamento de frases e outra de julgamento de aceitabilidade de respostas verbais. Diferentemente da literatura na área, a preferência pela ênclise foi menor. Analisando-se os contextos de próclise, não houve homogeneidade na evolução dos padrões de colocação, com aqueles envolvendo a negação realizados em maiores taxas já em níveis de proficiência mais baixos. Uma evolução na aceitabilidade de respostas verbais com o aumento da proficiência pôde ser constatada no segundo teste. Contudo, a análise estatística não demonstrou correlação estatística entre os resultados das tarefas. Concluiu-se que há indícios de validade da HRT, mas investigações futuras devem ser conduzidas para comprovar a natureza dessa reconfiguração.

**Palavras-chave:** Aquisição de língua segunda. Pronomes clíticos. Português europeu. Hipótese da Reconfiguração dos Traços (HRT).

The Feature Reassembly Hypothesis (FRH) suggests that the difficulties in second language (L2) acquisition happen when the learners have to reassemble the features present in their native language in order to present the L2's configuration. This study aimed to test the FRH regarding the acquisition of clitics pattern in European Portuguese (EP) L2 and the acceptability of verbal responses, following Martins (1994), who correlates both phenomena with EP's strong V-features. Divided into two different proficiency groups, 21 native speakers of Spanish executed two tasks: one of sentences completion and an acceptability judgment task of verbal responses. Differently to what has happened in other studies, preference for enclisis was lower. The analysis by proclitic context showed that the development of the placement patterns is not homogeneous, with negative contexts being produced in higher rates even at lower levels of proficiency. An evolution in the acceptance of verbal responses could be detected by the second task. Nevertheless, the statistical analysis did not show statistical correlation between results of the tasks. The conclusion is that there are signs that the FRH might be valid, but future research must be conducted in order to assess the nature of this reassembly.

---

\* Centro de Linguística da Universidade Nova de Lisboa (CLUNL), Portugal. ORCID: 0000-0003-2675-6378.

**Keywords:** Second language acquisition. Clitic pronouns. European Portuguese. Feature Reassembly Hypothesis (FRH).

## 1. Introdução

Há duas questões particularmente importantes na área de estudos de aquisição de língua segunda (L2). A primeira refere-se à influência que a língua materna (L1) dos aprendentes pode exercer sobre o processo de aquisição da L2. A segunda refere-se à trajetória pela qual o aprendente passa ao adquirir o novo idioma: ela pode ser igual ou não ao processo de desenvolvimento dos adquirentes desse idioma como L1.

Lardiere (2009) propõe que, num primeiro momento, o aprendente identifica formas da sua L1 e estabelece uma correspondência direta com formas da L2. Havendo falha nesse processo (ao identificar traços que não são equivalentes diretos), haverá uma reorganização interna deles, seja por reagrupamento de traços funcionais, seja por aquisição de novos traços, ou seja por abandono de traços previamente existentes e que não se aplicam à L2, havendo graus de dificuldade diferentes associados a cada uma destas alternativas. Tal hipótese foi denominada Hipótese da Reconfiguração de Traços (HRT).

Assim, este estudo propôs-se a testar a HRT no âmbito da aquisição dos padrões de colocação dos pronomes clíticos em português europeu (PE) por falantes nativos de espanhol em dois níveis diferentes de proficiência, baseando-se em uma tarefa de completamento de frases, além de uma tarefa de julgamento de aceitabilidade. De modo a dar-lhe fundamentação teórica, as próximas secções ocupam-se em abordar os temas pertinentes a este estudo. A Secção 2 versa sobre os clíticos em português e em espanhol e sobre uma teoria sintática que pretendeu explicar a diferença na colocação dos clíticos nesses idiomas, além de trazer os resultados de alguns estudos de aquisição de clíticos do PE. A Secção 3 foca-se na HRT, com os seus pressupostos e exemplos de alguns estudos prévios que a testaram. A Secção 4 apresenta as questões de investigação e hipóteses deste estudo, enquanto a Secção 5 detalha a metodologia utilizada e a Secção 6 os resultados. Por fim, a Secção 7 discute os resultados e as suas conclusões.

## 2. Pronomes clíticos

### 2.1. Padrão de colocação dos clíticos

Os pronomes pessoais clíticos no PE e no espanhol variam em pessoa e número, além de género e caso (no caso dos pronomes acusativos de 3ª pessoa), e assumem as seguintes formas que se apresentam na Tabela 1:

**Tabela 1: pronomes clíticos no PE e no espanhol.**

Pessoa/Número	PE	Espanhol
1ª/singular	Me	Me

2 <sup>a</sup> /singular	Te	Te
3 <sup>a</sup> /singular	O/A/Se/Lhe	Lo/La/Se/Le
1 <sup>a</sup> /plural	Nos	Nos
2 <sup>a</sup> /plural	Vos	Os
3 <sup>a</sup> /plural	Os/As/Se/Lhes	Los/Las/Se/Les

Fonte: Martins (2013), Soriano (2015).

Apesar da semelhança morfológica, quando se analisam os padrões de colocação dos clíticos nessas línguas, vê-se que funcionam de maneira diferente. O espanhol admite a colocação do clítico antes (próclise) ou depois (ênclise) do verbo, fenómeno sujeito à finitude do verbo. Em frases finitas (Exemplo 1) ou no imperativo negativo (Exemplo 2), a única opção de colocação é a próclise; a ênclise ocorre quando o verbo se encontra no infinitivo (Exemplo 3), gerúndio (Exemplo 4) ou no imperativo afirmativo (Exemplo 5) (Soriano 2015):

- (1) **Les** enseñé el dibujo a los niños.
- (2) No **le** des importancia a esas cosas.
- (3) Conviene hacer**lo**.
- (4) Ayudá**ndola** tanto no vas a lograr nada.
- (5) Ayúden**la** ustedes también.

No português, ambas as posições são possíveis, além da mesóclise<sup>1</sup>, com o clítico no interior da forma verbal, mas a colocação dá-se de acordo com diferentes contextos sintáticos. A colocação padrão é enclítica (Duarte 2003, p. 852), ocorrendo em frases simples (ou em orações principais de frases complexas e em orações coordenadas aditivas ou adversativas<sup>2</sup>), exemplificada no Exemplo 6<sup>3</sup>:

- (6) A Cíntia viu-**te** ontem no centro.

Os contextos para a ocorrência da próclise no PE estão associados “aos processos gramaticais da *negação*, da *quantificação*, da *focalização* e da *ênfase*, tomados singularmente ou não” (Martins 2013, p. 2239, grifos da autora). O traço negativo da frase pode ser assegurado pela presença das seguintes palavras em posição pré-verbal: *nem*, *nunca*, *jamais*, *ninguém*, *nenhum* e *nada*:

<sup>1</sup> Sintaticamente, a mesóclise ocorre nos mesmos contextos que a ênclise, mas com verbos conjugados no futuro do indicativo ou no condicional. Há evidências de que, presentemente, a mesóclise é uma estrutura de aquisição no âmbito do ensino formal, visto que perífrases ou tempos verbais alternativos são utilizados para exprimir o futuro do indicativo e o condicional (Santos 2002).

<sup>2</sup> Além disso, o padrão é a ênclise ainda que as frases sejam declarativas, imperativas, exclamativas ou interrogativas totais.

<sup>3</sup> Os exemplos utilizados nesta secção provêm do teste utilizado neste estudo, salvo menção contrária.

- (7) O Alberto não **me** viu na estação do metro.
- (8) Ninguém **me** ouviu durante a apresentação.

Outras classes de palavras que, quando em posição pré-verbal, estimulam a próclise são os quantificadores (Exemplo 9) e os advérbios (Exemplo 10), ainda que nem todos causem a ocorrência da próclise:

- (9) Todos **me** viram na televisão.
- (10) O Gustavo já **te** ensinou a fórmula de Bhaskara.

Ao serem analisadas as orações subordinadas finitas, normativamente preconiza-se o uso da próclise (Martins 2013, pp. 2267–2270)<sup>4</sup>:

- (11) O Lucas estava feliz porque **te** viu ontem.
- (12) Acho que **me** cortei com o papel.
- (13) Queremos que a Paula **se** sinta melhor.

Os contextos mencionados até aqui foram aqueles considerados no teste descrito na Secção 5. No entanto, a descrição apresentada não é exaustiva. Mais exemplos dos contextos abordados, além dos demais, podem ser encontrados em Martins (2013, 2016).

## 2.2. A categoria $\Sigma$

Para explicar o padrão de colocação dos pronomes clíticos, Martins (1994) elaborou uma teoria relacionada à categoria funcional  $\Sigma$  (Laka 1990). Tal categoria está intimamente relacionada a valores de polaridade da frase (afirmação, negação, modalidade) e, no português, contém traços-V fortes. Logo, o verbo, que se desloca inicialmente da sua posição de base para uma posição funcional mais baixa com o clítico em adjunção à sua esquerda, movimenta-se para  $\Sigma$ , a qual está numa posição entre Comp e AgrS. O resultado visível é um verbo com o clítico à sua direita, ou seja, em ênclise e com o sujeito na posição de especificador de  $\Sigma$ , como se pode ver pelo Exemplo 14 retirado de Martins (2002, p. 357):

- (14)
- a. O António viu-**o** ontem.
- b. [ <sub>$\Sigma$ P</sub> O António]<sub>j</sub> [ <sub>$\Sigma$ '</sub> [ <sub>$\Sigma$</sub>  viu<sub>i</sub>  $\Sigma$  [<sub>AgrSP</sub>  $t_j$  [<sub>AgrS'</sub> [<sub>AgrS</sub> **o** [<sub>AgrS</sub>  $t_i$  ...

<sup>4</sup> No entanto, há casos de ênclise em orações subordinadas, especialmente quando o verbo da oração está no indicativo.

Quanto à próclise em frases negativas, a ideia trazida por Martins (1994, 2002) é de que o marcador de negação predicativa ocupa a categoria  $\Sigma$ , impedindo a subida do verbo para tal posição (ou seja, mantendo-o em AgrS), resultando visivelmente a próclise:

(15)

a. O António não **o** viu ontem.

b. [ $\Sigma$ P O António]<sub>i</sub> [ $\Sigma'$  [ $\Sigma$  não  $\Sigma$  [ $\text{AgrSP}$  t<sub>j</sub> [ $\text{AgrS}'$  [ $\text{AgrS}$  **o** [ $\text{AgrS}$  viu ...

Em orações subordinadas, a próclise é, teoricamente, a única forma possível. Nesse tipo de orações,  $\Sigma$  se funde com a categoria C gerando o núcleo complexo [ $\text{C}$   $\Sigma$  C]:

(16)

a. O José disse que o António **o** viu ontem.

b. ...  $\text{C}'$ [ $\text{C}$   $\Sigma$  i [ $\text{C}$  que] [ $\Sigma'$  [ $\Sigma$  t<sub>i</sub> [ $\text{AgrSP}$  o António [ $\text{AgrS}'$  [ $\text{AgrS}$  **o** [ $\text{AgrS}$  viu ...

Por fim, a autora menciona que o português possibilita a instanciação de  $\Sigma$  como F(oco), ocupada por “[s]intagmas-qu, sintagmas focalizados, quantificadores (afirmativos e negativos) e certos advérbios, quando ocorrem na periferia esquerda da frase” (Martins 2002, p. 358), e que, ao contrário de  $\Sigma$ , não contém traços-V. Assim sendo, ao ser projetada, não proporciona a subida do verbo, tal como no caso das frases negativas e das orações subordinadas, obtendo-se, portanto, a próclise, como se mostra no seguinte exemplo de Martins (2002, p. 358):

(17)

a. Quem **lhe** deu o tiro?

b. [ $\text{FP}$  Quem [ $\text{F}'$  F [ $\text{AgrSP}$  [ $\text{AgrS}'$  [ $\text{AgrS}$  lhe [ $\text{AgrS}$  deu ...

Martins (1994), além disso, relaciona o traço-V do português com a possibilidade de resposta verbal a interrogativas totais (Exemplo 18):

(18)

a. Viste a Marisa ontem?

b. Vi.<sup>5</sup>

Isto explicar-se-ia novamente pelo facto de que  $\Sigma$  possui traços-V fortes. Dada a necessidade de o seu núcleo estar preenchido por conteúdo visível, o verbo é atraído a essa categoria durante uma frase-resposta mínima afirmativa, o que não ocorre em línguas cujo  $\Sigma$  possui traços-V fracos, como o espanhol:

<sup>5</sup> Exemplo de Martins (1994, p. 201).

(19)

a. Viste a Marisa ayer?

b. \*VÍ.

c. Sí, la ví.

Nestas, por não haver movimento do verbo de AgrS para  $\Sigma$ , o seu núcleo  $\Sigma$  é preenchido por uma palavra afirmativa (Exemplo 19c) (Martins 2002, p. 356). Assim, no espanhol, não existiria movimento do verbo em frases finitas para essa categoria, sendo, portanto, a próclise o resultado visível (Martins 2002, p. 357):

(20)

a. Antonio **lo** vio ayer.b. [ $\Sigma$ P Antonio]<sub>i</sub> [ $\Sigma$ :  $\Sigma$  [ $\Sigma$ SP t<sub>j</sub> [ $\Sigma$  AgrS' [ $\Sigma$  AgrS **lo** [ $\Sigma$  AgrS vio ...

### 2.3. Estudos prévios em aquisição de clíticos

Costa, Fiéis e Lobo (2015) estudaram sistematicamente a aquisição do padrão de colocação dos pronomes clíticos por crianças nativas de PE. Os autores realizaram uma tarefa de produção induzida do clítico reflexo *se*, pela sua menor taxa de omissão nas faixas etárias estudadas (Costa & Lobo 2007, 2009), com crianças de 5, 6 e 7 anos de idade.

Os resultados demonstraram que as crianças (e tampouco os adultos do grupo de controlo) não têm problemas em produzir a ênclise nos contextos que a exigem: houve somente uma ocorrência de próclise no grupo de 5 anos de idade. No entanto, encontraram outras alternativas de resposta. Por exemplo, no grupo de 5 anos a taxa de ênclise rondou os 69%, muito por encontrarem outras formas de responder às questões (25%), casos de omissão (5%) e um caso de redobro do clítico (0,4%).

Por outro lado, as taxas de produção de próclise nos contextos devidos foram expressivamente mais baixas (36% no grupo de 5 anos, 53,8% no grupo de 6 anos e 57,2% no grupo de 7 anos, taxas muito inferiores às do grupo de controlo: 88,8%). Analisando cada contexto proclítico individualmente, puderam notar que alguns contextos eram estabilizados primeiro, nomeadamente aqueles mais categóricos no *input*: frases negativas e com sujeitos negativos, além de orações completivas finitas. Costa *et al.* (2015) mencionam que traços como a negação são adquiridos cedo no processo de desenvolvimento, o que levaria à estabilização precoce dos contextos relacionados à negação – a criança tem de relacionar, portanto, a negação com a próclise. No caso da subordinação, a tarefa da criança é saber que todas as orações subordinadas induzem a próclise (*cf.* Martins 2013)

Nos outros contextos, a tarefa é mais pontual. Com advérbios, a tarefa é estabelecer quais induzem a próclise e quais não a induzem. O mesmo ocorre com as adverbiais finitas, visto que “na subordinação adverbial encontramos uma multiplicidade de conectores que podem funcionar como marcadores de um domínio de subordinação”

(Costa, Fiéis & Lobo 2016), mas os mesmos podem funcionar em contextos de coordenação, nos quais a ênclise deve ocorrer. Por fim, há quantificadores que categoricamente induzem a próclise e outros que possibilitam a alternância entre as duas formas. Desta forma, a tarefa é estabelecer quais são utilizados em cada contexto (o que envolve conhecimento lexical).

Note-se que, como as crianças, os adultos são mais propensos a produzir a próclise em contextos de negação e com subordinadas no conjuntivo (e um pouco menos com advérbios pré-verbais, mais ainda assim acima de 93%). Com orações adverbiais finitas e com quantificadores, contudo, chegam a 70,8% neste e a 77,6% naquele. Assim, as crianças adquirem mais tardiamente contextos “em que há uma maior dependência de conhecimento lexical” (Costa *et al.* 2016), os quais são justamente aqueles em que os adultos apresentam maior variação, expondo-as a um *input* muito variável.

No âmbito da aquisição de L2, Madeira, Crispim e Xavier (2006) realizaram uma tarefa de produção induzida aliada a um teste de julgamento de aceitabilidade com falantes de diferentes L1, os quais foram organizados em quatro subgrupos consoante a L1 (românica e germânica/urálica) e o tempo de aprendizagem do português (inferior a um ano e de um ano a dois anos). Os resultados mostraram uma tendência de aquisição iniciada por contextos de negativas e interrogativas-qu no teste de produção induzida. Já o teste de julgamento de aceitabilidade mostrou que os falantes do grupo de línguas românicas distinguem os contextos de ênclise e próclise mesmo no nível mais baixo de proficiência. Tal tendência não é observada pelos falantes de línguas não românicas. As autoras sugerem a existência de efeitos de influência da L1 a partir dos resultados obtidos, nomeadamente no ritmo de aquisição dos padrões de colocação.

Madeira e Xavier (2009) analisaram um *corpus* escrito produzido por dois diferentes grupos de L1: línguas românicas e línguas germânicas. Além disso, os grupos foram subdivididos em três níveis de proficiência: elementar, intermédio e avançado. Os resultados mostraram que, em ambos os grupos, a ênclise ocorre nos contextos em que deveria já nos estádios mais precoces. No entanto, é a forma preferida pelos falantes de nível elementar nos contextos em que a próclise deveria ocorrer (à exceção de frases negativas, em que a preferência pela próclise já ocorre), diminuindo à medida em que o nível de proficiência aumenta.

No mesmo estudo, as autoras aplicaram um teste de julgamento de gramaticalidade. Tendo de avaliar frases com diferentes contextos de colocação dos clíticos, observou-se que os participantes e o grupo de controlo foram mais precisos em contextos de ênclise em geral, e menos precisos em contextos de próclise (principalmente os que envolviam quantificadores pré-verbais e orações subordinadas finitas). Os testes estatísticos indicaram diferenças significativas entre a performance do grupo de controlo e os dois grupos experimentais em estudo, mas não entre os dois grupos experimentais.

Gu (2019) também se prestou a analisar a colocação dos clíticos em PE L2. No entanto, a amostra continha dois grupos compostos apenas por falantes chineses que realizaram uma tarefa de produção induzida, tendo os participantes sido divididos entre aqueles com dois ou menos anos de aprendizagem e aqueles com mais de dois anos. Os dados de produção demonstraram efeitos de desenvolvimento, constatando-se que a aquisição dos contextos de ênclise se deu mais facilmente do que a dos contextos de

próclise, havendo inclusive um processo gradual de aquisição desses contextos: negação > advérbio pré-verbal *também* > orações subordinadas adverbiais introduzida por *porque* > sujeito quantificado *todos*. Assim, de acordo com os estudos aqui apresentados, vê-se que os contextos de próclise parecem ser mais difíceis de serem adquiridos, havendo uma trajetória de desenvolvimento similar àquela descrita previamente para o PE L1.

### 3. A hipótese de reconfiguração dos traços

A extensa literatura sobre a aquisição de L2 mostra que diferentes formas de conhecimento da L1 podem ser transferidas para a L2. Lardiere (2009) considera que as dificuldades observadas na aquisição da L2 se dão pelo facto de que os aprendentes têm de reconfigurar os traços existentes na sua L1 de modo que passem a apresentar a configuração que a L2 possui. Tal visão assume que, embora todos os traços estejam disponíveis na Gramática Universal (GU), cada idioma seleciona traços específicos e os organiza de modos distintos, resultando em representações lexicais diferentes. Durante o processo de aquisição de uma L2, ocorre inicialmente uma correlação direta entre formas da L1 e da L2. Em seguida, o aprendente deve reorganizar os traços a fim de os estruturar da forma como são representados na L2, além de, eventualmente, ter de abandonar traços que estão presentes na L2 e não estão na L1 ou importar da GU traços que não estão presentes na L1, mas que estão na L2 (Lardiere 2009).

Logo, de acordo com tal proposta, o processo de aquisição ocorre em dois momentos. No primeiro, chamado *mapeamento*, o aprendente tenta criar relações entre os traços presentes na sua L1 com as realizações lexicais da L2. Ou seja, ele tenta encontrar “correspondências morfolexicais na L2 [...], presumivelmente baseado no significado semântico ou na função gramatical”<sup>6</sup> (Lardiere 2009). Esse processo pode, então, ser visto como uma tentativa inicial de estabelecer relações diretas entre as formas presentes na sua L1 e as formas da L2. No caso de este primeiro processo apresentar falhas, o aprendente deve passar à segunda etapa, chamada *reconfiguração*, a qual pode requerer diferentes tipos de operações, cada uma com o seu grau de dificuldade. É durante este estágio que o aprendente reorganizará os traços presentes na sua L1 para apresentar a configuração da L2 em questão. Além disso, poderá ser necessária a importação de traços da GU ou o abandono deles. Tal hipótese baseia-se, portanto, no pressuposto de que todos os traços podem ser eventualmente adquiridos a partir do momento em que o contraste é detetado na L2 (Lardiere 2009; Shimanskaya 2015).

Um exemplo de estudo que testou a HRT e que também lidou com a questão de aquisição de pronomes clíticos foi o de Shimanskaya e Slabakova (2015). As autoras basearam-se no facto de que os pronomes objeto em francês e inglês possuem traços semânticos e morfossintáticos distintos. Por exemplo, o francês possui clíticos, enquanto o inglês possui pronomes fortes na posição de objeto direto (a sua colocação em relação ao verbo também difere) e os pronomes objeto em inglês trazem o traço [ $\pm$ humano], enquanto os pronomes objeto do francês não fazem esta distinção. Além disso, os pronomes com traço [ $+$ humano] possuem morfologia específica de acordo com o género

---

<sup>6</sup> Do original: “morpholexical correspondences in the L2 [...], presumably on the basis of semantic meaning or grammatical function.”

biológico dos antecedentes; já os pronomes objeto do francês diferem em forma de acordo com o género gramatical do antecedente a que se referem.

Assim, aprendentes de francês L2 de três níveis distintos (iniciados, intermédios e avançados) e que tinham o inglês como língua materna passaram por dois testes: um de seleção de imagens e outros em que tinham de efetuar uma leitura automonitorizada. No primeiro, observou-se que os falantes de níveis mais baixos cometeram mais erros relativamente ao género gramatical, traço ausente na sua L1, o que vai ao encontro dos pressupostos da HRT, mais especificamente do primeiro estágio, o mapeamento. Os aprendentes de níveis iniciado e intermédio cometeram mais erros quando os pronomes se referiam a objetos inanimados; quando se referiam a humanos, houve interferência do género biológico (presente na L1) na interpretação off-line dos pronomes do francês pelos aprendentes (Shimanskaya & Slabakova 2015).

O mesmo não se observou no grupo de falantes de nível avançado (que teve performance similar à do grupo de controlo), o que indica que a reconfiguração de traços morfossintáticos é possível, visto que o género gramatical passou a fazer parte da gramática dos aprendentes, ainda que tal traço não exista na sua L1. Esse resultado é parecido ao obtido pela segunda tarefa. Nela, o grupo de falantes avançados tiveram resultados que apontam para a sua sensibilidade a violações de género gramatical durante a leitura. No entanto, tal sensibilidade ocorreu somente nos casos em que a violação ocorria com palavras relacionadas a humanos. As autoras discutem que a presença dos traços morfossintáticos da L1 são, portanto, detetáveis durante o processamento on-line da linguagem até mesmo em níveis mais altos de proficiência, diferentemente do que ocorre com o processamento off-line. As autoras finalizam concluindo que é possível considerar que adquirir uma representação e utilizá-la automaticamente on-line são dois processos que não ocorrem simultaneamente, havendo um atraso deste em relação àquele (Shimanskaya & Slabakova 2015).

#### 4. Questões de investigação e hipóteses

Levando-se em conta os pressupostos da HRT, bem como a análise proposta por Martins (1994) acerca da categoria  $\Sigma$ , estabeleceram-se as seguintes questões de investigação:

*Questão 1:* Os aprendentes de português L2 nativos de espanhol transferem os traços funcionais relevantes para a colocação dos clíticos nos estádios iniciais de aquisição?

*Questão 2:* Os aprendentes de português L2 nativos de espanhol em estádios mais avançados de aquisição são capazes de reconfigurar os traços envolvidos no processo de colocação dos clíticos?

*Questão 3:* Existe uma sequência de desenvolvimento para os contextos de próclise considerados no estudo?

A partir das questões acima, estipularam-se as seguintes hipóteses:

*Hipótese 1:* Espera-se que os falantes nativos de espanhol do grupo de proficiência mais baixo apresentem taxas de respostas alvo em proporção mais baixas que as dos de proficiência mais alta e do grupo de controlo em contexto de ênclise;

*Hipótese 2:* Espera-se que os falantes nativos de espanhol do grupo de proficiência mais alto apresentem taxas de respostas alvo mais próximas às dos nativos do que o grupo de proficiência mais baixo (em consequência da aquisição dos traços-V fortes da categoria  $\Sigma$ );

*Hipótese 3:* Espera-se que uma maior produção alvo de colocação no contexto de ênclise esteja associada à aceitação de respostas-verbais, visto que Martins (1994) atribui a possibilidade deste tipo de respostas no português àquela categoria e aos seus traços-V fortes.

*Hipótese 4:* Espera-se que os contextos relacionados com a negação sejam adquiridos primeiro, visto que é um traço adquirido precocemente (Costa *et al.* 2015) e disponível na gramática do nativo de espanhol, enquanto os contextos mais dependentes de conhecimento lexical sejam adquiridos mais tardiamente.

A metodologia utilizada será descrita a seguir.

## 5. Metodologia

### 5.1. Amostra

No total, 21 nativos de espanhol<sup>7</sup> participaram no estudo. Eles dominavam em diferentes níveis o português como língua estrangeira e habitavam em Lisboa ou viajavam frequentemente a essa cidade nos 12 meses antes da sua participação. Além deles, 30 nativos de português europeu serviram de grupo de controlo.

Todos os nativos de espanhol realizaram o mesmo procedimento, o qual incluía a assinatura de um termo de consentimento, o preenchimento de um questionário sociolinguístico e demográfico (ambos redigidos em espanhol), a realização de dois testes (preenchimento de frases e juízo de aceitabilidade) e, por fim, uma gravação de um curto monólogo em que falavam sobre três imagens, descrevendo-as ou contando uma anedota, por exemplo. Essa gravação teve como finalidade aferir o nível de proficiência dos falantes. As gravações foram ouvidas por duas falantes nativas<sup>8</sup> do PE e que nunca viveram noutro país, as quais julgaram de 0 a 10 as gravações de acordo com critérios como a pronúncia, a adequação das estruturas, a fluência e a adequação do vocabulário, sendo o valor 0 equivalente a “não nativo” e o valor 10 equivalente a “nativo”. Os critérios eram avaliados ao mesmo tempo por meio duma ficha de respostas. No entanto, as falantes podiam repetir os áudios se assim quisessem. Dentre as gravações havia, também, três amostras de áudio de falantes nativos. Contudo, tal informação não lhes foi fornecida.

---

<sup>7</sup> Note-se que, apesar de todos os falantes serem nativos de espanhol, são nativos de variedades nacionais distintas desse idioma; contudo, não há evidências de diferenças relativamente ao padrão de colocação dos clíticos nas diferentes variedades do espanhol faladas pelos participantes (*cf.* Soriano 2015).

<sup>8</sup> Uma com 20 anos de idade, estudante universitária, e outra com 44, licenciada. Ambas indicaram terem aprendido inglês e francês e eram naturais do concelho de Ponta Delgada (Açores).

Dos valores atribuídos pelas duas nativas realizou-se uma média aritmética, com exceção do valor da categoria “pronúncia.” Os sujeitos foram, então, agrupados de acordo com a classificação obtida após a aferição, resultando em dois grupos de proficiência denominados Baixo (M = 3,01, DP = 0,53) e Alto (M = 5,12, DP = 0,90). A Tabela 2 sintetiza os dados obtidos por meio do questionário sociolinguístico:

**Tabela 2: Perfil sociolinguístico dos nativos de espanhol divididos por nível.**

<b>Grupo</b>	<b>Baixo</b>	<b>Alto</b>
	Média = 33;4	Média = 36;11
Idade	Intervalo = 24 – 57	Intervalo = 27 – 50
	Masculino = 7	Masculino = 3
Género	Feminino = 4	Feminino = 7
		Catalão = 1
Outra L1 <sup>9</sup>		Quíchua = 1
	Inglês = 2	Nenhuma = 4
	Inglês e Italiano = 2	Francês = 1
L2	Inglês e Francês = 6	Inglês = 3
	Inglês, Francês e Chinês = 1	Inglês, Alemão e Catalão = 1
	Licenciatura = 6	Licenciatura = 4
Escolaridade	Pós-Graduação = 5	Pós-Graduação = 6
		Costa Rica = 1
	Cuba = 1	Cuba = 1
País de Origem	Espanha = 10	Espanha = 6
		México = 1
		Peru = 1
Idade de Início do	Média = 32:3	Média = 26:11
Contacto	Intervalo = 22 – 54	Intervalo = 22 – 34
	Média = 2:1	Média = 8:8
Tempo de		
Contacto (Anos)	Intervalo = 9 meses – 5 anos	Intervalo = 3 meses – 24 anos
	Média = 7 horas e 16 minutos	Média = 6 horas e 6 minutos
Uso Diário	Intervalo = 30 minutos – 24 horas	Intervalo = 2 horas – 24 horas
(Horas)		
	Íntimo = 1	Cotidiano = 2
	Cotidiano = 2	Profissional = 2
Situações	Profissional = 3	Cotidiano e Íntimo = 1
	Profissional e Íntimo = 1	Profissional, Íntimo e Cotidiano = 3
	Profissional e Cotidiano = 3	Escolar, Profissional e Cotidiano = 1
	Profissional, Íntimo e Cotidiano = 1	

O grupo de controlo realizou as mesmas tarefas que os falantes de espanhol e sob as mesmas condições. Contudo, não realizaram o teste de aferição de nível e o seu termo de consentimento e o questionário sociolinguístico foram redigidos em português. Os testes realizados são descritos a seguir.

<sup>9</sup> Note-se que o catalão possui clíticos e o seu padrão de colocação é similar ao do espanhol (Martins 2002, p. 356); já o quíchua não possui clíticos pronominais (Klee 1990). Uma análise individual dos resultados desses participantes não mostrou diferenças entre os seus resultados e os dos demais participantes do seu grupo.

## 5.2. Teste de completamento de frases

Os testes foram aplicados individualmente ou em grupo, de forma digital com o apoio de um impresso. No ecrã do computador, alternadamente, em forma de diapositivos, foram exibidos os itens de teste e os distratores aleatorizados. Cada item era constituído por uma frase incompleta em português, sendo que a parte em falta correspondia a um verbo e a um clítico. Quanto aos distratores, a parte em falta correspondia a um verbo e a um advérbio. Os sujeitos da pesquisa foram instruídos a ler a frase incompleta no ecrã e, dentre as duas opções existentes no impresso, escolher a que julgassem a mais adequada. As opções eram mostradas sempre em par, apresentando a ordem “verbo-clítico” e “clítico-verbo” ou “verbo-advérbio” e “advérbio-verbo” para os itens de teste (Figuras 1 e 2) e para os distratores (Figura 3), respetivamente. A apresentação de cada frase no ecrã teve um limite de tempo de sete segundos, de modo que se minimizasse o acesso ao conhecimento metalinguístico que os participantes pudessem ter. Os 40 itens de teste, os quais apresentavam diferentes contextos que exigiam ou o uso da ênclise, ou o uso da próclise, estavam distribuídos da seguinte forma: 12 itens de ênclise em orações simples (*cf.* Figura 1); 4 itens de próclise com negação (*cf.* Figura 2); 4 itens de próclise com sujeito negativo *ninguém*; 4 itens com sujeito quantificado *todos*; 4 itens com o advérbio pré-verbal *já*; 4 itens de próclise com orações subordinadas finitas do verbo *achar*; 4 itens de próclise com orações subordinadas finitas do verbo *querer*; e 4 itens de próclise com orações subordinadas adverbiais com *porque*. Quanto aos distratores, houve três para cada quatro itens de teste.

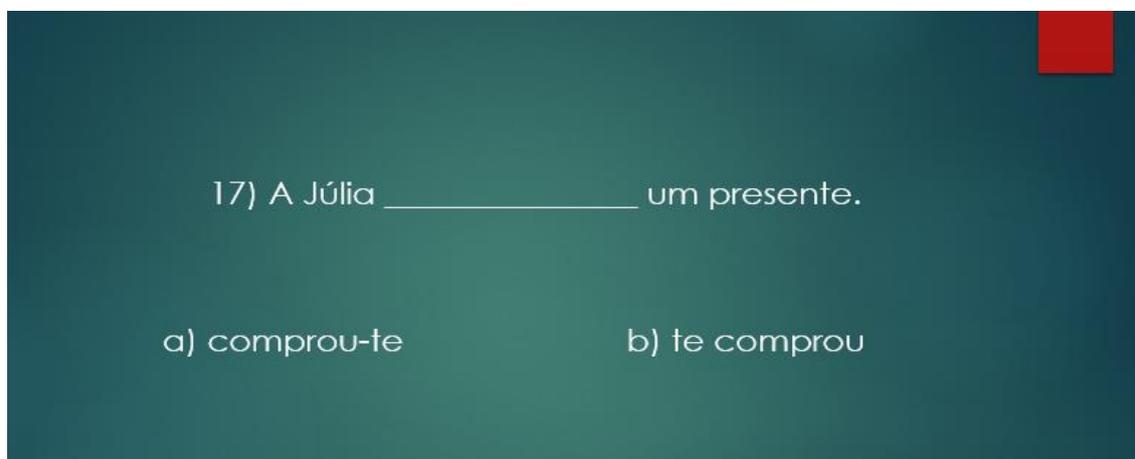


Figura 1. Exemplo de item de teste (contexto de ênclise).



	Inapropriada ←	①	②	③	→ Apropriada
1) – Viste a Marisa ontem? – Vi.		①	②	③	④
2) – Vamos ao teatro? – Sim, não quero ir.		①	②	③	④
3) – Já compraste o presente do Sílvio? – Sim, comprei-o.		①	②	③	④

Figura 4. Exemplo de itens para julgamento.

## 6. Resultados

### 6.1. Teste de completamento de frases

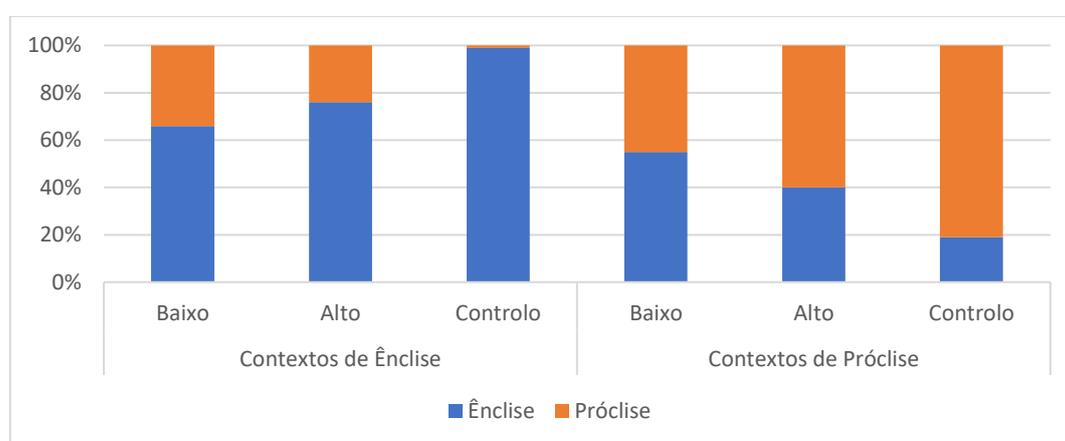


Gráfico 1. Percentagens de colocação do clítico por contexto e grupo.

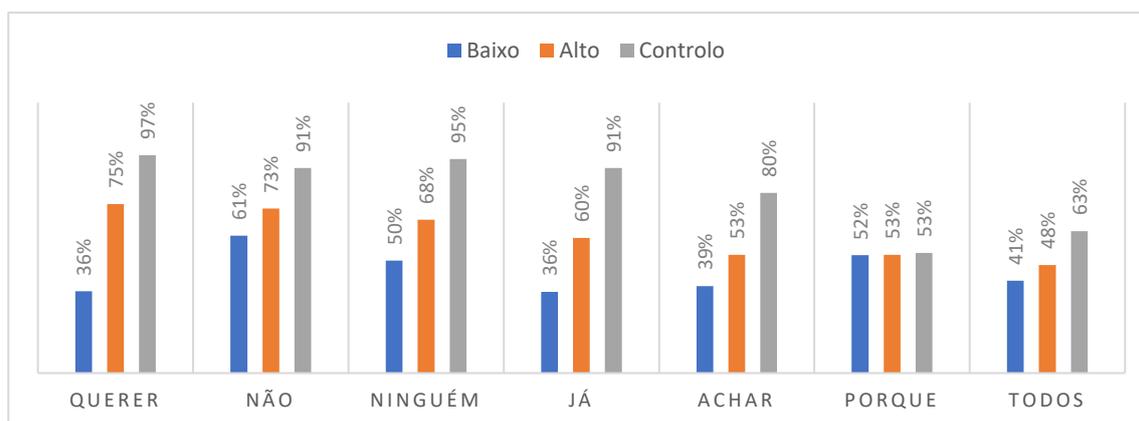


Gráfico 2. Percentagens de respostas alvo por contexto e grupo.

O Gráfico 1 mostra os resultados globais do teste de completamento de frases relativamente ao uso de ênclise ou de próclise. Quanto ao grupo de controlo, o resultado nos contextos de ênclise mostra um desempenho quase categórico (99% de respostas alvo). Os grupos Baixo e Alto tiveram percentagens menores: 66% no grupo Baixo e 76% no grupo Alto. Para a próclise, observaram-se percentagens menores de respostas alvo em todos os grupos, comparativamente à taxa de respostas alvo nos contextos de ênclise.

Os valores variam de 45% no grupo Baixo, passando a 60% no grupo Alto, sendo a taxa do grupo de controlo de 81%.

Ao serem analisados separadamente, os diferentes contextos de próclise não se apresentam de forma homogénea entre os grupos (Gráfico 2). No grupo de controlo, os contextos de orações subordinadas no conjuntivo com o verbo *querer*, sujeito negativo (ninguém), advérbio *já* e negativas são os que apresentam as maiores taxas de respostas alvo (97%, 95%, 91% e 88% respetivamente). O contexto de orações subordinadas no indicativo com o verbo *achar* vem em seguida com 80% de respostas alvo. Por fim, seguem-se os contextos de frases com o quantificador *todos* (63%) e adverbiais com o advérbio *porque* (53%).

No grupo Baixo, a percentagem em frases negativas chega a 61%, seguido por orações adverbiais com “porque” e por frases com o sujeito negativo *ninguém* (52% e 50%, respetivamente). As respostas alvo dos demais contextos ocorrem em taxas abaixo de 40% (subordinadas no indicativo introduzidas pelo verbo *achar* – 39%; subordinadas no conjuntivo introduzidas pelo verbo *querer* – 36%; e frases com o advérbio *já* pré-verbal – 36%), exceto com o sujeito quantificado *todos* (41%).

O grupo Alto apresentou taxas globalmente mais altas do que as do grupo Baixo. Os contextos de próclise com maiores taxas foram os de subordinadas no conjuntivo introduzidas pelo verbo *querer* e os de frases negativas (75% e 73%, respetivamente). Seguem-se nesta escala os contextos de frases com o sujeito negativo *ninguém* (68%), frases com o advérbio *já* pré-verbal (60%), subordinadas no indicativo introduzidas pelo verbo *achar* e adverbiais com *porque* (ambas com 53%) e frases com o sujeito quantificado *todos* (48%).

Para a análise estatística usou-se o programa SPSS (versão 25) para executar um Modelo de Equações Estimativas Generalizadas, sendo a correção da matriz de trabalho não estruturada, os erros padrão robustos e o valor de significância ( $p$ ) estipulado em  $<0,05$ . A análise entre grupos mostrou que são todos diferentes uns dos outros ( $X^2(2) = 139,51$ ,  $p = <0,001$ ), em quaisquer combinações, conforme se vê na Tabela 4:

**Tabela 4: Diferença estatística global entre grupos.**

Grupos		DM	EP	GL	$p$	IC (95%) Mínimo - Máximo	
Baixo	Alto	-0,242	0,085	1	0,013	-0,445	-0,038
Alto	Controlo	-0,264	0,079	1	0,002	-0,453	-0,076
Controlo	Baixo	0,506	0,043	1	$<0,001$	0,402	0,610

DM = diferença média; EP = erro padrão; GL = graus de liberdade; IC = intervalo de confiança.

Na análise dos resultados de cada grupo consoante cada contexto sintático, constatou-se que, à exceção do contexto de adverbiais com *porque*, todas as outras apresentam pelo menos uma diferença entre um par de grupos (Alto e Baixo, Baixo e Controlo, Controlo e Alto). Os contextos de subordinadas (no indicativo e no conjuntivo), de advérbio pré-verbal *já* e de sujeito negativo *ninguém* apresentaram diferença estatística entre todos os pares de grupos. O contexto de frases negativas apresentou diferença entre o grupo de controlo e os grupos Baixo e Alto. Os contextos de sujeito quantificado e de ênclise

mostraram diferença significativa entre o grupo Baixo e o de controlo. A Tabela 5 resume os resultados com diferenças estatísticas.

**Tabela 5: Diferença estatística entre grupos por contexto sintático.**

Grupos	Contextos	DM	EP	GL	p	IC (95%)	
						Mínimo	Máximo
Baixo – Alto	Achar	-0,262	0,980	1	0,021	-0,496	-0,029
	Já	-0,381	0,142	1	0,022	-0,722	-0,401
	Querer	-0,677	0,169	1	<0,001	-1,080	-0,273
	Ninguém	-0,385	0,130	1	0,009	-0,697	-0,074
Baixo – Controlo	Ênclise	0,118	0,048	1	0,040	0,004	0,233
	Achar	-0,552	0,701	1	<0,001	-0,720	-0,384
	Já	-0,820	0,077	1	<0,001	-1,004	-0,636
	Querer	-1,157	0,081	1	<0,001	-1,351	-0,963
	Ninguém	-0,792	0,068	1	<0,001	-0,955	-0,630
	Não	-0,395	0,066	1	<0,001	-0,552	-0,238
	Todos	-0,341	0,096	1	0,001	-0,517	0,083
Alto – Controlo	Não	-0,283	0,113	1	0,036	-0,553	-0,013
	Achar	-0,290	0,093	1	0,006	-0,512	-0,067
	Já	-0,438	0,130	1	0,002	-0,749	-0,127
	Querer	-0,480	0,159	1	0,008	-0,862	-0,099
	Ninguém	-0,407	0,118	1	0,002	-0,690	-0,124

## 6.2. Teste de julgamento de aceitabilidade

Os resultados do grupo de controlo demonstram uma maior aceitabilidade de respostas verbais ( $M = 3,67 - DP = 0,47$ ) e para respostas que envolviam a palavra afirmativa *sim* seguida de um verbo com o objeto direto em forma de clítico ( $M = 3,21 - DP = 0,60$ ). Os participantes rejeitaram as frases com respostas que não condiziam com a pergunta ( $M = 1,15 - DP = 0,33$ ).

Para o grupo Baixo, o contexto mais bem avaliado foi o que envolvia respostas com a palavra afirmativa *sim* ( $M = 3,25 - DP = 0,59$ ). Segue-se o contexto de respostas verbais, com uma média de 3,02 ( $DP = 0,60$ ) e, por fim, o contexto com respostas inapropriadas, com uma média de 1,25 ( $DP = 0,33$ ).

O grupo Alto foi o que proporcionou avaliações mais altas em todos os contextos. O valor mais alto foi obtido com as respostas verbais ( $M = 3,84 - DP = 0,20$ ), seguido de respostas com *sim* ( $M = 3,46 - DP = 0,47$ ) e do contexto com respostas inapropriadas ( $M = 1,36 - DP = 0,36$ ). A Tabela 6 a seguir resume todas as médias:

**Tabela 6: Resultados do teste de aceitabilidade.**

Grupo	Contexto das Respostas		
	Verbais	Sim	Inapropriadas
Baixo	3,02 (DP = 0,60)	3,25 (DP = 0,59)	1,25 (DP = 0,33)
Alto	3,84 (DP = 0,20)	3,46 (DP = 0,47)	1,36 (DP = 0,36)

Controlo	3,67 (DP = 0,47)	3,21 (DP = 0,60)	1,15 (DP = 0,33)
----------	------------------	------------------	------------------

Realizou-se uma análise estatística com o programa SPSS (versão 25) para determinar se há correlação entre as taxas de respostas alvo de ênclise e uma maior aceitação de respostas verbais. A análise global, feita com um teste de Spearman, à partida, não demonstrou nenhuma correlação entre as duas variáveis ( $r(1) = 0,244$ ,  $p = 0,084$ ). No entanto, este valor pode ser considerado marginalmente significativo. Todos os resultados serão discutidos na próxima secção.

## 7. Discussão e conclusão

De acordo com Lardiere (2009), a reconfiguração dos traços é o ponto culminante de um processo iniciado pelo estágio de mapeamento. O aprendente da L2 estabelece uma correspondência entre formas da sua L1 e formas da L2, determinando que estas se caracterizam pelo conjunto de propriedades funcionais da forma presente na sua L1. Tais propriedades podem ser realizadas tanto como material lexical (palavras ou morfemas), quanto por meio da ordem das palavras (Shimanskaya 2015). No entanto, esse processo produz falhas quando (i) os traços estão presentes numa língua e ausentes na outra, ou ainda quando (ii), ainda que presentes em ambas, são expressos de diferentes formas. Assim, é necessário ter lugar o processo de reconfiguração, que ocorre de maneiras distintas consoante a relação de presença-ausência dos traços no par de línguas. No caso de um traço estar presente na L2 e ausente na L1, o aprendente tem de buscá-lo no arcabouço de traços da GU; no contrário, deve “abandonar” o traço quando se expressa na L2. No segundo caso, o aprendente deve construir novas relações entre traços e formas.

No caso dos pronomes clíticos no par de línguas espanhol-PE, vê-se a expressão dos traços por meio do padrão de colocação pronominal, facto que Martins (1994, 2002) sugere ter que ver com a categoria  $\Sigma$ , a qual apresenta traços-V fortes nesta e fracos naquela. Portanto, a tarefa derradeira dos falantes nativos de espanhol e aprendentes de PE, consoante os clíticos, é a de adquirir o traço-V forte da categoria  $\Sigma$ , o que resultaria na colocação pronominal do português.

Durante o mapeamento, é possível que os falantes de nível de proficiência mais baixo correlacionem os clíticos em espanhol com os clíticos em PE – são línguas tipologicamente próximas, com uma morfologia muito similar. Portanto, é de se esperar que os falantes nativos de espanhol relacionem os seus clíticos morfologicamente com os do PE, mas baseados nos padrões sintáticos de colocação do seu idioma materno. Isso significaria que os aprendentes transfeririam as propriedades da categoria  $\Sigma$  para a L2. O resultado seria próclise em todos os contextos, pois os utilizados neste trabalho são contextos de próclise no espanhol.

No entanto, o que se observa vai contra o proposto no parágrafo anterior – há taxas altas de ênclise no grupo Baixo deste estudo em todos os contextos (de ênclise e de próclise – ainda que menores do que as taxas do grupo Alto e do grupo de controlo, o que confirma a Hipótese 1). Isso não refuta de todo a HRT. Não se sabe exatamente em que momento o aprendente “nota” que o mapeamento falhou (dando início à reconfiguração) e quando o aprendente finalmente reconfigura os traços. A HRT está orientada para

afirmar que todos os traços podem ser eventualmente adquiridos a partir do momento em que o contraste é detetado na L2 (Lardiere 2009; Shimanskaya 2015). É possível que os aprendentes possam detetar precocemente a diferença na colocação dos clíticos em espanhol e português: a ênclise é, possivelmente, uma estrutura saliente, pois é agramatical em espanhol nos contextos em questão. Apesar da presença no *input* de ambas as colocações, não significa que os aprendentes efetivamente se apercebam dos contextos em que a próclise é usada. Isso é evidenciado de forma mais marcada pelos resultados do grupo Baixo, com ocorrências mais equilibradas de cada tipo de colocação, independentemente do contexto avaliado.

Considerando o processo de aquisição de clíticos em PE L1, vê-se que é a ênclise que as crianças generalizam nas fases iniciais do desenvolvimento, inclusive com ênclise em contextos de próclise, enquanto a ocorrência de próclise em contextos de ênclise é marginal (Costa *et al.* 2015, 2016). Neste estudo encontram-se não só taxas consideráveis de próclise em contextos de ênclise, mas também menos ênclise em contextos de próclise do que o observado com as crianças no estudo de Costa *et al.* (2015) e nos estudos em aquisição de L2, o que indica que o processo em questão difere quanto ao processo de aquisição dos padrões de colocação em português L1. Ora, a HRT prediz que o ponto de partida da aquisição é a L1 do aprendente. Sendo assim, observam-se indícios de transferência inicial da gramática do espanhol. Porém, como citado anteriormente, o *input* recebido pelos aprendentes apresenta uma evidência clara de ênclise. Isto sugere que o processo de reestruturação da gramática mental dos aprendentes se dá em fases iniciais do processo de aquisição, o que explica a alternância entre as duas possibilidades de colocação, com mais ênclise nos contextos de ênclise do que nos de próclise, e a aquisição gradual dos contextos em que a próclise deve ocorrer.

Portanto, há evidências de transferência dos traços funcionais relevantes para a colocação dos clíticos nos estádios iniciais de aquisição, mas é necessário estabelecer a natureza desse processo. Salienta-se aqui que é provável que os sujeitos que compõem o grupo Baixo desta pesquisa já não estejam no nível inicial de aquisição, pois, pelo método de aferição de nível, o agrupamento não possibilitou uma amostra muito homogênea.<sup>10</sup> Assim, observam-se indícios do estágio de mapeamento (i.e., transferência), mas as suas gramáticas já devem estar em reestruturação.

Quando se comparam os resultados do grupo Alto aos resultados do grupo Baixo, vê-se que as taxas de respostas alvo obtidas são todas superiores. Aliás, a análise estatística demonstrou que o grupo Alto tem um comportamento diferente do grupo Baixo ( $p = 0,013$ ), apesar de ele variar por contextos. Isso demonstra que, ainda que haja variabilidade (reflexo dos diferentes estádios de desenvolvimento individuais), o grupo Alto apresenta produções mais próximas das produções do grupo de controlo, confirmando a Hipótese 2. A aparente reconfiguração dos traços possibilita responder parcialmente de modo afirmativo à Questão 2: há evidências na ênclise, mas não na próclise. Novamente, o facto de os participantes em cada grupo não se apresentarem de

---

<sup>10</sup> Os desvios padrão dos resultados do teste de aferição de nível de ambos os grupos parecem corroborar esta hipótese, ainda que, isoladamente, não sejam indicadores suficientes.

forma homogénea (além do seu número reduzido) relativamente ao seu nível de proficiência pode ter causado interferências nos resultados.

A Hipótese 3 previa uma relação entre a produção alvo em contexto de ênclise e a aceitabilidade de respostas verbais, pois ambas estão relacionadas com os traços-V fortes da categoria  $\Sigma$  (Martins 1994). Porém, não é possível ter conclusões muito fundamentadas. Conforme a Tabela 6, nota-se uma evolução na aceitação do grupo Baixo para o grupo Alto, da mesma forma que as taxas de respostas alvo no contexto de ênclise (vd. Gráfico 1). Contudo, a análise estatística demonstrou uma relação marginal entre a aceitação e a colocação alvo ( $p = 0,084$ ). Por outro lado, por ser pequena a dimensão da amostra, e também baixo era o número de itens de teste para julgamento de aceitabilidade, o método estatístico utilizado não teve um poder de análise muito robusto. Logo, não se pode nem refutar, nem confirmar o proposto por Martins (1994). São necessárias investigações futuras para que se obtenham dados mais contundentes a que se refere a correlação entre a ênclise e as respostas verbais no português.

A última questão de investigação estava relacionada com os contextos de próclise, mais especificamente, se existia um percurso de aquisição deles. Como se pode ver no Gráfico 2, o grupo Baixo apresenta taxas de respostas alvo nos contextos de próclise visivelmente mais baixas do que as do grupo Alto. Naquele grupo, os contextos relativos à negação (frases negativas e sujeito negativo *ninguém*) e contexto de adverbiais causais com *porque* apresentaram-se com taxas superiores às dos outros contextos, de maneira similar ao que ocorre em crianças nativas de PE: os contextos de negativas emergem primeiro, enquanto os de adverbiais causais e de sujeito quantificado surgem mais tardiamente (Costa, Fiéis & Lobo 2015).

Apesar de semelhantes – a tendência mostra-se clara relativamente a uma hierarquia na aquisição – uma comparação *vis-à-vis* dos resultados não é aconselhada, visto que os processos de aquisição de L1 e de L2, apesar de similares, não são idênticos (Ipek 2009). Assim, comparam-se os resultados aos obtidos por Madeira *et al.* (2006), Madeira e Xavier (2009) e Gu (2019). A Tabela 7 compila os dados desses estudos e desta pesquisa por contexto de próclise e níveis de proficiência (ou idade, no caso de Costa *et al.* 2015).

Madeira *et al.* (2006) obtiveram uma sequência de aquisição parecida à deste estudo (apesar de as percentagens terem sido diferentes) (vd. Tabela 7). Notam-se os contextos com interrogativas-qu, com negação e com advérbios pré-verbais no topo da escala. Entretanto, os contextos analisados não foram de todo idênticos. Como mencionado, as autoras consideram contextos com palavras-qu, além de orações infinitivas introduzidas por preposição, as quais não são consideradas neste estudo. Por outro lado, aqui os contextos de orações subordinadas são divididos em subordinadas no indicativo ou no conjuntivo; em Madeira *et al.* (2006) não há essa diferenciação. Outro ponto que dificulta uma comparação *vis-à-vis* entre os estudos é a metodologia empregada, pois Madeira *et al.* (2006) utilizaram um teste de produção induzida. Ademais, os valores da amostra foram baixos (*i.e.*, <10 indivíduos por grupo), o que não é aconselhável em termos estatísticos.

A tendência global mencionada anteriormente repete-se em Madeira e Xavier (2009): os aprendentes de nível elementar durante a produção espontânea colocam o clítico na posição alvo em proporções menores do que os falantes de nível avançado,

sejam eles falantes nativos de línguas românicas ou germânicas (*vd.* Tabela 7). Entretanto, mais uma vez, uma comparação integral dos dois estudos não é possível. Concernente à metodologia, Madeira e Xavier (2009) recolheram os dados de um *corpus* de produção escrita, enquanto este trabalho se apoia em um teste de completamento de frases por seleção de opções – ou seja, os participantes acabam por eleger uma resposta ou outra, sem casos de omissão, a qual foi constatada pelas autoras, por exemplo, no contexto de quantificadores pré-verbais no grupo Avançado de línguas germânicas, em que não se produziu nenhum clítico nesse contexto. Note-se que, novamente, os contextos avaliados não são exatamente os mesmos.

Tabela 7. Percentagens de colocação alvo por diferentes contextos de próclise.

	Este Estudo		Costa et al. (2015)				Madeira et al. (2006)				Madeira e Xavier (2009)				Gu (2019)	
	B	A	5	6	7	R1	R2	G1	G2	IR	AR	IG	AG	≤2	>2	
Negação	61%	68%	61%	88%	84%	83,3%	89%	52%	76%	71%	82%	60%	100%	42%	78%	
Ninguém	50%	68%	48%	69%	71%											
Subordinadas (conjuntivo)	36%	75%	47%	75%	66%											
Interrogativas-qu						75%	92%	56%	61%	100%	93%	33%	90%			
Advérbio “Já”	36%	60%	36%	58%	66%											
Advérbio “Também”														22%	68%	
Advérbios (em geral)						43%	83%	43%	58%	N/A	75%	50%	100%			
Subordinadas (indicativo)	39%	53%														
Subordinadas (em geral)						20%	46%	13%	41%	42%	72%	100%	91%			
Adverbiais Causais	52%	53%	14%	20%	33%									20%	54%	
Infinitivas com preposição						27%	44%	22%	21%							
Quantificador “Todos”	41%	48%	10%	12%	21%									13%	24%	
Quantificadores (em geral)						23%	39%	17%	18%	N/A	100%	N/A	N/A			
Não finitos										N/A	31%	75%	22%			

Nota. As médias foram arredondadas. B = Baixo; A = Alto; 5 = 5 anos de idade; 6 = 6 anos de idade; 7 = 7 anos de idade; R1 = L1 espanhol, italiano, francês com tempo de aprendizagem de 7 meses; R2 = L1 espanhol e italiano com tempo de aprendizagem de 1 a 2 anos; G1 = L1 alemão com tempo de aprendizagem de 2 a 9 meses; G2 = L1 alemão e inglês com tempo de aprendizagem de 1 a 2 anos; I.R = Iniciados Língua Românica; A.R = Avançados Língua Românica; I.G = Iniciados Língua Germânica; M.A = Avançados Língua Germânica; ; ≤2 = igual ou menos de dois anos de estudo; >2 = mais de dois anos de estudo; N/A = não foi produzido nenhum clítico neste contexto.

Ressalta-se que os dois estudos previamente mencionados não foram homogêneos quanto à L1 dos participantes. Neste estudo, houve tal uniformidade, de modo que todos os sujeitos investigados tivessem o espanhol como língua materna. Como mencionado anteriormente, o ponto de partida para a aquisição da L2, segundo a HRT, é a L1. Se os sujeitos participantes nas pesquisas tiverem L1 diferentes, esperam-se percursos de aquisição distintos – o que é perceptível em Madeira e Xavier (2009), mesmo que os grupos só sejam uniformes na tipologia das L1.

Assim, o estudo de Gu (2019) traz dados mais concisos sobre a uniformidade dos falantes investigados: a autora trabalhou com falantes nativos de chinês. Gu (2019), por meio do seu teste de produção induzida, constatou um aumento nas produções alvo de próclise com o aumento do nível de proficiência. Já nos contextos de ênclise, os participantes foram categóricos na colocação enclítica nos contextos que a exigiam, o que não se observou neste estudo (*vd.* Tabela 8). Tal facto pode estar relacionado com o estágio de desenvolvimento pelo qual o grupo de nível de proficiência mais baixo está a passar. Como mencionado anteriormente, as taxas de próclise em contexto de ênclise produzidas pelo grupo Baixo podem ser indícios de transferência inicial da gramática do espanhol. O chinês, por outro lado, possui pronomes fortes e, logo, essa transferência inicial, pelo menos no que tange às propriedades dos clíticos, não teria lugar.

**Tabela 8: Percentagens globais de colocação enclítica e próclítica.**

	Presente Estudo				Gu (2019)			
	B		A		≤2		>2	
	Ênclise	Próclise	Ênclise	Próclise	Ênclise	Próclise	Ênclise	Próclise
Alvo	66%	45%	76%	60%	100%	24,29%	100%	56,25%
Não Alvo	34%	55%	24%	40%	0%	75,71%	0%	43,75%

*Nota.* B = Baixo; A = Alto; ≤2 = dois anos ou menos de estudo; >2 = mais de dois anos de estudo.

Quanto aos diferentes contextos de próclise, vê-se a seguinte sequência de aquisição: primeiramente adquire-se o contexto de frases negativas, em seguida com advérbio pré-verbal *também*, com adverbiais causais com *porque* e por último com sujeito quantificado *todos*. A progressão é idêntica em ambos os grupos avaliados (≤2 anos de estudo e >2 anos de estudo). Comparativamente a este estudo, cujos contextos são em maior número, observa-se a tendência que as frases negativas têm em ocupar posições mais altas, enquanto na ponta inversa estão as adverbiais causais e o sujeito quantificado, como exemplificado pela Tabela 7.

Portanto, vê-se que os contextos relacionados à negação foram adquiridos primeiro, confirmando a hipótese 4, o que possibilita responder parcialmente à questão 3, acerca da existência de uma sequência de desenvolvimento dos diferentes contextos. De acordo com Martins (1994), a ocorrência da próclise deve-se às características da categoria  $\Sigma$ , que possui traços-V fortes. Em frases negativas, há presença de material lexical em  $\Sigma$ , o que impede o movimento do verbo para tal categoria. Com quantificadores ou com certos advérbios pré-verbais, ocorre a instanciação de  $\Sigma$  como F(oco), que não possui traços-V e, portanto, não ocorre o movimento do verbo. Já em orações subordinadas, a fusão da categoria C com  $\Sigma$  forma o núcleo complexo  $[C \Sigma C]$ , que também possui material lexical. Aparentemente a natureza dessas operações interfere nas produções. Pode-se postular

que, para os falantes nativos de espanhol, a próclise em contexto de negativas se consolida mais precocemente porque o traço de negação já existe naquele idioma (ou porque a próclise em frases negativas corresponde a uma configuração estrutural idêntica em PE e em espanhol). A tarefa do aprendente seria de adquirir o traço-V forte para a categoria  $\Sigma$ .

No grupo dos quantificadores e de certos advérbios pré-verbais, é possível que a tarefa envolva também a decodificação de quais quantificadores e de quais advérbios ocorrem em F. Aqui a dependência recai sobre o conhecimento lexical (como apontado por Costa *et al.* 2015) – enquanto, segundo a norma, todas as frases negativas levam à próclise, somente alguns quantificadores e advérbios são responsáveis pela sua ocorrência (Martins 2013).

No caso das subordinadas, teoricamente, o material lexical encontrado no núcleo complexo  $[C \Sigma C]$  deveria ser o bastante para que não houvesse movimento do verbo para  $\Sigma$  e a próclise derivasse. Contudo, veem-se comportamentos distintos de acordo com o modo verbal das orações subordinadas e também com a conjunção introdutora da oração, ocorrendo a próclise em maiores taxas no modo conjuntivo do que no indicativo, do mesmo modo que subordinadas introduzidas por *que* possuem maiores taxas do que as introduzidas por *porque* (ainda que neste trabalho tenha havido mais próclise com *porque* do que com subordinadas no indicativo introduzidas pelo verbo *achar* no grupo Baixo). Isto pode ter que ver com o *input* recebido pelos falantes. Como visto anteriormente, existe variabilidade nos contextos de próclise entre adultos. Portanto, é exetável que as formas que ocorram mais consistentemente no sistema sejam adquiridas primeiro, enquanto aquelas que apresentam maior variabilidade tendam a ter uma evolução mais lenta. Esse movimento é visível: as taxas mais altas de próclise no grupo Alto coincidem, em geral, com as taxas mais altas do grupo de controlo. Sob a perspetiva da HRT, isso implica estar em um ambiente que forneça um maior *input* aos aprendentes, de modo que a reconfiguração dos traços tenha dados mais robustos para ocorrer consistentemente.<sup>11</sup>

Outro ponto a ser levado em consideração é o conhecimento metalinguístico dos participantes. Os estudos de aquisição de clíticos em PE L2 supramencionados utilizaram aprendentes formais de português e a colocação pronominal é um dos tópicos ensinados em sala de aula (Costa 2018). Isso não deve ser ignorado. Por exemplo, durante um ensaio de aplicação do teste com falantes de PE L2 cujas L1 não eram o espanhol, um dos participantes disse lembrar-se de que a próclise era utilizada em frases negativas ou que continham palavras-qu – dois contextos com altas taxas de produção nos estudos mencionados anteriormente. Logo, os participantes podem ter feito uso desse conhecimento para produzir respostas alvo, principalmente no corpus de Madeira e Xavier (2009), por terem sido obtidas produções escritas. No presente estudo, apenas quatro sujeitos mencionaram ter tido alguma vez aulas de português, pelo que os efeitos do conhecimento metalinguístico devem ter sido minimizados e a aquisição deu-se num ambiente naturalístico.

---

<sup>11</sup> Um parecerista anónimo comenta que, no caso das subordinadas, é provável que traços relacionados ao modo verbal também tenham um papel no padrão de colocação dos clíticos no PE, visto que a variação na colocação em subordinadas é mais comum quando o verbo está no indicativo. Neste sentido, a tarefa envolveria a aquisição de mais traços além dos traços-V, sendo, portanto, um ponto que pode ser abordado em investigações futuras.

Apesar da variedade de dados, os resultados obtidos durante esta pesquisa e pelos estudos prévios não surpreendem. A HRT prevê que os falantes tracem trajetórias de aquisição diferentes, visto que a disponibilidade de traços consoante a L1 do aprendente é diferente dos traços existentes no português. Os trabalhos aqui revistos trouxeram amostras com diferentes L1. Gu (2019), no entanto, trabalhou somente com falantes de chinês. Não obstante, apenas uma parte dos contextos analisados pela autora coincidiram com aqueles analisados neste estudo, o que dificulta uma comparação mais precisa. Ainda, sob a luz da HRT, para que a análise fosse mais robusta, seria necessária a descrição dos traços existentes no chinês de modo que se estabelecessem hipóteses em relação aos estádios de mapeamento e de reconfiguração dos traços.

O processo de aquisição de L2 é complexo e envolve diversas variáveis como o nível de desenvolvimento individual, a qualidade do *input* e as propriedades gramaticais das L1 dos aprendentes. Em síntese, a HRT proposta por Lardiere (2009) mostrou indícios de ser válida relativamente à aquisição dos padrões de colocação dos pronomes clíticos do PE por falantes de espanhol L1. Isso é evidenciado pelos sinais de transferência inicial, ainda que parcial, da gramática do espanhol e pelo movimento em direção à produção do grupo de controlo quando se comparam os resultados do grupo Baixo com o do grupo Alto. Contudo, a forma de organização dos participantes nos grupos supramencionados não foi ideal, o que dificulta uma análise mais precisa do fenómeno. Portanto, investigações futuras devem ser conduzidas de modo que se comprove a natureza dessa reconfiguração, seja ela associada aos traços da categoria funcional  $\Sigma$ , como proposto por Martins (1994), ou associada a outro traço disponível no português.

**Agradecimentos:** Agradece-se à Prof<sup>a</sup> Dr<sup>a</sup> Ana Madeira do Centro de Linguística da Universidade Nova de Lisboa (CLUNL) pelo auxílio em todas as etapas da realização deste estudo, à Dr<sup>a</sup> Luísa Pilz do Laboratório de Cronobiologia e Sono do Hospital de Clínicas de Porto Alegre pelo auxílio na realização dos testes estatísticos e aos revisores anónimos, cujas sugestões permitiram aperfeiçoar este artigo. Agradece-se também a todos os participantes pela sua disponibilidade em participar no estudo.

## Referências

- Costa, A. (2018). Variação linguística e ensino de Português, Língua Não Materna: O caso dos pronomes átonos. *Medi@ções, Revista OnLine*, 6(1), pp. 147–161.
- Costa, J. & Lobo, M. (2007). Clitic omission, null objects or both in the acquisition of European Portuguese? In S. Baauw, F. Drijkoningen, & M. Pinto (Eds.), *Romance languages and linguistic theory 2005* (pp. 59–72). Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.291.06cos>
- Costa, J. & Lobo, M. (2009). Clitic omission in the acquisition of European Portuguese: Data from comprehension. In A. Pires, & J. Rothman (Eds.), *Minimalist inquiries into child language acquisition. Case studies across Portuguese* (pp. 63–84). Berlin/Nova Iorque: Mouton de Gruyter. <https://doi.org/10.1515/9783110215359.1.63>
- Costa, J., Fiéis, A. & Lobo, M. (2015). Input variability and late acquisition: Clitic misplacement in European Portuguese. *Lingua*, 161, pp. 10–26. <https://doi.org/10.1016/j.lingua.2014.05.009>

- Costa, J., Fiéis, A. & Lobo, M. (2016). A aquisição dos pronomes clíticos no português L1. In A. M. Martins, & E. Carrilho (Eds.), *Manual de linguística portuguesa* (pp. 365–386). Berlin: De Gruyter. <https://doi.org/10.1515/9783110368840-018>
- Duarte, I. (2003). Padrões de colocação dos pronomes clíticos. In M. H. M. Mateus *et al.* (Eds.), *Gramática da língua portuguesa* (pp. 847–867). Lisboa: Caminho.
- Gu, W. (2019). Aquisição de pronomes clíticos de português europeu por falantes de chinês: Dados sobre a colocação. *Revista da Associação Portuguesa de Linguística*, 5. <https://doi.org/10.26334/2183-9077/rapln5ano2019a14>
- Ipek, H. (2009). Comparing and constrasting first and second language acquisition: Implication for language teachers. *English Language Teaching*, 2(2), pp. 155–163. <https://doi.org/10.5539/elt.v2n2p155>
- Klee, C. A. (1990). Spanish-Quechua language contact: The clitic pronoun system in Andean Spanish. *Word*, 41(1), pp. 35–46. <https://doi.org/10.1080/00437956.1990.11435812>
- Laka, I. (1990). *Negation in syntax. On the nature of functional categories and projections* (Tese de Doutoramento, Massachusetts Institute of Technology, Cambridge).
- Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, 25, pp. 173–227. <https://doi.org/10.1177/0267658308100283>
- Madeira, A. M. & Xavier, M. F. (2009). The acquisition of clitic pronouns in L2 European Portuguese. In A. Pires, & J. Rothman (Eds.), *Minimalist inquiries into child and adult language acquisition* (pp. 273–299). Berlin/New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110215359.2.273>
- Madeira, A. M., Crispim, M. L. & Xavier, M. F. (2006). Clítico pronominal em português L2. In J. Barbosa & F. Oliveira (Orgs.), *APL - Textos selecionados. XXI Encontro Nacional da Associação Portuguesa de Linguística*, pp. 495–510. Lisboa: Colibri.
- Martins, A. M. (1994). Enclisis, VP-deletion and the nature of Sigma. *Probus*, 6, 173–206. <https://doi.org/10.1515/prbs.1994.6.2-3.173>
- Martins, A. M. (2002). Tipologia e mudança linguísticas: os pronomes pessoais do português e do espanhol. In E. Raposo, & H. Sharrer (Eds.), *Santa Barbara Portuguese studies VII* (pp. 340–386).
- Martins, A. M. (2013). A posição dos pronomes pessoais clíticos. In E. P. Raposo, M. F. Bacelar, M. A. Mota, L. Segura, & A. Mendes (Eds.), *Gramática do português* (pp. 2231–2302). Lisboa: Fundação Calouste Gulbenkian.
- Martins, A. M. (2016). A colocação dos pronomes clíticos em sincronia e diacronia. Em A. M. Martins, & E. Carrilho (Eds.), *Manual de Linguística Portuguesa* (pp. 401–430). Berlin/Boston: De Gruyter.
- Santos, M. F. (2002). *Os pronomes pessoais átonos no português europeu. Descrição de problemas que ocorrem no 3º ciclo e proposta de actividades didáticas* (Dissertação de Mestrado, FLUL, Lisboa).
- Shimanskaya, E. M. (2015). *Feature reassembly of semantic and morphosyntactic pronominal features in L2 acquisition* (Tese de Doutoramento, Universidade de Iowa, Iowa).
- Shimanskaya, E. & Slabakova, R. (2015). Re-assembling objects: A new look at the L2 acquisition of pronominal clitics. *Bilingualism: Language and Cognition*, 20(3), 1–18. <https://doi.org/10.1017/S1366728915000784>
- Soriano, O. F. (2015). Clíticos. In J. Gutiérrez-Rexach (Ed.), *Enciclopedia de lingüística hispánica* (pp. 423–436). Londres: Routledge.

[recebido em 15 de março de 2021 e aceite para publicação em 27 de fevereiro de 2022]

## ACEITABILIDADE DE SENTENÇAS DE TÓPICO-COMENTÁRIO EM PB E EM INGLÊS: UM ESTUDO PSICOLINGUÍSTICO

### THE ACCEPTABILITY OF TOPIC-COMMENT SENTENCES IN BP AND IN ENGLISH: A PSYCHOLINGUISTIC STUDY

Lorrane da Silva Neves Medeiros Ventura\*  
lorranesnm@hotmail.com

O presente estudo investiga estruturas de tópico-comentário e sujeito-predicado no Português do Brasil (PB), sob a luz da psicolinguística experimental, a fim de trazer suporte empírico que sustente a hipótese de que o PB seria uma língua mista, orientada tanto para a sentença quanto para o discurso, diferente do inglês, que é, indiscutivelmente, uma língua com proeminência de sujeito, orientada para a sentença (*cf.* Li & Thompson 1976). Investigamos estruturas de tópico geradas na base (tópico estilo-chinês), pois tais estruturas são prototípicas de uma língua onde a noção de tópico é importante e ausentes em uma língua onde a noção de sujeito é a dominante (*cf.* Yuan 1995), além de estruturas de tópico geradas via movimentação sintática (topicalização). O experimento de julgamento de aceitabilidade demonstrou que o comportamento dos falantes de PB e de inglês foi significativamente diferente durante a tarefa proposta, o que pode indicar que as línguas não se encaixam na mesma tipologia, quanto a sua orientação. O resultado do teste com o grupo bilíngue confirmou, ainda mais, as suposições iniciais, de que PB e inglês não enxergam o tópico da mesma maneira. Tais resultados sugerem que o PB deva ser caracterizado como uma língua onde ambos, tópico e sujeito, são noções importantes (*cf.* Pontes 1987, entre outros).

**Palavras-chave:** Tópico. Português do Brasil. Inglês. Julgamentos de aceitabilidade. Psicolinguística

The aim of this study is to investigate topic-comment and subject-predicate structures in BP, through the experimental psycholinguistics approach, in order to bring empirical support to the hypothesis that BP would be a mixed language, in which both, topic and subject ideas are important. In other words, BP would be a discourse-oriented and sentence-oriented language, unlike English, that is, undoubtedly, a sentence-oriented language (*cf.* Li & Thompson 1976). Base-generated topic structures (Chinese-style topic) will be investigated, because such structures are prototypical in a topic-prominent language, as Chinese, in which the topic notion is important, and absent in a language in which the subject notion is predominant, as English (*cf.* Yuan 1995), in addition to the topic structures generated through syntactic movement (topicalization). The acceptability judgment test has demonstrated that the native speakers of BP and the native speakers of English presented a different behavior during the task, which can be an indicative that the languages are not the same, considering its typology. These results suggest that BP should be considered as a language in which both, topic and subject, are important notions (*cf.* Pontes, 1987, and others).

---

\* Doutora em Linguística pelo Programa de Pós-graduação em Linguística da UFRJ, Faculdade de Letras, Universidade Federal do Rio de Janeiro, Brasil. ORCID: 0000-0003-1409-3597

**Keywords:** Topic. Brazilian Portuguese. English. Acceptability judgments. Psycholinguistic

•

## 1. Introdução

O tópico–comentário é um mecanismo bastante produtivo nas línguas naturais (*cf.* Li & Thompson 1976), por meio do qual destaca-se um elemento na periferia esquerda da sentença, e sobre ele, faz-se um comentário. Esse comentário pode ser de dois tipos: pode ser formado de uma sentença completa com sujeito, verbo e objeto (como “Massa, eu adoro lasanha.”) ou pode estabelecer alguma relação sintática com o constituinte que foi topicalizado (como, “Lasanhai, minha mãe comprou \_\_i no mercado”).

De acordo com Li e Thompson (1976), a ordenação de constituintes via tópico-comentário é um fenômeno presente em todas as línguas naturais, apesar de cada língua particular ter uma distribuição diferente dessas construções. As estruturas de tópico-comentário diferenciam-se das estruturas de sujeito-predicado, por conterem traços característicos que as definem, como a saliência sintática, prosódica e discursiva, por exemplo. A sentença presente em 1 apresenta a ordenação de tópico-comentário, em contraste, tem-se a estrutura de sujeito-predicado (Exemplo 2), no PB:

- (1) *Aquele aluno*, o professor reprovou ontem.
- (2) O professor reprovou *aquele aluno* ontem.

Na sentença (1), temos o elemento “Aquele aluno” ocupando a periferia esquerda da sentença, uma posição não argumental na estrutura, sendo o tópico. Seguindo este elemento topicalizado, temos o comentário “o professor reprovou ontem”, que refere-se ao tópico, além de apresentar uma relação sintática entre as duas posições, já que o elemento que foi topicalizado pode ser realocado na posição pós verbal, no comentário. Em contraste, temos (2), uma sentença na ordenação sujeito-predicado, com o elemento “aquele aluno”, ocupando a posição pós-verbal, canônica, argumental, da estrutura.

O PB parece fazer uso bem assíduo de construções de tópico-comentário, sendo tal fenômeno muito produtivo na língua, conforme apontam diversas obras sobre o tema (entre outros *cf.* Berlinck, Duarte & Oliveira 2015; Duarte 1996; Galves 2001; Kato 2006; Kenedy 2002, 2014; Maia 1997; Negrão 1999; Orsini 2003; Pontes 1987). Tal uso tão assíduo na língua fez com que nascesse uma discussão bem ativa na literatura sobre qual seria a tipologia do PB, ao se considerar a caracterização proposta por Li e Thompson (1976), e isso vem despertando grande interesse entre estudiosos do PB ao longo dos últimos 30 anos, pelo menos desde o trabalho seminal de Pontes (1987).

## 2. Suporte teórico

Em estudo clássico, Li e Thompson (1976) lançam a base de uma tipologia linguística baseada na relação gramatical entre sujeito-predicado e tópico-comentário. Como exemplificação do que seria uma estrutura de sujeito-predicado e uma estrutura de tópico-comentário, os autores dão dois exemplos de sentenças em Inglês:

- (3) John.Sujeito hit Mary.Predicado  
'O John bateu na Mary'.
- (4) As for education,Tópico John prefers Bertrand Russell's ideas.Comentário  
'Quanto a educação, o John prefere as ideias de Bertrand Russell.'

(Exemplos de Li & Thompson 1976, p. 460)

De acordo com os autores, em línguas de sujeito, a estrutura mais básica seria similar ao Exemplo 3, onde temos uma estrutura composta de Sujeito + Verbo + Complemento, enquanto em línguas de tópico, a estrutura mais básica seria similar a frase em 4, composta de Tópico + Comentário, sendo o comentário composto de uma sentença completa, SVO.

Baseando-se no que seria a estrutura mais básica nas línguas naturais em relação as duas estruturas em questão, os autores propõem que as línguas sejam divididas em quatro grandes tipos, em relação à ordenação mais frequente (básica) dos constituintes em uma sentença, de acordo com sua orientação, sendo para o discurso ou para sentença. São eles, a saber: i) línguas com proeminência de sujeito, onde a ordenação mais básica se daria em termos de Sujeito+Verbo+Complemento; ii) línguas com proeminência de tópico, onde a ordenação mais frequente seria via Tópico+Comentário; iii) línguas com proeminência de tópico e de sujeito, ou mistas, onde as duas estruturas seriam igualmente básicas e iv) línguas sem proeminência de tópico e de sujeito, onde ambas as estruturas se mesclam e não é possível diferenciá-las.

O trabalho de Pontes (1987) inaugura o interesse da questão sobre as construções de tópico no PB. No trabalho, Pontes analisa o português oral, falado no dia-a-dia, da classe culta. Pontes verifica que as estruturas de tópico-comentário não são apenas altamente recorrentes, mas também estão presentes em variados tipos. A autora, então, analisa as características das línguas com proeminência de tópico, mostradas por Li e Thompson (1976) e relacionando essas características com o PB, conclui que o PB deveria ser considerado, no mínimo, como uma língua mista e não como uma língua com proeminência de sujeito, conforme vinha sendo considerada.

Enquanto parece ser um consenso na literatura linguística descritiva a língua ser de tópico proeminente, ou no mínimo mista, (Berlinck *et al.* 2015; Orsini 2003; Orsini & Vasco 2007; Pontes 1987; entre outros), a literatura psicolinguística parece considerar que a língua seria, no entanto, orientada para a sentença, tendo o sujeito proeminente, a partir da análise de dados de compreensão (Kenedy 2011, 2014; Kenedy & Mota 2012; Silva 2015; Silva 2017) e produção (Silva 2015). O presente estudo tem como objetivo

trazer dados empíricos, oriundos de experimentação psicolinguística, que sustentem a hipótese de que o PB seria, de fato, uma língua mista.

### 2.1. As estruturas de tópico estilo-chinês e a orientação para o discurso

Li e Thompson (1976) consideram que uma das características mais importantes para se verificar a proeminência de uma língua, em relação a sua orientação, para a sentença e/ou para o discurso, é a ocorrência de Duplo sujeito. Para os autores, tais estruturas seriam prototípicas de uma língua de tópico proeminente e pouco (ou nada) comuns em línguas de sujeito proeminente, conforme o Exemplo 5:

- (5) Pihengki - nun 747 - ka khu - ta (koreano)  
 avião - top. - subj. grande - stative  
 ‘Avião (tópico), o 747 é grande.’

(Exemplo adaptado de Li & Thompson 1976, p. 469)

Para Li e Thompson (1976), estruturas assim seriam os casos mais claros de estruturas de tópico-comentário, pois, nelas é possível fazer uma distinção entre o sujeito e o tópico facilmente; o tópico não apresenta uma relação seletional com o verbo; tais estruturas seriam geradas na base e não geradas de outras estruturas, através de uma regra de movimento e todas as línguas de tópico analisadas pelos autores tinham sentenças deste tipo, enquanto nenhuma língua de sujeito proeminente as tinha. Para os autores, a construção de sujeito duplo é um traço característico de uma língua de tópico proeminente.

No PB, tal tipo de construção pode ser encontrado, também sob o rótulo de Anacolutos. Pontes (1987) identificou dezenas de estruturas do tipo em sua análise de *corpora* de língua oral. Alguns deles podem ser observados nos exemplos abaixo:

- (6) Essa bolsa as coisas somem, aqui dentro.  
 (7) Essa bolsa aberta aí, eu podia te roubar a carteira.  
 (8) A última prisão dele, sabe o que ele fez?  
 (9) Eu agora, cabe desculpa de concurso né?

(Exemplos retirados de Pontes 1987, p. 15)

O trabalho de Orsini e Vasco (2007) faz uma análise das distribuições das construções de tópico nas falas culta e popular no PB. As construções de Anacoluto aparecem tanto na fala culta quanto na fala popular, apesar de não serem tão frequentes quanto o Deslocamento à esquerda e Topicalização. Abaixo podemos observar algumas das estruturas encontradas neste estudo:

- (10) Doce eu gosto de gelatina, gosto de pudim.
- (11) A seleção brasileira, quando começou a Copa do Mundo, um campeonato que é pra valer mesmo a coisa muda de figura.
- (Exemplos retirados de Orsini & Vasco 2007, p. 84)

As estruturas de anacoluto ou duplo sujeito, também conhecidas como construções de tópico estilo-chinês, vêm desafiando análises na Linguística Teórica e Descritiva desde que trabalhos clássicos as identificaram na década de 1970 (cf. Li & Thompson 1976). Tal construção parece indicar a possibilidade já denominada de “estilo cognitivo” na literatura relevante, em que ao invés de se estabelecer a relação gramatical entre um sujeito e um predicado, estabelece-se uma relação do tipo tópico-comentário, em que um tema é lançado, realizando-se sobre ele comentários, sem a necessidade de uma posição sintática de onde o tópico poderia haver se movido para a periferia esquerda da oração. Essa construção coloca potencialmente em cheque teorias importantes sobre a representação gramatical e também tem impacto para a caracterização tipológica do PB.

Nas construções de tópico estilo-chinês, o tópico não é argumento do predador, o que ocorre é apenas um elo semântico entre ele e um constituinte no comentário. O tópico, nessas construções, é desprovido de marcação de caso (cf. Cinque 1977). Em tais estruturas, o contexto permite fazer a devida interpretação do enunciado. Tais ocorrências seriam, de acordo com Li e Thompson (1976), estruturas prototípicas de línguas com orientação para o discurso.

Em inglês, que é uma língua de sujeito, orientada para a sentença (cf. Li & Thompson 1976), parece que tal tipo de construção de tópico não é produtiva. Yuan (1995) argumenta que o inglês não permite tópicos gerados na base, isto é, tópicos que não são gerados via movimentação sintática, como os tópicos estilo-chinês, sem a presença de elemento introdutor, conforme ocorre em línguas com proeminência de tópico. O inglês seria uma língua de sujeito proeminente e as topicalizações presentes na língua seriam um tipo de construção marcada. Sendo assim, construções de tópico no inglês são um fenômeno periférico e geralmente são geradas via movimentação sintática.

Para Xu (1999), ao estudar dados em chinês, a principal diferença entre o que ele chama de tópico estilo-inglês e o tópico estilo-chinês é a questão da relação que o elemento tópico tem com o seu comentário, em cada um dos casos: no primeiro, o tópico é correferencial (nulo ou pleno) com algum elemento no comentário, já no segundo caso, o tópico tem apenas uma relação de “*aboutness*” com o seu comentário, não sendo realizada através de um elemento correferente. O autor mostra essa diferença utilizando dois exemplos:

- (12) Zheben xiaoshwo Zhangsan Kanwan le Ø  
 este romance Zhangsan ler terminar perf. Ø  
 ‘Este romance, o Zhangsan terminou de ler.’
- (13) Neike shu, yezi da.  
 aquela árvore folhas grandes

‘Aquele árvore, as folhas são grandes.’

(Retirados e adaptados de Xu 1999)

O autor analisa as estruturas da seguinte maneira: em 12 há uma categoria vazia que é controlada pelo tópico. Em contraste, não há nenhuma posição sintática parecida em 13. Em vez disso, em 13 tópico e comentário se relacionam pragmaticamente. Neste caso, o tópico é considerado como gerado na base (*base generated*), na periferia esquerda da sentença, e tem uma relação de *aboutness* com o restante da sentença. Xu (1999) argumenta que, em construções de tópico-comentário mais comuns no chinês, há uma relação de *bridging* entre os elementos. O termo *bridging* seria algo como uma ponte de inferência e refere-se à capacidade de inferir que dois objetos ou eventos, que são introduzidos numa sentença, são relacionados de uma maneira que não foi explicitamente dita. Assim, o *bridging* ocorre se não há anáfora explícita presente dentro do comentário e exige uma capacidade de inferência. Em 13, por exemplo, para que ocorra a interpretação da sentença, é necessário que o leitor faça as seguintes inferências: “folhas” são partes de uma planta maior; as folhas que são grandes são parte de *x*, onde *x* é o tópico “Aquele árvore”. O conceito contido é inferido pelo conhecimento lexical sobre folhas, inferindo que folhas são partes de uma árvore.

Yuan (1995) argumenta que uma das principais diferenças entre o chinês e o inglês é que o chinês permite tópicos gerados na base sem qualquer tipo de introdução, que são tópicos que não são gerados via movimentação sintática, enquanto o inglês não permite tópicos gerados na base (a não ser que haja um elemento introdutor que facilite a interpretação do elemento topicalizado). Essa diferença pode ser observada nos Exemplos 14 e 15, com suas contrapartes agramaticais em inglês.

- (14) Zhongguo de da chengshi, wo zhi qu-guo Beijing. (chinês)  
 \*‘Big cities in China, I have only been to Peking.’ (inglês)  
 ‘Cidades grandes na China, eu só estive em Pequim’
- (15) ta jia li de ren, wo zhi jian-guo ta mama. (chinês)  
 \*‘People in her family, I have only met her mother.’ (inglês)  
 ‘Pessoas na família dela, eu só conheço a mãe dela.’

(Retirados e adaptados de Yuan 1995, p. 4)

Os tópicos “*Zhongguo de da chengshi*” e “*ta jia li de ren*” são tópicos sem lacuna, gerados na base. Para que as sentenças acima fossem aceitas em inglês, de acordo com Yuan (1995), elas deveriam ser introduzidas por expressões como “*As for*”, “*of*” ou “*Speaking of*”, que funcionariam como uma espécie de facilitador para interpretação da estrutura que, do contrário, seria agramatical.

Olhando para o PB, se traduzirmos as sentenças acima, temos 16 e 17. Tais estruturas parecem gramaticais no PB, assim como no chinês:

- (16) Cidades grandes na China, eu só estive em Pequim. (PB)

(17) Pessoas na família dela, eu só conheço a mãe dela. (PB)

O fato de o PB parecer permitir tópicos gerados na base (com ou sem elemento introdutor) e tais estruturas serem encontradas em diferentes *corpora* do PB, pode ser um indício de que o PB não seja uma língua de sujeito puro, orientada para a sentença, como o inglês. Além destas construções de tópico, no PB também há ocorrência de tópicos gerados via movimento sintático.

Yuan (1995) testou a aceitabilidade de tópicos estilo-chinês com falantes nativos de inglês, aprendizes de chinês como segunda língua, em diferentes estágios de proficiência, com o objetivo de verificar como as estruturas de tópico geradas na base são interpretadas por esses falantes. Ao aprender chinês, os falantes nativos de inglês têm *input* de tópicos gerados na base, o que não é aceito em sua língua materna. O estudo mostrou que falantes nativos de inglês, aprendizes de chinês, tiveram dificuldade em aceitar a estruturas de tópico chinês tanto no nível básico e intermediário, quanto no nível intermediário-avançado de proficiência. Os participantes só apresentaram aceitabilidade destas estruturas no nível avançado de proficiência em chinês.

Um outro estudo testando tópico no chinês e no inglês (*cf.* Chen 2011), identificou que falantes nativos de chinês, aprendizes de inglês, produzem estruturas de tópico e tendem a transferir a proeminência de tópico da sua língua materna para as construções da sua segunda língua. Porém, a medida em que o nível de proficiência vai aumentando, os aprendizes de inglês vão tentando produzir menos estruturas de tópico-comentário.

Além do fato de no inglês as estruturas de tópico chinês clássico (sem elemento introdutor) não serem aceitas, as topicalizações, estruturas de tópico geradas via movimentação sintática, só são permitidas em contextos contrastivos, onde o elemento topicalizado puxa uma situação contrastiva entre outros elementos do discurso prévio (*cf.* Weskott, Horning & Webelhuth 2019). Já no PB, a contrastividade é pouco requisitada (*cf.* Orsini 2003).

A ocorrência de tópicos gerados na base parece ser uma característica não presente em línguas de sujeito proeminente. Por outro lado, tais construções parecem fazer parte da gramática de uma língua onde a noção de tópico é importante. Ao considerarmos o PB, temos a interessante observação de que tais estruturas não somente podem ocorrer, mas também não precisam possuir um elemento introdutor, como é necessário em inglês. Os estudos das construções de tópico estilo-chinês na literatura em PB, entretanto, ainda são escassos. Principalmente estudos acerca do processamento dessas construções, pelos falantes da língua. Pretende-se, neste estudo, preencher esta lacuna com dados oriundos de experimentação psicolinguística, com testes verificando a aceitabilidade dessas estruturas pelos falantes.

### 3. Suporte metodológico

Neste experimento testamos estruturas de tópico estilo-chinês sem elemento introdutor (em 18), tópico estilo-chinês com elemento introdutor (em 19) e topicalização (em 20):

- (18) Comida japonesa, eu faço sushi pelo menos uma vez por mês.
- (19) Quanto a comida japonesa, eu faço sushi pelo menos uma vez por mês.
- (20) Comida japonesa, eu faço pelo menos uma vez por mês.

As questões propostas foram investigadas através de um teste de aceitabilidade *off-line*, utilizando uma escala *Likert* de 5 pontos, na plataforma *on-line* PCIBex<sup>1</sup>.

O teste foi aplicado em três grupos distintos: i) falantes nativos de PB, testados com sentenças em PB; ii) falantes nativos de inglês, testados com sentenças equivalentes em inglês e iii) falantes nativos de PB, com elevado conhecimento linguístico da língua inglesa (inglês L2), aferido através de questionário, testados com sentenças em inglês. Sendo assim, trata-se de três experimentos separados, testando três grupos distintos, porém, investigando as mesmas estruturas. É importante salientar que apesar de o grupo de falantes de inglês L2 ter sido testado, o objetivo dessa testagem foi de caráter exploratório, para observação de uma possível influência da sua língua materna (PB), durante o julgamento de sentenças em inglês, já que o foco desta pesquisa não é processamento de L2. O crucial aqui, no entanto, foi analisar os grupos de falantes nativos de PB e de inglês, julgando suas respectivas línguas maternas.

### 3.1. Hipóteses e previsões

A hipótese adotada é a de que PB e inglês não seriam línguas do mesmo tipo. Isto é, a hipótese é a de que o PB seria uma língua mista, orientada tanto para a sentença quanto para o discurso. Por outro lado, o inglês seria uma língua prototípica de sujeito, orientada para sentença, de acordo com a tipologia proposta por Li e Thompson (1976).

Se estivéssemos considerando que PB e inglês fossem línguas de sujeito proeminente, isto é, orientadas para a sentença, esperaríamos que as línguas exibissem o mesmo padrão de aceitabilidade por construções de tópico-comentário, pelo fato de as sentenças de sujeito-predicado serem a ordem padrão e preferida pelos falantes nas duas línguas e por línguas orientadas para sentença não aceitarem tópicos gerados na base. No entanto, estamos considerando aqui que PB e inglês apresentarão comportamento diferenciado em relação à aceitabilidade de construções de tópico, por não pertencerem ao mesmo agrupamento tipológico, em relação a sua orientação, para o discurso ou para a sentença. Sendo assim, o PB aceitaria tópicos gerados na base, o que não ocorreria com o inglês.

Nesse sentido, de acordo com a hipótese adotada, tem-se as seguintes previsões:

- a) Em relação ao grupo de falantes nativos de inglês: espera-se um alto índice de recusa para as construções de tópico chinês sem elemento introdutor, pois tópicos gerados na base não seriam aceitos no idioma, já que o inglês tem sido analisado como uma língua de sujeito proeminente. Já em relação a construções de topicalização, espera-

---

<sup>1</sup> *PennController for IBEX* é um programa gratuito, utilizado para montar e rodar experimentos *on-line*, desenvolvido por Florian Schwarz e Jeremy Zehr, da University of Pennsylvania/USA.

se um índice de aceitabilidade relativamente mais alto, pois a língua permite construções de tópico gerados via movimento sintático, que seriam geradas da estruturação mais básica da língua, SVO. Para as construções de tópico chinês com elemento introdutor, espera-se um índice de aceitabilidade melhor que a mesma estrutura sem elemento introdutor, pois tal elemento atuaria como um facilitador para a interpretação deste tipo de construção de tópico, que seria inaceitável sem esse facilitador. No entanto, comparando-se com o PB, espera-se que o grau de aceitabilidade das três estruturas testadas no Inglês seja menor, pelo fato de o PB ser uma língua onde tanto tópicos quanto sujeitos são proeminentes e o inglês ser uma língua orientada para a sentença;

- b) Em relação ao grupo de falantes nativos de PB: as construções de topicalização e de tópico chinês seriam aceitas, já que tanto tópicos gerados na base quanto tópicos gerados via movimento sintático são possíveis na língua. A hipótese é a de que, assim como construções SVO, as construções de tópico-comentário em PB são aceitas e muito produtivas. Comparando-se apenas as condições em que há um tópico chinês, com ou sem um elemento introdutor, espera-se um maior índice de aceitabilidade na condição onde há um elemento introdutor, já que não haveria um maior custo entre esse tipo de construção e uma topicalização, por exemplo, mas ainda assim, espera-se que o tópico chinês sem o elemento introdutor não seja recusado;
- c) Em relação ao grupo de falantes nativos de PB, fluentes em inglês: espera-se que a língua materna atue, de alguma maneira, sobre a L2, facilitando a aceitabilidade de construções de tópico chinês. Considerando que o inglês não permite tópicos gerados na base, mas o PB permite, espera-se que as construções de tópico chinês neste grupo sejam mais aceitas, em comparação com o grupo de falantes nativos de inglês, e menos aceitas, em comparação com o grupo de falantes nativos de PB, atestando assim, uma possível transferência linguística. Assim, teríamos evidências de que a realidade psicológica de construções de tópico geradas na base em PB seria forte, de modo que esta realidade atrapalhasse a interpretação das mesmas estruturas ao julgar o inglês, que não as permite, mesmo que o falante seja fluente e que domine a gramática da língua em um grau elevado.

### 3.2. Participantes

Dados de 81 participantes foram utilizados na análise, sendo 27 participantes em cada grupo. No experimento em inglês com falantes nativos, participaram como voluntários alunos da University of Pennsylvania/EUA, com idade entre 18 e 22 anos, do curso de Psicologia, que estavam cursando a disciplina *Language and Thought*, na ocasião. Os alunos receberam créditos de horas complementares pela sua participação e todos se auto declararam como falantes nativos de inglês, tendo o inglês como sua língua materna. No experimento de PB com falantes nativos, participaram como voluntários alunos do curso de Letras da UFRJ, com idade entre 19 e 36 anos. Os alunos receberam 2 horas complementares, pela sua participação na atividade e se auto declararam como falantes

nativos de PB, tendo o português do Brasil como sua língua materna. No experimento com falantes nativos de PB, fluentes em inglês (L2), os participantes tinham idades entre 21 e 69 anos. Todos os participantes declararam ser fluentes em língua inglesa, com nível de inglês avançado na ocasião do teste. A maioria deles realizou testes de proficiência em língua inglesa (fora do âmbito da pesquisa) e obtiveram alto grau de proficiência no idioma. Os participantes deste grupo eram graduandos, graduados e alguns deles, pós-graduados. Oito voluntários deste grupo já moraram nos EUA durante algum período de suas vidas. Todos os participantes deste grupo se auto declararam como falantes nativos de PB, tendo o português do Brasil como sua língua materna e sendo fluentes em inglês, tendo aprendido o inglês como segunda língua. Os participantes deste grupo que eram graduandos em Letras da UFRJ, receberam certificado de 2 horas AACC pela sua participação.

### 3.3. Materiais e design

Cada experimento apresenta um *design* 3, onde manipulou-se como **variável independente** o fator **Tipo de construção**, com três níveis, gerando as seguintes condições experimentais: Tópico chinês **com** elemento introdutor (**TopChinCom**), Tópico chinês **sem** elemento introdutor (**TopChinSem**) e Topicalização (**Top**). Trata-se de um teste auditivo, portanto, as frases foram gravadas e apresentadas aos participantes em forma de áudios. Todas as frases também contavam com um contexto prévio. A Tabela 1 ilustra um conjunto de frases experimentais utilizadas no experimento:

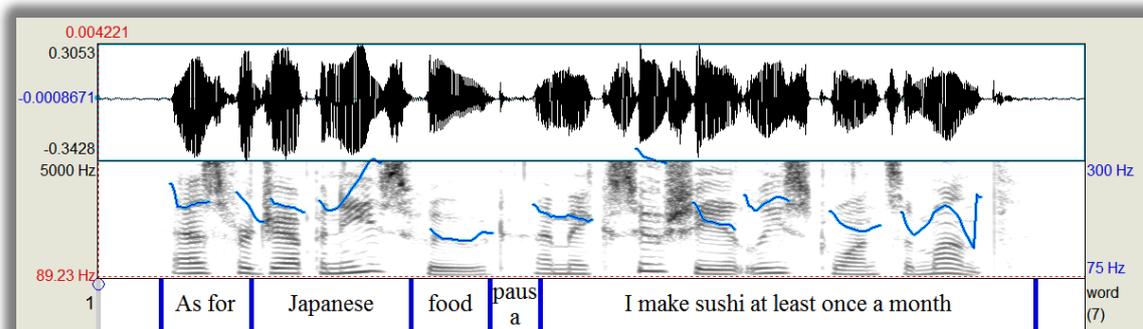
**Tabela 1. Exemplo dos materiais utilizados no experimento auditivo.**

<b>Frases experimentais – Experimento PB (grupo falantes nativos PB)</b>	
Contexto	Eu sempre faço comida italiana e comida japonesa, porque eu amo as duas culinárias.
	Quanto a comida japonesa, eu faço sushi pelo menos uma vez por mês. (TopChinCom)
	Comida japonesa, eu faço sushi pelo menos uma vez por mês. (TopChinSem)
	Comida japonesa, eu faço pelo menos uma vez por mês. (Topic)
<b>Frases experimentais – Experimento Inglês (grupo falantes nativos Inglês e Inglês L2)</b>	
Contexto	I always make Italian and Japanese food, because I love both.
	As for Japanese food, I make sushi at least once a month. (TopChinCom)
	Japanese food, I make sushi at least once a month. (TopChinSem)
	Japanese food, I make at least once a month. (Topic)

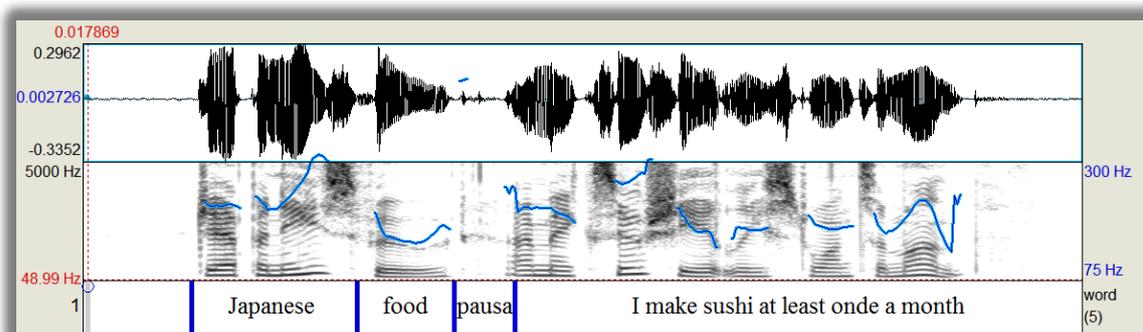
A variável dependente foi o índice de aceitabilidade das três condições, experimentais, pelos participantes. Optamos por apresentar estímulos auditivos, pois o tópico em PB, no geral, é considerado como um fenômeno da oralidade (*cf.* Pontes 1987, Kenedy 2014, entre outros), já que na língua escrita este fenômeno é deixado à margem, apesar de sua grande produtividade na língua. Também foram capturados os tempos dessa decisão (*RT*) em milésimos de segundo, como medida *off-line* cronométrica.

As frases experimentais e distratoras do experimento foram gravadas por falantes nativos de PB e de inglês. As frases continham curva entoacional característica de uma

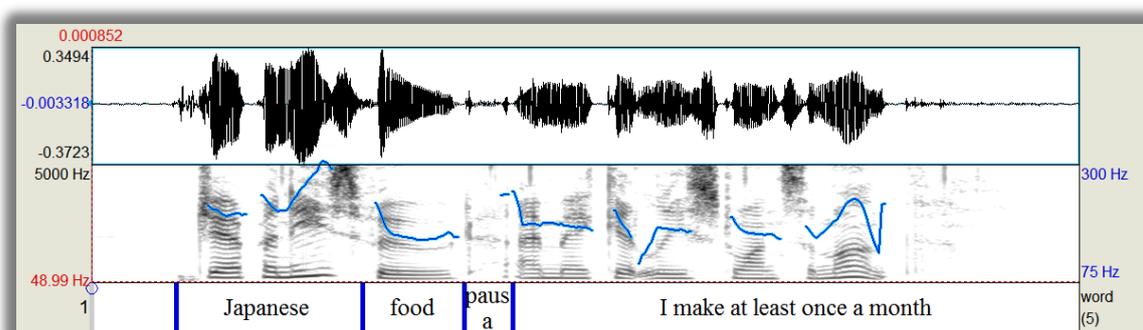
estrutura de tópico-comentário, onde foi dada uma espécie de saliência prosódica ao constituinte topicalizado, de forma que a estrutura fosse dividida em dois sintagmas entoacionais (I). Abaixo, podemos observar os espectrogramas de um dos itens experimentais, nas três condições. As Figuras 1, 2 e 3 referem-se às sentenças em inglês e as Figuras 4, 5 e 6 referem-se às sentenças em PB:



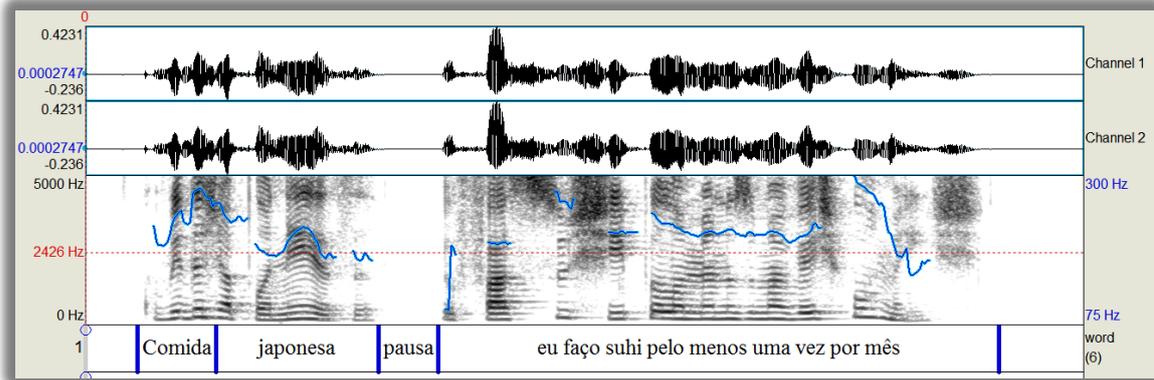
**Figura 1. Espectrograma de uma frase na condição “Tópico chinês com elemento introdutor” - Áudio em inglês**  
**“As for Japanese food, I make sushi at least once a month”**



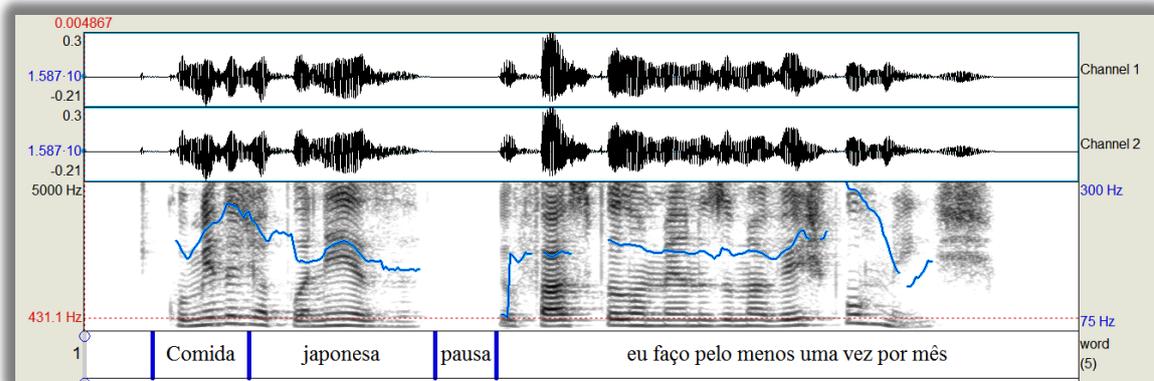
**Figura 2. Espectrograma de uma frase na condição “Tópico chinês sem elemento introdutor” – Áudio em inglês**  
**“Japanese food, I make sushi at least once a month”**



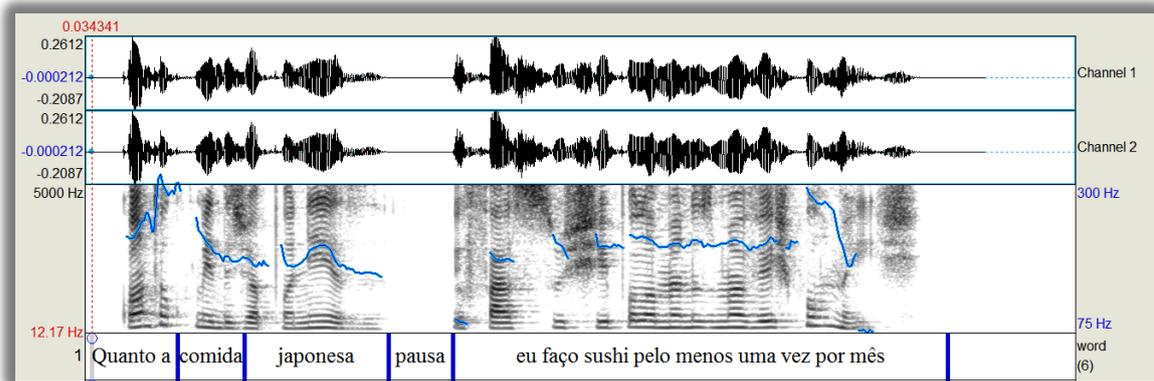
**Figura 3. Espectrograma de uma frase na condição “Topicalização” - Áudio em inglês**  
**“Japanese food, I make at least once a month”**



**Figura 4. Espectrograma de uma frase na condição “Tópico chinês com elemento introdutor” - Áudio em PB**  
 “Quanto a comida japonesa, eu faço sushi pelo menos uma vez por mês”



**Figura 5. Espectrograma de uma frase na condição “Tópico chinês sem elemento introdutor” – Áudio em PB**  
 “Comida japonesa, eu faço sushi pelo menos uma vez por mês”



**Figura 6. Espectrograma de uma frase na condição “Topicalização” - Áudio em PB**  
 “Comida japonesa, eu faço pelo menos uma vez por mês”

Durante a confecção do material experimental, tomamos o cuidado de criar os contextos experimentais com ideia contrastiva, pois no inglês o tópico só é licenciado mediante tal contexto (*cf.* Yuan 1995). Em PB, por outro lado, o contexto contrastivo não é necessariamente solicitado (*cf.* Orsini 2003), mas também ocorre.

Foram elaborados 12 conjuntos de frases experimentais para cada experimento, cada conjunto contendo 1 frase contexto + 3 frases experimentais, uma em cada condição experimental e 24 frases distratoras. As frases foram randomizadas automaticamente pelo programa PCIBex seguindo a configuração em quadrado latino.

### 3.4. Procedimentos

Cada *trial* tinha um total de 3 telas. A tela 1 (Figura 7) continha a frase contexto. Após 1 segundo da entrada da tela 1, o áudio tocava automaticamente. Na tela 2, o participante ouvia a frase experimental, que também aparecia 1 segundo após a tela 2 ser carregada. Após ouvirem a frase, os participantes clicavam no botão “Next/Próximo” para responderem a escala *Likert* de 5 pontos, de acordo com a interpretação dada. Tal escala era apresentada na tela 3 (Figuras 8), onde havia uma pergunta sobre a interpretação da frase ouvida (Quão boa/aceitável esta frase soou para você? / *How good/acceptable does this sentence sound?*). Abaixo da pergunta havia uma escala que ia do número 1 ao 5, seguindo os respectivos valores: 1= Altamente inaceitável/ *Strongly unacceptable*; 2= Inaceitável/*Unacceptable*; 3= Neutro/*Neutral*; 4= Aceitável/*Acceptable*; 5= Altamente aceitável/*Strongly Acceptable* (Figura 9). O participante deveria então clicar sobre a opção (frase) que melhor respondia a pergunta, de acordo com a sua opinião, baseado no áudio ouvido. Na tela de apresentação foi deixado claro que o participante deveria julgar o áudio ouvido na tela 2 e que ele deveria fazer este julgamento conforme a frase soasse aos seus ouvidos, numa escala de muito ruim até muito boa, na sua língua.



Figura 7. Tela 1 (contexto) do experimento



Figura 8. Tela 2 (frase experimental) do experimento

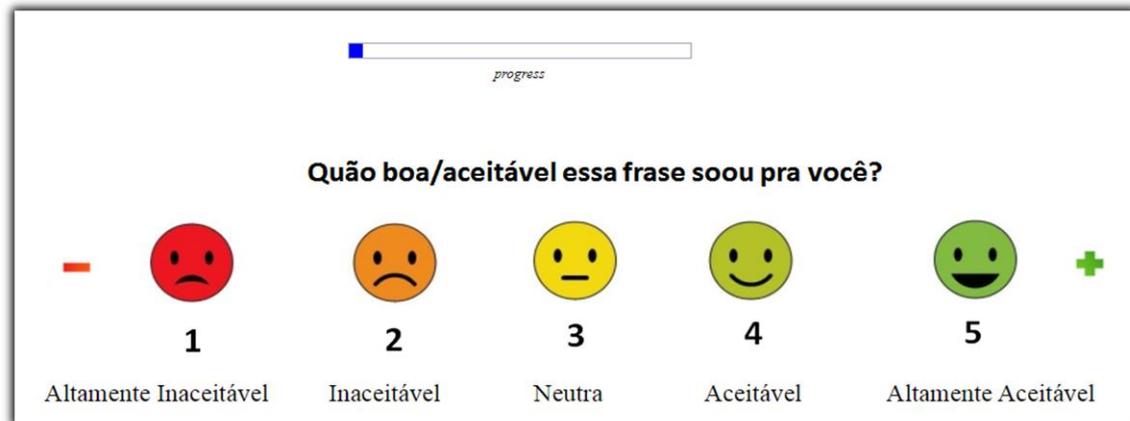


Figura 9. Tela 3 do experimento – escala *Likert*

O experimento foi montado e aplicado de maneira totalmente *on-line*. Após as instruções, o participante fazia uma pequena prática, para se familiarizar com a tarefa. A prática tinha 3 *trials*. O teste tinha duração de 15 minutos.

### 3.5. Resultados

Após confecção de tabelas e planilhas com os dados obtidos para verificação da análise descritiva, realizamos um teste de regressão logística ordinal (do tipo *cumulative link mixed model*, utilizando o elemento “Participante” como fator de randomização), já que a variável dependente testada nos experimentos é categórica ordenada. Tomou-se a condição ‘Topicalização’<sup>2</sup> como referência, para verificarmos a ocorrência de diferenças estatísticas significativas. Após o teste de regressão ordinal comparando os dados com a condição de referência, realizamos um teste *post hoc Tukey*, para verificarmos diferença estatística entre os demais pares. Inicialmente foi feita uma análise estatística isolada por grupo, para verificarmos o comportamento do fenômeno linguístico nos falantes de cada língua separadamente. Posteriormente, conduzimos uma análise entre grupos, para assim podermos verificar o comportamento do fenômeno entre as línguas analisadas e para verificarmos se havia diferença entre a preferência das estruturas por um grupo específico e se essa preferência era, de alguma forma, diferenciada no grupo bilíngue. Na análise comparativa entre grupos, a condição de referência era *Topicalização* e o grupo de referência era o Grupo 1 – falantes nativos de PB, sendo assim, as demais condições e grupos foram comparados diretamente com estes referenciais. Para análise dos tempos de reação durante o julgamento das sentenças por cada grupo foi conduzido um teste ANOVA.

<sup>2</sup> Selecionamos a condição *Topicalização* como a condição de referência, pois tal estrutura é produtiva em ambas as línguas, sendo o tópico clássico. O grupo 1 – falantes nativos de PB foi tomado como referência em testes posteriores, pois trata-se da língua-alvo desta pesquisa.

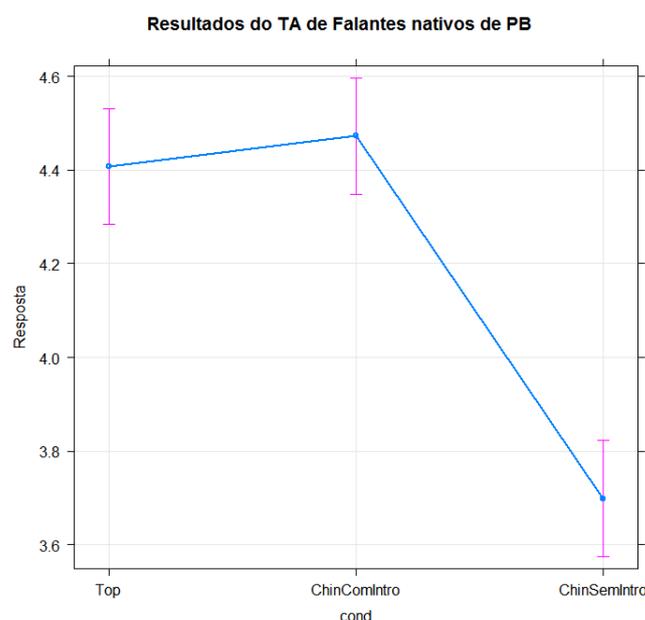
### 3.5.1. Grupo 1 – Falantes nativos de PB

A Tabela 2 mostra o número de respostas em cada um dos níveis, de 1 a 5 para cada uma das condições, dentro das 648 ocorrências no total.

**Tabela 2. N.º de ocorrências de acordo com o índice de aceitabilidade por condição Grupo 1 – falantes nativos de PB**

Condição	Índice de aceitabilidade					Total
	1	2	3	4	5	
<b>Topicalização</b>	2	8	20	56	130	216
<b>TopChinCOM</b>	0	4	22	58	132	216
<b>TopChinSEM</b>	4	35	48	64	65	216

Podemos verificar que no grupo de falantes nativos brasileiros, parece haver uma prevalência no nível mais alto de aceitabilidade, do que para a não aceitabilidade das estruturas em questão. O Gráfico 1 mostra as médias do teste de aceitabilidade (TA) em cada uma das condições testadas:



**Gráfico 1. Médias de aceitabilidade por condição**

A condição *Topicalização* teve uma média de aceitabilidade de **4,40**. A condição *Tópico chinês com elemento introdutor* teve uma média de aceitabilidade de **4,47** e a condição *Tópico chinês sem elemento introdutor* obteve uma média de aceitabilidade de **3,69**, conforme esperado. Em outras palavras, pode-se dizer que as construções de tópico-comentário em PB obtiveram uma média de aceitabilidade alta, no geral, sendo *Topicalização* e *Tópico chinês com introdução* com médias acima de 4 (numa escala até 5). A condição *Tópico chinês sem introdução* obteve aceitabilidade acima da média, já que passou da nota de corte, que é 3.

A Tabela 3 mostra os valores da média de aceitabilidade das estruturas testadas, obtidos a partir da somatória do número de ocorrências:

**Tabela 3. Média de aceitabilidade pelo n° de ocorrências  
Grupo 1 – falantes nativos de PB**

Índice de aceitabilidade	N° de ocorrências	Percentual de ocorrências	Média das respostas
1- Altamente Inaceitável	6	0,92%	Min.: 1.000
2- Inaceitável	47	7,25%	1st Qu.: 4.000
3- Neutra	90	13,88%	Median.: 5.000
4- Aceitável	178	27,46%	Mean.: 4.193
5- Altamente Aceitável	327	50,46%	3rd Qu.: 5.000
<b>Total: 648</b>			Max.: 5.000

Conforme podemos verificar na Tabela 3, a média geral de aceitabilidade obtida para o Grupo 1 – falantes nativos de PB foi de **4,19**, o que indica que as construções de tópico-comentário, neste grupo, tiveram uma alta taxa de aceitabilidade, considerando uma escala de 5 pontos.

Conduzimos um teste de regressão ordinal, para verificarmos as diferenças estatísticas entre as condições.

**Tabela 4. Resultado estatístico - Teste de regressão ordinal  
Grupo 1 – falantes nativos de PB**

Condições	Estimate	Std. Error	z value	Pr(> z )
Topicalização X TopChinComIntro	0.1050	0.2058	0.510	0.61
Topicalização X TopChinSemIntro	-1.6924	0.2041	-8.291	< 2e-16 ***

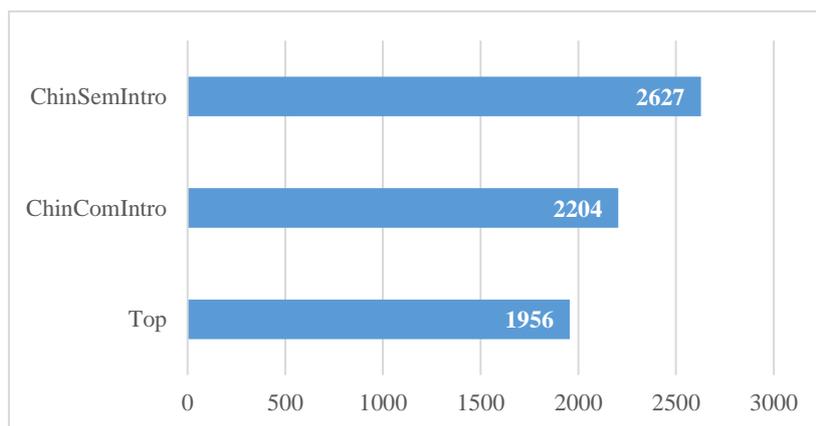
Podemos verificar na Tabela 4 que entre a condição [Topicalização] vs [Tópico chinês com elemento introdutor], não houve diferença estatística significativa, ou seja, para falantes nativos de PB, as duas estruturas têm o mesmo grau de aceitabilidade. Já na comparação [Topicalização] vs [Tópico chinês sem elemento introdutor], foi observado diferença altamente significativa, sendo a condição [Topicalização] mais aceita que a condição [Tópico chinês sem elemento introdutor], o que já era esperado.

Após comparação da condição *Intercept* com as demais condições, realizamos um teste *post hoc Tukey*, a fim de compararmos também o par [Tópico chinês sem elemento Introdutor] vs [Tópico chinês com elemento introdutor]. O resultado foi uma diferença altamente significativa, indicando que a condição *Tópico chinês sem elemento introdutor* foi menos aceitável que a condição *Tópico chinês com elemento introdutor*. Tal comparação pode ser verificada na Tabela 5:

**Tabela 5. Resultado estatístico - Teste *post hoc* Tukey**  
**Grupo 1 – falantes nativos de PB**

Condições	Estimate	Std. Error	z.ratio	p value
TopChinCom X TopChinSem	1.797	0.205	8.782	<.0001

Também foram aferidos os tempos de reação das decisões dos participantes, ao julgarem as sentenças de tópico-comentário. O resultado pode ser verificado no Gráfico 2:



**Gráfico 2. RT's das decisões por condição**

O teste estatístico ANOVA indicou efeito principal de tipo de estrutura ( $F(2,430) = 9,77$   $p < 0,000071$ ). Os testes-t indicaram diferença estatística significativa entre as comparações [Top]vs[Chin\_Sem\_Intro]  $t(215)=4,38$   $p < 0,0001$  e [Chin\_Sem\_Intro]vs[Chin\_Com\_Intro]  $t(215)=2,55$   $p < 0,0113$ . Já entre a comparação [Top]vs[Chin\_Com\_Intro] não ocorreu diferença estatística significativa ( $t(215)=1,76$   $p < 0,0790$ ). Tal resultado indica que a condição *Tópico chinês sem elemento introdutor* obteve os tempos de resposta mais elevados, isto é, os participantes demoraram mais para decidirem se a sentença era aceitável ou não, em comparação com as demais.

Em suma, os dados do teste de aceitabilidade com falantes brasileiros, analisando as construções de tópico-comentário em PB, mostram que os participantes demonstraram uma alta taxa de aceitabilidade pelas construções de tópico-comentário, no geral. Os valores das médias de aceitabilidade passaram do grau neutro nas três condições testadas, sendo acima de **4,40** para *Topicalização* e *Tópico chinês com elemento introdutor*. A taxa de aceitabilidade da condição *Tópico chinês sem elemento introdutor* foi menor em relação as demais, no entanto, também fica acima da média neutra, em **3,69**, indicando aceitabilidade. Tais resultados vão em direção às nossas previsões. Importante salientar que a ocorrência do julgamento das não-aceitáveis nas três condições testadas (Inaceitável e Altamente inaceitável) foi baixíssimo.

Tais resultados sugerem que os participantes deste grupo demonstraram boa aceitabilidade para as construções de tópico via movimento sintático e também para as geradas na base.

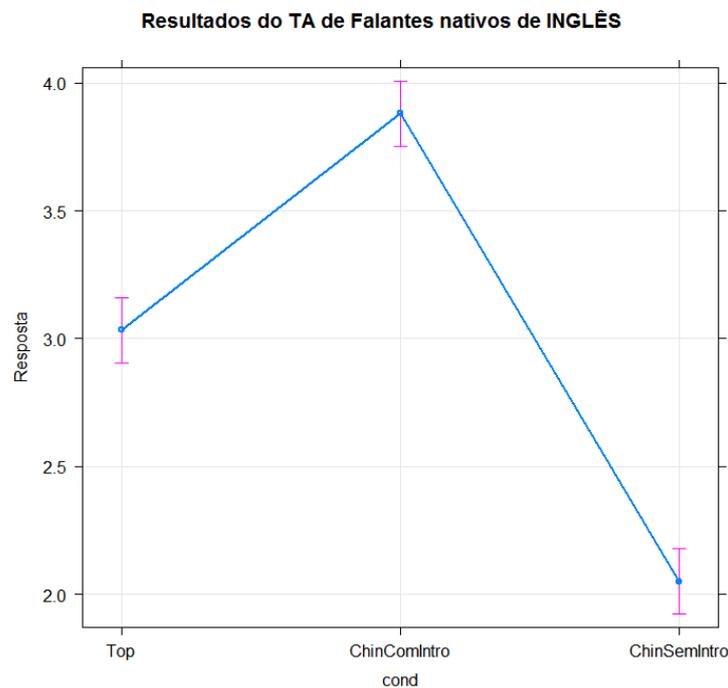
### 3.5.2. Grupo 2 – Falantes nativos de inglês

A Tabela 6 mostra o número respostas em cada um dos níveis, de 1 a 5 para cada uma das condições, dentro das 648 ocorrências no total.

**Tabela 6. N.º de ocorrências de acordo com o índice de aceitabilidade por condição  
Grupo 2 – falantes nativos de Inglês**

Condição	Índice de aceitabilidade					Total
	1	2	3	4	5	
<b>Topicalização</b>	18	49	71	64	14	216
<b>TopChinCOM</b>	1	12	52	98	53	216
<b>TopChinSEM</b>	67	92	37	19	1	216

Podemos verificar que no grupo de falantes nativos de inglês, há uma prevalência no nível mais alto de aceitabilidade apenas na condição *Tópico chinês com elemento introdutor*. Na condição *Topicalização*, o índice de aceitabilidade ficou na zona neutra. Já na condição *Tópico chinês sem elemento introdutor*, ocorre um baixo índice de aceitabilidade. O Gráfico 3 mostra as médias do teste de aceitabilidade (TA) em cada uma das condições testadas:



**Gráfico 3. Médias de aceitabilidade por condição**

Conforme observado, a condição *Topicalização* teve uma média de aceitabilidade de **3,03**, o que não era esperado, já que, aparentemente, em inglês estruturas geradas via movimentação sintática em contextos contrastivos seriam aceitas. No entanto, uma média de 3,03 não indica aceitabilidade nem rejeição, mas indica neutralidade/dúvida no

juízo. Sendo assim, não podemos afirmar que os falantes deste grupo rejeitaram as topicalizações, nem que as aceitaram. A condição *Tópico chinês com elemento introdutor* teve uma média de aceitabilidade de **3,87** e a condição *Tópico chinês sem elemento introdutor* obteve uma média de aceitabilidade de **2,05**. Tais médias demonstram que quando o tópico gerado na base não possui elemento introdutor, como ocorre em línguas de tópico prominentes, os falantes nativos de inglês rejeitam a estrutura (média: 2,05). Porém, quando o tópico gerado na base possui uma introdução que facilita a interpretação do SN topicalizado, há aceitabilidade da estrutura em inglês (média: 3,87). Tais dados confirmam achados anteriores da literatura (Yuan 1995).

A Tabela 7 mostra os valores da média de aceitabilidade das estruturas testadas, obtidos a partir da somatória do número de ocorrências:

**Tabela 7. Média de aceitabilidade pelo n° de ocorrências**  
Grupo 2 – falantes nativos de Inglês

Índice de aceitabilidade	N° de ocorrências	Percentual de ocorrências	Média das respostas
1- Altamente Inaceitável	86	13,27%	Min.: 1.000
2- Inaceitável	153	23,61%	1st Qu.: 2.000
3- Neutra	160	24,69%	Median.: 3.000
4- Aceitável	181	27,93%	Mean.: 2.988
5- Altamente Aceitável	68	10,50%	3rd Qu.: 4.000
<b>Total: 648</b>			Max.: 5.000

A média geral de aceitabilidade obtida para o Grupo 2 – falantes nativos de inglês foi de **2,98**, o que indica que as construções de tópico-comentário, neste grupo, tiveram uma taxa de aceitabilidade média-baixa, considerando uma escala de 5 pontos, tendo o *Tópico chinês sem introdução* uma média de aceitabilidade bem baixa, ficando abaixo da linha de corte.

Realizamos um teste de regressão logística ordinal, para verificarmos as diferenças estatísticas entre as condições.

**Tabela 8. Resultado estatístico - Teste de regressão ordinal**  
Grupo 2 – falantes nativos de Inglês

Condições	Estimate	Std. Error	z value	Pr(> z )
Topicalização X TopChinComIntro	1.7827	0.1950	9.14	< 2e-16 ***
Topicalização X TopChinSemIntro	-2.0181	0.1965	-10.27	< 2e-16 ***

Podemos verificar na Tabela 8 que entre a condição [Topicalização] vs [Tópico chinês com elemento introdutor], houve diferença estatística altamente significativa, ou seja, para falantes nativos de inglês, a estrutura *Tópico chinês com elemento introdutor* é mais aceitável que a estrutura de *Topicalização*. Na comparação [Topicalização] vs [Tópico chinês sem elemento introdutor], também foi observada uma diferença estatística

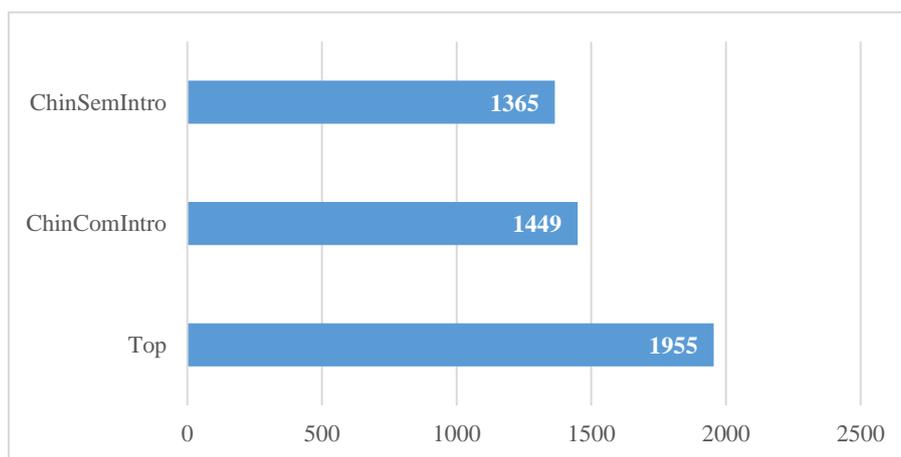
altamente significativa, sendo a condição *Tópico chinês sem elemento introdutor* menos aceita que a *Topicalização*.

Após comparação da condição *Intercept* com as demais condições, realizamos um teste *post hoc Tukey*, a fim de compararmos também o par [Tópico chinês sem elemento Introdutor] vs [Tópico chinês com elemento introdutor]. O resultado foi uma diferença altamente significativa, indicando que a condição *Tópico chinês sem elemento introdutor* foi menos aceita que a condição *Tópico chinês com elemento introdutor*, conforme esperado. Tal comparação pode ser verificada na Tabela 9:

**Tabela 9. Resultado estatístico - Teste post hoc Tukey**  
**Grupo 2 – falantes nativos de Inglês**

Condições	Estimate	Std. Error	z.ratio	p value
TopChinCom X TopChinSem	3.80	0.230	16.534	<.0001

Os tempos de reação das decisões dos participantes ao julgarem as sentenças de tópico-comentário, podem ser verificados no Gráfico 4:



**Gráfico 4. RT's das decisões por condição**

O teste estatístico ANOVA não indicou efeito principal de tipo de estrutura ( $F(2,430) = 2,54$   $p < 0,079998$ ), apesar de o índice do tempo de reação ter sido mais elevado na condição topicalização, visualmente. Os testes-t não indicaram diferença estatística significativa em nenhuma das comparações: [Top]vs[Chin\_Sem\_Intro]  $t(215)=1,75$   $p < 0,0814$ ; [Top]vs[Chin\_Com\_Intro]  $t(215)=1,49$   $p < 0,1386$  e [Chin\_Sem\_Intro]vs[Chin\_Com\_Intro]  $t(215)=0,80$   $p < 0,4269$ . Tal resultado demonstra que não há diferença entre os tempos de decisão sobre a aceitabilidade das estruturas de tópico-comentário testadas, pelo grupo de falantes nativos de inglês.

Em suma, os dados do teste de aceitabilidade com falantes nativos de inglês, analisando as construções de tópico-comentário em inglês, mostram que os participantes demonstraram uma baixa taxa de aceitabilidade pelas construções de tópico-comentário, principalmente nos casos de *Tópico chinês sem elemento introdutor* e de *Topicalização*. A condição *Tópico chinês sem elemento introdutor* teve suas médias de aceitabilidade

abaixo da linha neutra, ficando mais próxima da não-aceitabilidade (Inaceitáveis e Altamente inaceitáveis).

Os valores das médias de aceitabilidade passaram do grau neutro apenas quando o tópico chinês possuía um elemento introdutor, ficando em **3,87**. A taxa de aceitabilidade das condições *Topicalização* e *Tópico chinês sem elemento introdutor* foram bem menores, sendo esta última a mais baixa, ficando na linha das inaceitáveis, com média de **2,05**. A aceitabilidade das construções de *Topicalização* foi melhor que a *Tópico chinês sem elemento introdutor*, porém, não chegou a atingir uma boa aceitabilidade, ficando com média **3,03**, o que não era esperado. Tal resultado segue na direção esperada, exceto pelo fato de as nossas previsões esperarem que *Topicalizações* seriam tão aceitáveis quanto o *Tópico chinês com elemento introdutor* em inglês. Esses resultados sugerem que os participantes deste grupo demonstraram baixa aceitabilidade para as construções de tópico via movimento sintático e também para as geradas na base, quando esta não possui um elemento introdutor que funcione como um facilitador para a sua interpretação.

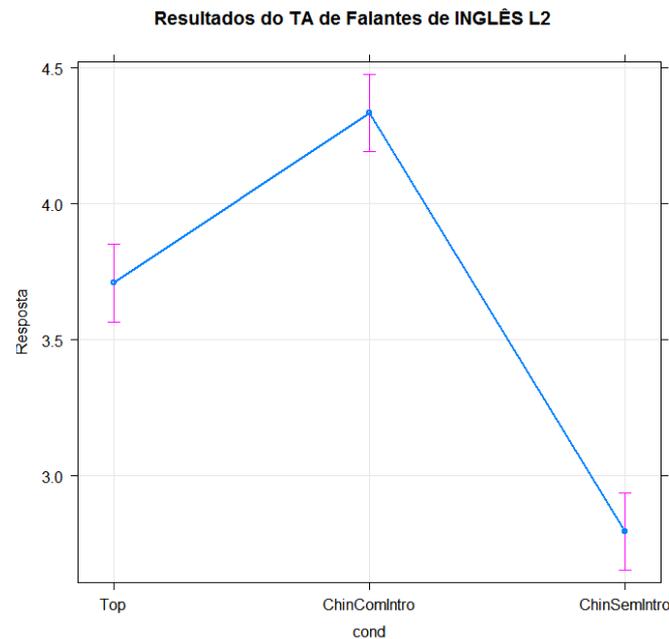
### 3.5.3. Grupo 3 – Falantes nativos de PB, fluentes em inglês (L2)

A Tabela 10 mostra o número de respostas em cada um dos níveis, de 1 a 5 para cada uma das condições, dentro das 648 ocorrências no total.

**Tabela 10. N.º de ocorrências de acordo com o índice de aceitabilidade Grupo 3 – falantes nativos de PB, fluentes em inglês (L2)**

Condição	Índice de aceitabilidade					Total
	1	2	3	4	5	
<b>Topicalização</b>	4	31	51	68	62	216
<b>TopChinCOM</b>	0	6	30	66	114	216
<b>TopChinSEM</b>	32	72	44	44	24	216

Podemos verificar que no grupo de falantes nativos brasileiros, com fluência em inglês, há uma prevalência no nível mais alto de aceitabilidade das condições *Topicalização* e *Tópico chinês com elemento introdutor*. Já na condição *Tópico chinês sem elemento introdutor*, ocorre um nível mais baixo de aceitabilidade. O Gráfico 5 mostra as médias do teste de aceitabilidade (TA) em cada uma das condições testadas:



**Gráfico 5. Médias de aceitabilidade por condição**

Conforme observado acima, a condição *Topicalização* teve uma média de aceitabilidade de **3,70**. A condição *Tópico chinês com elemento introdutor* teve uma média de aceitabilidade de **4,33** e a condição *Tópico chinês sem elemento introdutor* obteve uma média de aceitabilidade de **2,79**. A Tabela 11 mostra os valores da média de aceitabilidade das estruturas testadas, obtidos a partir da somatória do número de ocorrências:

**Tabela 11. Média de aceitabilidade pelo n° de ocorrências  
Grupo 3 – falantes nativos de PB, fluentes em Inglês (L2)**

Índice de aceitabilidade	N° de ocorrências	Percentual de ocorrências	Média das respostas
1- Altamente Inaceitável	36	5,55%	Min.: 1.000
2- Inaceitável	109	16,82%	1st Qu.: 3.000
3- Neutra	125	19,29%	Median.: 4.000
4- Aceitável	178	27,46%	Mean.: 3.613
5- Altamente Aceitável	200	30,86%	3rd Qu.: 5.000
<b>Total: 648</b>			Max.: 5.000

Conforme a Tabela 11 mostra, a média geral de aceitabilidade obtida para o Grupo 3 – falantes nativos de PB, fluentes em Inglês (L2) foi de **3,61**, o que indica que as construções de tópico-comentário, neste grupo, tiveram uma taxa de aceitabilidade média-alta, considerando uma escala de 5 pontos. Em linhas gerais, pode-se dizer que as construções de tópico-comentário em inglês, para falantes nativos brasileiros, obtiveram uma aceitabilidade média-alta, nos casos *Tópico chinês com introdução*, sendo acima de **3,0**, que é a média-corte e uma taxa de aceitabilidade alta, nos casos de *Topicalização*, que foi acima de **4,0**, passando do corte. Em relação ao *Tópico chinês sem introdução*, observamos uma média de aceitabilidade baixa, ficando abaixo da linha de corte.

Conduzimos um teste de regressão logística ordinal, para verificarmos as diferenças estatísticas entre as condições. Selecionamos a condição *Topicalização* como o *Intercept*, sendo assim, as demais condições foram comparadas a esta.

**Tabela 12. Resultado estatístico - Teste de regressão ordinal  
Grupo 3 – falantes nativos de PB, fluentes em Inglês (L2)**

Condições	Estimate	Std. Error	z value	Pr(> z )
Topicalização X TopChinComIntro	1.6063	0.2036	7.891	3e-15 ***
Topicalização X TopChinSemIntro	-1.9308	0.1990	-9.703	< 2e-16 ***

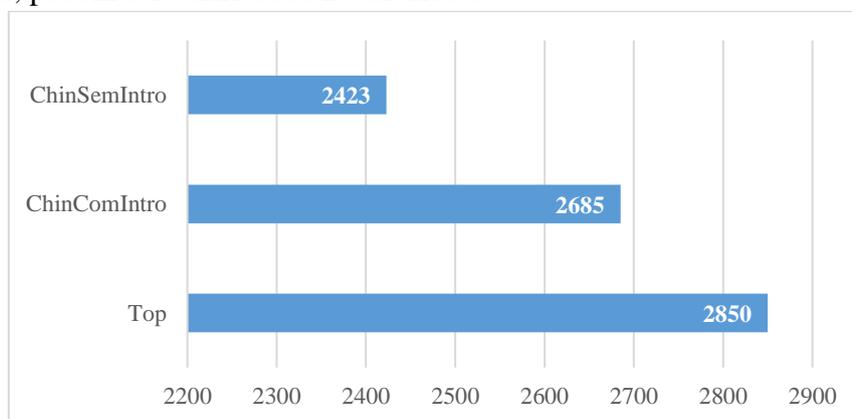
Podemos verificar na Tabela 12 que entre a condição [Topicalização] vs [Tópico chinês com elemento introdutor], houve diferença estatística altamente significativa, ou seja, para falantes nativos de PB, fluentes em inglês (L2), a estrutura de *Tópico chinês com elemento introdutor* foi mais aceita que a *Topicalização*. Na comparação [Topicalização] vs [Tópico chinês sem elemento introdutor], também foi observado uma diferença altamente significativa, sendo a condição [Topicalização] mais aceita que a condição [Tópico chinês sem elemento introdutor].

Após comparação da condição *Intercept* com as demais condições, realizamos um teste *post hoc Tukey*, a fim de compararmos também o par [Tópico chinês sem elemento Introdutor] vs [Tópico chinês com elemento introdutor]. O resultado foi uma diferença altamente significativa, indicando que a condição *Tópico chinês sem elemento introdutor* foi menos aceitável que a condição *Tópico chinês com elemento introdutor*, conforme Tabela 13.

**Tabela 13. Resultado estatístico - Teste *post hoc Tukey*  
Grupo 3 – falantes nativos de PB, fluentes em Inglês (L2)**

Condições	Estimate	Std. Error	z.ratio	p value
TopChinCom X TopChinSem	3.54	0.235	15.027	<.0001

Os tempos de reação das decisões dos participantes ao julgarem as sentenças de tópico-comentário, podem ser verificados no Gráfico 6:



**Gráfico 6. RT's das decisões por condição**

O teste estatístico ANOVA não indicou efeito principal de tipo de estrutura ( $F(2,430) = 0,287$   $p < 0,750538$ ). Os teste-t não indicaram diferença estatística significativa em nenhuma das comparações: [Top]vs[Chin\_Sem\_Intro]  $t(215)=1,49$   $p < 0,1370$ ; [Top]vs[Chin\_Com\_Intro]  $t(215)=0,24$   $p < 0,8102$  e [Chin\_Sem\_Intro]vs[Chin\_Com\_Intro]  $t(215)=0,41$   $p < 0,6856$ . Tal resultado demonstra que não há diferença entre os tempos de decisão sobre a aceitabilidade das estruturas de tópico-comentário testadas, pelo grupo de falantes nativos de PB, fluentes em inglês.

Em suma, os dados do teste de aceitabilidade com falantes brasileiros, fluentes em inglês (L2), analisando as construções de tópico-comentário em inglês, mostram que os participantes demonstraram uma taxa de aceitabilidade média-alta pelas construções de tópico-comentário, no geral. Os valores das médias de aceitabilidade passaram do grau Neutro nas condições de *Topicalização*, sendo **3,70** e acima do grau Aceitável *Tópico chinês com elemento introdutor*, sendo **4,33**. A taxa de aceitabilidade da condição *Tópico chinês sem elemento introdutor* foi menor em relação as demais, ficando próxima da média neutra, em **2,79**. Tal resultado segue na direção esperada e sugere que os participantes deste grupo demonstraram boa aceitabilidade para as construções de tópico via movimento sintático e para as geradas na base, quando há elemento introdutor. Nos casos de tópico chinês sem elemento introdutor, essa taxa de aceitabilidade tem uma queda, porém se aproxima mais da zona neutra, do que das não aceitáveis.

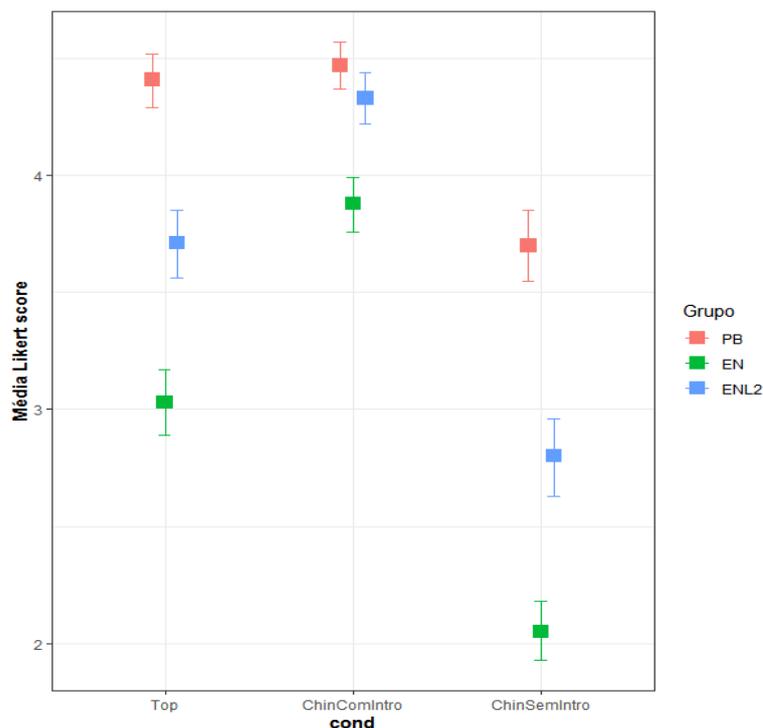
### 3.5.4. Resultados da comparação entre os grupos

Os resultados abaixo mostram a comparação entre os três grupos testados. Na Tabela 14 podemos observar a média de aceitabilidade de cada um dos grupos, para as condições testadas:

**Tabela 14. Médias de aceitabilidade por grupo e por condição**  
Comparação dos três grupos

Grupo	Condição	Média de aceitabilidade	Taxa de aceitabilidade
PB	Top	4,40	4,19
	TopChinCOM	4,47	
	TopChinSEM	3,69	
Inglês	Top	3,03	2,98
	TopChinCOM	3,87	
	TopChinSEM	2,05	
Inglês L2	Top	3,70	3,61
	TopChinCOM	4,33	
	TopChinSEM	2,79	

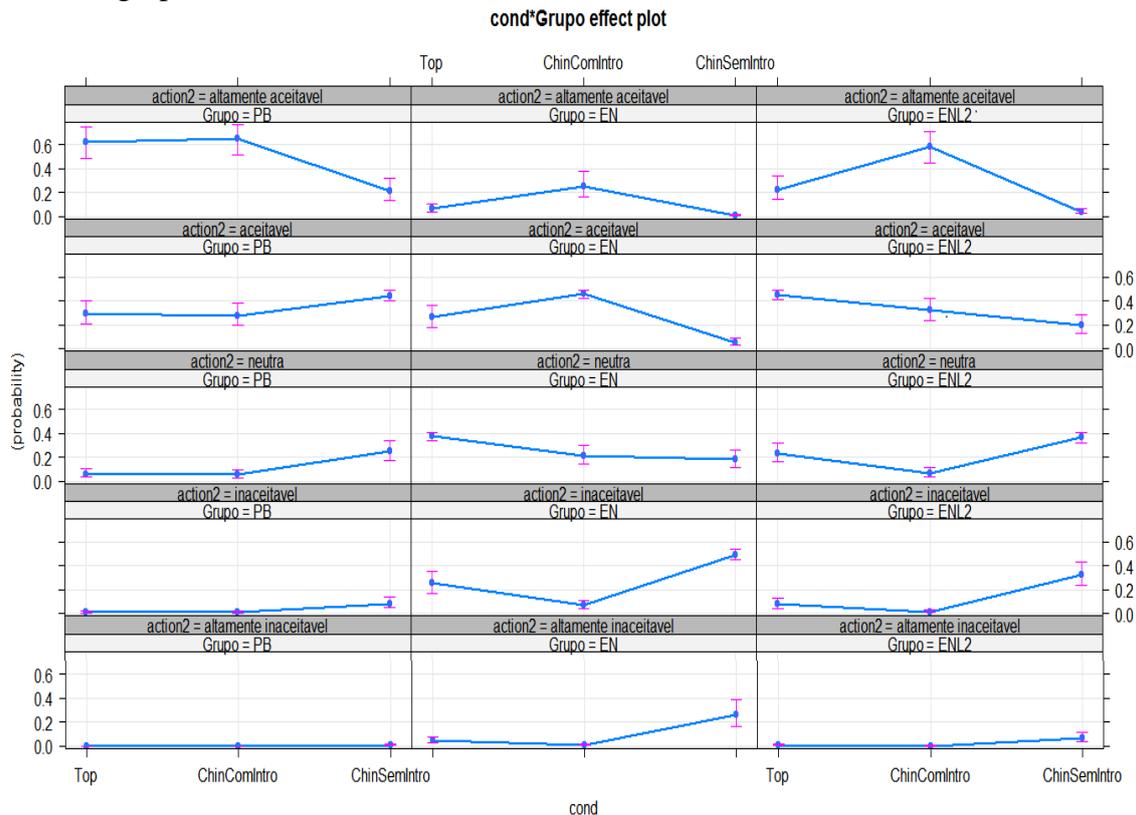
De acordo com as médias acima, podemos observar que o Grupo 1 - Falantes nativos de PB - obteve altas médias de aceitabilidade nas três condições testadas, em comparação com os outros grupos, ficando com uma taxa de aceitabilidade mais próxima das Aceitáveis e Altamente aceitáveis, inclusive para a condição de *Tópico chinês sem elemento introdutor*. O Grupo 2 - falantes nativos de inglês, obteve uma taxa alta de aceitabilidade apenas para a condição *Tópico chinês com elemento introdutor*. Já na condição *Topicalização*, a aceitabilidade foi média, dentro da linha neutra. A condição *Tópico chinês sem elemento introdutor*, por sua vez, obteve a média mais baixa de aceitabilidade, inclusive entre os três grupos, ficando bem abaixo da linha neutra, na média das Inaceitáveis. O Grupo 3 – falantes nativos de PB, fluentes em inglês (L2), obteve uma média alta para a condição *Tópico chinês com elemento introdutor*, ficando próxima da Altamente aceitável. Na condição *Topicalização*, a média ficou acima da zona neutra, mais próxima das Aceitáveis. Já na condição *Tópico chinês sem elemento introdutor*, verificamos uma queda na aceitabilidade, porém, se aproximando da zona neutra. No Gráfico 7 é possível observar os dados acima em mais detalhes:



**Gráfico 7. Média dos índices de aceitabilidade por grupo e por condição**

O Gráfico 7 ilustra a média de aceitabilidade das condições de tópico-comentário testadas, nos três grupos. Pode-se verificar que as médias de aceitabilidade do PB são sempre mais altas, independente da condição. Em contrapartida, as médias de aceitabilidade do inglês são sempre mais baixas. Interessante observar neste gráfico, que o Grupo 3 – falantes nativos de PB, fluentes em inglês (L2), tem suas médias localizadas na interseção entre as duas línguas, PB e inglês, principalmente nas condições *Topicalização* e *Tópico chinês sem elemento introdutor*, o que pode sugerir uma possível interferência da sua língua nativa ao julgar as sentenças em inglês, causada pela diferença tipológica entre as línguas.

O Gráfico 8 mostra a probabilidade de aceitabilidade em cada umas das condições, entre os grupos:



**Gráfico 8. Probabilidade de aceitabilidade - Condição x Grupo**

Podemos observar que o Grupo 1 - falantes nativos de PB, possui uma baixa probabilidade de julgamentos inaceitáveis, em comparação aos outros grupos, principalmente ao Grupo 2 – falantes nativos de inglês, nas três condições testadas. Já olhando para a probabilidade de julgamentos aceitáveis, o Grupo 1- falantes nativos de PB, tem uma probabilidade alta de escolhas aceitáveis, enquanto o Grupo 2- falantes nativos de inglês, tem uma alta probabilidade de escolhas inaceitáveis, também nas três condições testadas. O padrão observado no Grupo 3 – falantes nativos de PB, fluentes em inglês (L2), aproximou-se mais do Grupo 1. Também observações podem ser confirmadas nos testes estatísticos a seguir.

Realizamos um teste *Tukey* para verificarmos diferenças estatísticas entre os pares relevantes, a partir da comparação das condições experimentais entre os grupos testados. O resultado do teste estatístico pode ser observado na Tabela 15:

**Tabela 15. Resultado estatístico - Teste *post hoc* Tukey  
Comparação entre Grupos 1, 2 e 3 – Condição Topicalização**

Condições	Estimate	Std. Error	df	z. ratio	p.value
PB_Topicalização – Inglês_Topicalização	3.2159	0.414	Inf	7.767	<.0001
PB_Topicalização - InglêsL2_ Topicalização	1.7551	0.409	Inf	4.292	0.0006
Inglês_Topicalização – InglêsL2_Topicalização	-1.4609	0.402	Inf	-3.639	0.0084

A Tabela 15 mostra os resultados da comparação entre os grupos, referente a condição *Topicalização*. A estatística mostrou uma diferença significativa entre todas as comparações. Sendo assim, a *Topicalização* foi mais aceita pelo Grupo 1 – falantes nativos de PB, do que pelos outros dois grupos. Comparando-se o Grupo 2 – falantes nativos de inglês com o Grupo 3 – falantes nativos de PB, fluentes em inglês (L2), a *Topicalização* foi mais aceita pelo Grupo 3. Os falantes nativos de inglês foram os que apresentaram a menor taxa de aceitabilidade das topicalizações.

A Tabela 16 mostra a comparação dos grupos julgando a aceitabilidade das sentenças de *Tópico chinês com elemento introdutor*:

**Tabela 16. Resultado estatístico - Teste *post hoc* Tukey**  
**Comparação entre Grupos 1, 2 e 3 – Condição Tópico chinês com introdução**

Condições	Estimate	Std. Error	df	z. ratio	p.value
PB_ChinComIntro – Inglês_ChinComIntro	1.6874	0.411	Inf	4.110	0.0013
PB_ChinComIntro - InglêsL2_ChinComIntro	0.2861	0.413	Inf	0.693	0.9989
Inglês_ChinComIntro – InglêsL2_ChinComIntro	-1.4014	0.408	Inf	-3.435	0.0173

Podemos observar que esta condição foi mais aceita igualmente pelos grupos 1 e 3, sendo o Grupo 1 – falantes nativos de PB e Grupo 3 – falantes nativos de PB, fluentes em inglês (L2). As comparações estatísticas não mostram diferença significativa da condição *Tópico chinês com elemento introdutor* nesses grupos. Já em relação ao Grupo 2 – falantes nativos de inglês, a estatística mostra uma diferença significativa, indicando que tal construção foi mais aceita pelo Grupo 1 – falantes nativos de PB. Comparando-se o Grupo 2 – falantes nativos de Inglês com o Grupo 3 – falantes nativos de PB, fluentes em Inglês (L2), observa-se uma diferença estatística significativa, o que indica que a estrutura de *Tópico chinês com elemento introdutor* foi mais aceita pelo Grupo 3. Sendo assim, o Grupo 2 foi o que menos aceitou a construção de *Tópico chinês com elemento introdutor*.

A Tabela 17 mostra a comparação entre os grupos, a partir dos julgamentos de aceitabilidade para a condição *Tópico chinês sem elemento introdutor*:

**Tabela 17. Resultado estatístico - Teste *post hoc* Tukey**  
**Comparação entre Grupos 1, 2 e 3 – Condição Tópico chinês sem introdução**

Condições	Estimate	Std. Error	df	z. ratio	p.value
PB_ChinSemIntro – Inglês_ChinSemIntro	3.3446	0.409	Inf	8.176	<.0001
PB_ChinSemIntro - InglêsL2_ChinSemIntro	1.8107	0.401	Inf	4.513	0.0002
Inglês_ChinSemIntro – InglêsL2_ChinSemIntro	-1.5339	0.403	Inf	-3.808	0.0044

A diferença estatística significativa na comparação entre os três grupos mostra que a condição *Tópico chinês sem elemento introdutor* foi mais aceita pelo Grupo 1 – falantes nativos de PB e que foi menos aceita pelo Grupo 2 - falantes nativos de inglês, em todas

as comparações. Assim, o Grupo 2 – foi o que menos aceitou tal tipo de estrutura. Tal resultado sugere que o inglês não aceita estruturas de tópico-comentário geradas na base, sendo estas estruturas presentes numa língua de tópico proeminente, o que confirma os achados na literatura que consideram o inglês como uma língua de sujeito puro. Essa constatação, conseqüentemente, afasta o PB do inglês, nesse sentido, já que observamos aceitabilidade dessas construções no PB.

#### 4. Conclusão

Concluimos, a partir dos resultados obtidos, que o comportamento dos participantes dos três grupos analisados não se deu de forma semelhante, pois PB e inglês não pertencem a mesma tipologia linguística, em relação a sua orientação para a sentença ou para o discurso. Tais achados dão sustentação a hipótese que norteou este estudo, de que o PB seria uma língua mista, onde as duas noções, tópico e sujeito, seriam importantes, diferentemente do inglês. Isto é, o inglês seria uma língua de sujeito proeminente, cuja orientação se dá para a sentença, enquanto o PB compartilharia características das línguas de tópico proeminente. Por este motivo, o inglês prefere construções de tópico chinês com elemento introdutor, já que a introdução facilita a interpretação do constituinte topicalizado, confirmando achados anteriores da literatura (*cf* Yuan 1995). Quando não havia elemento introdutor, os participantes do grupo de falantes nativos de inglês, rejeitaram tal construção, conforme observado pela média de aceitabilidade de **2,05** (Inaceitável). Construções de tópico chinês sem elemento introdutor representam construções de tópicos puros, predominantes em línguas orientadas para o discurso, sendo estruturas geradas na base, não aceitas em línguas orientadas para a sentença. Em relação aos tópicos gerados via movimentação sintática, os falantes nativos de inglês não as consideraram como inaceitáveis, mas também não julgaram essas estruturas como aceitáveis, ficando na linha neutra, com média de **3,03**.

Em contrapartida, no grupo de falantes nativos de PB, observamos uma alta aceitabilidade para as construções de tópico-comentário, no geral, mesmo quando o tópico chinês não possui um elemento introdutor que atue como um facilitador de sua interpretação. Ou seja, os falantes nativos de PB foram capazes de interpretar e aceitar como boas, estruturas de tópico geradas via movimentação sintática e também aquelas geradas na base, presentes em línguas de tópico proeminente. Agora, tomando o grupo de falantes nativos de PB, fluentes em inglês, fazemos uma observação interessante: os participantes parecem ficar “no meio do caminho” ao julgarem a aceitabilidade das sentenças. Os participantes deste grupo aceitam menos essas estruturas em relação ao PB, porém, aceitam mais as mesmas estruturas, em relação ao inglês. Aqui, podemos constatar uma possível influência da sua língua materna, no julgamento das estruturas na sua segunda língua, no caso o inglês. Como os participantes deste grupo julgaram estruturas de tópico-comentário em inglês, uma língua que apresenta uma aceitabilidade baixa para esses tipos de estruturas, e como a sua língua materna aceita bem essas estruturas (não só as aceita, como também são muito produtivas), houve uma interferência nesse julgamento. Portanto, os dados obtidos na análise do Grupo 3 – falantes nativos de PB, fluentes em inglês (L2) - reforçam a ideia de não considerar o PB como uma língua

orientada para sentença, mas sim como uma língua com orientação mista, em concordância com estudos anteriores de Pontes (1987), Orsini (2003), entre outros.

Como fruto de investigação futura seria interessante verificarmos a aceitabilidade destas estruturas por falantes nativos de chinês e/ou bilíngues, tendo o inglês como L1 e o PB como L2. Dessa forma, poderíamos observar a questão dentro de outros grupos, para comparação com os dados dos grupos já testados - falantes nativos de inglês, falantes nativos de PB e bilíngues (PB-L1 e inglês-L2).

**Financiamento:** Esta pesquisa foi financiada pelo CNPq e pela CAPES.

**Agradecimentos:** Ao Prof. Marcus Maia, orientador da tese de Doutorado que deu origem a esta pesquisa. À Prof.<sup>a</sup> Aline Fonseca, pela orientação e suporte durante a análise estatística desta pesquisa. Ao Prof. John Trueswell, pela supervisão do experimento na University of Pennsylvania, durante a pandemia. À Prof.<sup>a</sup> Katy Carlson, pela gravação dos áudios em inglês para o teste auditivo que compôs esta pesquisa.

## Referências

- Berlinck, R., Duarte, M. E. L., & Oliveira, M. (2015). Predicação. In M. A. Kato, M. Nascimento (Orgs.), *Gramática do Português Culto falado no Brasil: A construção da sentença* (Vol II, pp. 81–149). São Paulo: Ed. Contexto.
- Chen, C. (2011). *The influence of Chinese topic prominence construction on English acquisition – a study on non-English majors and English majors at a Chinese college*. Kristianstad University, Sweden. Consultado em <http://www.diva-portal.org/smash/get/diva2:429120/FULLTEXT01.pdf>
- Cinque, G. (1977). The movement nature of left dislocation. *Linguistic Inquiry*, 8, 397–412.
- Duarte, I. (1996). A topicalização no português europeu: Uma análise comparativa. In I. Duarte & I. Leiria (Eds.), *Actas do Congresso Internacional sobre o Português* (pp. 327–360). Lisboa: APL/Colibri.
- Galves, C. (2001). *Ensaio sobre as gramáticas do português*. Campinas: Editora da Unicamp.
- Kato, M. (2006). *Comparando o português da América com o português de Portugal e com outras línguas*. Museu da língua Portuguesa. Consultado em <https://www.museudalinguaportuguesa.org.br/wp-content/uploads/2017/09/Comparando-o-port-da-Ame%CC%81rica-com-port-Portugal.pdf>
- Kenedy, E. (2002). *Aspectos estruturais da relativização em português: Uma análise baseada no modelo raising* (Dissertação de Mestrado, UFRJ, Rio de Janeiro).
- Kenedy, E. (2011). Tópicos e Sujeitos no PB: Uma abordagem experimental. *Revista da Anpoll*, 1(31), 69–88. <https://doi.org/10.18309/anp.v1i31.201>
- Kenedy, E., & Mota, C. (2012). Orientações de anáforas nulas e pronominais para sujeitos e tópicos no PB. *Revista Linguística*, 8(2), 159–172. <https://doi.org/10.31513/linguistica.2012.v8n2a4556>
- Kenedy, E. (2014). O status tipológico das construções de tópico no português do Brasil: uma abordagem experimental. *Revista da ABRALIN*, 13(2), 151–183. <https://doi.org/10.5380/rabl.v13i2.39614>
- Li, C-N., & Thompson, S. (1976). Subject and topic: A new typology of language. In C-N. Li (Ed.), *Subject and topic* (pp. 457 – 489). New York: Academic Press.
- Maia, M. (1997). The processing of object anaphora in Brazilian Portuguese. *Révue Linguistique de Vincennes*, 26, 151–172.
- Negrão, E. V. (1999). *O português brasileiro: Uma língua voltada para o discurso*. (Tese de livre docência, USP, São Paulo).

- Orsini, M. T. (2003). *As construções de tópico no português do Brasil: Uma análise sintático-discursiva e prosódica*. (Tese de Doutorado, UFRJ, Rio de Janeiro).
- Orsini, M., & Vasco, S. (2007). Português do Brasil: Língua de tópico e de sujeito. *Revista Diadorim*, 2, 83–98. <https://doi.org/10.35520/diadorim.2007.v2n0a3852>
- Pontes, E. (1987). *O tópico no Português do Brasil*. Campinas: Ed. Pontes.
- Silva, C.G.C. (2015). *A interface prosódia-sintaxe na produção e no processamento das estruturas de tópico e SVO* (Tese de Doutorado, UFJR, Rio de Janeiro).
- Silva, A. (2017). *Processamento prosódico na compreensão e produção de estruturas de tópico e sujeito no Português brasileiro* (Dissertação de Mestrado, UFJF, Juiz de Fora).
- Weskott, T., Horning, R., & Webelhuth, G. (2019). On the contextual licensing of English locative inversion and topicalization. In S. Featherston, R. Hörnig, S. von Wietersheim & S. Winkler (Eds.), *Experiments in focus* (pp. 153–182). Berlin: De Gruyter <https://doi.org/10.1515/9783110623093-007>
- Xu, H. (1999). *DRT-Analysis for topic-comment constructions in Chinese*. University of the Saarland, Germany. Consultado em <http://www.coli.uni-saarland.de/publikationen/softcopies/Xu:1999:DAT.pdf>
- Yuan, B. (1995). Acquisition of base-generated topics by English-speaking learners of Chinese. *Language Learning*, 45, 567–603. <https://doi.org/10.1111/j.1467-1770.1995.tb00455.x>

[recebido em 26 de outubro de 2021 e aceite para publicação em 14 de março de 2022]

## DIFICULDADES DE LEITURA DE ESTUDANTES UNIVERSITÁRIOS COM TDAH: UM ESTUDO DA INFLUÊNCIA DA MEMÓRIA DE TRABALHO NA COMPREENSÃO LEITORA

### READING DIFFICULTIES OF COLLEGE STUDENTS WITH ADHD: A STUDY OF THE INFLUENCE OF WORKING MEMORY ON READING COMPREHENSION

José Ferrari Neto\*  
joseferrarin1974@gmail.com

Gustavo Lopez Estivalet\*\*  
gustravoestivalet@hotmail.com

Priscilla de Albuquerque Almeida\*\*\*  
prisca.albuquerque@gmail.com

O Transtorno do Déficit de Atenção com Hiperatividade (TDAH) é um transtorno do neurodesenvolvimento que tem como algumas de suas comorbidades os transtornos de linguagem. No âmbito acadêmico, a queixa mais comum é a dificuldade de compreensão leitora de adultos com este transtorno. Este artigo propõe analisar como se dá o processo de leitura em estudantes universitários com TDAH, conferir se os prejuízos na compreensão leitora são decorrentes do TDAH, investigar a influência da memória de trabalho na compreensão da leitura dos participantes com TDAH e comparar a velocidade do desempenho na tarefa de decisão lexical entre os participantes com TDAH e sem TDAH. Participaram nesse estudo 15 estudantes universitários com diagnóstico de TDAH e 15 estudantes universitários sem o referido transtorno. Os instrumentos utilizados foram: o Cloze, a Tarefa de Atenção Linguística (TAL) e a Versão Computadorizada do Reading Span Test para o Português Brasileiro (VCRST-PB). Os resultados apontaram que a capacidade de memória de trabalho dos participantes com TDAH medida no VCRST-PB foi reduzida e exerceu influência negativa na compreensão leitora, assim como a TAL mostrou uma influência maior do que a VCRST-PB na capacidade de leitura. Isso pode ser justificado devido ao processo de leitura requerer uma maior demanda de sustentação da atenção para a emergência da compreensão do conteúdo lido. Sendo assim, isso é mais custoso para os universitários com TDAH devido as dificuldades que os mesmos apresentam no controle executivo.

**Palavras-chave:** Transtorno de Déficit de Atenção/Hiperatividade. Compreensão leitora. Memória de trabalho. Atenção linguística.

---

\* Centro de Ciências Humanas Letras e Artes, Universidade Federal da Paraíba, João Pessoa, Brasil. ORCID: 0000-0003-2734-7197.

\*\* Centro de Ciências Humanas Letras e Artes, Universidade Federal da Paraíba, João Pessoa, Brasil. ORCID: <http://orcid.org/0000-0003-3462-4156>

\*\*\* Centro de Ciências Humanas Letras e Artes, Universidade Federal da Paraíba, João Pessoa, Brasil. ORCID: 0000-0002-9702-365X.

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder which has language disorders as some of its comorbidities. Reading comprehension difficulty in adults with this disorder has been the most common complaint within the academic scope. This paper aims to analyze how the reading process occurs in undergraduate students with ADHD, check whether impairments in reading comprehension are due to ADHD, investigate the influence of working memory on the reading comprehension of participants with ADHD, and compare the speed performance on the lexical decision task among participants with and without ADHD. Fifteen undergraduate students diagnosed with ADHD and 15 undergraduate students without this disorder participated in this study. The experiments used for data collection were: the Cloze, the Task for Linguistic Attention (TAL) and the Computerized Version of Reading Span Test for Brazilian Portuguese (VCRST-PB). The results indicated that the working memory capacity of participants with ADHD measured by the VCRST-PB was reduced a negative influenced the reading comprehension, as well as the TAL showed a greater influence than the VCRST-PB reading ability. This can be explained by the fact that the reading process requires a greater demand of sustaining attention for the emergence of the understanding of the content read. Therefore, this takes a great deal of effort for undergraduate students with ADHD due to the difficulties they present in the executive control.

**Keywords:** Attention-Deficit/Hyperactivity Disorder. Reading comprehension. Working memory. Linguistic attention.

•

## 1. Introdução

A leitura é uma competência fundamental para ampliar os nossos conhecimentos, para nos tornar mais críticos da realidade em que vivemos, para adquirir cultura e para nos tornar mais hábeis na arte da comunicação. Para os letrados, o processo da leitura é algo automático. No entanto, quando os sujeitos possuem o Transtorno de Déficit de Atenção com Hiperatividade (TDAH), este processo pode ser algo angustiante e mais custoso e menos automático, pois os sujeitos com TDAH demonstram apresentar dificuldades linguísticas manifestadas na leitura (Albuquerque, 2012).

Aproximadamente 14% dos casos diagnosticados com TDAH apresentam como comorbidades transtornos de linguagem. Consequentemente, as dificuldades de caráter linguístico decorrentes desse transtorno podem contribuir para um baixo rendimento acadêmico (Rotta, Ohlweiler & Riesgo 2016). O Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-V) apresenta tanto o TDAH como o Transtorno da Aprendizagem e o Transtorno da Linguagem no mesmo eixo referente ao Transtorno do Neurodesenvolvimento, mostrando a interligação existente entre esses transtornos (American Psychiatric Association [APA] 2013).

Em razão disso, os adultos com TDAH quando ingressam na universidade, quer na graduação ou na pós-graduação, encontram dificuldades em organizar as suas ideias, realizar leituras, análises e interpretações de conteúdos científicos. As dificuldades de caráter linguístico podem trazer diversos prejuízos na vida acadêmica, pessoal e laboral,

consequentemente aumentando o risco de evasão no Ensino Superior (Albuquerque 2012). O ambiente acadêmico exige dos discentes o pleno funcionamento cognitivo e desempenho acadêmico. Contudo, muitos estudantes universitários com TDAH apresentam dificuldades de aprendizagem associadas a leitura. Sugere-se que as dificuldades de leitura estejam relacionadas a um componente fundamental do processo de leitura: a memória de trabalho (MT). Barkley (1997) afirma que o TDAH interfere na capacidade da memória de trabalho (CMT), dificultando a retenção e recuperação de informações. Enfim, sabe-se que o processo de entrada e saída de informações no sistema de memória é primordial para que ocorra a compreensão da leitura.

Diante disso, surgem as seguintes questões que exigem uma melhor investigação: Os adultos com TDAH possuem uma maior lentificação em tarefas de leitura em comparação aos universitários que não possuem este transtorno? A menor CMT proveniente do TDAH pode estar associada a dificuldades na compreensão leitora? Sendo assim, as hipóteses que assumimos nessa pesquisa são que a menor CMT está associada a dificuldades na compreensão leitora e que os estudantes universitários com TDAH podem apresentar uma maior lentificação na tomada de decisão.

Diante desta problemática, o presente trabalho investigou como se dá a compreensão leitora em estudantes universitários diagnosticados com TDAH. Para isso, utilizaram-se testes psicolinguísticos *on-line* e *off-line*. Assim, buscou-se conferir os prejuízos na compreensão leitora decorrentes do TDAH, bem como analisar a influência da MT na compreensão de leitura dos participantes com TDAH, comparando-se a velocidade do desempenho da tarefa de decisão lexical entre os participantes do grupo experimental com TDAH com a do grupo controle sem TDAH.

Estudos psicolinguísticos sobre a correlação entre o desempenho da CMT na compreensão leitora de crianças com TDAH revelaram que os participantes do grupo experimental tiveram interferência de falhas de MT no processamento linguístico, pois a velocidade de leitura mostrou ser reduzida e a compreensão de sentenças também foi menor (Albuquerque *et al.* 2012; Klein & Lamprecht 2012). O estudo de Yeari, Avramovich e Schiff (2016) com adolescentes diagnosticados com TDAH sobre as inferências preditivas, explicativas e evocativas de informações textuais na compreensão leitora por meio de medida *on-line* mostraram que os participantes do grupo experimental conseguiram acertar as perguntas dos textos quando estes eram relativamente curtos, porém, apresentaram dificuldades para gerar inferências preditivas e explicativas, assim como reter informações relevantes do texto na MT durante o processo de leitura. Observou-se também que o processamento de texto por meio de técnica online foi mais lento e a compreensão inferencial foi menor. Esses estudos apresentados aqui nos ajudam a perceber que as dificuldades advindas desse transtorno emergem nos anos iniciais da Educação Básica, porém, podem permanecer ao longo de toda formação escolar e acadêmica, dificultando o desempenho de pessoas diagnosticadas com TDAH. Portanto, o presente trabalho se propôs a investigar como a compreensão leitora é prejudicada pela CMT de jovens adultos com TDAH no âmbito do Ensino Superior.

O artigo está estruturado do seguinte modo. Na seção 2, são apresentados alguns estudos sobre o TDAH, a MT e a compreensão leitora. Na seção 3, é apresentada a

metodologia adotada nesta pesquisa, o conjunto experimental, os resultados e a discussão. Enfim, na seção 4, são apontadas as conclusões do trabalho.

## **2. Reflexões acerca das relações entre o TDAH, a MT e a compreensão leitora**

O TDAH é um transtorno neurobiológico, de causas genéticas. Estudos com ressonância magnética mostram que os sujeitos com TDAH apresentam um menor volume na região frontal ântero-superior direita e ântero-inferior direita, assim como disfunção pré-frontal-estriatal direita. Observa-se que essas áreas são responsáveis pelas funções executivas como a atenção, a MT, o planejamento, o controle dos impulsos e das emoções. Assim, os indivíduos com esse transtorno possuem uma menor quantidade dos neurotransmissores dopamina e noradrenalina na região frontal do cérebro, responsáveis pela manutenção da atenção, da MT e da motivação (Cypel 2007; Filipek *et al.* 1997; Hynd 1993). Ainda, Aarts *et al.* (2015) afirmam que os efeitos positivos dos psicoestimulantes no tratamento desse transtorno, que são inibidores de recaptação da dopamina, evidenciam a hipótese de uma disfunção nos sistemas dopaminérgico e adrenérgico.

Estudos genéticos e de neuroimagem funcional e estrutural têm contribuído para o entendimento de que o TDAH é um transtorno cerebral com sintomas cognitivos e comportamentais que se alteram com o passar dos anos. As queixas mais frequentes de ordem cognitiva são: dificuldades para iniciar tarefas, organizar o trabalho, planejar/executar metas e monitorar o tempo/os prazos, assim como procrastinação, perda da motivação, baixa tolerância à frustração, inconstância e lentidão no desempenho. Observa-se que esses sintomas do TDAH podem ocorrer em todas as fases do desenvolvimento (Neto 2010).

O indivíduo com o referido transtorno possui um padrão persistente de déficit de atenção e/ou hiperatividade-impulsividade. Segundo o DSM-V (APA 2013), o déficit de atenção se manifesta como divagação nas atividades, desorganização e dificuldade de sustentar o foco. Na infância, a hiperatividade diz respeito à agitação motora, remexer-se na cadeira, bater em objetos repetidas vezes ou conversar em excesso; já na fase adulta, a hiperatividade manifesta-se como extrema inquietude ou exaustão dos seus pares com sua atividade. A impulsividade se refere à intromissão social, tomada de decisões sem planejamento prévio, desejo de recompensas imediatas e procrastinação.

Segundo Barkley (2006), o TDAH afeta cerca de 3 a 6% ou mais das crianças em idade escolar, estimando-se que em uma sala com 20 alunos, haverá pelo menos um com TDAH. Os prejuízos advindos desse transtorno podem perdurar ao longo da vida, logo, a estimativa da persistência dos sintomas na fase da adolescência ocorre em cerca de 70%, já na fase adulta em torno de 66% dos casos diagnosticados na infância. Os sintomas do TDAH modificam-se ao longo dos anos, por exemplo, a hiperatividade mais frequente na infância diminui na fase adulta. Porém, mesmo adultos com este transtorno continuam apresentando prejuízos funcionais que interferem na vida acadêmica, laboral e afetiva (Barkley 2006). Devido aos sujeitos com TDAH apresentarem uma menor CMT - componente fundamental para a compreensão leitora - este estudo analisou as possíveis

implicações da menor CMT na compreensão leitora em estudantes universitários com TDAH.

A memória é um sistema cognitivo que tem como função a aquisição, a manutenção e a evocação das informações. Ela é dividida em três subsistemas: memória sensorial, quando o estímulo é percebido no canal sensorial e dura de milissegundos a segundos; logo mais, a informação é enviada para a MT, habilidade cognitiva associada com o armazenamento e processamento da informação; enfim, a memória de longo prazo permite a retenção da informação. Esse processo é fundamental para que ocorra a aprendizagem. A informação pode ser armazenada na memória de longo prazo por dias ou até anos. Assim, a MT está intimamente relacionada com a memória de longo prazo, pois é por meio da ativação da MT que as informações consolidadas na memória de longo prazo são evocadas e utilizadas (Pasquali 2019).

Baddeley e Hitch (1974) afirmaram que a MT é um sistema cognitivo composto por três componentes de armazenamento e processamento de informações em domínios específicos e com competência de modificar, inibir e atualizar os conhecimentos. Esses autores propuseram que a CMT é limitada e estabeleceram um modelo de MT conhecido como Modelo Multicomponente composto por três componentes: a alça fonológica, o bloco de esboço visuoespacial e o executivo central. Posteriormente, Baddeley (2000) adicionou um quarto componente, o *buffer* episódico.

A alça fonológica é responsável pelo armazenamento da informação verbal. Ela organiza o estímulo linguístico de modo temporal e sequencial, sustentando-o por um curto período de tempo e, se for reverberada pelo subcomponente denominado alça articulatória, a informação continuará sendo mantida no armazenador fonológico. A alça fonológica exerce um papel fundamental para a aquisição de novas palavras e para a compreensão leitora (Baddeley & Hitch 1974).

O bloco de esboço visuoespacial tem a função de armazenar e processar informações visuais e espaciais. Assim como na alça fonológica, o tempo de duração do armazenamento do bloco de esboço visuoespacial é curto. Esse componente está relacionado na aquisição do conhecimento semântico relativo à imagem mental quanto à forma dos objetos (Mourão-Júnior & Melo 2011).

O executivo central (EC) é um sistema atencional que controla o fluxo de entrada e saída das informações. Suas principais funções são: a) coordenar o desempenho em tarefas simultâneas, como por exemplo, o armazenamento e o processamento da informação; b) selecionar a informação relevante e inibir a informação irrelevante; c) escolher entre uma tarefa, estratégia ou operação; e, d) ativar e recuperar as informações armazenadas na memória de longo prazo (Baddeley 1996). Portanto, esse componente está relacionado com a atenção, o planejamento de estratégias e a tomada de decisão.

Quando uma dada tarefa possui alta complexidade, exigindo maior demanda cognitiva, o EC regula o fluxo de informações e seleciona a entrada dessas nos subcomponentes alça fonológica e bloco de esboço visuoespacial. Barkley (2002) afirma que os indivíduos com TDAH são mais sujeitos a uma ruptura ou perda de atenção no componente EC que em qualquer outro componente da MT. Possivelmente, o baixo desempenho do EC está associado às dificuldades de compreensão leitora dos indivíduos com TDAH (De Jong *et al.* 2009; Dias & Mecca 2019; Whipple & Nelson 2016).

O *buffer* episódico é um componente de armazenamento com capacidade limitada, ele faz a conexão das informações entre os componentes verbal e visual, como também entre a memória de longo prazo e a MT em uma representação episódica unitária. O *buffer* também está relacionado à conexão das informações, como por exemplo, arquivar e evocar a informação temporária de qual lugar foi guardado um determinado objeto (Zanella & Valentini 2016).

A CMT exerce uma função importante na compreensão leitora. Daneman e Carpenter (1980) afirmam que existem diferentes tipos de MT correspondentes aos do sistema linguístico relacionados às representações fonológicas, morfológicas, sintáticas e semânticas.

A compreensão da leitura se dá quando o leitor consegue armazenar as informações sintáticas, semânticas e pragmáticas lidas no conteúdo do texto anterior e as usa nas relações com o conteúdo do texto posterior. Essa é uma atividade complexa que envolve processos cognitivos de ordem geral e especificamente linguísticos (Smith & Geva 2000). Diante disso, os prejuízos na compreensão leitora dos estudantes universitários com TDAH podem estar associados também às dificuldades da CMT.

O processamento linguístico da leitura envolve a identificação das palavras por meio de decodificação fonológica ou do acesso lexical direto. Quando a palavra lida possui representação completa no léxico ortográfico, o leitor utiliza a rota lexical. Nesta rota, o item é reconhecido ortograficamente e a sua forma ortográfica ativa a representação semântica antes de ativar a forma fonológica. Já na leitura de palavras e pseudopalavras que não estão no léxico ortográfico, a leitura é feita através da decodificação grafofonêmica, que é a rota fonológica. Nesta rota, o acesso ao significado é obtido mais tarde, quando a forma fonológica da palavra ativa o sistema semântico (Capovilla *et al.* 2004; Ellis & Young 1988).

O leitor proficiente é dotado de conhecimentos grafofonológicos, grafomorfológicos, morfossintáticos, semânticos e pragmáticos. O processo de leitura se inicia através do *input* visual ortográfico e, a partir da decodificação fonológica, a palavra é identificada, permitindo o acesso lexical e a ativação da informação semântica associada à palavra (Giraldello 2016, Perfetti & Verhoeven 2017).

A finalidade principal da leitura é compreender o conteúdo que o texto traz, ou seja, que o conteúdo do texto proporcione um sentido para o leitor. Contudo, como discutido anteriormente, exigem-se diversos processos cognitivos simultâneos, tais como de ordem linguística e da CMT que possibilitam a incorporação e recuperação dos conteúdos. Para que isso ocorra, os leitores precisam ter esses processos cognitivos funcionando adequadamente (Crenitte *et al.* 2010; Dias & Meca 2019).

Embora os universitários com TDAH possuam a competência linguística de reconhecer vocábulos e não apresentem um comprometimento significativo na consciência fonológica ou nas habilidades de decodificação envolvidas na leitura, eles ainda podem apresentar uma compreensão leitora prejudicada. Logo, os prejuízos de desempenho na compreensão leitora dos indivíduos com TDAH podem ser provenientes dos déficits executivos da MT e de atenção sustentada e seletiva, bem como percepção visual e velocidade de processamento (DuPaul & Stoner 2007; Samuelsson, Lundberg & Herkner 2004).

Durante a leitura, seja silenciosa, seja em voz alta, é frequentemente observado nos leitores com TDAH, tanto em crianças como em adultos, algumas substituições e omissões de palavras. Isso é um indicativo de falhas no acesso lexical, mais precisamente na seleção do item lexical, o que pode influenciar no comprometimento da compreensão leitora. O estudo de Albuquerque (2012) confirmou isso durante a realização de triagem para seleção dos participantes. Os participantes do grupo experimental substituíram algumas palavras apresentadas no texto proposto, como por exemplo: *bebia* por *bebida* e *precauções* por *preocupações*.

A lentificação da velocidade de leitura é uma das queixas mais comuns entre os universitários com TDAH. Atividades que envolvem a leitura normalmente requerem uma maior demanda de tempo para estes, pois é necessário fazer várias retomadas do texto. Logo, durante a leitura, ocorre uma disputa de recursos entre armazenamento e processamento das informações, ou seja, quanto mais o leitor precisa armazenar informações, menor capacidade de processamento ele terá, ou quanto mais ele precisa da capacidade de processamento, menos consegue armazenar informações. Assim, o EC faz o controle desse sistema, mandando mais recursos atencionais para o armazenamento ou para o processamento na alça fonológica ou no bloco de esboço visuoespacial da MT (Baddeley 2000; Fergusson & Horwood 1992; Lima, Tambaquim & Ciasca 2010).

### 3. Procedimentos metodológicos

Nesse sentido, o presente estudo se caracteriza como uma pesquisa experimental de caráter quantitativo e comparativo, pois as hipóteses levantadas foram analisadas a partir de análises estatísticas comparativas entre um grupo experimental e um grupo controle. O estudo foi aprovado pelo Comitê de Ética em Pesquisa do Centro de Ciências da Saúde da Universidade Federal da Paraíba (CEP/CCS/UFPB), conforme a Resolução n.510/2016 do CNS/MS, emitida em 26 de março de 2020.

Participaram no estudo 30 estudantes da Universidade Federal da Paraíba (UFPB) da cidade de João Pessoa-PB, sendo 15 com diagnóstico de TDAH para o grupo experimental, sendo 4 do sexo feminino e 11 do sexo masculino, com idade média de 24a6m (DP = 3,7a) do grupo geral, com média de idade por sexo: 25a5m (DP = 7,1a) do sexo feminino e 26a3m (DP = 2,1a) do sexo masculino. Para o grupo controle foram 15 universitários sem o referido transtorno, sendo 4 do sexo feminino e 11 do sexo masculino, com idade média de 22a9m (DP = 1,7a) do grupo geral, com média de idade por sexo: 22a (DP = 0,8a) do sexo feminino e 24a6m (DP = 1,9a) do sexo masculino. Assumiu-se como critério de exclusão participantes sob o efeito de medicações psicoestimulantes no momento da coleta dos dados. Foram aplicados três experimentos: Experimento 1 - Cloze, com duração média de 15min; Experimento 2 – Tarefa de atenção linguística (TAL), com duração média de 20min; e Experimento 3 – Versão Computadorizada do Reading Span Test do português brasileiro (VCRST-PB) para avaliar a CMT, com duração média de 30min.

Para a seleção dos participantes do grupo controle, uma psicóloga capacitada utilizou a Bateria Psicológica para Avaliação da Atenção (BPA) (Javier & Monteiro 2013) que objetiva avaliar os três tipos de atenção: Atenção Concentrada, Dividida e

Alternada, possibilitando identificar os indivíduos que não possuem déficit atencional, assim como foi aplicada a Escala de Transtorno do Déficit de Atenção e Hiperatividade (ETDAH-AD) (Benczic 2013). Estes instrumentos auxiliam na identificação da presença ou não dos vários sintomas que descrevem o TDAH, sendo compreendida por cinco fatores: Déficit de Atenção, Impulsividade, Aspectos Emocionais, Autorregulação da Atenção, da Motivação e da Ação, Hiperatividade.

Foram selecionados os participantes do grupo controle que obtiveram os resultados médio superior ou superior da BPA (2013) e que também obtiveram os resultados inferior, médio-inferior ou médio da ETDAH-AD (2013). Esses resultados foram um indicativo de que os participantes não apresentam sintomas associados ao TDAH. O grupo experimental não precisou submeter-se a estes testes, pois os mesmos já possuíam o diagnóstico de TDAH dado por um médico psiquiatra ou neurologista. Foram selecionados 10 participantes com TDAH que receberam atendimentos psicopedagógicos na Clínica-Escola de Psicopedagogia da UFPB, entre os anos de 2014 a 2016, todos com queixas de dificuldades de compreensão leitora. Estes participantes indicaram outros 5 colegas que também possuem o diagnóstico de TDAH. Participantes com Dislexia não foram incluídos na pesquisa.

### 3.1. Experimento 1 - Teste Cloze

O teste Cloze é um instrumento desenvolvido por Taylor (1953) e é utilizado para avaliar a compreensão leitora a partir de três níveis de desempenho: frustração, quando o participante consegue acertar até 44% das lacunas, indicando que o mesmo obteve baixo desempenho na compreensão do texto apresentado; instrucional, quando o participante obtém um percentual de acertos entre 44% a 57%, indicando que esse obteve uma compreensão leitora suficiente, contudo, necessita de um auxílio externo como a ajuda de um professor ou de uma pessoa que se dispõe a ler em voz alta para o aluno; independente, quando o participante consegue acertar 57% ou mais, mostrando assim que possuiu uma compreensão leitora desejável e autônoma (Bormuth 1968). O teste Cloze foi aplicado com a finalidade de avaliar a compreensão leitora dos participantes de ambos os grupos. Nossa hipótese é que o grupo experimental apresentaria uma capacidade de compreensão leitora inferior ao grupo controle.

Utilizamos a crônica Desentendimento (Veríssimo 1995), adaptada por Santos *et al.* (2002) com a técnica Cloze. Este texto é composto por 369 palavras. Omitiu-se os quintos vocábulos de cada frase do texto para os participantes preencherem com livre escolha. O total de lacunas a preencher é 40. A primeira e as últimas orações do texto foram mantidas para que os participantes tivessem um entendimento mais completo do tema trabalhado. Para a correção, utilizamos um crivo das respostas corretas. A pontuação do experimento 1 foi o resultado da somatória de lacunas preenchidas com palavras exatas do texto. Cada lacuna equivalia a 1 ponto. As variáveis independentes foram: tamanho do texto (369 palavras); quantidade de vocábulo omitido (40). A variável dependente foi: respostas corretas.

#### 3.1.1. Procedimentos

Os participantes fizeram primeiramente uma leitura silenciosa do texto com as omissões dos termos, pois isso contribui e é essencial para a compreensão do conteúdo do texto. Depois preencheram as lacunas com o termo que os mesmos julgaram com mais sentido para a compreensão das frases do texto. Os participantes foram informados que os termos omitidos poderiam ser: artigo, preposição, verbo, substantivo, advérbio, adjetivo ou pronome.

### 3.1.2. Resultados e discussão

Os resultados obtidos no teste Cloze entre os grupos são mostrados no Gráfico 1.

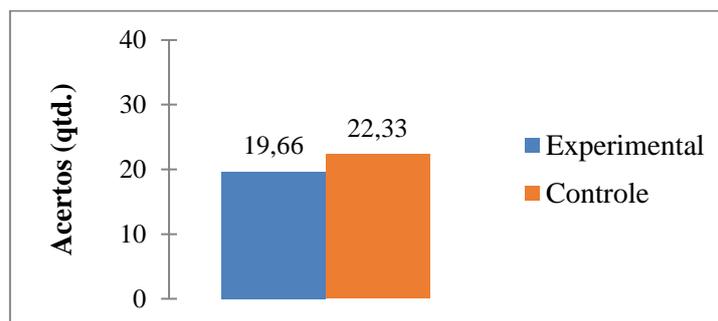


Gráfico 1. Quantidade de acertos do Teste Cloze entre os grupos

Observa-se que os participantes com TDAH acertaram menos em comparação aos participantes que não possuem o transtorno ( $GE = 19,66$  e  $GC = 22,33$ ) ( $t(14) = 1,82^1$ ,  $p$ -valor  $< 0,05$ ). O desempenho inferior do grupo experimental no teste Cloze confirmou que os participantes com TDAH apresentam dificuldade de compreensão leitora.

Este resultado pode ser justificado pelo fato dos sujeitos com TDAH apresentarem comprometimento nas funções executivas associadas a processos cognitivos necessários para que ocorra a compreensão leitora, como: estado de alerta, atenção seletiva e sustentada, velocidade de processamento e fluência. Esse comprometimento acaba inibindo o acesso ao significado das palavras do texto e, por sua vez, dificultando o objetivo final da leitura que é a compreensão (Crenitte *et al.* 2010; Lima, Tambaquim & Ciasca 2010; Samuelsson, Lundberg & Herkner 2004).

### 3.2. Tarefa de atenção linguística (TAL)

Este instrumento avalia o desempenho do processamento do acesso lexical durante a leitura na tarefa de decisão lexical dos participantes. A hipótese proposta foi que o grupo experimental possuiria um desempenho da decisão lexical mais lento em comparação com os participantes do grupo controle. A lentificação na tomada de decisão dos universitários com TDAH justificaria a dificuldade de compreensão leitora, pois esses sujeitos requerem uma maior demanda no tempo de leitura. Para isso, adaptou-se o experimento da TAL que é baseado na junção de uma tarefa de *Simon* com decisão

<sup>1</sup> Foi aplicado um teste de normalidade por participantes aos dados e constatada a normalidade do mesmo, permitindo a aplicação do teste paramétrico Teste-t de *Student*.

lexical. O experimento foi implementado no programa *Paradigm Experiments* (*Perception Research Systems Incorporated*, EUA), sendo aplicado através de um notebook LENOVO B490.

Esta tarefa consiste de um total de 384 estímulos, sendo 192 palavras e 192 pseudopalavras. As palavras foram selecionadas por meio do Léxico do Português Brasileiro (Estivalet & Meunier 2015) em uma análise comparativa dos corpora SUBTLEX-BR e WordLex-BR. Metade das 192 palavras contêm muitos vizinhos (>10) e metade das palavras contêm poucos vizinhos (<4) ortográficos/fonológicos. Utilizou-se o paradigma da tarefa de *Simon* na apresentação dos estímulos quanto a posição (esquerda, centro e direita).

Os estímulos experimentais foram controlados pela frequência de superfície ortográfica, frequência de superfície fonológica, número de letras, número de fonemas, vizinhança ortográfica, vizinhança fonológica, número de sílabas, frequência de sílaba, frequência de bigrama, frequência de trígama e homógrafos. As pseudopalavras foram criadas através da transposição da primeira letra dos estímulos experimentais, sendo construídas pela ferramenta de criação de pseudopalavras disponibilizada no Léxico do Português Brasileiro (Estivalet & Meunier 2017).

Foram adotados três critérios para equilibrar quanto à posição e à ordem de apresentação dos estímulos: máximo de três estímulos da mesma condição em sequência, apresentação não consecutiva de dois estímulos iniciando com a mesma letra e existência de pelo menos oito estímulos entre dois estímulos da mesma condição. Para isso, utilizou-se o programa *Mix* (*Mix Software, Richardson, TX, EUA*) (Van Casteren & Davis 2006).

As variáveis independentes foram: Congruência (congruente, incongruente e neutro), Vizinhança ortográfica (o número de palavras que são formadas ao se mudar uma letra da palavra alvo enquanto a posição das demais letras é mantida) muitos vizinhos ou poucos vizinhos. Já as variáveis dependentes foram: tempo de reação e acurácia.

Quanto as condições experimentais: *congruente* é quando o estímulo pseudopalavra aparece na mesma posição da tela e da tecla (à esquerda) ou quando o estímulo for palavra aparece na mesma posição da tela e tecla (à direita), *incongruente* é quando os estímulos (pseudopalavra ou palavra) aparecem na tela do computador em posição contrária da tecla, *neutro* é quando o estímulo aparece no centro da tela. Para pseudopalavras, o participante pressionava a tecla Q com a mão esquerda e para palavras, pressionava a tecla P com a mão direita.

### 3.2.1. Procedimentos

Os estímulos foram apresentados da seguinte forma: primeiro, uma tela entre estímulos é exibida com tempo aleatório entre 1000 e 1500ms; em seguida, um ponto de fixação é exibido no centro da tela por 1000ms; em seguida, o estímulo alvo é exibido no centro da tela por 3000ms ou até a resposta do participante; a apresentação de um novo estímulo começa com a apresentação da tela entre estímulos. O tempo da tela entre estímulos varia (entre os tempos 1000, 1100, 1200, 1300, 1400, 1500ms) de forma contrabalanceada e pseudoaleatória ao longo do experimento para que os participantes não prevejam o início da sequência do próximo estímulo e não entrem em um ritmo de processamento

automatizado. Os estímulos foram apresentados em uma das três posições (esquerda, centro, direita) da tela do computador, em letras maiúsculas, fonte *Courier New*, tamanho 20, brancas contra um fundo preto. A tarefa dos participantes era identificar se as palavras apresentadas existiam ou não no português, eles deveriam responder o mais corretamente e rapidamente possível.

### 3.2.2. Resultados e discussão

Para a análise dos tempos de reação, utilizou-se uma análise da variância (ANOVA)<sup>2</sup> com o tempo de reação como variável dependente e grupo, congruência e vizinhança como variáveis independentes. Os resultados obtidos do tempo de reação entre os grupos são mostrados no Gráfico 2.

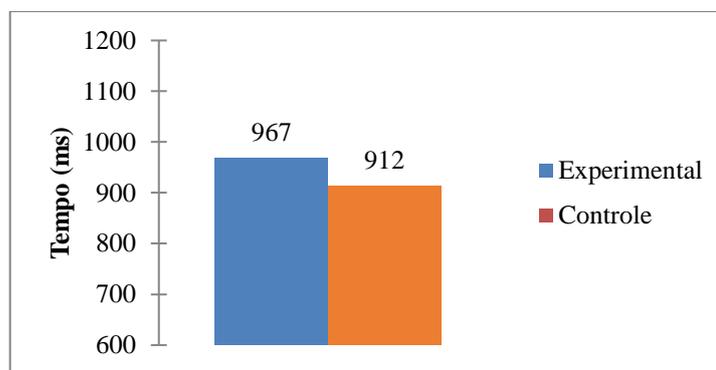


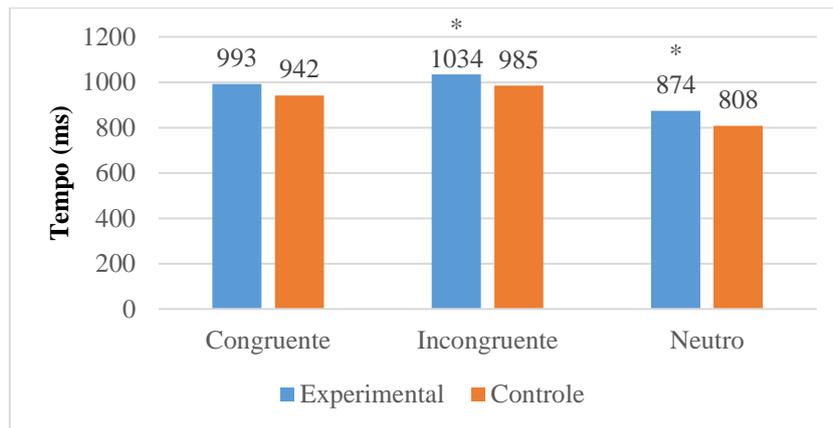
Gráfico 2. Tempo de reação entre o GE e GC na TAL

Conforme o Gráfico 2 mostra, os participantes com TDAH foram mais lentos nas respostas na TAL em comparação com os participantes que não possuem o TDAH (GE = 967ms e GC = 912ms) ( $f = 7,51836$ ,  $p\text{-valor} < 0,05$ ). Esse resultado pode ser justificado pelo fato dos sujeitos com o referido transtorno possuírem déficit no EC, que é o sistema de controle atencional do modelo multicomponete da MT (Baddeley & Hitch 1974).

Pesquisadores afirmam que tarefas que envolvem tomada de decisão lexical mais rápida, requerem dos participantes uma alta demanda cognitiva de recuperação de informação fonológica, atenção seletiva e percepção visual. Por isso, este resultado sugere que a tarefa complexa da TAL reflete as dificuldades de controle executivo dos participantes com TDAH que, por sua vez, pode impactar o desempenho de leitura e a compreensão de textos escritos (Barkley 1997; Whipple & Nelson 2016).

O tempo de reação quanto a congruência entre os grupos é apresentado no Gráfico 3.

<sup>2</sup> Foi aplicado um teste de normalidade por participantes aos dados e constatada a normalidade dos mesmos, permitindo a aplicação do teste paramétrico ANOVA. Foi realizada uma ANOVA para testarem-se os efeitos principais e de interação das variáveis de interesse.

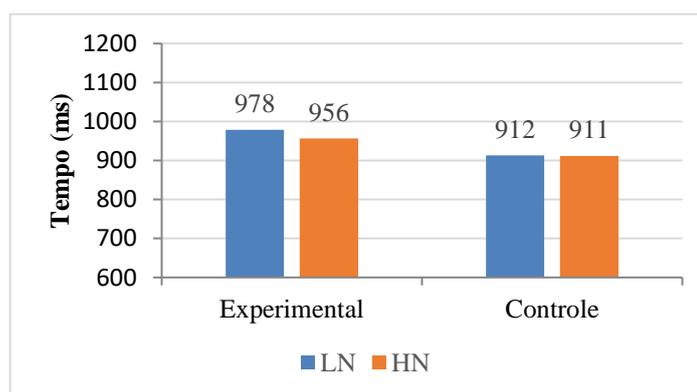


**Gráfico 3. Tempo de reação quanto à condição congruência**

O Gráfico 3 mostra que também houve efeito significativo na condição congruência ( $f = 25,09951$ ,  $p\text{-valor} < 0,001$ ). O gráfico 3 nos revela que na condição congruência neutra, o grupo experimental teve o tempo de reação mais rápido (874ms) em relação com as condições congruente (993ms) e incongruente (1034ms). Podemos notar que a mais lenta foi a incongruente, conforme nossa hipótese.

Conforme citado anteriormente, a TAL contém a tarefa de *Simon*, cuja variável é a congruência: congruente e incongruente. Esta tarefa verifica o controle executivo da atenção do sujeito na tomada de decisão sobre as palavras e pseudopalavras. De fato, a condição incongruente é a condição que mais requer a atenção do participante e por isso ela se torna mais custosa, principalmente para os sujeitos com TDAH que normalmente apresentam falhas no EC. Estudos também apontam que dentre as dificuldades advindas deste transtorno estão: o déficit na decisão lexical, na MT visuoespacial e atenção, justamente por demandarem mais controle EC (Barkley 2002; De Jong *et al.* 2009).

Os resultados do tempo de reação entre os grupos quanto à vizinhança ortográfica/fonológica são mostrados no Gráfico 4.



**Gráfico 4. Resultado do tempo de reação quanto a vizinhança ortográfica: muitos vizinhos (HN) e poucos vizinhos (LN).**

Como mostra o Gráfico 4, não houve diferença significativa entre os tempos de reação dos grupos no tocante à vizinhança ortográfica ( $f = 0,33746$ ,  $p\text{-valor} = 0,56$ ). Este resultado pode sugerir que os participantes com TDAH não tiveram dificuldades para identificar estímulos com muitos ou poucos vizinhos.

A literatura indica que crianças com TDAH em idade escolar apresentam alterações no processamento fonológico (Rotta, Ohlweiler & Riesgo 2006), contudo, esse resultado não foi encontrado na nossa amostra, possivelmente porque essa dificuldade é superada na fase adulta.

**Tabela 1. Análise dos coeficientes (grupo, condição e vizinhança)**

Preditor	Estimativa	Estat.t	P-valor
Grupo experimental	55,35	2,74	<0,01
Condição IN	42,00	1,69	0,09
Condição NE	126,28	5,10	<0,001
Vizinhança LN	11,72	0,58	0,56

Observando a Tabela 1, a estimativa ( $t(14) = 2,74$ ,  $p$ -valor  $< 0,01$ ), nota-se que houve diferença entre os grupos, os participantes com TDAH foram mais lentos em aproximadamente 55ms. Na condição incongruente ( $t(14) = 1,69$ ,  $p$ -valor = 0,09) em comparação com a congruente, a condição incongruente foi mais lenta em aproximadamente 42ms. Na condição neutra ( $t(14) = -5,10$ ,  $p$ -valor  $< 0,001$ ), em comparação à congruente foi aproximadamente 126ms mais rápida. Na variável vizinhança ortográfica ( $t(14) = 0,58$ ,  $p$ -valor = 0,56), percebe-se que não houve efeito de vizinhança. Para ambos os grupos, a palavra alvo ter muitos ou poucos vizinhos não fez diferença. Quanto à quantidade de acertos dos grupos neste experimento (Teste de *Wald* = -5,19,  $p$ -valor  $< 0,001$ , estimativa = -0,61), percebe-se que o grupo experimental cometeu mais erros em comparação com o grupo controle, conforme nossas hipóteses, pois esse experimento exige mais esforço das funções executivas, deficitária nos sujeitos com TDAH.

DuPaul e Stoner (2007) afirmam que os indivíduos com o referido transtorno apresentam baixo desempenho em testes cognitivos correspondentes com déficits associados à leitura, como por exemplo, em testes de velocidade de processamento perceptivo.<sup>3</sup> De fato, o baixo desempenho acadêmico tem sido uma característica muito presente na vida dos universitários com TDAH, principalmente em atividades acadêmicas que envolvem leitura pelo fato de demandarem mais controle EC. Isso justifica a crescente demanda de atendimentos psicopedagógicos em estudantes universitários com o referido transtorno.

### 3.3. Versão Computadorizada do Reading Span Test do português brasileiro (VCRST-PB)

O VCRST-PB mensura a CMT dos participantes (Vasconcelos *et al.* 2019). Nossa hipótese foi que a CMT dos participantes com TDAH seria inferior em comparação com os que não possuem o referido transtorno. Tal investigação se fez necessária para poder compreender se a CMT dos sujeitos com TDAH interfere na compreensão leitora. O

<sup>3</sup> Testes que avaliam o foco atencional do participante ao realizar tarefas simples automatizadas. Por exemplo: Associar rapidamente os números com símbolos e nomeação rápida de estímulos visuais.

VCRST-PB foi desenvolvido nos programas Paradigm Experiments (*Perception Research Systems Incorporated*, EUA).

Adaptou-se uma nova versão do RST computadorizada e padronizada de Van den Noort *et al.* (2008) para o português brasileiro. Essa VCRST-PB possui adaptações quanto ao controle dos materiais e à aplicação do teste, tais como: controle do tamanho das sentenças por meio do número de palavras, sílabas e letras; controle da frequência das palavras-alvo no final das frases; controle de concretude das palavras-alvo; controle de plausibilidade das sentenças utilizadas; apresentação aleatória dos conjuntos de 2 a 6 frases e pontuação calculada a partir do total de palavras recordadas no experimento.

A VCRST-PB é constituída por 100 frases distribuídas em 5 séries de 20 frases. Em cada série, foram construídos conjuntos de 2, 3, 4, 5 ou 6 frases. Em cada série as frases foram controladas quanto ao número de: palavras, sílabas e letras. As palavras-alvo de cada frase foram controladas quanto ao número de sílabas, letras, vizinhos ortográficos e frequência das palavras-alvo. Visando reduzir possíveis interferências e estratégias de memorização das palavras-alvo, as séries foram organizadas em ordem pseudorandomizada, evitando-se a sequência de palavras iniciadas com a mesma letra e a sequência de frases longas e curtas no mesmo conjunto.

### 3.3.1. Procedimentos

Cada frase foi apresentada individualmente no centro da tela do computador durante oito segundos ou até que o participante apertasse a tecla ESPAÇO. Os conjuntos foram apresentados aos participantes de forma aleatória em cada uma das cinco séries do experimento.

Os participantes iniciavam o experimento com uma tela de instruções informando que deveriam ler em voz alta as frases apresentadas o mais rapidamente possível, memorizar a última palavra de cada frase e, ao final de cada conjunto de frases, quando solicitado, recordar e dizer em voz alta as palavras memorizadas. Em seguida, o participante realizou um treino com três conjuntos de frases (2, 3 e 4 frases) com a finalidade de se habituar à tarefa. Finalmente, o participante realizou o teste completo contendo as 100 frases experimentais.

### 3.3.2. Resultados e discussão

Os resultados da VCRST-PB foram analisados de acordo com os critérios de medidas utilizados por Van den Noort *et al.* (2008). Estes pesquisadores fizeram a soma de todas as palavras lembradas no RST completo. Segundo os mesmos, esses critérios possibilitam uma melhor observação do desempenho dos participantes do RST.

Comparou-se o *Span* de memória do grupo experimental com o do grupo controle. Os resultados obtidos são mostrados no Gráfico 5.

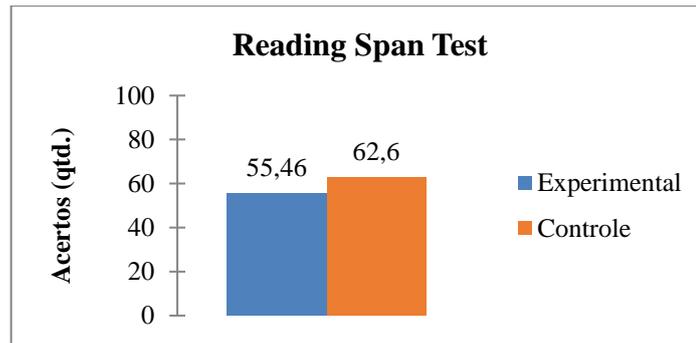


Gráfico 5. Percentual de acertos do VCRST-PB entre os grupos.

Conforme o gráfico 5 mostra, o grupo experimental acertou 55,46 palavras apresentadas no VCRST-PB, já o grupo controle acertou 62,6 palavras. Pode-se perceber que o span de memória dos participantes com TDAH foi menor que o span dos participantes que não têm o transtorno (teste de *Mann-Whitney*,  $U = 71,5$ , valor crítico de  $U$  em  $p < 0,05$  é 72).

Sabe-se que a MT é fundamental para que ocorra a compreensão leitora, pois ela serve como mediadora das informações adquiridas pelo *input* visual, mantendo-as ativas até que haja a compreensão do texto lido. Dias e Mecca (2019) corroboram essa afirmação ao propor que quanto maior a CMT, melhor desempenho da competência leitora o sujeito pode ter. Conforme o resultado obtido no VCRST-PB, os participantes com TDAH apresentaram mais dificuldades para memorizar as palavras apresentadas nesse experimento, confirmando assim a hipótese que assumimos de que o comprometimento no processo de leitura em estudantes universitários com TDAH é também proveniente da baixa CMT.

### 3.4. Discussão geral

Para a realização de uma análise estatística geral, padronizaram-se os dados dos três experimentos porque as respectivas escalas não eram equivalentes.

Tabela 2. Modelo linear múltiplo da discussão geral (Cloze, TAL, VCRST-PB)

Preditor	P-valor
Grupo	0,994
VCRST-PB	0,011
TAL	0,049

Na Tabela 2, percebe-se que não houve efeito de grupo ( $p$ -valor = 0,994). Nesta modelagem, usou-se a atenção, a medida na TAL e a CMT como preditoras da capacidade de leitura do resultado do teste Cloze. A nossa hipótese é que quando ocorre alguma mudança da CMT e de atenção linguística, a compreensão leitora é afetada. Como não houve diferença nos dois grupos, sugerimos que ambos foram igualmente afetados por mudanças no VCRST-PB e na TAL.

No VCRST-PB houve efeito significativo, confirmando que a CMT é uma variável que afeta a capacidade leitora. O mesmo se deu na TAL, sugerindo que um efeito significativo de maior demanda de atenção do EC foi um fator que também afetou a capacidade de leitura medida pelo teste Cloze. Esses resultados também confirmam as

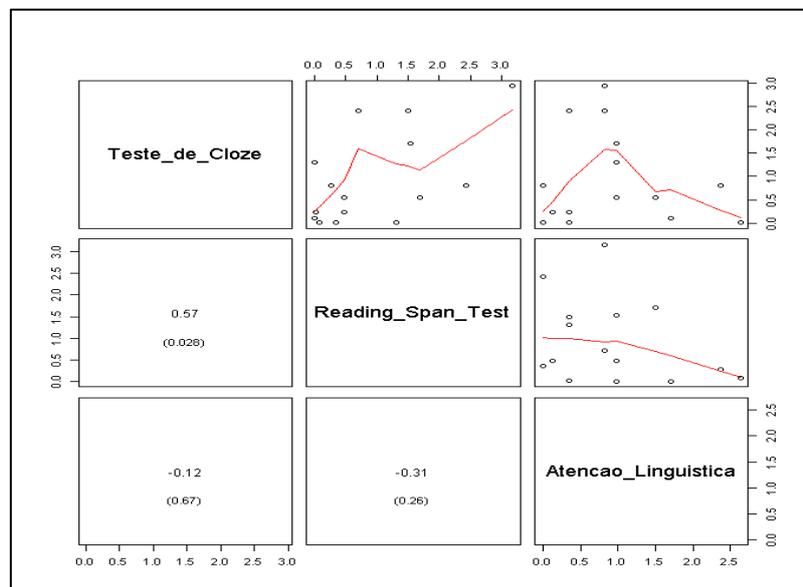
hipóteses assumidas neste estudo. Os resultados dos coeficientes são mostrados na Tabela 3:

**Tabela 3. Resultado dos coeficientes (Cloze, TAL, VCRST-PB)**

Preditor	Estimativa	P-valor
Grupo	-0,002	0,993
VCRST-PB	0,390	0,006
TAL	0,239	0,049

Como mostra a Tabela 3, o grupo experimental e o grupo controle não diferem significativamente (Estimativa = -0,002, p-valor = 0,993). Esse resultado nos remete a estudos que indicam que o déficit observado em sujeitos com TDAH não é de ordem cognitiva e sim decorrente de uma maior duração no desempenho das tarefas. Os indivíduos com esse transtorno conseguem obter os mesmos resultados que os indivíduos sem TDAH, porém, necessitam de um tempo maior para ter os mesmos resultados em comparação aos que não possuem o referido transtorno, principalmente em atividades que envolvem leitura (Albuquerque *et al.* 2012; Barkley 1998).

No VCRST-PB, o coeficiente foi 0,39 indicando que o aumento da CMT influencia a capacidade leitora, fazendo com que esta também seja aumentada. O mesmo ocorreu para a atenção medida na TAL, seu coeficiente foi 0,23, revelando que quando a atenção linguística aumenta, a capacidade de compreensão leitora aumenta também. Os dados de correlação entre o grupo experimental e o grupo controle podem ser observados no Gráfico 6.



**Gráfico 6. Correlação das tarefas (Cloze, TAL, VCRST-PB)**

Percebe-se que houve uma correlação positiva entre o teste Cloze e o VCRST-PB para o grupo controle. Este resultado indica que para o grupo controle, quando aumenta a capacidade de memória, também é aumentada a capacidade leitora. Já para o grupo

experimental, a capacidade leitora teve uma correlação maior com a atenção. Quando aumentou a atenção, aumentou também o resultado de leitura medido pelo teste Cloze.

Este resultado faz sentido para os sujeitos com TDAH que têm mais falhas no EC, pois estudos revelam que quanto maior for o grau de déficit atencional do indivíduo, mais prejuízos este terá na compreensão leitora (Fergusson & Horwood 1992). Diferentemente dos participantes sem o referido transtorno que possuem uma capacidade de atenção desejada, comprovada pela BPA durante a seleção dos participantes para o grupo controle. Esses últimos precisaram ter mais capacidade de memória para conseguir ter um bom desempenho no teste Cloze.

#### 4. Conclusão

O presente estudo se propôs a analisar como se dá o processamento da leitura em estudantes universitários diagnosticados com TDAH. Para isso, buscou-se também conferir se os prejuízos na compreensão leitora são preponderantes do TDAH, investigar a influência da MT na compreensão de leitura dos participantes com TDAH e comparar a velocidade do desempenho da tarefa de decisão lexical entre os participantes do grupo experimental e do grupo controle.

Os resultados apresentados nos três experimentos aplicados revelaram que a CMT dos participantes com TDAH medida no VCRST-PB é reduzida em comparação ao grupo controle e esta exerceu influência na compreensão leitora dos mesmos. Contudo, a MT não foi um fator determinante no desempenho do processo de leitura e sim a atenção linguística verificada pela TAL. Isso pode ser justificado devido ao processo de leitura requerer uma maior demanda de sustentação da atenção para poder obter a compreensão do conteúdo lido, sendo mais custoso para os adultos com TDAH por possuírem uma ruptura do controle executivo.

Este resultado confirmou que os sujeitos com TDAH necessitam de mais tempo durante a realização de tarefas de leitura, pois demandam mais recursos do controle executivo. Destaca-se que esse tempo adicional necessário já é concedido no Exame Nacional do Ensino Médio e o professor universitário que detém o conhecimento dessas dificuldades terá a autonomia de conceder um tempo maior durante a realização das atividades, fortalecendo as políticas educacionais inclusivas no Ensino Superior.

Devido à natureza exploratória dos estudos reportados neste artigo, a continuidade da pesquisa faz-se necessária. Pretendemos investigar, por exemplo, se a compreensão leitora pode ser afetada de maneira diferenciada nos sujeitos predominantemente desatentos, predominantemente hiperativo-impulsivos e entre aqueles que manifestam esses sintomas em combinação. Para isso, pretendemos fazer uma análise correlacional entre a MT e o processamento linguístico com base no processamento da anáfora pronominal através da análise dos movimentos sacádicos captados por uma metodologia baseada em rastreamento ocular (*eye-tracker*).

## Referências

- Aarts, E., Holstein, M. V., Hoogman, M., Onnink, M., Kan, C., Franke, B., Buitelaar, J., & Cools, R. (2015). Reward modulation of cognitive function in adult attention-deficit/hyperactivity disorder. *Behavioural Pharmacology*, 26, 227–240. <https://doi.org/10.1097/FBP.0000000000000116>
- Albuquerque, G., Maia, M., Franca, A., Mattos, P., & Pastura, G. (2012). Processamento da linguagem no transtorno do déficit de atenção e hiperatividade (TDAH), *DELTA*, 28(2), 245–280. <https://doi.org/10.1590/S0102-44502012000200003>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-V)*. Arlington, VA: American Psychiatric Association.
- Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Org.), *The psychology of learning and motivation* (Vol. 8, pp. 47–91). London: Academic Press.
- Baddeley, A. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology Section A*, 49(1), 5–28. <https://doi.org/10.1080/713755608>
- Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends Cognitive Sciences*, 4(11), 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- Barkley, R. A. (1997). *ADHD and the nature of self-control*. New York: Guilford Press.
- Barkley, R. A. (2002). *Transtorno do déficit de atenção/hiperatividade – TDAH: guia completo para pais, professores e profissionais da saúde*. Porto Alegre: Artmed.
- Barkley, R. A. (2006). *Transtorno de Déficit de Atenção e Hiperatividade – manual para diagnóstico e tratamento* (3ª ed.). Porto Alegre: Jones & Bartlett.
- Bormuth, J. R. (1968). Cloze Test readability: Criterion references scores. *Journal of Educational Measurement*, 5, 189–196. <https://doi.org/10.1111/j.1745-3984.1968.tb00625.x>
- Capovilla, A. G. S., Joly, M. C. R. A., Ferracini, F., Caparrotti, N. B., Carvalho, M. R., & Raad, A. J. (2004). Estratégias de leitura e desempenho em escrita no início da alfabetização. *Psicologia Escolar e Educacional*, 8(2), 189–197. Disponível em [http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1413-85572004000200007&lng=pt&tlng=pt](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1413-85572004000200007&lng=pt&tlng=pt)
- Crenitte, P. A. P. et al. (2010). Aspectos linguísticos no transtorno de déficit de atenção e hiperatividade. In S. M. Ciasca, S. D. Rodrigues & C. A. Salgado (Orgs.), *TDAH: Transtorno de Déficit de Atenção e Hiperatividade* (pp. 93–109). Rio de Janeiro: Revinter.
- Cypel, S. (2007). *Déficit de Atenção e Hiperatividade e as funções executivas: Atualização para pais, professores e profissionais da saúde* (3ª ed.). São Paulo: Lemos Editorial.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal Behavior*, 19(4), 450–466. [https://doi.org/10.1016/S0022-5371\(80\)90312-6](https://doi.org/10.1016/S0022-5371(80)90312-6)
- De Jong, C. G., Van De Voorde, S., Roeyers, H., Raymaekers, R., Oosterlaan, J., & Sergeant, J. A. (2009). How distinctive are ADHD and RD? Results of a double dissociation study. *Journal of abnormal child psychology*, 37(7), 1007–1017. <https://doi.org/10.1007/s10802-009-9328-y>
- Dias, N. M., & Mecca, T. P. (2019). *Avaliação neuropsicológica cognitiva: memória de trabalho* (Vol. 4). São Paulo: Memnon.
- DuPaul, G. J., & Stoner, G. (2007). *TDAH nas escolas* (1ª ed.). São Paulo: M. Books do Brasil Editora Ltda.
- Ellis, A., & Young, A. W. (1988). *Human cognitive neuropsychology* (6ª ed.). Porto Alegre: Artes Médicas.
- Estivalet, G. L., & Meunier, F. (2015). The Brazilian Portuguese lexicon: An instrument for psycholinguistic research. *PLoS ONE*, 10(12), e0144016. <https://doi.org/10.1371/journal.pone.0144016>
- Estivalet, G. L., & Meunier, F. (2017). Corpus psicolinguístico Léxico do Português Brasileiro. *Revista SOLETRAS*, 33(1), 212–229. <https://doi.org/10.12957/soletras.2017.29702>

- Fergusson, D. M. & Horwood, L. J. (1992). Attention Deficit and Reading Achievement. *The Journal of Child Psychology and Psychiatry*, 3(2), 375–385. <https://doi.org/10.1111/j.1469-7610.1992.tb00873.x>
- Filipek, P. A., Semrud-Clikeman, M., Steingard R. J., Renshaw, P. F., Kennedy, D. N., & Biederman, J. (1997). Volumetric MRI analysis comparing subjects having attention-deficit hyperactivity disorder with normal controls. *Neurology*, 48(3), 589–601. <https://doi.org/10.1212/WNL.48.3.589>
- Giraldello, A. P. (2016). Aspectos psicolinguísticos da leitura. *Unoesc & Ciência - ACHS*, 7(1), 23–30.
- Hynd G. W., Hern K. L., Novey E. S., Elipulos, D., Marshall, R., & Gonzalez, J. J. (1993). Attention deficit disorder and asymmetry of the caudate nucleus. *Journal of Child Neurology*, 8(4), 339–347. <https://doi.org/10.1177/088307389300800409>
- Klein, A. I., & Lamprecht, R. R. (2012). A compreensão em leitura e a consciência fonológica em crianças com Transtorno de Déficit de Atenção/Hiperatividade. *Signo*, 37(63), 25–54. <https://doi.org/10.17058/signo.v37i63.2969>
- Lima, R. F., Tambaquim, M. de L. M., & Ciasca, S. M. (2010). Sistema atencional e funções executivas na infância. In S. M. Ciasca, S. D. Rodrigues & C. A. Salgado (Orgs.), *TDAH: Transtorno de Déficit de Atenção e Hiperatividade*. Rio de Janeiro: Revinter.
- Mourão-Júnior, C. A., & Melo, L. B. R. (2011). Integração de três conceitos: Função executiva, memória de trabalho e aprendizado. *Psicologia: Teoria e Pesquisa*, 27(3), 309–314. <https://doi.org/10.1590/S0102-37722011000300006>
- Neto, M. R. L., et al. (2010). *TDAH ao longo da vida*. Porto Alegre: Artmed Editora.
- Pasquali, L. (2019). *Os processos cognitivos* (1.ª ed.). São Paulo: Vetor Editora.
- Perfetti, C. A., & Verhoeven, L. (2017). *Read across languages and writing systems*. Cambridge: Cambridge University Press.
- Rotta, N. T., Ohlweiler, L., & Riesgo, R. S. (2006). *Transtornos da aprendizagem: Abordagem neurobiológica e multidisciplinar*. Porto Alegre: Artmed.
- Samuelsson, S., Lundberg, I., & Herkner, B. (2004). ADHD and reading disability in male adults: Is there a connection? *Journal of Learning Disabilities*, 37(2), 155–168. <https://doi.org/10.1177/00222194040370020601>
- Santos, A. A. A., Prime, R., Taxa, F. O. S., & Vendramini, C. M. M. (2002). O teste Cloze na avaliação da compreensão em leitura. *Psicologia: Reflexão e Crítica*, 15(3), 549–560. <https://doi.org/10.1590/S0102-79722002000300009>
- Smith, E. E., & Geva, A. (2000). Verbal working memory and its connections to language processing. In Y. Grodzinsk, L.P., Shapin & D. Winney (Eds.), *Language and the brain* (pp. 123–141). San Diego: Academic Press.
- Taylor, W. L. (1953). Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, 30(4), 415–433. <https://doi.org/10.1177/00222194040370020601>
- Van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, 38(4), 584–589. <https://doi.org/10.3758/BF03193889>
- Vasconcelos, L. F., Almeida, P. A., Estivalet, G. L., & Ferrari-Neto, J. (2019). Teste de memória de trabalho de leitura: Versão computadorizada do reading span test para o português brasileiro. In C. A. Prolo & L. H. M Oliveira (Eds.), *Proceedings STIL19* (pp. 303–311). Disponível em <http://comissoes.sbc.org.br/ce-pln/stil2019/proceedings-stil-2019-Final-Publicacao.pdf>
- Van den Noort, M., Bosch, P., Haverkort, M., & Hugdahl, K. (2008). A standard computerized version of the Reading Span Test in different languages. *European Journal of Psychological Assessment*, 24(1), 35–42. <https://doi.org/10.1027/1015-5759.24.1.35>
- Veríssimo, L. F. (1995). Desentendimento. *Ícaro – Revista de Bordo da Varig*, 136, 11.
- Whipple, B. D., & Nelson, J. M. (2016). Naming speed of adolescents and young adults with attention deficit hyperactivity disorder: Differences in alphanumeric versus color/object naming. *Archives of Clinical Neuropsychology*, 31, 66–78. <https://doi.org/10.1093/arclin/acv061>
- Yeari, M., Avramovich, A., & Schiff, R. (2016). Online inferential and textual processing by adolescents with attention-deficit/hyperactivity disorder during reading comprehension:

evidence from a probing method. *Journal of Clinical and Experimental Neuropsychology*, 39(5), 1–17. <https://doi.org/10.1080/13803395.2016.1236906>

Zanella, L. W., & Valentini, N. C. (2016). Como funciona a memória de trabalho? Influências na aprendizagem e crianças com desordem coordenativa desenvolvimental. *Medicina (Ribeirão Preto)*, 49(2), 160–174. <https://doi.org/10.11606/issn.2176-7262.v49i2p160-174>

[recebido em 30 de outubro de 2021 e aceite para publicação em 23 de março de 2022]

**ROEPER, TOM; MAIA, MARCUS; PILATI, ELOISA**  
***EXPERIMENTANDO LINGUÍSTICA NA ESCOLA:***  
***CONHECIMENTO GRAMATICAL, LEITURA E ESCRITA***

São Paulo: Pontes, 2020.

Heloisa Maria M. Lima Salles\*  
hsalles@unb.br

A obra *Experimentando Linguística na Escola: Conhecimento Gramatical, Leitura e Escrita*, por sua abordagem e pelos temas desenvolvidos, chega ao público de especialistas, linguistas, neurocientistas, educadores, estudantes, com o objetivo de redirecionar as discussões sobre o ensino de língua (materna) e sobre a educação em geral. Nos três capítulos, intitulados “Como podemos enriquecer a experiência e a compreensão da linguagem nas escolas”, de Tom Roeper, “Contribuições das ciências cognitivas e do conhecimento gramatical para o ensino da leitura”, de Marcus Maia, e “Contribuições do conhecimento gramatical para o processo da escrita”, de Eloisa Pilati, estão reunidas propostas educacionais inovadoras. Na discussão, os autores tomam como fundamento teórico a concepção da biolinguística, originalmente formulada por Noam Chomsky, segundo a qual o conhecimento linguístico é um atributo mental inato da espécie humana, referido como língua interna (Língua-I) ou Gramática Universal (GU), a ser distinguido da chamada língua externa (língua-E), que corresponde aos enunciados produzidos nas situações de uso.<sup>1</sup>

Ao apresentar resultados experimentais e propostas de aplicação do conhecimento gerado no âmbito da linguística a problemas educacionais específicos, os estudos atendem a uma demanda dos profissionais da educação por metodologias que propiciem decisões fundamentadas, em um momento histórico que desafia a sociedade brasileira, em particular, alcançando também o cenário internacional, em que se sobressai a necessidade de superação de graves problemas sociais, particularmente a pobreza, em grande parte decorrente da falta de acesso à educação de qualidade. A adoção do ponto de vista da biolinguística permite avançar muitas questões colocadas para o ensino de língua(s), como o entendimento de que existe um conhecimento linguístico prévio, a ser explicitado e explorado, em benefício do desenvolvimento acadêmico do estudante. Definir e qualificar o papel do aluno e do professor na realização das atividades são tarefas

---

\* Universidade de Brasília, Brasília, Brasil. ORCID: 0000-0002-4190-1804

<sup>1</sup> A abordagem da biolinguística está desenvolvida em várias obras de Noam Chomsky, entre as quais destaca-se Chomsky (1995), que lança as bases do Programa Minimalista de investigação linguística, na formulação do modelo da Gramática Universal (GU).

primordiais, discutidas nos três capítulos em que se estrutura a obra. É o que podemos apreender mediante a análise de cada um deles.

O Capítulo 1, intitulado “Como podemos enriquecer a experiência e a compreensão da linguagem nas escolas?”, de Tom Roeper, inicia-se com a discussão de quatro tópicos centrais para a caracterização da natureza do conhecimento linguístico, que se configuram como propriedades da Gramática Universal, nomeadamente, a recursividade, as regras de longa distância, a elipse e a quantificação, sob o argumento de que a explicitação desses processos pode enriquecer a compreensão da linguagem humana, assim como a experiência linguística, na sala de aula. O autor observa que, em exemplos como *Ela é muito grande*, ou *muito, muito, muito grande*, a criança pode ser confrontada com a recursividade, ao ser indagada quanto ao número de vezes que poderia repetir a palavra *muito*, o que a levaria a descobrir a noção de infinito, ao concluir que “esse tipo de construção poderia continuar para sempre”, exatamente como os números naturais.

Operações que alteram a ordem dos termos podem ser igualmente usadas, como em *Ontem sanduíche comi, depois futebol joguei*. Segundo Roeper, atividades como essa oferecem a oportunidade para mencionar línguas que adotam o padrão em que o objeto antecede o verbo (referida como ordem Objeto-Verbo, ou OV), como algumas línguas indígenas faladas no Brasil, o que permite examinar a estrutura oracional e “estimular o respeito tanto por essas línguas quanto pelas pessoas que as falam” (p. 24). Além da explicitação dessas propriedades como recurso metodológico, a discussão de Tom Roeper faz referência ao conceito de criatividade linguística pelo ponto de vista do fundamento ético que subjaz ao caráter inato dessa capacidade. Com esses procedimentos, a criança seria levada a perceber que, “além de ser uma pessoa única, é criativa de maneira única” (p. 25).

Nesse sentido, o entendimento da linguagem como um fenômeno inato, pelo qual o ser humano expressa seu pensamento em enunciados sempre novos, por meio de recursos finitos e sistemáticos, oferece uma perspectiva inovadora para superar a metodologia baseada no ensino de formas ditas ‘corretas’, pela possibilidade de explorar fenômenos centrais da gramática das línguas naturais, com exemplos simples, que permitem identificar padrões, tornando explícito o conhecimento intuitivo da estrutura linguística. Além disso, segundo o autor, o estudo das línguas em perspectiva comparada é um meio de “lutar contra o preconceito linguístico no nível mais profundo” (p. 59), ao demonstrar-se que todas as línguas permitem ao falante expressar sua criatividade individual, devendo, portanto, ser não só respeitadas, mas também reconhecidas como legítimas.

O Capítulo 2, intitulado “Contribuições das ciências cognitivas e do conhecimento gramatical para o ensino da leitura”, de Marcus Maia, propõe uma abordagem do tema da leitura como um fenômeno interdisciplinar, tendo por referência os resultados das ciências que investigam a representação e o processamento da linguagem. Segundo o autor, esses estudos podem trazer uma contribuição mais efetiva “para um dos maiores desafios da educação no Brasil e no mundo – formar leitores competentes e cidadãos com pensamento crítico” (p. 64).

Fazendo uma retrospectiva de pesquisas recentes sobre a leitura, Maia destaca os estudos do rastreamento ocular, que se baseiam na constatação de que durante a leitura,

em vez de deslizar suavemente sobre as linhas do texto, o nosso olhar se move constantemente, alternando entre fixações e “saltos”. A análise dos parâmetros desta movimentação permite estudar a leitura e outros processos visuais. (pp. 71–72). A técnica inclui o uso dos chamados mapas de calor, “que indicam cromaticamente como o olhar é distribuído na visualização de um estímulo”, e os mapas de movimentação e fixação ocular, “que indicam a localização, a ordem, o número e o tempo das fixações em um estímulo” (p. 74). Segundo Maia, essa técnica permite estabelecer padrões de leitura proficiente, com implicações para o desenvolvimento de atividades metacognitivas a serem aplicadas no contexto educacional, pelo desenvolvimento da consciência do educando quanto à eficiência do processo, pela análise dos próprios mapas.

Análises produzidas no Laboratório de Psicolinguística Experimental da Universidade Federal do Rio de Janeiro (LAPEX/UFRJ) voltadas para a percepção de anomalias morfosintáticas e semânticas ilustram a metodologia, com a conclusão de que “a percepção de anomalias gramaticais é um processo de computação mais reflexo do que a análise de anomalias semânticas, que não impacta, em um primeiro momento, a leitura da frase, pois as regressões só ocorrem ao final de sua leitura completa” (pp. 77, 78).

Na confirmação da hipótese de trabalho, a técnica do rastreamento ocular demonstra a relevância da obtenção de dados *on-line*, por oferecer meios de avaliar os processos e subprocessos cognitivos da leitura não acessíveis às análises *off-line*. Tendo em vista a complexidade das ações envolvidas para se alcançar o objetivo final da leitura, que é o de compreender a essência do texto que se lê (o chamado *gist* da leitura), a exposição prossegue com a discriminação das principais etapas da cognição da leitura, “desde os microprocessos que iniciam com a decodificação das letras até os macroprocessos que levam à extração do *gist*” (p. 93). Nesse contexto, é examinado o processo de parseamento (ou processamento) de frases em sequências do tipo *garden-path*, como em *Enquanto Maria costurava as meias a roupa secava ao sol* e *Enquanto Maria costurava as meias secavam ao sol*, demonstrando-se que a primeira recebe menos fixações e fixações menores do que a segunda, o que se explica pela necessidade de desfazer, nesta última, a leitura do verbo ‘costurar’ como transitivo, ativada pelo sintagma contíguo “as meias”, a ser lido como sujeito da oração seguinte. Esse resultado confirma “a ideia de que a fase de parseamento estrutural é automática e reflexa e precederia os processos interpretativos que levam à extração do *gist*” (p. 100), em que atuam, de forma integrada, a decodificação, as análises morfológicas e sintáticas, agregando-se então semântica, pragmática, experiências prévias, conhecimento do mundo, para a construção dos sentidos. Essa discussão dá ensejo à apreciação do **período** como uma unidade fundamental de análise, o que, segundo o autor, tem sido negligenciado no ensino de línguas nas escolas, diante da primazia do texto (cf. Maia 2019).<sup>2</sup>

---

<sup>2</sup> Um parecerista anônimo observa que os rótulos ‘frase’, ‘sentença’, ‘período’ parecem designar o mesmo objeto sintático, cabendo esclarecimento sobre sua ocorrência no texto. De fato, são conceitos intercambiáveis em alguns casos, embora haja fatores que os distingam, a depender dos objetivos da análise. Tradicionalmente, o conceito de ‘frase’ refere-se a uma unidade mínima de sentido, que se institui no nível discursivo. No nível sintático, ‘frase’ pode ser sinônimo de ‘sintagma’, correspondendo, portanto, aos grupos sintáticos ou constituintes da sentença. O termo ‘sentença’ (também referido como ‘oração’ ou ‘cláusula’) corresponde a uma unidade sintática que inclui sujeito e predicado. Quanto ao termo ‘período’,

O capítulo 3, intitulado *Contribuições do conhecimento gramatical para o processo da escrita: a proposta da gramaticoteca*, de Eloisa Pilati, parte da observação de que “a depender do tipo de texto e da complexidade do tema, a escrita é tarefa trabalhosa e desafiadora” (p. 122). A autora observa que o vasto conhecimento da língua falada por uma criança de três anos, sem que tenha recebido instrução formal, o que se explica pela Faculdade de Linguagem, contrasta com a leitura e a escrita, que não se desenvolvem em todas as comunidades humanas, manifestando-se como artefatos. No entanto, existem pontos em comum, particularmente pela compatibilidade entre a fala e a escrita, referidas como “domínios diferentes do mesmo objeto”, como propõe Lobato (2003/2010). Com esse pressuposto, Pilati argumenta que, “no nível da geração de estruturas, tanto na fala quanto na escrita, partimos da mesma matéria prima mental, a gramática” (p. 124).

Antes de se deter nas questões teórico-metodológicas, a autora dedica-se a três aspectos externos relacionados: (1) a situação educacional brasileira pelo ponto de vista das avaliações nacionais e internacionais, (2) a dicotomia texto-gramática como métodos opostos (3) as competências de escrita e de gramática na Base Nacional Curricular Comum do Ensino Médio (BNCC). Na análise dos resultados em avaliações, em que se constata que “menos da metade dos jovens brasileiros não desenvolve habilidades mínimas necessárias em áreas centrais: leitura, escrita, conhecimento científico e matemático” (p. 129), a autora destaca ser necessário considerar as condições de desigualdade que, historicamente, afetam a sociedade brasileira, e por conseguinte as crianças, como a pobreza, o abandono, a violência, o estresse, fatores que interferem negativamente no processo de aprendizagem. Em relação aos métodos, alerta para os problemas causados pela primazia das práticas orientadas para o texto, em detrimento da reflexão gramatical, destacando ser necessário trilhar o caminho do equilíbrio e da complementaridade. Em relação à BNCC, ressalta que as competências sugeridas, baseadas em atividades cognitivas de baixo nível de complexidade como ‘lembrar, compreender e aplicar’, “não orientam o sistema educacional brasileiro sobre as possibilidades de o estudante usar seus conhecimentos gramaticais de forma crítica e criativa” (p. 137).

Dessa forma, a autora propõe uma mudança de caráter epistemológico e metodológico em que os estudantes sejam apresentados ao processo de geração de sentenças, nos moldes da GU, pelo emprego de materiais concretos, a constituir a chamada *Gramaticoteca*. Uma consequência dessa abordagem é a mudança na concepção de conceitos gramaticais em relação ao ensino tradicional, passando-se a focar a centralidade do verbo e sua estrutura argumental, assim como a noção de sintagma, cuja explicitação é necessária, embora constituam conhecimento intuitivo. Essa metodologia promove o conhecimento aprofundado do fenômeno, pelo ponto de vista da manipulação criativa das estruturas – e não da memorização de rótulos e conceitos teóricos (cf. Pilati 2017).

---

refere-se a um enunciado delimitado pelo ponto final, o que o vincula ao texto escrito, abarcando ainda o conceito de ‘sentença’, por ser qualificado como simples, se constituído de uma única sentença, ou composto, se constituído de duas ou mais sentenças. Para uma discussão desses termos, em perspectiva teórica, confira-se Perini (1996), entre muitos outros.

Além da base conceitual e teórica e da fundamentação experimental e prática, brevemente apresentadas nesta exposição, cada capítulo inclui uma seção dedicada a sugestões de atividades didáticas, o que constitui um canal oportuno de comunicação com o público leitor, interessado nas aplicações do conhecimento científico. Essa abordagem confirma ainda que os estudos formais das línguas naturais propiciam o desenvolvimento de propostas educacionais consistentes e consequentes.

Ao percorrer brevemente o roteiro desta obra, verificamos que os temas desafiadores são aqueles que não se esgotam. Em matéria de educação linguística, muitos caminhos podem ser trilhados no desenvolvimento do potencial crítico e criativo dos educandos. Nesse processo, cabe reconhecer que a educação formal nas sociedades complexas é indispensável ao projeto de nação. Para tanto, é essencial a presença da escola pública, e seu corolário, a educação como aspiração de todos, como ressalta o notável educador brasileiro Anísio Teixeira (1977, p. 17), para quem “o ensino se tem de fazer pelo trabalho e pela ação, e não somente pela palavra e pela contemplação”. Entre tantos atores, o professor tem a responsabilidade de liderar a criação do projeto a ser vivenciado pelo estudante na sala de aula. Nesse sentido, o planejamento é uma etapa fundamental, devendo nele constar os saberes tradicionais, as bases científicas e artísticas, os fundamentos éticos e afetivos que sustentam a busca do conhecimento, em consonância com os objetivos educacionais estabelecidos. Esta obra é um exemplo de compromisso com esses requisitos. Que seja bem acolhida e amplamente ressignificada!

## Referências

- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Lobato, L. (2010). O que o professor de ensino básico deve saber sobre gramática? *Linguística e Ensino de Língua*. Brasília: Editora da UnB. (original publicado em 2003)
- Maia, M. (2019). Rastreamento de períodos compostos e consciência sintática. In M. Maia (Org.), *Psicolinguística, metacognição na escola* (1.<sup>a</sup> edição). Campinas, São Paulo: Editora Mercado de Letras.
- Perini, M. (1996). *Gramática Descritiva do Português* (2.<sup>a</sup> edição). São Paulo: Ática.
- Pilati, E. (2017). *Linguística, gramática e aprendizagem ativa*. Campinas: Pontes Editores.
- Teixeira, A. (1977). *Educação não é privilégio*. São Paulo: Companhia Editora Brasileira.

[recebido em 16 de novembro de 2021 e aceite para publicação em 7 de abril de 2022

**MARTÍN AIZPURU, LEYRE**

***LA ESCRITURA CANCELLERESCA DE FERNANDO III, ALFONSO X, SANCHO IV Y FERNANDO IV.***

***ESTUDIO PALEOGRÁFICO Y GRÁFICO-FONÉTICO DE LA DOCUMENTACIÓN REAL DE 1230 A 1312.***

Peter Lang, 2020, 364 pp.

Pedro Dono López\*  
pdono@elach.uminho.pt

A presente publicación deriva da tese de doutoramento da profesora Leyre Martín Aizpuru, investigadora integrante do Grupo de Estudio de Documentos Históricos y Textos Antiguos (<http://campus.usal.es/~gedhytas/>), da Universidad de Salamanca, integrado na Red CHARTA (<https://www.redcharta.es/>), e supón un importante fito nunha produción científica xa extensa e consistente no ámbito da edición e estudo grafemático e lingüístico do castelán notarial medieval. A obra constitúe unha aproximación aos perfís da escrita na chancelaría real desde a incorporación do romance ata o final do reinado de Fernando IV, partindo dun extenso corpus de 516 diplomas. O abano cronolóxico, de 1230 a 1312, abrangue os reinados en Castela e León dos monarcas aludidos no título, un amplo segmento temporal que fala ben da ambición e alcance da publicación.

Para alén da preceptiva introdución (“Parte introductoria”, pp. 31–38), onde, ademais de presentar o contido dos diferentes capítulos da obra, se define e contextualiza o obxecto de estudo relativamente á literatura e se informa dos obxectivos e metodoloxía da investigación, o volume está organizado en seis capítulos, distribuídos en dúas “partes”: unha primeira, designada “Marco teórico”, na que se atende, nos dous capítulos que a integran, por un lado, ao concepto de “norma lingüística” do castelán medieval e, por outro, á caracterización da chancelaría real no período en foco. A segunda parte, designada “Presentación del corpus y estudio” (capítulos 3 a 6), está dedicada ao estudo paleográfico e grafofonético dos documentos, precedido dunha presentación do corpus que tamén se caracteriza dun punto de vista diplomático e prosopográfico (isto é, no relativo aos funcionarios que interveñen na elaboración dos documentos).

No primeiro capítulo, titulado “El concepto histórico de norma lingüística medieval” (pp. 41–53), aborda o proceso de incorporación do romance á escrita da chancelaría, que se ve impulsado a partir da unión dos reinos de León e Castela (1230)

---

\* Departamento de Estudos Românicos, Escola de Letras, Artes e Ciências Humanas, Universidade do Minho (Braga, Portugal). ORCID: 0000-0003-0780-6982.

no reinado de Fernando III (1217–1252), durante o cal o castelán se vai afianzando progresivamente nos diplomas da corte, proceso que culmina no reinado de Afonso X. A autora non podía deixar de revisar o concepto de “norma lingüística” castelá asociado á figura de Afonso X, asunto que xa conta cunha extensa e intensa bibliografía e foi obxecto de múltiples abordaxes. Destácase neste punto que non convén proxectar unha visión actual de conceptos como norma ou ortografía sobre o pasado (pp. 52–53), entendendo que a *variatio*, a existencia de varias tradicións de escrita, é propia da escrita “normativa” da chancelaría do Rei Sabio.

En “La cancellería real castellana (1230–1312)” (pp. 55–85), que constitúe o capítulo segundo da obra, aténdese á súa orixe, organización, funcionamento e evolución no período estudado, ben como ás diferentes fases do proceso de produción dos diplomas (xénese documental) e ao funcionariado da chancelaría (chanceleres, notarios e escribáns) e os seus cometidos, para o que se basea en Kleine (esp. 2015), cuxos traballos son de referencia ao abordar esta dimensión, principalmente durante o reinado de Afonso X. No caso dos escribáns, é de sinalar a relevancia atribuída ao factor da especialización (ou non) en determinadas tipoloxías documentais (máis ou menos solemnes) na abordaxe dos redactores que actuaban na chancelaría, de acordo coa investigación da xa referida profesora Marina Kleine. Pecha o capítulo cunha aproximación á caracterización e tipoloxía dos diplomas expedidos pola oficina real, clasificados fundamentalmente polo seu grao de solemnidade (pp. 80–85).

No capítulo 3, “Corpus de documentación de la cancellería real castellana” (pp. 89–104), informa sobre o proceso de constitución do corpus, discutindo o concepto central, neste ámbito, de “representatividade”, e o tratamento editorial aplicado aos documentos. Do extenso corpus chanceleresco constituído e editado no ámbito do aludido grupo GEDHYTAS, o estudo selecciona, como xa foi dito, máis de quíntos diplomas, para cuxa análise, no plano gráfico, botou man da ferramenta Lyneal<sup>1</sup> do profesor Hiroto Ueda, que permite a procura e extracción de datos estatísticos sobre textos en español. Xa na epígrafe titulado “Caracterización diplomática y prosopográfica del corpus” (pp. 105–137) inclúe unha análise diplomática ou pragmática do corpus, no sentido de establecer e exemplificar as diferentes partes ou constituíntes dos documentos. Estúdanse, asemade, os diferentes intervinientes na preparación dos diplomas, comezando por analizar a presenza ou ausencia da subscripción e as variadas formas que pode adoptar, para a seguir, tendo moi en conta a obra referencial de Kleine (2015), realizar unha abordaxe prosopográfica da chancelaría nos diferentes reinados, na que distingue os *iussores* do amplo elenco de *redactores* que interveñen na escrita das cartas.

Os dous capítulos seguintes constitúen a cerna da investigación de Leyre Martín: a análise do corpus nos planos paleográfico e grafonético. A riqueza e pormenor de “Descripción paleográfica: la escritura y otros signos gráficos” (pp. 139–228) non nos permiten atender, por mor da brevidade, a cada unha das dimensións abordadas, polo que aludiremos unicamente a algúns aspectos de relevo dentro deste apartado,

---

<sup>1</sup> <http://shimoda.llf.uam.es/ueda/lyneal/>

comezando por citar as importantes cuestións de investigación ás que pretende dar resposta a autora:

por un lado, qué implicacións tiene a escritura individual de cada uno dos redactores identificados –siendo conscientes de que non sempre se pode asegurar a execución por parte do individuo mencionado na suscripción canceleresca– y, por outro, cómo afecta a crecente, aunque moderada, cursivización de las últimas décadas do século XIII e principios do XIV a usos estritamente gráficos, tengan ou non reflexión fonética. (p. 143)

Unha das achegas de alcance á que nos referiamos máis arriba ten que ver coa minuciosa análise do *tipo de letra* empregado nos diplomas. Leyre Martín clasifica a maior parte dos documentos do seu corpus en tres grupos (“formada”, “usual”, “corriente”) segundo o carácter máis ou menos caligráfico da gótica documental empregada nos manuscritos, designada como fracturada. Ademais de presentar os trazos caracterizadores de cada un destes tipos de letra, deixa constancia da problemática asociada á devandita taxonomía, pois non está exenta de dificultades (debido, principalmente, ás múltiples modalidades de transición recoñecibles) na adscrición dos diplomas a unha ou outra das escritas documentais definidas. A análise paleográfica do corpus lévaa a concluir que

apenas hay mostramos de letra corriente (11 [diplomas]) y que una mayoría destacada de los documentos del corpus están redactados con la letra gótica fracturada usual (269), seguida de un buen número de diplomas redactados con el tipo formado (190). [...] Es significativo el cambio de tendencia en la década de 1270: hasta entonces, predomina el empleo del tipo formado y a partir de esa fecha decae el número de documentos que lo emplean a favor de la letra gótica fracturada usual. (p. 158)

Constata tamén que o tipo de letra depende en boa medida da tipoloxía diplomática, ou dito noutros termos, da solemnidade do documento, aínda que se rexistran particularidades das que se deixa constancia na abordaxe da presenza das diferentes góticas fracturadas en cada un dos tipos textuais do corpus (pp. 162–168). Trata igualmente a relación entre clases de letra e redactores (pp. 168–171).

Este capítulo inclúe aínda un estudo en profundidade da braquigrafía, cuxa análise cuantitativa permite á autora chegar a conclusións interesantes relativamente ao recurso a procedementos abreviativos, advertindo, durante o período, “una clara progresión ascendente, especialmente, a partir de la década de 1270” (p. 185). Os seus datos tamén veñen confirmar unha menor incidencia da abreviación nos diplomas máis solemnes e nos escritos en letra gótica fracturada formada. Realiza, igualmente, un estudo moi pormenorizado das diferentes marcas e procedementos de abreviación usados nos diplomas do seu corpus, ben como dos segmentos elididos e respectiva posición na cadea gráfica. A alografía, pola súa parte, protagoniza o último apartado do estudo paleográfico dos diplomas reais, centrado nos usos das variantes *i/j* e *u/v* con valor vocálico ou consonántico en posición inicial, interior e final de palabra. Constata nas ocorrencias de *i/j* a incidencia do contexto paleográfico (a seguir ás chamadas “letras de palo”), sendo moito máis difusas as tendencias de uso de *u/v*.

O capítulo 6 (“Estudio gráfico-fonético”, pp. 229–337), pola súa parte, está organizado en tres bloques que procuraremos presentar con brevidade. No primeiro deles, “Algunas cuestiones previas de grafemática” (pp. 229–290), aborda cuestións como a representación de /i/ e /ts/ ou a situación dos manuscritos canto ao uso de *h*, de grafías de referencia nasal (m,n) ante bilabial (p,b) ou de grafías simples e dobres. Para varios dos trazos analizados, a autora chega a definir a década de 1270 como punto de viraxe nos usos gráficos. No segundo, “Vocalismo. Apócope” (pp. 291–316), o título apunta un asunto que se desdobra na abordaxe dos dous fenómenos coñecidos como apócope normal e extrema. No plano cuantitativo, constata unha presenza constante da apócope normal no período, aínda que se intensifica no segmento 1270–1300; xa para a apócope extrema, concéntranse os exemplos nos anos de 1240 a 50, mingando ata unha práctica desaparición desde 1290. Tal como para os diferentes fenómenos gráficos e paleográficos estudados, aquí tamén se contempla a variable do redactor, advertindo que determinados documentos e autores concentran un volume importante de ocorrencias. Na terceira e última parte do capítulo, dedicada ao “Consonantismo” (pp. 316–337), achega o contributo do seu corpus a tres aspectos de interese na configuración grafofonética do castelán medieval (sistema de sibilantes, uso das grafías *b/v* e os resultados de F- latino). Da análise da situación dos diplomas en causa, só aludiremos a unha interesante constatación relativa á representación das fricativas apicoalveolares e africadas dentoalveolares: nas “confusións” (permítasenos o termo por mor da concisión) de grafías cobra relevancia o factor do redactor, pois o fenómeno evidénciase para ambas as parellas de sibilantes en determinados escribáns (p. 326).

O traballo péchase coa “Parte conclusiva” (pp. 341–344), na que, indo ao encontro do obxectivo principal da investigación desenvolvida, que non é outro que o de achegar os perfís da escrita da chancelaría real castelá no período estudado, conclúe que existen dúas normas diferenciadas empregadas nos reinados de Fernando III e Afonso X, por unha parte, e Sancho IV e Fernando IV, pola outra. Pon de relevo a autora, igualmente, as súas achegas no que ten que ver coa incidencia de factores como a cronoloxía, reinado, redactor, tipoloxía diplomática ou de letra sobre a escrita da época, recoñecendo o importante papel na súa investigación de Kleine (2015) e, no plano cuantitativo, da ferramenta Lyneal.

Despois deste rápido repaso polo contido da obra de Leyre Martín, a primeira constatación que cabería realizar é que constitúe un magnífico exemplo da contribución que pode realizar ao ámbito da lingüística histórica o recurso a corpora dixitais: ao tempo que permite confirmar certos perfís da escrita da época apuntados en estudos precedentes, achega informacións valiosas ao analizar os datos recollidos de acordo con variables como a tipoloxía diplomática, o tipo de letra ou os redactores. Son virtudes destacadas do texto, na nosa opinión, unha exposición clara, precisa e sintética das diferentes dimensións exploradas á volta da documentación chanceleresca, o que proporciona unha visión de conxunto moi completa do obxecto de estudo e constituirá unha investigación de referencia para todo aquel que pretenda aproximarse aos instrumentos medievais en castelán (e non só) da Chancelaría Real (e non só). Neste punto destaca a presentación dos resultados obtidos en cada vertente da investigación, contextualizados e contrastados coa bibliografía e investigacións relevantes.

[recebido em 01 de março de 2021 e aceite para publicação em 18 de fevereiro de 2022]

## PAISAGEM LINGUÍSTICA DA “ZONA VELHA” DA CIDADE DO FUNCHAL: CONSIDERAÇÕES HISTÓRICAS E SOCIOLINGUÍSTICAS

### LINGUISTIC LANDSCAPE OF THE “ZONA VELHA” OF THE CITY OF FUNCHAL: HISTORICAL AND SOCIOLINGUISTIC CONSIDERATIONS

Bruna Pereira \*

brunamicaela.freitas@uvic.cat

A Paisagem Linguística (PL) tem sido uma temática bastante estudada nos anos recentes, no entanto muitas cidades e países europeus ainda não usufruíram de estudos relevantes no seu ambiente sociocultural. Esta investigação pretende analisar o espectro linguístico da chamada “Zona Velha” do Funchal, a capital da Madeira e a primeira localidade da ilha a ser habitada. A ocupação britânica na Ilha da Madeira a partir do séc. XIX, bem como a forte afluência turística nas zonas baixas junto ao mar, tem tornado a cidade do Funchal um ponto de diversidade cultural e linguística. O impacto da língua inglesa manifesta-se especialmente através dos sinais linguísticos comerciais, da restauração e é também um reflexo da relação previamente estabelecida aquando da ocupação mencionada. Partindo deste princípio, este trabalho irá focar-se nos efeitos dessa realidade através do registo e análise de 216 sinais linguísticos presentes numa das zonas mais emblemáticas da cidade, a chamada “Zona Velha”, localizada na freguesia de Santa Maria Maior, apresentando as diferenças entre os sinais públicos e governamentais e os privados e industriais. A metodologia deste trabalho assenta na abordagem *top down* e *bottom up* de Ben-Rafael *et al.* (2006) em que a presença de sinais linguísticos *top down* demonstra a influência e ocupação britânica no plano sociopolítico, cultural e económico, e os sinais linguísticos *bottom up* apontam para o fenómeno turístico.

**Palavras-chave:** Paisagem Linguística. Zona Velha-Funchal. Ocupação britânica.

The Linguistic Landscape (LL) has been a theme extensively studied in recent years, however many European cities and countries have not enjoyed yet relevant studies in their socio-cultural environment. This investigation intends to analyze the linguistic spectrum of Funchal's “Zona Velha”, the capital of Madeira Island and the first town to be inhabited. The British occupation of Madeira Island from the century XIX, as well as the strong tourist influx in the low areas by the sea, has made the city of Funchal a point of cultural and linguistic diversity. The impact of the English language is manifested especially through commercial linguistic signs, catering and is also a reflection of the relationship previously established during the mentioned occupation. Based on this principle, this work will focus on the effects of that reality through the registration and analysis of 216 linguistic signs located in one of the most emblematic areas of the city, “Zona Velha” in the parish of Santa Maria Maior, presenting the differences between public and governmental signals and private and industrial signals. The methodology of this work is based

---

\* Departament de Traducció, Interpretació i Llengües Aplicades. Facultat d'Educació, Traducció, Esports i Psicologia. Universitat de Vic - Universitat Central de Catalunya. Barcelona, España. ORCID: 0000-0001-8568-6176

on the *top down* and *bottom up* approach of Ben-Rafael *et al.* (2006) in which the presence of *top down* linguistic signs demonstrates the British influence and occupation in the socio-political, cultural and economic spheres, and the *bottom up* linguistic signs demonstrate the tourist phenomenon.

**Keywords:** Linguistic Landscape. Zona Velha-Funchal. British Occupation.

•

## 1. Introdução

Falar de paisagem linguística (PL) envolve abordar o reflexo em sinais linguísticos de todo um conjunto complexo de percepções, fatores culturais e sociais: a forma como um individuo identifica a si próprio. A PL tornou-se um elemento tão característico das cidades que agora é possível falar em *cityscape*, uma vez que as cidades (*cities*) tornaram-se as protagonistas da grande maioria dos estudos de PL, devido à proliferação de grandes áreas comerciais onde abundam mensagens em múltiplas línguas. Até ao momento, não existe uma definição que seja completamente satisfatória para todos os campos que a PL passou a integrar, no entanto, abaixo abordaremos os principais conceitos e estudos que contribuíram para o progresso desta área do saber até os dias de hoje.

Gorter (2006) afirma que o substantivo “paisagem” tem dois significados. Em primeiro lugar, o seu significado mais literal, sendo uma extensão de território que a vista consegue alcançar e, em segundo lugar, uma imagem, desenho ou representação de uma paisagem natural, como por exemplo, uma imagem do mar. Por abranger diversas realidades culturais e envolver ciências como a sociologia, a psicologia e a geografia, o conceito de PL tem sido utilizado em sentidos distintos. Na literatura, o conceito tem sido empregue para descrever e analisar a situação linguística de uma determinada localidade, como no trabalho de Sciriha & Vassallo (2001) em que os autores tratam da realidade de Malta, uma minúscula ilha no coração do Mediterrâneo e uma ex-colónia da Grã-Bretanha até a sua independência em 1964. O conceito de PL tem sido também utilizado para tratar a presença de múltiplas línguas em grandes áreas geográficas, como é o caso do estudo de Kreslins (2003), em que o autor analisa as línguas presentes no Báltico e esboça críticas sobre trabalhos anteriores na mesma região.

A PL compreende muitos itens tais como os sinais de trânsito, postes publicitários, nomes de ruas e de lugares, sinais comerciais e ainda os sinais públicos em edifícios governamentais. Gorter (2006, p. 2) precisa: “the use of language in its written form in the public sphere. It refers to language that is visible in a specified area”. Ben-Rafael, Shohamy, Amara & Trumper-Hecht (2006, p. 7) acrescentam que a PL pode também compreender “any sign or announcement located outside or inside a public institution or a private business in a given geographical location”. Os investigadores mencionam ainda “linguistic objects that mark the public space” (Ben-Rafael *et al.* 2006, p. 14). Daí que estudar a PL seja um verdadeiro desafio: analisar a partir de localidades multilingues muito diversas entre si, espaços bilingues ou monolingues, zonas com minorias

linguísticas históricas ou mesmo lugares onde o tempo parece não ter passado e permanecem iguais por décadas.

As línguas rodeiam-nos em qualquer lugar e de muitas maneiras: estão presentes nas ruas, nos nomes dos edifícios públicos, restaurantes, roupas, lojas, cafés, placas de trânsito, grafites, sinais publicitários e cartazes. O espaço interdisciplinar da PL envolve um amplo leque de disciplinas científicas tais como a sociologia, literatura, antropologia, geografia e política (Carr, 2019, p. 1) e procura entender as razões, mensagens e usos da linguagem no espaço público, focando-se nas representações da linguagem no ambiente social, sendo que o seu objeto de estudo é qualquer sinal linguístico exposto e o modo como as pessoas interagem com esses sinais. O motivo desta interdisciplinaridade está relacionado com a “linguagem escrita manifestada no ambiente físico-social e que revela a identidade linguístico-cultural de uma população, suas experiências de contato linguístico-cultural, os valores e as ideologias dos indivíduos (...)” (Teis, Seide & Lucas 2018, p. 18).

Em geral, a principal tarefa da pesquisa relativa à paisagem linguística é refletir a situação sociolinguística da área. Para tal efeito, pretende-se descrever a diversidade linguística e cultural, refletir o grau de uso de determinadas línguas e ligar estes fenômenos à política da linguagem e a um senso de identidade linguística. (Bielenin-Lenczowska, 2021, p. 5276)

Na atualidade, todos os cidadãos vivem num mundo em constante mudança, tendo que enfrentar novas realidades e adaptar-se a elas. Aquando do surgimento desta área científica, não existiam painéis interativos digitais nas ruas ou sinalização digital, entre outros elementos que hoje em dia constituem o *corpus* de análise. Um dos contributos que a investigação em PL pode fornecer está precisamente relacionado com o seu principal aspeto de análise: a visibilidade. A PL é interpretada, por muitos, como um reflexo da língua na sociedade; por vezes, direto, indireto ou deturpado. De facto, Gorter, Marten & Mensel (2012, p. 11) compara esta situação a um espelho, mas um espelho de circo (distorcido): “It would be a mistake to conceive of the LL as a mirror of the language relationships in a city, region or country, because at most its distortions can seem like a carnival mirror”.

Apesar da PL poder ser um reflexo da sociedade na qual existe, não há uma correlação direta entre as duas. Questões relacionadas com o passado histórico social e ideológico podem afetar a perceção que um indivíduo tem da sua língua e do estatuto e poder que ela adquire numa dada comunidade. Gorter *et al.* (2012, p. 3) afirma que “the Linguistic Landscape can inform us about and exhibit some of the underlying ideas, ideologies, conflicts and power struggles between different stakeholders.”

Este trabalho pretende analisar o espectro linguístico da Zona Velha do Funchal, mais especificamente na freguesia de Santa Maria Maior. Parte da história da cidade está relacionada com a ocupação e exploração comercial britânica. Com o passar dos anos, além dos habitantes ingleses que tornaram a Madeira um ponto de referência para a residência e turismo britânico, a Ilha assistiu também a uma grande afluência de outros povos. Desta forma, esta investigação pretende analisar qual o impacto da ocupação britânica na PL da Zona Velha e de que forma esta se manifesta. Assim, partindo do

princípio de que a língua portuguesa é a língua oficial na área geográfica em estudo, as hipóteses da investigação são as seguintes:

- I. A Zona Velha, por ser uma zona muito turística, rica em espaços comerciais e de restauração, será abundante em mensagens multilingues com muitas ofertas de produtos e serviços aos turistas.
- II. A língua inglesa também irá manifestar-se no espaço público em contextos distintos dos comerciais.

E tentará responder às seguintes perguntas:

Quais são as línguas que surgem no panorama linguístico da cidade do Funchal e qual a sua importância? Quais são as características dos sinais bilíngues ou multilingues? Em que tipo de sinais linguísticos surge a língua inglesa?

## 2. Paisagem Linguística

O estudo da PL ganhou destaque especial nos últimos anos e tem-se assistido a um cada vez maior número de publicações na área. Viver em sociedades cada vez mais abertas ao diálogo entre culturas e que colocam as línguas em contacto, reflete a importância da PL enquanto mecanismo de aproximação ou afastamento de grupos etnolinguísticos. Mas o que é a PL e por que razão ou razões é o seu estudo tão importante? Uma das investigações a dar bases de análise aos seguintes surgiu em 1997, quando Rodrigue Landry e Richard Bourhis definiram o conceito de PL na publicação “Linguistic Landscape and Ethnolinguistic Vitality An Empirical Study”:

The language of public road signs, advertising billboards, street names, place names, commercial shop signs, and public signs on government buildings combines to form the linguistic landscape of a given territory, region or urban agglomeration. The linguistic landscape of a territory can serve two basic functions: an informational function and a symbolic function. (Landry & Bourhis, 1997, p. 23)

A partir dessa definição, é possível observar duas funções básicas da PL: a função informativa e a função simbólica. Segundo os autores, a função informativa serve como um marcador distintivo de uma determinada área geográfica ocupada por um grupo humano concreto com uma linguagem comum. Aqui, eles enfatizam que o uso de uma linguagem única e consistente é um fator que delimita as fronteiras com outros grupos linguísticos que passam por lá. “Thus, the linguistic landscape serves to inform in-group and out-group members of the linguistic characteristics, territorial limits, and language boundaries of the region they have entered” (Landry & Bourhis, 1997, p. 25). Uma vez que os sinais linguísticos podem ser monolíngues, bilíngues ou multilingues, eles também podem fornecer informações sociais sobre os grupos que vivem lá, ou se uma língua aparece mais predominantemente em relação às outras, essa corresponderá à língua dominante.

A prevalência de marcadores linguísticos da comunidade dominante pode simbolizar a força ou vitalidade de sua linguagem perante outras comunidades linguísticas. “Public signs in the in group language can also symbolize the vitality of the

ethnolinguistic group in other institutional support domains such as cultural production and commercial and religious activities” (Landry & Bourhis, 1997, p. 28).

Cenoz & Gorter (2006, p. 68) chamam à atenção sobre a relevância do “contexto sociolinguístico”:

On the other hand, the linguistic landscape contributes to the construction of the sociolinguistic context because people process the visual information that comes to them, and the language in which signs are written can certainly influence their perception of the status of the different languages and even affect their own linguistic behaviour. The linguistic landscape or parts of the linguistic landscape can have an influence on language use.

O estudo da PL tem por objetivo refletir sobre as políticas linguísticas e sociais adotadas por uma determinada região e qual ou quais as suas consequências na vida quotidiana dos cidadãos. A utilização de uma determinada língua por instituições governamentais é diferente da utilização dessa mesma língua por entidades privadas. Quando utilizada por canais governamentais, a língua adquire um estatuto diferente de quando não o é utilizado, ou seja, adquire mais valor perante o falante pois este reconhece que a sua língua pertence a toda ou a uma parte considerável da comunidade e não é exclusiva de um grupo minoritário. Como afirmam Cotterill, Muir, Joinson & Dewdney (2015, p. 2) “the effect of hierarchy and power on linguistic choices has always been of interest to linguists and sociologists”.

The configuration of languages present in the linguistic landscape therefore can provide important information about the diglossic nature of a particular bilingual or multilingual setting. Thus, before communicating interpersonally with a single inhabitant, one can use the linguistic landscape as an indicator of the power and status relationship that exists between the various language groups present within a given administrative or geographical region (Landry & Bourhis, 1997, p. 26).

Ben-Rafael *et al.* (2006, p. 7) também abordam “power relations that eventually exist behind choices of patterns where sociopolitical forces share relevant incompatible interests”. Neste estudo, os investigadores afirmam que os estudos das relações de poder têm a capacidade de desvendar as realidades sociais de uma determinada comunidade. A ordem em que as línguas surgem, quais as línguas que surgem e em que contexto podem ajudar a compreender qual o poder e estatuto de uma língua sobre outra ou várias. O facto de uma determinada língua surgir na porta de um determinado edifício, por exemplo, governamental, mostra que essa língua é aceite nos meios governativos e políticos.

Levando em consideração as abordagens teóricas e metodológicas apresentadas e o contexto do presente estudo, entendemos a PL como qualquer letra, palavra, frase ou símbolo presente no espaço público tais como nomes de restaurantes, placas de informação, horários de funcionamento, cartazes promocionais e menus que têm por objetivo transmitir uma informação ao leitor e cuja interpretação simboliza algo tanto ao emissor como ao recetor. A escrita de um menu exclusivamente em português não transmite devidamente a informação que os não falantes da língua deveriam receber e poderá também ser um indicador de que estes terão problemas de comunicação se realmente optarem por entrar naquele estabelecimento.

Os primeiros estudos em PL centravam-se na descrição das várias línguas que surgiam no espaço público, descurando uma abordagem mais diacrónica e histórica:

Se os primeiros estudos se centraram na descrição e análise da presença das diferentes línguas em determinados espaços públicos, geralmente urbanos, muitas vezes numa perspetiva quantitativa e sincrónica que incluía o inventário das línguas presentes, os investigadores têm chamado à atenção para a necessidade de se ir além de uma abordagem centrada na contagem de línguas num determinado momento, para abraçar perspetivas mais dinâmicas das evoluções das PL no tempo e no espaço, considerando, desta feita, uma perspetiva diacrónica e histórica (Melo-Pfeifer & Lima-Hernandes 2020, p. 1028).

O estudo da PL do Funchal, analisando as relações atuais entre a utilização do inglês e do português no espaço público é pertinente uma vez que a Ilha da Madeira já foi ocupada política e economicamente pelos britânicos por duas vezes aquando do século XIX. Como destaca Blommaert (2010, p. 4), “we now see that the mobility of people also involves the mobility of linguistic and sociolinguistic resources”.

### 3. Estudo de caso

O estudo da PL tem vindo a crescer consideravelmente nos últimos anos, no entanto, e devido a ser uma área de estudo relativamente recente existem muitas cidades europeias e no Mundo que ainda não beneficiaram de estudos sociolinguísticos no seu meio. Em Portugal, existem poucas teses de mestrado e de doutoramento que contemplem este campo do saber e as que foram realizadas baseiam-se, essencialmente, na influência do turismo ou são realizadas em contextos educativos como é o caso do estudo de Clemente (2018). Uma das razões para esta lacuna linguística está relacionada com a própria génese social de Portugal, uma vez que a língua portuguesa é, sem dúvida, a língua de eleição para os portugueses e é exclusivamente a língua oficial. Por outro lado, não descuidando a grande variedade de línguas presentes em pequenas comunidades de emigrantes por Portugal Continental e ilhas, estas não são objeto de estudo deste trabalho.

No que diz respeito à Ilha da Madeira, esta tem sido também alvo de diversas afluências de emigrantes ao longo dos anos, com especial destaque para os chineses, ucranianos, russos, moldavos e, mais recentemente, por venezuelanos que têm também um passado histórico muitas vezes associado à emigração de pais, avós e outros familiares e veem a Madeira como uma forma de escapar à repressão política e social que a Venezuela vive.

O estudo da ocupação britânica e o seu reflexo na PL do Funchal surge da necessidade de contribuir para a expansão desta área do conhecimento no Mundo e na Madeira de um tema que é tão comum e presente na mente dos madeirenses, mas que nunca foi devidamente explorado. No que diz respeito a trabalhos desenvolvidos na Ilha, é possível destacar o trabalho de Macedo (2019) onde a investigadora analisou a PL de espaços públicos e comerciais nas cidades do Funchal, Santa Cruz e Ribeira Brava, Macedo (2020) em que a autora compara os nomes de restaurantes da Zona Velha do Funchal com os de Santa Cruz de Tenerife e ainda os trabalhos de Rebelo (2021a) onde se observa a toponímia madeirense como PL e património linguístico e outro onde

apresenta uma revisão bibliográfica dos principais trabalhos na área da PL (Rebelo 2021b).

Desta forma, o estudo da influência britânica na PL do Funchal é uma necessidade linguística, histórica e cultural, visto que irá apresentar informações concretas acerca do estado atual da PL e providenciar dados pertinentes para os gestores da cidade e dos serviços governamentais.

A presença dos britânicos na Ilha da Madeira, especialmente na cidade do Funchal, e na Zona Velha é um marco histórico característico da Região. De uma forma ou outra, e ainda que distanciado no tempo, todos os residentes identificam-se como associados aos ingleses. A presença destes foi muito importante na medida em que definiu e estabeleceu diversos aspetos culturais e comerciais que imperam até os dias de hoje. De seguida, iremos contextualizar a situação social que caracterizava a Madeira desde o período da ocupação britânica até os dias de hoje, que foi fundamental na criação de laços entre ingleses e madeirenses e foi responsável pelo grande desenvolvimento económico da Ilha.

Até a década de sessenta do século XX, quase todos os funchalenses são reféns desta situação. Em “Dia de S. Vapor, a cidade altera-se, completamente. Há múltiplos serviços e produtos a oferecer aos visitantes e a cidade rejubila e ganha um movimento e comércio inusitados”. Há uma economia familiar, alimentada pelos bomboteiros. Há, ainda, o contrabando que estes protagonizam, que vai dando o necessário para a sobrevivência do núcleo familiar. Há, por parte dos passageiros em trânsito, uma atenção especial às frutas da ilha, que são transportadas em cestas para bordo ou vendidas pelos bomboteiros<sup>1</sup>. Morangos, nêspers, laranjas e bananas, captam a sua atenção e são o desenhoo para as ementas de bordo. (Vieira 2018, p. 11)

Ainda acerca do ambiente e espírito vivenciado pelo povo madeirense, Raúl Brandão, apresenta a sua visão:

Agora conheço melhor a Madeira passado o primeiro entusiasmo, vejo tudo a frio. Esta ilha é um cenário e pouco mais - cenário deslumbrante com pretensões a vida sem realidade e desprezo absoluto por tudo que lhe não cheira a inglês. Letreiros em inglês, tabuletas em inglês e tudo preparado e maquinado para inglês ver e abrir a bolsa. (Brandão 1927, p. 55)

Vivia-se um clima de emoção, agitação e esperança de que os ingleses trouxessem um futuro melhor, dinheiro, prosperidade e acabassem com a fome. Na referência acima, Brandão (1927) explora a questão da dependência e servência madeirense. Quando os barcos ingleses chegavam, as tabuletas, ou melhor, os sinais linguísticos eram escritos em inglês e dirigidos aos visitantes, pois eram eles os clientes que iam gastar dinheiro. Por outro lado, os grandes proprietários e comerciantes da Ilha eram ingleses e a PL refletia essa realidade. Graças a esta estreita relação comercial, viram-se reunidas as condições para a emigração madeirense:

---

<sup>1</sup> Bamboteiros era o nome que designava os comerciantes do Bambote (barco de pequenas dimensões que levava iguarias regionais aos barcos ingleses que aportavam no Funchal). A palavra “bambote” é um neologismo derivado da aglutinação das palavras inglesas “bum boat” que os madeirenses ouviam. Bazenga, Ribeiro & Sequeira (2012, p. 48).

Na ilha da Madeira existia uma colónia de mercadores ingleses e a ínsula era um importante interposto comercial britânico, nas rotas atlânticas das embarcações para as Índias Ocidentais. Na sequência da crise de mão-de-obra nas Antilhas Britânicas, com os proprietários a procurarem colonos para substituírem os antigos escravos, emancipados em 1834, os madeirenses eram vistos como bons candidatos, descritos como excelentes trabalhadores e artífices competentes. (Teixeira 2009, p. 54)

O Funchal assumiu um papel crucial no abastecimento à navegação, com pontos de abastecimento de água, alimentos, vinho e carvão. Segundo Biddle (1896, p. 23) o Funchal era “uma importante estação de abastecimento de carvão para a maior parte das linhas dos navios de Inglaterra e do continente europeu para a África do Sul”. A ocupação e influência britânica refletiram-se também na construção de importantes instituições, tais como centros hospitalares, companhias de transporte e de eletricidade. Neste trabalho, o autor destaca as principais famílias que se deslocaram desde Inglaterra atraídas pelas condições climáticas propícias tanto ao usufruto pessoal das mesmas (clima ameno mais agradável que Inglaterra) como ideais para a prática da agricultura e plantação de cana de açúcar, uvas (que viriam a ser fundamentais na produção de vinho, especialmente vinho Madeira), entre outras.

Um dos motivos pelos quais a Ilha da Madeira era tão desejada está relacionado com a sua posição geográfica estratégica e a sua antiga tradição comercial. Os acontecimentos que antecedem a primeira ocupação britânica estão relacionados com a longa e profunda inimizade e rivalidade entre Inglaterra e França. Por essa altura, Napoleão Bonaparte lançava várias investidas à metrópole e ao Império britânico e preparava-se para investir contra a Madeira. Nesse momento, Londres teve conhecimento dessa intenção e enviou forças para protegerem a Ilha.

Decidida a ocupação, o processo que conduziu as tropas britânicas à Ilha da Madeira desenvolveu-se com relativa rapidez, entre a última semana de Junho e a última de Julho de 1801, período em que foram remetidas todas as instruções e no fim do qual se procedeu ao desembarque, embora de início com a ressalva de que o destino da expedição, que foi organizada sob o comando do coronal William Henry Clinton, se deveria manter secreto. Como veremos, a rapidez com que tudo se verificou acabou por se repercutir na eficácia geral da acção e na surpresa das autoridades portuguesas, mas também, a curto prazo, por provocar alguns inconvenientes e dificuldades entre os britânicos. (Rodrigues 1999, p. 145)

A Ilha da Madeira acabava por não pertencer a si própria. Por um lado, vivia sob o comando britânico e por outro estava na mira do império francês que reconhecia a sua posição estratégica, solo com qualidade e bons produtos naturais.

A presença britânica em grande parte do espaço atlântico é esmagadora e o seu domínio naval nada tem a temer. No caso português esta presença centrava-se em quatro pontos: Lisboa, Madeira (Funchal), Porto e Açores (Angra e Ponta Delgada), onde os navios da Royal Navy eram uma constante e a intervenção britânica, directa ou indirectamente, sustentava grupos, partidos, correntes políticas, num ambiente em que o Estado central era fraco, estava ausente ou dava sinais evidentes de ineficácia e incapacidade. (Rodrigues 2007, p. 20).

Atualmente, a presença britânica continua a fazer parte da vida madeirense. Seja pelas linhagens que sucedem até hoje as grandes famílias inglesas do século passado ou seja pelo papel económico que desempenham decorrente das empresas que fundaram e se mantêm em atividade. Na genealogia dos madeirenses de uma classe social alta, os apelidos estrangeiros cruzam-se com os portugueses.

A comunidade inglesa representa, ainda hoje, um grande marco na economia regional. Durante décadas, o povo insular viveu de cabeça curvada. Não em relação ao Continente mas a tudo o que vinha de fora do país. Por alguma razão, na gíria diária, o turista é sempre inglês, mesmo que seja alemão. (Vxmag 2019)

O início da colonização na Ilha da Madeira teve lugar em Santa Maria Maior, mais especificamente, na Zona Velha. Esta freguesia, inicialmente com pouca concentração habitacional, contribuiu para o desenvolvimento da cidade do Funchal, através da evolução do comércio do trigo, açúcar e vinho, dinamizados também pela presença e comércio ingleses. Esta zona compreende diversos pontos de atração turística, tais como o Forte de Santiago, o Mercado dos Lavradores, a Igreja do Socorro, a Capela de Nossa Senhora da Oliveira, e diversos restaurantes, bares e locais de animação noturna. Para além de ser uma área residencial é dos maiores pontos de concentração histórica e turística da Ilha da Madeira.

De acordo com as informações disponibilizadas no site da Junta de freguesia de Santa Maria Maior (2020), após o início do povoamento da Madeira, o território dividiu-se em duas áreas. A primeira ia desde as ribeiras de Santa Luzia e João Gomes (também conhecida por ribeira de Santa Maria) até o Forte de São Tiago e a segunda zona, ia desde a ribeira de São João até ao Convento de Santa Clara:

O aglomerado populacional que se estendeu da Ribeira de João Gomes até o Corpo Santo, não foi mais de que um núcleo de população formado de maneira espontânea que foi crescendo ao longo da praia de calhau ali existente. Deste facto advém-lhe, muito provavelmente, o nome dessa mesma zona: Santa Maria do Calhau, nome dado também ao pequeno templo que no ano de 1430 começou a servir o povoado da zona leste da baía do Funchal, também conhecido por Nossa Senhora da Conceição de Baixo, ou simplesmente por Santa Maria Maior. (Junta de freguesia de Santa Maria Maior 2020).

Abaixo analisaremos as questões metodológicas relacionadas com este estudo.

#### **4. Metodologia de recolha e classificação dos dados**

No que diz respeito à estratégia metodológica, o primeiro passo foi a definição da unidade de análise. Este estudo empírico foi baseado na coleção fotográfica de todos os sinais linguísticos encontrados no espaço selecionado, neste caso, a Zona Velha do Funchal. A codificação ou categorização dessas amostras é uma das etapas mais importantes e, ao mesmo tempo, das mais difíceis deste processo, pois surgiram algumas questões relacionadas com decisões no plano metodológico. Gorter (2006) abordou também algumas dessas dúvidas aquando do seu trabalho de investigação: Onde tirar fotos e quantas? A amostra coletada é representativa de toda uma cidade, área ou ponto de pesquisa?

No passado, alguns investigadores assumiram, por exemplo, uma loja como um sinal linguístico, desconsiderando publicidades, papéis escritos à mão e colados na janela, grafites, entre outros. A falta de consistência por parte do investigador pode trazer vários problemas na interpretação de resultados e provocar conclusões erradas. Desta forma, para este estudo, a unidade de análise é o sinal linguístico individual. Isto significa que se numa montra de uma loja tiver o nome dessa loja, um cartaz promocional ou indicações de horário, estas corresponderão a distintos sinais linguísticos. A vantagem deste método é que os sinais serão facilmente contáveis e categorizados.

O método e os critérios utilizados por diferentes investigadores ao longo dos anos são variáveis, levando em consideração as cidades ou países sob análise e a variedade linguística de cada local. Huebner (2006) e Backhaus (2006), por exemplo, limitaram o seu estudo às zonas de Bangucoque (Tailândia) e Tóquio (Japão) e, novamente, limitaram essas mesmas cidades a áreas que consideravam representativas. Mais uma vez, após a seleção desses espaços, eles procederam à seleção das ruas. Huebner (2006) coletou amostras de 15 bairros urbanos e rurais e Backhaus (2006) analisou 28 estradas. Como Huebner (2006) destaca, a amostra não precisa necessariamente expor toda a composição linguística da cidade, mas deve apresentar pelo menos a variedade que a localidade selecionada exhibe. Aqui entra a importância da seleção prévia de locais pelo investigador. Deve estudar-se quais as zonas que podem ser mais representativas e que permitem uma melhor e mais fácil compreensão e comparação. Neste caso, as ruas escolhidas dentro da Zona Velha do Funchal não foram aleatórias, e correspondem às ruas mais frequentadas tanto por residentes como por turistas e que fornecerão uma maior amostra linguística.

Outra questão também variável e controversa está relacionada com o objeto do estudo em si. De acordo com a definição de Landry e Bourhis (1997), a PL pode ser entendida como qualquer objeto ou signo linguístico que marca o espaço público. Ainda assim, essa definição pode levar a interrogações de resposta difícil como o que poderá constituir a unidade de análise. Estará na mesma linha de comparação, um sinal de trânsito, um cartaz e um veículo com texto que está passando na rua? Neste sentido é importante desenvolver uma classificação que separe por categorias os sinais linguísticos registados e para análise. Backhaus (2006) considerou apenas os sinais multilingues, e dos 12000 sinais que registou, selecionou apenas cerca de 2400. Cenoz e Gorter (2006) escolheram uma rua com grandes áreas comerciais e Ben Rafael *et al.* (2006) conduziram o seu estudo em zonas religiosas, em localidades de atividade comercial e instituições públicas.

Tendo em conta a metodologia levada a cabo por estes investigadores, é possível concluir que a seleção das áreas de estudo está relacionada com o propósito do trabalho em si. As duas primeiras categorias presentes na maioria das investigações estão relacionadas com o conteúdo dos sinais: *top down* (de cima para baixo) e *bottom up* (de baixo para cima), que correspondem a sinais públicos ou governamentais e privados, respetivamente. Calvet (1994) referiu-se aos sinais governamentais e privados como componentes *in vitro* e *in vivo* da PL. Estes termos fazem uma distinção global entre o que é escrito pela autoridade (os nomes das ruas ou sinais de regras de trânsito pelo

governo) e aqueles que são escritos pelos cidadãos (a denominação de lojas, grafites, publicidade, entre outros).

Em estudos mais recentes, como por exemplo, em Clemente (2018) foi realizado um levantamento fotográfico documental da PL de Aveiro e realizou-se uma análise de conteúdo e de discurso. Foram selecionadas as áreas mais significativas de acordo com a autora e foram também realizadas entrevistas aos proprietários dos espaços. Em Berger e Lecheta (2019), ao trabalhar sobre a PL de um campus universitário fronteiriço no Brasil e eventuais relações de poder, as investigadoras procederam à recolha de fotografias dos muros e paredes dos espaços de circulação pública do campus e examinaram esses elementos de forma a determinar reflexos da necessidade de visibilidade de certos grupos. Os investigadores Da Silva, Santos e Jung (2016) captaram imagens fotográficas de distintos locais públicos da cidade de Foz do Iguaçu, analisando as funções informativas e simbólicas dos sinais que descobriram e sugeriram tratar-se de uma PL com influência dos processos de globalização. Em Rodrigues (2020) verifica-se a utilização de uma metodologia semelhante. A investigadora recolheu imagens de placas e letreiros de estabelecimentos comerciais localizados perto da fronteira entre o Brasil e a Colômbia, e apresentou um estudo de caso com base nos conceitos de PL de diversos autores.

Visto que este estudo pretende analisar a possível influência da ocupação britânica na Ilha da Madeira através da análise da PL na Zona Velha do Funchal, inserida na freguesia de Santa Maria Maior, a metodologia mais adequada encontrada foi a definida por Ben-Rafael *et al.* (2006, pp. 19–21) como mencionado anteriormente.

Na Tabela 1, apresentamos as categorias descritas acima:

**Tabela 1. Categorização *top down* e *bottom up* de acordo com a proposta de Ben-Rafael *et al.* (2006)**

<b>Categoria</b>	<b>Tipo de item</b>
<i>Top down</i>	Sinais institucionais, governamentais, municipais, culturais, educacionais e/ou outros
	Nomes de ruas
<i>Bottom up</i>	Sinais comerciais (vestuário, restauração, e/ou outras lojas)
	Sinais de negócios privados (empresas e/ou outros)
	Sinais de anúncios privados (vendas, aluguer de carros/casas e/ou outros)

Na Tabela 2, apresentamos o modelo utilizado para a categorização *top down* e *bottom up* por número de itens encontrados por língua:

**Tabela 2. Modelo usado na categorização *top down* e *bottom up* por língua (número de itens)**

<b>Categoria/Língua</b>	<b>Português</b>	<b>Inglês</b>	<b>Outras</b>
<i>Top down</i>			
<i>Bottom up</i>			

São alguns os constrangimentos em abraçar um projeto desta natureza, seja pela tensão social decorrente de se fotografar espaços públicos e privados perante a presença de proprietários, funcionários e clientes e o desconhecimento destes ao objetivo desse registro e possíveis consequências, seja pela dificuldade em categorizar sinais linguísticos com base em critérios do passado que não compreendem a realidade do estudo linguístico na análise paisagística atual. Por exemplo, o que pode ser considerado como um “sinal de interesse público”? Ou um “anúncio público”? Como se categorizam autocolantes de iniciativas ou candidaturas políticas coladas em paredes ou sinais de trânsito? (cf. Figura 1 direita). Se, por um lado, há sinais imediatamente identificáveis na categoria correspondente, como, um sinal religioso (cf. Figura 1 esquerda), outros carecem de maior consideração e análise.



**Figura 1. Exemplos de sinais linguísticos presentes na Zona Velha: “Igreja Evangélica” (esquerda) e “Vote Edgar Silva” (direita).**

Surgiu, assim, a necessidade de definir o que foi considerado para cada item das tabelas utilizadas. Na categoria *top down*, em instituições públicas (religiosas, municipais, culturais e educacionais) foram consideradas igrejas, escolas, lojas do cidadão, museus, postos de correio e estações policiais, entre outras entidades governamentais; em sinais de interesse público procedeu-se à análise de mapas presentes nas ruas ou identificadores de pontos interessantes a visitar ou próximos (cf. Figura 2); em anúncios públicos registaram-se os cartazes de votação eleitoral regional.

No que diz respeito à categoria *bottom up* (vestuário, restauração e/ou outras lojas), em sinais comerciais, analisou-se os sinais presentes em lojas de roupa e restaurantes; em sinais de negócios privados (empresas ou escritórios) considerou-se os sinais em cabeleireiros, empresas de advogados, empresas de passeios turísticos e outras lojas comerciais de venda de bens e produtos e sinais de anúncios privados (vendas, aluguer de carros/casas e/ou outros).



**Figura 2. Sinal *top down* (sinal de interesse público) - direções para pontos de interesse na “R. Brigadeiro Oudinot”.**

Foi também realizado um esquema de codificação baseado em Cenoz e Gorter (2006) e Gorter (2007), que desenvolveram uma grelha com uma categorização especial, distinguindo o número de línguas nos sinais, a frequência de cada língua no sinal, classificação *top down* e *bottom up*, a composição dos sinais, o tamanho das línguas nos sinais bi e multilingues e apresentaram um quadro com todas as línguas encontradas. Estes resultados serão discutidos na secção seguinte.

## 5. Apresentação e discussão de resultados

Foram registados 216 sinais linguísticos neste trabalho, sendo 28 *top down* e 188 *bottom up*. No que diz respeito aos sinais *top down* foram encontrados 6 sinais relativos a instituições públicas religiosas, municipais, culturais e educacionais, 14 sinais de interesse público, 8 anúncios públicos e 3 nomes de ruas. Destes, 24 estavam na língua portuguesa, 3 na língua inglesa e 1 na língua francesa. Já na categoria de *bottom up* estavam presentes 114 sinais de lojas: roupa, comida ou joalheria; 38 sinais de negócios privados: escritórios, fábricas, empresas e agências e 33 sinais de anúncios privados: “procura-se”, publicidade, vendas, arrendamentos de carros ou apartamentos. Destes, 120 sinais estavam na língua portuguesa, 61 na língua inglesa e 7 em outras línguas. Analisando os sinais *top down* e *bottom up*, verifica-se uma diferença acentuada decorrente da necessidade de informar os locais (madeirenses e portugueses) e os turistas. É possível verificar que o item em que mais elementos surgem é *bottom up*, ou seja, elementos não governamentais - maioritariamente publicitários e da restauração. Este facto pode ser explicado pela atração turística decorrente da oferta gastronómica e vida noturna característicos da Zona Velha.

Na tabela 3, é possível verificar o número de itens encontrados nas categorias *top down* e *bottom up*.

**Tabela 3. Categorização *top down* e *bottom up* por número de itens na Zona Velha.**

<i>Top down</i>	Instituições públicas: religiosas, municipais, culturais e educacionais.	6	28
	Sinais de interesse público.	14	
	Anúncios públicos	8	
	Sinais de nomes de ruas	3	
<i>Bottom up</i>	Sinais de lojas: roupa, comida ou joalheria.	114	188
	Sinais de negócios privados: escritórios, fábricas, empresas e agências.	38	
	Anúncios privados: “procura- se”, publicidade, vendas, arrendamentos de carros ou apartamentos.	33	
	Total	216	

Na Tabela 4 verificamos o número de elementos *top down* e *bottom up* por língua, onde foram identificados 24 itens em português, 3 em inglês e 1 em francês, referentes à categoria *top down*. A categoria *bottom up* contou com 120 elementos em português, 61 em inglês e os restantes 7 noutras línguas.

**Tabela 4. Categorização *top down* e *bottom up* por língua na Zona Velha.**

<i>Top down</i>	Português	24
	Inglês	3
	Francês	1
<i>Bottom up</i>	Português	120
	Inglês	61
	Outras	7
Total		216

Em relação aos sinais, foi possível recolher uma amostra abrangente. Verifica-se uma preocupação com a apresentação dos sinais comerciais de restauração que aparecem na língua inglesa ou outra língua estrangeira. Habitualmente são impressos e afixados com cuidado. Uma grande maioria dos restaurantes e bares apresenta inclusive fotografias dos pratos e das bebidas disponíveis no menu (*cf.* Figura 3). Quando surgem línguas estrangeiras, os espaços comerciais tendem a apresentar o português ou o inglês em primeiro lugar.



Figura 3. Sinais *bottom up* - sinais comerciais de restauração no Restaurante “Madeira Story Centre” (esquerda e centro) e no Restaurante “O Brinquinho” (direita).

Habitualmente, estas superfícies comerciais colocam as bandeiras dos países para indicar o menu e os preços com a língua respetiva dessa bandeira. Estes menus são apresentados em grandes placas metálicas ou de plástico com suporte em formato de quadro de grandes dimensões para atrair os turistas e para tentarem destacar-se dos concorrentes (*cf.* Figura 4 direita). Junto às ofertas gastronómicas, alguns dos espaços optam também por juntar as ofertas culturais, muito frequentemente em inglês (*cf.* Figura 4 esquerda).



**Figura 4.** Sinais *bottom up* – informação de espetáculo ao vivo em língua inglesa no Restaurante “Almirante” (esquerda) e menu com quatro línguas no Restaurante “Donna Maria” (direita).

No que diz respeito à apresentação dos menus em português, esta zona demonstra uma preocupação menor perante estes, apresentando-os em placas escritas a giz e cartão. Estes menus são também, na sua maioria, referentes a “pratos do dia”, refeições de almoço com tudo incluído (prato, bebida e café), mais baratas que o menu habitual ou *à la carte* (cf. Figura 5). Por esse motivo e pela constante mudança da oferta do dia ao cliente madeirense, é necessário alterar diariamente os cartões e as placas de ardósia e essa alteração não implica gastos desnecessários com impressões e apresentações atrativas.



Figura 5: Sinais *bottom up* - sinais comerciais de restauração no Restaurante “As Escadinhas” em placas de ardósia (esquerda) e em cartão (direita) na língua portuguesa.



Figura 6: Sinal *bottom up* relativo a um serviço de reprografia.

Na Figura 6 é possível verificar a afixação de um cartaz de tamanho considerável com informações em inglês e, ao lado, uma folha impressa colada na parede com as mesmas informações, mas desta vez em português. O tamanho reduzido e o tipo de

afixação provocam a percepção de que terá sido afixado posteriormente e sem tanto cuidado e preocupação, uma vez que é dirigido ao público madeirense.



Figura 7. Sinal *bottom up* em língua inglesa.

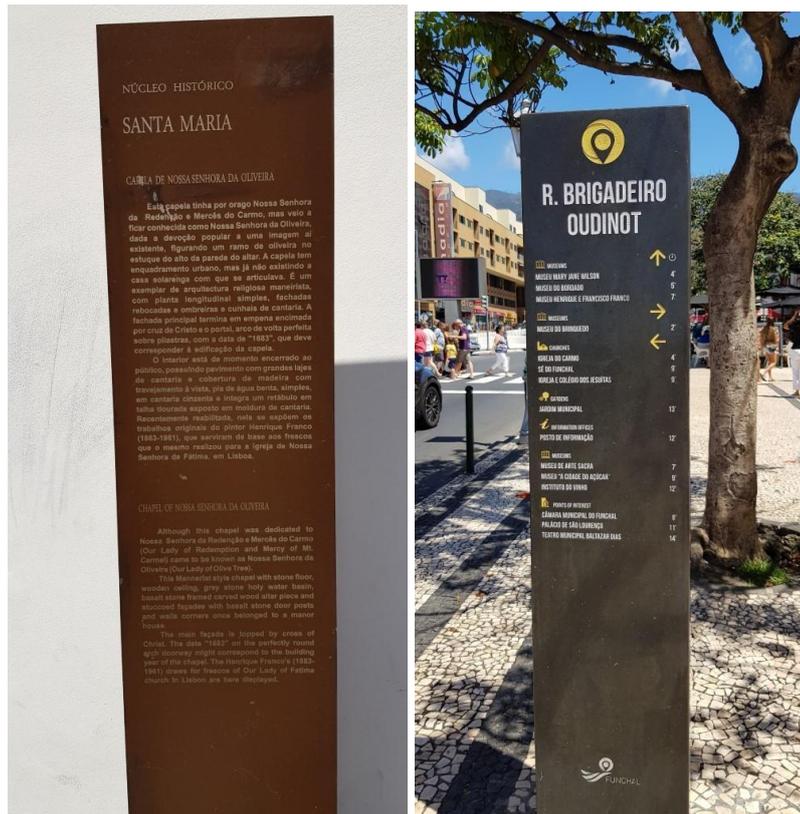
Na Figura 7 é possível observar um sinal *bottom up* referente à instituição privada Madeira Medical Center e que se apresenta em inglês, assumindo-se como uma instituição que está capacitada a dialogar e tratar dos seus pacientes através da comunicação em inglês e, provavelmente, é também dirigida aos visitantes estrangeiros que dela possam necessitar. Visto tratar-se de uma instituição privada e localizar-se na Zona Velha que uma área bastante turística, entendemos que este espaço é dirigido tendencialmente aos turistas que ali afluem. Um outro fator pertinente é o facto deste mesmo grupo empresarial ter aberto um outro hospital privado localizado nas zonas superiores do Funchal, mais residenciais e pouco turísticas, a que designou de “Hospital Particular da Madeira”.

Dos 216 sinais linguísticos registados, 28 deles foram identificados como pertencentes à categoria *top down* e apenas 3 se apresentaram na língua inglesa, sendo os restantes 24 sinais em português e 1 em francês, demonstrando a forte dominância do português perante estes. Este sinal linguístico em língua francesa trata-se da *Federation Internationalé Maisons de L'Europe*, uma organização politicamente independente que tem por objetivo consolidar o diálogo e parcerias entre instituições e cidadãos europeus. O surgimento desta língua numa instituição pública indica que os autores do sinal a reconhecem este local como uma grande zona turística.



**Figura 8. Sinal linguístico *top down* em língua francesa:  
*Federation Internationále Maisons de L'Europe.***

Quanto aos sinais ingleses referem-se a painéis de informação local com dados históricos e indicações, como é possível verificar nos exemplos abaixo.



**Figura 9. Sinais linguísticos *top down*: “Núcleo Histórico de Santa Maria” em português e inglês (esquerda) e “Ponto de informação geográfico” em língua inglesa (direita).**

Ao circular pela Zona Velha é fácil observar a existência de algumas placas informativas com dados históricos sobre a freguesia Santa Maria (cf. Figura 9 esquerda) e outras com indicações de tempo da caminhada a pé até espaços comerciais e culturais de interesse (cf. Figura 9 direita).

Quanto à categoria *bottom up* e porque esta se trata de uma zona de comércio e turística, foram registados 188 elementos, sendo 61 deles em inglês. A diferença entre as categorias demonstra que a língua inglesa é utilizada com muita relevância nos espaços comerciais e hoteleiros, mas com pouco destaque nos espaços governamentais e instituições, o que significa que os sinais linguísticos são maioritariamente dirigidos ao público inglês e aos estrangeiros de passagem, turistas e visitantes e não exclusivamente a residentes.

A Zona Velha, por ser um local bastante frequentado por turistas, é abundante em mensagens multilingues especialmente na oferta de menus e produtos comerciais, confirmando a hipótese 1 de que esta zona iria apresentar uma paisagem multilingue. Por outro lado, a língua inglesa manifestou-se no espaço público em contextos diferentes dos comerciais e das outras línguas, como por exemplo em painéis governamentais históricos, comprovando a hipótese 2.



**Figura 10. Sinais linguísticos *bottom up* em língua inglesa num espaço de reatuação (esquerda) e num supermercado (direita).**

Os espaços comerciais e da restauração da Rua Santa Maria, que é conhecida por pelo seu movimento turístico onde abundam restaurantes de uma ponta a outra e de um lado e de outro da rua, mostraram preferência pela utilização da língua inglesa, excluindo por vezes a língua portuguesa por completo (cf. Figura 10 esquerda) ou apresentando o inglês em primeiro lugar e o português em segundo: “Sunday we open at 10:00 o’clock” e “ao domingo abrimos às 10:00” (cf. Figura 10 direita). Atente-se que a informação em português não está tão completa como a informação em inglês, uma vez que nesta última utilizaram o termo “o’clock” que teria o seu equivalente em português pela palavra “horas”, mas que não surge na tradução. Alomoush (2015, p. 105) descreve esta situação como pertencente a “sinais fragmentados”, ou seja, sinais que não aparecem completos na tradução. Este facto sugere que este supermercado deve ser muito frequentado por estrangeiros e os produtos que vende são destinados a estes.

Na Figura 11 (esquerda) é possível ver que a primeira língua em que surge o menu é a inglesa, seguida da portuguesa e na loja da Figura 11 (direita) a escolha da língua recaiu também para a inglesa, demonstrando a preferência por esta.



Figura 11. Sinais linguísticos *bottom up* de anúncios comerciais em língua inglesa no Restaurante “The Snug” (esquerda) e “Lillie Ceramics” (direita).

Verificou-se ainda a criação e afixação de alguns anúncios com a informação de que aquele estabelecimento “fala” a língua indicada e os clientes poderão comunicar com os empregados na sua língua mãe (*cf.* Figura 12).



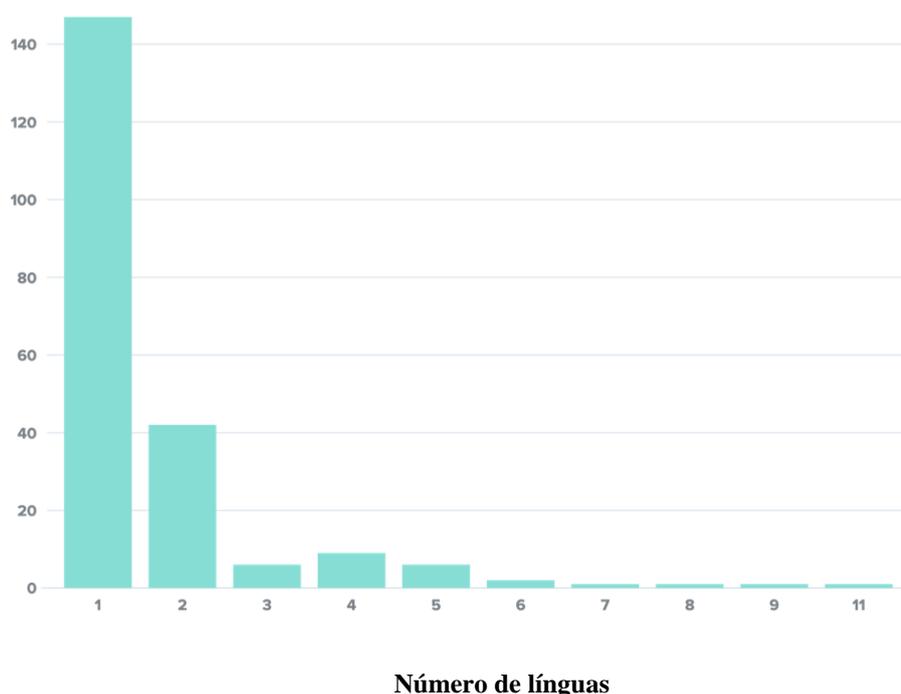
Figura 12. Sinais linguísticos *bottom up* de um restaurante e alojamento em inglês, francês, alemão e espanhol.

Tabela 5. Lista das diferentes línguas encontradas na Zona Velha.

1. Português	2. Inglês	3. Francês	4. Italiano
5. Alemão	6. Norueguês	7. Russo	8. Espanhol
9. Holandês	10. Sueco	11. Finlandês	12. Polaco

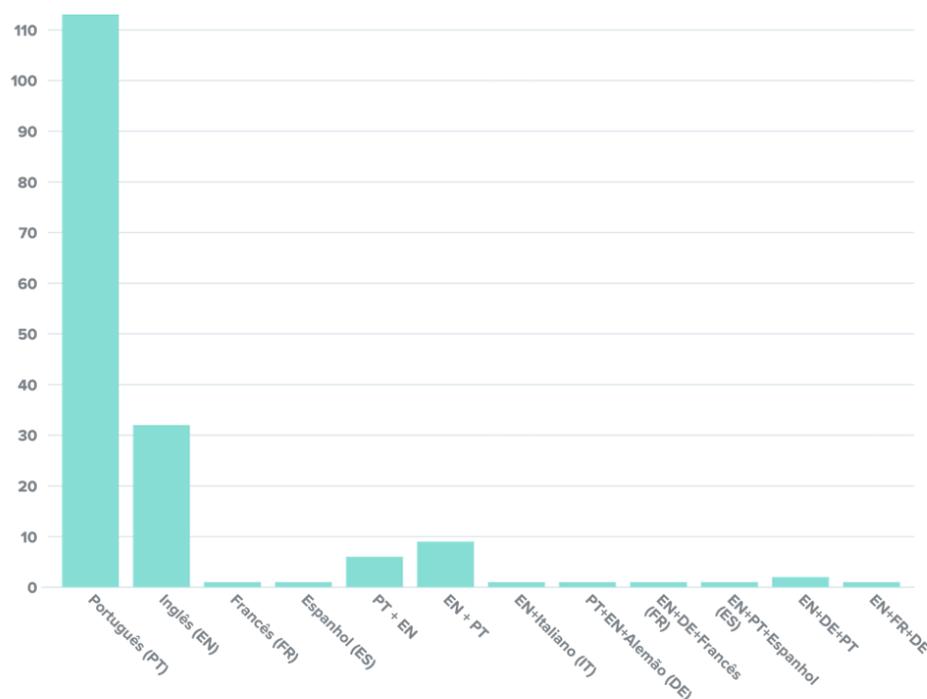
Foram encontradas 12 línguas no espaço analisado, sendo elas o português, o inglês, o francês, o italiano, o alemão, o norueguês, o russo, o espanhol, o holandês, o sueco, o finlandês e o polaco (*cf.* Tabela 5).

Os números de línguas encontradas na PL da Zona Velha (*cf.* Gráfico 1) não deixaram dúvidas que se trata de uma paisagem dominada pela presença da língua portuguesa, sendo que, entre os 147 identificados em apenas numa língua, 113 estavam escritos em português, 32 em inglês, 1 em francês e 1 em espanhol (*cf.* Gráfico 2).



**Gráfico 1. Número de línguas nos sinais linguísticos da Zona Velha.**

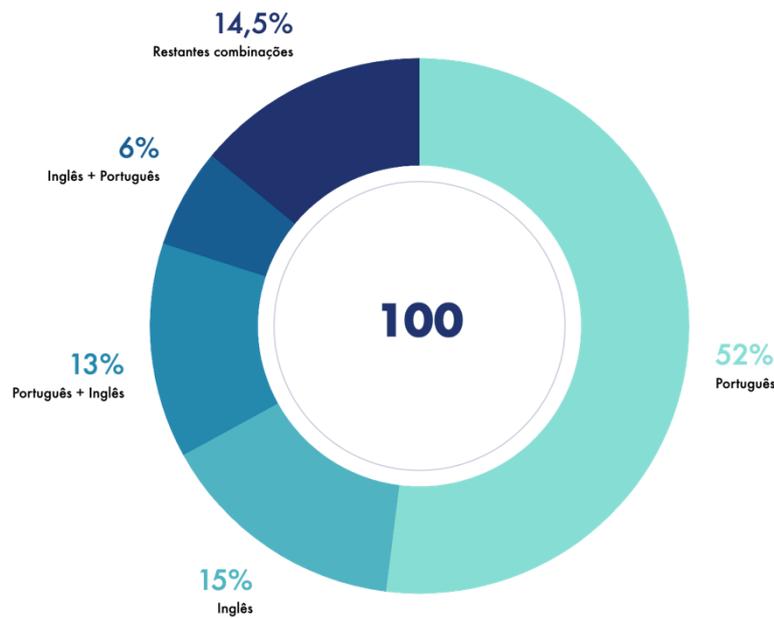
Verificou-se que outros restaurantes com uma oferta mais simples para turistas, optaram por cartazes simples exclusivamente em inglês para esse público, desconsiderando tanto o português como outras línguas. Os dados mostram que a tendência é utilizar uma língua e depois duas línguas, neste caso, português-inglês ou vice-versa e que as outras combinações de línguas ocorrem com baixa frequência. As 12 línguas identificadas apontam para a presença do turismo gastronômico, uma vez que estas línguas surgem exclusivamente em espaços da restauração. O português é claramente a língua mais utilizada (*cf.* Gráfico 2.) Nos sinais multilingues encontrados, a tendência é incluir o português ou o inglês. Foram encontradas algumas combinações distintas (*cf.* Gráfico 2) unicamente em menus de restaurantes, entre as línguas francesa, italiana, alemã, norueguesa, russa, espanhola, holandesa, sueca, finlandesa e polaca.



**Número de línguas e combinações encontradas**

**Gráfico 2. Frequência de cada língua no sinal linguístico.**

O facto de apenas uma língua surgir em mais de metade da amostra sugere que a Zona Velha é uma área que para além de turística, é também característica do comércio tradicional, com edifícios como o Mercado dos Lavradores, a Escola Secundária Jaime Moniz, cabeleireiros, escritórios de advogados, centros de cópia, cafés e restaurantes que oferecem um menu económico e optam por preferir utilizar apenas a língua portuguesa uma vez que é esse o público-alvo e, por sua vez, justificando o aparecimento de 113 sinais nessa língua, que corresponde a 52% (*cf.* Gráfico 3) da amostra.



**Gráfico 3. Frequência de cada língua no sinal linguístico (%).**

No que diz respeito à combinação português + inglês, esta surgiu 28 vezes, correspondendo a 13% da amostra e a combinação inglês + português, 13 vezes, correspondendo a 6% (*cf.* Gráfico 3), demonstrando que há uma maior tendência em utilizar o português em primeiro lugar tanto nos sinais monolíngues como bilingues. Esta situação pode ser explicada pela atividade comercial desenvolvida nesta área, como lojas de sapatos, de lojas de chineses, bancas de venda de fruta e peixe e pelo facto de alguns espaços de restauração oferecerem um prato do dia mais barato para os locais e tentar ainda aliciar os turistas. Quanto ao tamanho de uma língua em relação a outra, não foi verificada nenhuma situação flagrante, no entanto, há a destacar a diferença entre diversos menus em português e inglês, sendo que os primeiros são apresentados tendencialmente em quadros de ardósia para que possam ser facilmente modificados no dia seguinte, e em folhas separadas afixadas com pouco cuidado nas paredes. Quanto aos menus apresentados em inglês, verificou-se um cuidado especial, incluindo inclusive fotografias alusivas à oferta gastronómica. Alguns dos restaurantes, colocaram também anúncios em destaque com as línguas que aquele estabelecimento fala para comunicar com os seus clientes, como, por exemplo, “We speak English” e “Nous Parlous Français”.

Por outro lado, a língua inglesa assumiu nos últimos anos um estatuto de modernidade e sofisticação. Como cidadãos no e do mundo, é fácil apercebermo-nos do seu estatuto como língua de comunicação internacional e a forma como esta prolifera na comunicação entre pares e na própria cidade em si. O inglês tem vindo a refletir-se especialmente através do nome de restaurantes e bares, em nomes de bebidas especiais ou pratos distintos.

There has been no other language having such a broad coverage as a lingua franca in the history of human being. The composition of English users has also shifted tremendously. The number of the so-called non-native speakers of English has reached three times the number of native speakers. As a matter of fact, the population of the so-called non-native English speakers continuously increases while that of the first language speakers decreases. This has brought English into its new role as a “global currency”. (Dewi 2013, p. 3)

Este fenómeno explica a utilização do inglês para a nomenclatura de novos bares e restaurantes (cf. Figura 13).



Figura 13. Bar com nome em inglês: “The Old City Pub”.

É fácil observar que os novos espaços de restauração e comércio na cidade do Funchal optam, muitas vezes, pela designação inglesa. Este fenómeno poderá ser um reflexo do estatuto que a língua inglesa passou a adquirir. Ao passar pelos restaurantes mais antigos da zona histórica, o visitante repara que estes têm nomes portugueses, tais como “A Tropicana”, “A Taberna Ruel”, “Taberna Madeirense”, “O Regional” e “Restaurante do Forte”. Estes restaurantes, mais antigos e tradicionais, são conhecidos por serem frequentados especialmente por turistas.

No que diz respeito aos espaços hoteleiros mais recentes, estes têm, muitas vezes, nomenclaturas inglesas e na maior parte deles, o público-alvo é o mais jovem e local. Ou seja, estes espaços sugerem ser propriedade de donos madeirenses jovens cujos clientes são tendencialmente os seus conterrâneos. Alguns exemplos destes espaços são os

conhecidos “The Flair Spot”, “Revolution Bar”, “The Old Town Pub”, “Mini Eco Bar” e “The Town”.

É de referir ainda os bares e espaços com designação em língua portuguesa, frequentados tanto por locais como por estrangeiros – devido à sua localização estratégica – e que optaram pela língua mãe. Alguns destes bares são alusivos aos nomes dos proprietários, como “Tasca do Zé” e “Dona Maria”. Existem, por fim, algumas designações distintas, ainda que em minoria, alusivas aos países que esses restaurantes e bares representam, como, por exemplo, nomes de restaurantes japoneses, italianos e nepaleses (cf. Figura 14).



Figura 14. Restaurante com nome em italiano: “CyberC@fé Chévere”.

## 6. Conclusão

As condições económico-sociais que imperavam na Ilha da Madeira, aliadas à sua posição estratégica geográfica, condições climáticas e produção agrícola foram um fator impulsionador para a presença e ocupação britânica. Esta foi responsável pelo desenvolvimento da Região e definiu um conjunto de hábitos culturais e linguísticos no povo madeirense. Se no passado a PL a era um reflexo da ocupação governamental britânica e estava mais presente em sinais *top down*, hoje é um reflexo de uma presença maioritariamente turística. Os laços genealógicos, sociais e comerciais permitiram a continuidade de uma ligação entre os madeirenses e os ingleses, tornando a Madeira um ponto de referência para o mundo britânico. Estes acontecimentos favoreceram a existência de um número muito superior de sinais linguísticos na categoria *bottom up*, e

apresentaram-se em quantidade elevada nos sinais de restauração, lojas e espaços comerciais. Entre estes, é possível destacar especialmente os sinais dos restaurantes.

Através da análise das representações da linguagem no espaço público, a PL explora o sinal linguístico exposto e o modo como as pessoas interagem com esses mesmos sinais. Desta forma, a utilização do inglês nos sinais linguísticos *bottom up* é um reflexo dos clientes a quem este é dirigido, neste caso, aos turistas. Grande parte dos estudos em PL têm sido dirigidos a comunidades multilingues, tentando analisar a presença ou a inexistência de sinais linguísticos sem levar em consideração as condições de construção dos mesmos ou com um grande foco no contexto histórico. Este trabalho pretendeu estudar não apenas a PL da Zona Velha, mas também todo o seu enquadramento histórico relacionado com a ocupação britânica na Ilha da Madeira e, mais especificamente, na cidade do Funchal.

A exploração e investigação dos sinais linguísticos a partir de fotografias demonstrou ser um método eficaz para a determinação e categorização dos sinais encontrados de acordo com a classificação de Ben-Rafael *et al.* (2006, pp. 19–21). Os resultados demonstraram ser ferramentas úteis para compreender o atual panorama linguístico da Zona Velha, na medida em que demonstram uma presença preponderante da língua inglesa perante outras e, como tal, é uma informação importante para os gestores da cidade, dos serviços turísticos e da hotelaria. Por outro lado, são também expressões de uma raiz histórica e cultural a que os madeirenses não são indiferentes, consolidando a sua memória e identidade.

Através dos dados recolhidos é possível concluir que a influência da ocupação britânica já não é muito sentida no panorama linguístico da cidade e hoje a utilização do inglês poderá ser um reflexo das relações comerciais e profissionais anteriormente travadas que abriram portas para uma grande afluência do turismo britânico. Por outro lado, é preciso ter em mente o estatuto de modernidade do inglês e a sua utilização como língua global. Desta forma, comprovamos as hipóteses iniciais de que a Zona Velha, por ser uma zona muito turística, rica em espaços comerciais e da restauração é abundante em mensagens multilingues com muitas ofertas de produtos e serviços aos turistas e que a língua inglesa também se manifesta no espaço público em contextos distintos dos comerciais.

No que diz respeito às questões da investigação, podemos responder que a língua inglesa surge tanto em sinais *top down* como *bottom up*, ou seja, em sinais governamentais e sinais comerciais, tendo maior expressão nestes últimos devido ao fenómeno turístico britânico. Surgiram outras dez línguas estrangeiras, sendo elas o francês, o italiano, o alemão, o norueguês, o sueco, russo, o espanhol, o holandês, o finlandês e o polaco. O aparecimento destas em sinais exclusivamente *bottom up* e, mais especificamente, em menus de restaurantes, aponta para a existência de turismo proveniente dos países onde essas línguas são faladas. Este é um fator importante para demonstrar que a presença do inglês surge em contextos diferentes das outras línguas.

Os dados demonstraram que a tendência é utilizar uma língua e depois duas línguas, neste caso, português-inglês ou inglês-português e que as outras combinações de línguas ocorrem com baixa frequência. Estes sinais bilingues têm como característica a apresentação da língua portuguesa habitualmente em primeiro lugar e a inglesa em

segundo lugar, uma vez que estes sinais surgem em ofertas de pratos do dia dirigidos aos trabalhadores locais da zona que falam português. Por outro lado, os sinais bilingues que surgem com a língua inglesa em primeiro lugar dizem respeito a bares e espaços comerciais dirigidos aos ingleses, e onde se pode observar uma informação complementar em português. Por fim, surgiram algumas combinações distintas nos sinais multilingues que foram encontrados unicamente em sinais da restauração e cujas combinações não sugeriram ter nenhuma razão evidente para além de uma escolha arbitrária de cada restaurante quanto à seleção da primeira língua a apresentar.

Como sugestão de estudo futuro, pretendemos realizar um questionário aos proprietários e autores dos diferentes sinais linguísticos de forma a compreender qual a sua intenção ao apresentá-los da forma como foram registados.

**Financiamento:** Esta pesquisa não recebeu financiamento.

## Referências

- Alomoush, O. (2015). *Multilingualism in the linguistic landscape of urban Jordan*. (Tese de Doutoramento não editada, University of Liverpool, Inglaterra).
- Bazenga, A., Ribeiro, J., & Sequeira, M. (2012). *The British presence in Madeira Island: historical overview and linguistic outcomes*. Comunicação apresentada no British Scholar Society Conference Edinburgh, UK.
- Backhaus, P. (2006). Multilingualism in Tokyo: A look into the linguistic landscape. *International Journal of Multilingualism*, 3(1), 52–66. <https://doi.org/10.1080/14790710608668385>.
- Berger, R., & Lecheta, M. (2019). A paisagem linguística de um campus universitário fronteiriço: língua e poder em perspectiva. *Entrepalavras*, 9(2), 396–414. <https://doi.org/10.22168/2237-6321-21486>.
- Ben-Rafael, E., Shohamy, E., Amara, M., & Trumper-Hecht, N. (2006). Linguistic landscape as symbolic construction of the public space: the case of Israel. *International Journal of Multilingualism*, 3(1), 7–30. <https://doi.org/10.1080/14790710608668383>.
- Biddle, A. (1896). *The Madeira Islands*. Philadelphia: Drexel Biddle & Bradley Publishing Company.
- Bielenin-Lenczowska, K. (2021). A paisagem sócio-linguística: A política, a diversidade e a migração no espaço público. *Fórum Linguístico*, 17(4), 5275–5288. <https://dx.doi.org/10.5007/1984->
- Blommaert, J. (2010). *The Sociolinguistics of Globalization*. Cambridge: Cambridge University Press.
- Bourhis, R., & Landry, R. (1997). Linguistic landscape and ethnolinguistic vitality—An empirical study. *Journal of Language and Social Psychology*, 16(1), 23–49. <https://doi.org/10.1177/0261927X970161002>.
- Brandão, R. (1927). *As ilhas desconhecidas*. Lisboa: Edições QETZAL.
- Calvet, J. (1994) *Les voix de la ville: Introduction à la sociolinguistique urbaine*. Paris: Collection Petite Bibliothèque Payot.
- Carr, J. (2019). *Linguistic landscapes*. <https://doi.org/10.1093/OBO/9780199772810-0251>.
- Cenoz, J., & Gorter, D. (2006). Linguistic landscape and minority languages. *International Journal of Multilingualism*, 3(1), 67–80. <https://doi.org/10.1080/14790710608668386>.
- Clemente, M. (2018). *Paisagem linguística urbana: O caso de Aveiro e a sua relevância educativa*. (Tese de doutoramento não editada, Universidade de Aveiro, Portugal).
- Cotterill, R., Muir, K., Joinson, A., & Dewdney, N. (2015). Identifying linguistic correlates of social power. *International Journal of Computational Linguistics and Applications*, 6(1),

- 9–26. Disponível em [https://www.researchgate.net/publication/304497624\\_Identifying\\_Linguistic\\_Correlates\\_of\\_Social\\_Power](https://www.researchgate.net/publication/304497624_Identifying_Linguistic_Correlates_of_Social_Power)
- Da Silva, I., Santos, M., & Jung, N. (2016). Multilinguismo e política linguística: análise de uma paisagem linguística transfronteiriça. *Domínios de Linguagem*, 10(4), 1257–1277. <https://doi.org/10.14393/DL27-v10n4a2016-4>.
- Dewi, A. (2013). English as an international language: An overview. *Journal of English and Education*, 6, 1–11. <https://doi.org/10.20885/jee.vol6.iss2.art1>
- Gorter, D. (2006). Introduction: The study of the linguistic landscape as a new approach to multilingualism. *International Journal of Multilingualism*, 3(1), 1–6. <https://doi.org/10.1080/14790710608668382>.
- Gorter, D. (2007, fevereiro). *The linguistic landscape in Rome: Aspects of multilingualism and diversity*. (Working Paper of the IPRS). [https://www.academia.edu/2122945/The\\_linguistic\\_landscape\\_in\\_Rome\\_Aspects\\_of\\_multilingualism\\_and\\_diversity](https://www.academia.edu/2122945/The_linguistic_landscape_in_Rome_Aspects_of_multilingualism_and_diversity)
- Gorter, D., Marten, H. F., & Mensel, L. (2012). *Minority languages in the linguistic landscape*. <https://doi.org/10.1057/9780230360235>.
- Huebner, T. (2006). Bangkok's linguistic landscapes: Environmental print, codemixing and language change. *International Journal of Multilingualism*, 3(1), 31–51. <https://doi.org/10.1080/14790710608668384>.
- Junta de freguesia Santa Maria Maior. (2020). *História da Junta de Freguesia de Santa Maria Maior*. Consultado em <https://www.funchal.pt/juntas-de-freguesiapt/junta-de-freguesia-de-santa-maria-maior.html>
- Kreslins, J. (2003). Linguistic landscapes in the baltic, *Scandinavian Journal of History*, 2, 165–174. <https://doi.org/10.1080/03468750310003659>.
- Landry, R., & Bourhis, R. (1997). Linguistic landscape and ethnolinguistic vitality an empirical study. *Journal of Language and Social Psychology*, 16(1), 23–49. <https://doi.org/10.1177/0261927X970161002>
- Macedo, M. (2019). Produção de uma paisagem linguística em espaços públicos e comerciais na ilha da Madeira Comparação com dados relativos ao Algarve e Aveiro. Universidade da Madeira. Consultado em: <https://www.researchgate.net/publication/330741275>
- Macedo, M. (2020). *Ementas e nomes de restaurantes em zona turística do Funchal e de Santa Cruz de Tenerife. Um olhar sobre a paisagem linguística*. (Dissertação de Mestrado não editada, Universidade da Madeira, Portugal).
- Melo-Pfeifer, S. & Lima-Hernandes, M. (2020). Paisagens Linguísticas: Ideologias, discursos e práticas multilingues nos espaços sociais. *Domínios de Linguagem*, 14(4), 1024–1058. Universidade Federal de Uberlândia, Brasil. <https://doi.org/10.14393/DL44-v14n4a2020-1>
- Rebelo, H. (2021a). A toponímia como paisagem linguística e património linguístico: Ruivós, no cruzamento de rotas para a Madeira? *SABUCALE-Revista Anual do Museu do Sabugal*, 12, 113–132. Consultado em <http://hdl.handle.net/10400.13/3416>
- Rebelo, H. (2021b). Da paisagem à paisagem linguística como património ou da prática à teoria: Para uma tipologia da paisagem linguística. *Confluência, Revista do Instituto de Língua Portuguesa*, 60, 198–221. <https://doi.org/10.18364/rc.v1i60.395>.
- Rodrigues, L. (2020) Paisagem linguística em contexto fronteiriço: Estudo de caso em Tabatinga (Bra) e Letícia (Col). *Revista científica TRAMA*, 16(37), 149–160. <https://doi.org/10.48075/rt.v16i37.23694>.
- Rodrigues, P. (1999) *A política e as questões militares na Madeira – o período das guerras napoleónicas*. Funchal: Centro de Estudos de História do Atlântico (CEHA).
- Rodrigues, P. (2007) *A Madeira entre 1820 e 1842: Relações de poder e influência britânica*. (Tese de doutoramento não editada, Universidade da Madeira, Portugal).
- Scirihha L., & Vassallo M. (2001) *Malta, a linguistic landscape*. Malta: Malta Caxton.

- Teis, T., Seide, S., & Lucas, P. (2018). Os topônimos na paisagem linguística da Av. Zelina, em São Paulo: Um encontro na interdisciplinaridade. *Revista do GELNE*, 20(2), 16–29. <https://doi.org/10.21680/1517-7874.2018v20n2ID14089>.
- Teixeira, V. (2009). *Entre a Madeira e as Antilhas: A emigração para a Ilha de Trindade: século XIX*. (Dissertação de Mestrado não editada). Universidade da Madeira, Portugal.
- Vieira, A. (2018). Da Madeira a Cape Town, República da África do Sul. *Cadernos de divulgação do CEHA*, 9, 1–90. Consultado em [https://www.academia.edu/37784782/VIEIRA\\_ALBERTO\\_DA\\_MADEIRA\\_A\\_CAPE\\_TOWN\\_REP%C3%A9BLICA\\_DA\\_%C3%81FRICA\\_DO\\_SUL](https://www.academia.edu/37784782/VIEIRA_ALBERTO_DA_MADEIRA_A_CAPE_TOWN_REP%C3%A9BLICA_DA_%C3%81FRICA_DO_SUL).
- Vxmag. (2019, abril). *1807: Quando a Madeira pertenceu aos ingleses durante 7 anos*. Consultado em <https://www.vortexmag.net/1807-quando-a-madeira-pertenceu-aos-ingleses-durante-7-anos/>.

[recebido em 01 de março de 2021 e aceite para publicação em 03 de agosto de 2021]

## EMPRÉSTIMOS DO ITALIANO EM NOMES EMPRESARIAIS EM PORTUGAL

### ITALIAN LOANWORDS IN COMPANY NAMES IN PORTUGAL

Rosa Lídia Coimbra\*  
rlcoimbra@ua.pt

Lurdes de Castro Moutinho\*\*  
lmoutinho@ua.pt

Juntamente com os movimentos dos povos, as línguas também viajam, deixando e/ou transportando marcas evidentes desse movimento pelos espaços geográficos por onde transitam. De entre as marcas linguísticas mais visíveis, destaca-se o campo do léxico. A entrada de italianismos no português, diretamente ou via outras línguas românicas, como o espanhol e o francês, tem sido situada principalmente na época do Renascimento, período em que a Itália era vista como uma referência cultural importante na Europa. No estudo que aqui nos propomos apresentar, referiremos a presença de marcas lexicais italianas no português atual, sob uma perspetiva sincrónica. Com este objetivo, partimos de um corpus de 151 italianismos, já dicionarizados e compilados, que fazem parte do nome de 1087 empresas em Portugal. O corpus foi obtido num diretório de empresas online, tendo as buscas sido efetuadas por palavra-chave com os referidos italianismos. O corpus obtido foi analisado em categorias formais (nomes, adjetivos) e semânticas (música, arquitetura, artes, gastronomia). Foram, assim, apurados dados relativos a estes nomes, bem como ao ramo de atividade a que se referem e sua localização geográfica. O cruzamento dos dados e os gráficos obtidos permitirão uma leitura objetiva da sua utilização em denominações de empresas sediadas em Portugal

**Palavras-chave:** Empréstimos. Italianismos. Nomes de empresas. Portugal.

Languages travel along with peoples, leaving and/or carrying evident linguistic marks on their way. The lexicon is the most evident of these language marks. Italianisms entered the Portuguese language either directly or via other Romance languages such as Spanish and French. This entry has occurred mainly in the Renaissance, a period when Italy was seen as an important cultural reference in Europe. In the present study we refer to the presence of Italian lexical marks in the current Portuguese lexicon, from a synchronic perspective. For this purpose, we gathered, in a Portuguese dictionary, a corpus of 151 Italianisms which are part of 1087 business names in Portugal. The corpus of these names was obtained in an online business directory, and searches have been carried out by keyword with those Italianisms. The resulting corpus was analysed according to formal categories (names, adjectives) and semantic categories (music, architecture, arts, gastronomy). The resulting data included field of activity and

---

\* Departamento de Línguas e Culturas / Centro de Línguas, Literaturas e Culturas, Universidade de Aveiro, Aveiro, Portugal. ORCID: 0000-0002-8578-673X

\*\* Centro de Línguas, Literaturas e Culturas, Universidade de Aveiro, Aveiro, Portugal. ORCID: 0000-0003-0869-9912.

geographical location. The crossing of the data and the graphs obtained allowed an objective reading of the use of Italianisms Portuguese business denominations.

**Keywords:** Loans. Italianisms. Business names. Portugal.

•

## 1. Introdução

Nenhuma língua evolui isoladamente. Mesmo em épocas remotas, sem tecnologias da comunicação e em que o mundo não apresentava a atual globalização digital, as fronteiras são sempre atravessadas e, com a circulação de pessoas, seja por motivos militares, comerciais ou outros, circulam também as respectivas línguas e léxicos. Algumas palavras, provenientes de uma determinada língua, acabam por ser incorporadas em outra língua como estrangeirismos e, posteriormente, passado um processo de adaptação ortográfica e fonética, como empréstimos lexicais. A importação de palavras é “um fenómeno incontornável e muitas vezes difícil de evitar, dada a inexistência de léxico autóctone com idêntico valor referencial” (Villalva 2008). A língua portuguesa não é exceção e, no seu léxico, vamos encontrar palavras provenientes de diversas origens para colmatar estas lacunas lexicais. A que nos ocupa na presente pesquisa é a origem italiana. Particularmente na época do renascimento, entraram no português uma quantidade de italianismos:

A influência italiana marcou a fase renascentista. Dessa língua, o português recebeu unidades lexicais relativas à poesia (*soneto, terceto...*), à música (*arpejo, contralto...*), à náutica (*barcaça, bússola, fragata, galera, piloto...*), à milícia (*acampar, batalhão...*), ao teatro (*arlequim, bufo...*) (Alves & Maraneze 2004, p. 29).

De facto, nessa altura, diversas foram as áreas de atividade que importaram, direta ou indiretamente, vocabulário de origem italiana:

A contribuição do italiano manifesta-se principalmente nas terminologias artísticas. Em termos de música: *piano, contralto, ópera*; de arquitectura: *balcão, fachada*; de pintura: *aguarela, pitoresco*; de literatura: *soneto, burlesco*; de conceitos militares: *brigada, infantaria, escopeta*; de comércio: *banco, fiasco*; de indumentária: *calção*. Trata-se, de um modo geral, de termos de civilização que, na época do Renascimento, em que a Itália desempenhou o papel cultural-histórico que se sabe, se propagaram a quase todas as línguas românicas e outras europeias. Nem todos os italianismos do português foram importados directamente, tendo às vezes servido de intermediários o espanhol ou o francês (Piel 1989).

A incorporação de estrangeirismos e empréstimos faz parte do processo evolutivo de qualquer língua viva, como bem refere Peterson (2008):

Qualquer sistema linguístico contém regras sistemáticas que operam de modo dinâmico e com grau de variabilidade, indicando ser a língua não uma, mas várias, assim como vários são seus falantes. Segundo os linguistas, os empréstimos fazem parte de um processo dinâmico comum de criação vocabular, não se constituindo em ameaça à sintaxe da

língua. Empréstimos estrangeiros são esquecidos, aportuguesados, ou mantidos à medida que os falantes os usam ao longo do tempo (Peterson 2008, p. 325).

Sobre a mesma temática, considerando ainda a distinção que, em Vilela (1995), é estabelecida entre léxico e vocabulário, podemos afirmar que, uma vez sancionado pelo uso, o empréstimo passa a fazer parte tanto do léxico da língua, como do vocabulário dos falantes. Ou seja, por um lado, estes termos são incorporados no léxico, passando a fazer parte do “conjunto das palavras fundamentais, das palavras ideais duma língua” (Vilela 1995, p. 13), surgindo, seguidamente, dicionarizados na língua de chegada. Por outro lado, passam também a fazer parte do vocabulário que, “num determinado lugar e num determinado tempo” (idem, p. 3), é utilizado pelos falantes da comunidade linguística em diversos contextos de uso quotidiano, contribuindo, assim, para a renovação de uma língua.

Apesar de presentes em todos os domínios da atividade humana, o campo da publicidade é particularmente propenso à utilização de palavras de origem estrangeira. Por exemplo, um estudo de Valadares e Bertozzi (2015) apurou que 103 em 192 anúncios publicitários brasileiros continham estrangeirismos.

O presente estudo, sobre a presença de italianismos em nomes empresariais em Portugal, vem na senda de pesquisas anteriores sobre este género textual (Moutinho & Coimbra 1997, 1998, 2019; Coimbra & Moutinho 2009, 2012a, 2012b).

O nome empresarial é adotado por um empresário ou uma empresa, é protegido juridicamente e tem como objetivo identificar a entidade no decurso das suas atividades, pelo que surge em todos os elementos textuais produzidos nesse âmbito: nos cartões de visita, nos letreiros identificativos no exterior dos edifícios, no papel timbrado e em geral em todos os materiais publicitários, administrativos ou informativos. Raposo e Nascimento incluem estas denominações na classe dos nomes próprios, definindo-as como “nomes próprios de base descritiva” (2013, p. 1007), bem como as siglas e os acrónimos formados a partir delas.

A importância estratégica da escolha destes nomes, já que identificam a empresa e também de certo modo a descrevem perante o público alvo, justifica o interesse do seu estudo.

## 2. Obtenção do *Corpus*

O *corpus* que serviu de base ao presente estudo é constituído por um total de 151 lexemas de origem italiana. Todos estes termos estão dicionarizados como italianismos no *Dicionário Universal da Língua Portuguesa*, Editora Priberam, e fazem parte de uma lista de mais de 435 italianismos compilados, a partir deste dicionário *online*, por Sabbatini (2007). Os elementos lexicais constantes da lista de Sabbatini (2007) foram pesquisados num diretório *online* de empresas sediadas em Portugal<sup>1</sup>, entretanto descontinuado. Em resultado desta pesquisa, foram encontrados 151 italianismos diferentes (que se apresentam na Secção 3, Tabela 1) presentes num total de 1087 designações de empresas. Foram eliminados da busca: nomes comuns referentes à

---

<sup>1</sup> <http://www.empresaslista.com>

atividade ou produto (*café, banho, crédito, loja, informática*), componentes de topónimos (*Centro Comercial do Campo Alegre*) ou de antropónimos (*Augusto Alegre Artes Gráficas*), por se considerar que, nestes casos, a escolha do italianismo estaria à partida condicionada.

Na Figura 1, podemos observar dois exemplos de estabelecimentos comerciais portugueses, o primeiro em Águas Santas e o segundo na Figueira da Foz, em cujas denominações se encontram lexemas provenientes do italiano.



**Figura 1. Dois exemplos de nomes comerciais com italianismos constantes do *corpus* analisado: *Bufete Bagatela* e *Bergantim Discoteca*.**

Fonte: Facebook de Bufete Bagatela<sup>2</sup> e Facebook de Bergantim Discoteca<sup>3</sup>.

O presente estudo tem como foco o léxico de origem italiana já incorporado na língua de chegada por um processo de aportuguesamento. Assim, a procura na base de dados teve por base a grafia aportuguesada do termo: aguarela (e não *acquarello*), capricho (e não *capriccio*), extravagância (e não *stravaganza*), embora estas versões originais também se encontrem na base de dados. Interessou-nos, por isso, o italianismo não enquanto estrangeirismo, mas enquanto empréstimo, ou seja, “quando o elemento estrangeiro é utilizado numa determinada língua e passa a ser codificado por ela” (Alves 1984, p. 120).

Apurados os 1087 nomes empresariais com empréstimos de proveniência italiana, procedemos à sua categorização. Em primeiro lugar, dividimo-los em categorias morfológicas. Seguidamente, observámos a sua distribuição por ramo de atividade, bem como a distribuição geográfica no país. Estas duas últimas informações são fornecidas nas fichas da base de dados pesquisada. Ver exemplos, na Figura 2, de resultados obtidos na pesquisa.

<sup>2</sup> <https://www.facebook.com/Bufete-Bagatela-228381687185036>

<sup>3</sup> <https://www.facebook.com/bergantimdisco>

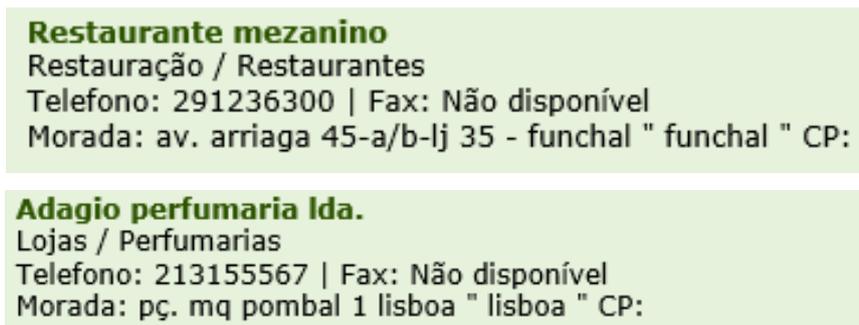


Figura 2. Exemplos de fichas obtidas na base de dados *online* pesquisada

Quanto à categorização segundo a localização geográfica do estabelecimento, fizemos uma pesquisa automática socorrendo-nos dos indicativos telefónicos constantes das fichas recolhidas. A fim de identificarmos o significado do prefixo telefónico, utilizámos a lista de indicativos fornecida *online*.<sup>4</sup> Os resultados obtidos são apresentados e analisados na secção seguinte.

### 3. Apresentação e análise dos resultados

A lista completa dos italianismos encontrados na nossa pesquisa encontra-se na Tabela 1. Na primeira coluna, apresentam-se os nomes encontrados na base de dados pesquisada e, na segunda coluna, a frequência absoluta (FA) de cada um.

Tabela 1. O *corpus* de italianismos recolhido e a frequência absoluta (FA) de cada um.

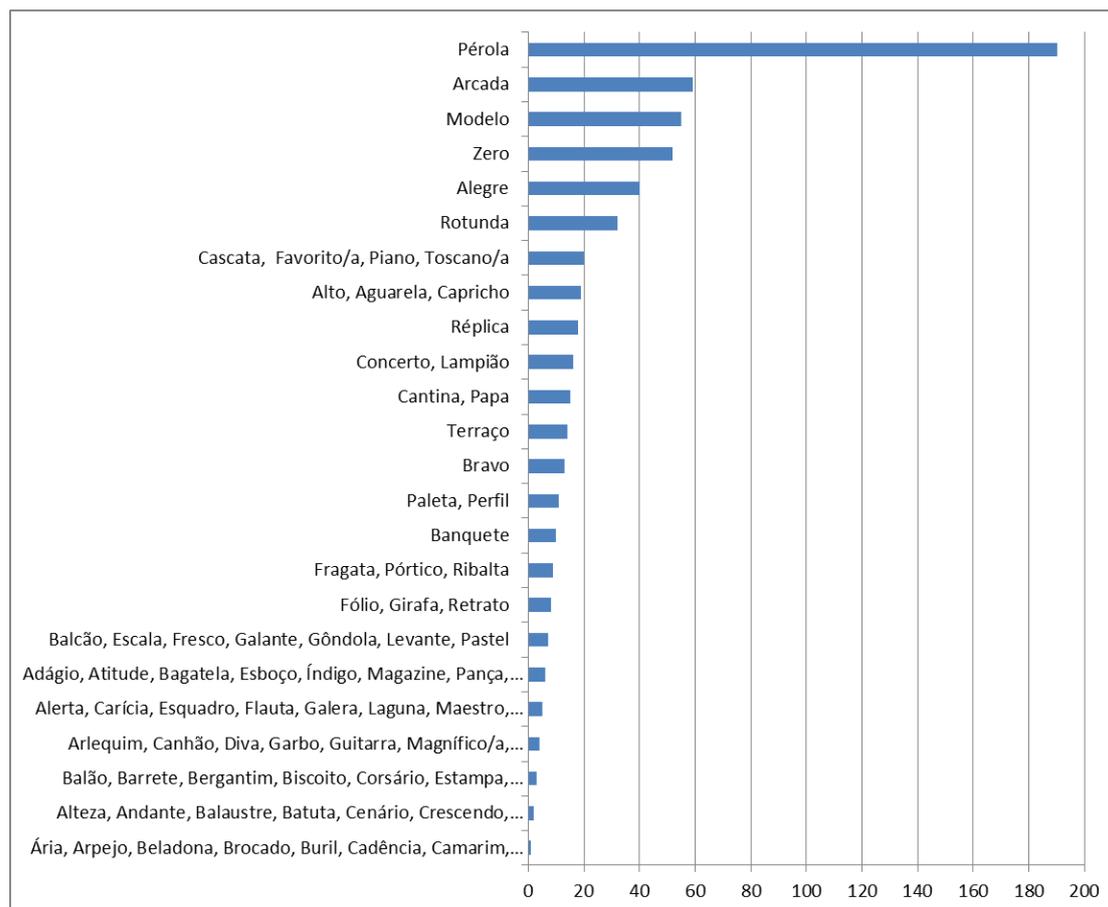
Italianismos nos nomes das empresas	FA
Pérola	190
Arcada	59
Modelo	55
Zero	52
Alegre	40
Rotunda	32
Cascata, Favorito/a, Piano, Toscano/a	20
Alto, Aquarela, Capricho	19
Réplica	18
Concerto, Lampião	16
Cantina, Papa	15
Terraço	14
Bravo	13
Paleta, Perfil	11
Banquete	10
Fragata, Pórtico, Ribalta	9
Fólio, Girafa, Retrato	8
Balcão, Escala, Fresco, Galante, Gôndola, Levante, Pastel	7
Adágio, Atitude, Bagatela, Esboço, Índigo, Magazine,	6

<sup>4</sup> Ver, por exemplo, em <https://conteudos.meo.pt/meo/Documentos/Listagens/Prefixos-regionais.pdf> (acedido em 2020).

Pança, Serenata, Sinfonia, Siroco	
Alerta, Carícia, Esquadro, Flauta, Galera, Laguna, Maestro, Palhaço, Quarteto, Regata, Sonata, Trampolim	5
Arlequim, Canhão, Diva, Garbo, Guitarra, Magnífico/a, Namorado/a	4
Balão, Barrete, Bergantim, Biscoito, Corsário, Estampa, Quinteto, Reduto, Sentinela, Tartaruga, Viola	3
Alteza, Andante, Balaustre, Batuta, Cenário, Crescendo, Embrulho, Girândola, Grafite, Magenta, Manifesto, Máscara, Medalha, Mezanino, Piloto, Pipa, Polichinelo, Primadona, Sarrabulho, Soneto, Taça, Taco, Tarântula, Têmpera, Terracota, Tostão, Tutifruti, Violino	2
Ária, Arpejo, Beladona, Brocado, Buril, Cadência, Camarim, Canção, Capuchino, Caricatura, Carola, Carrossel, Cebola, Cicerone, Concordata, Comédia, Confeti, Corredor, Credência, Espadachim, Esparguete, Estravagância, Lancha, Madrigal, Malandro/a, Merengue, Milanesa, Moscatel, Novela, Ópera, Parapeito, Pedante, Penacho, Pérgula, Pícolo, Postilhão, Ravióli, Salame, Simpático, Surdina, Talharim, Tarantela, Tchou, Trio, Urro, Zibelina	1

Fonte: elaborado pela própria.

Os italianismos que apresentam a mesma FA são agrupados, na mesma célula da Tabela 1, em ordem decrescente. Podemos verificar que, nas células da segunda metade da tabela, encontramos muito mais lexemas do que na primeira. Isto significa que a esmagadora maioria destes lexemas não surge muitas vezes nos nomes empresariais, apenas entre 1 e 9 ocorrências. Pelo contrário, uma minoria de nomes, mais precisamente 23 de entre os 151, ou seja apenas 15% do total, apresenta uma frequência entre 10 e 190 ocorrências. O Gráfico 1 mostra visualmente como esta desproporção se configura.



**Gráfico 1. Italianismos nos nomes de empresas por frequências absolutas (ver lista completa na Tabela 1).**

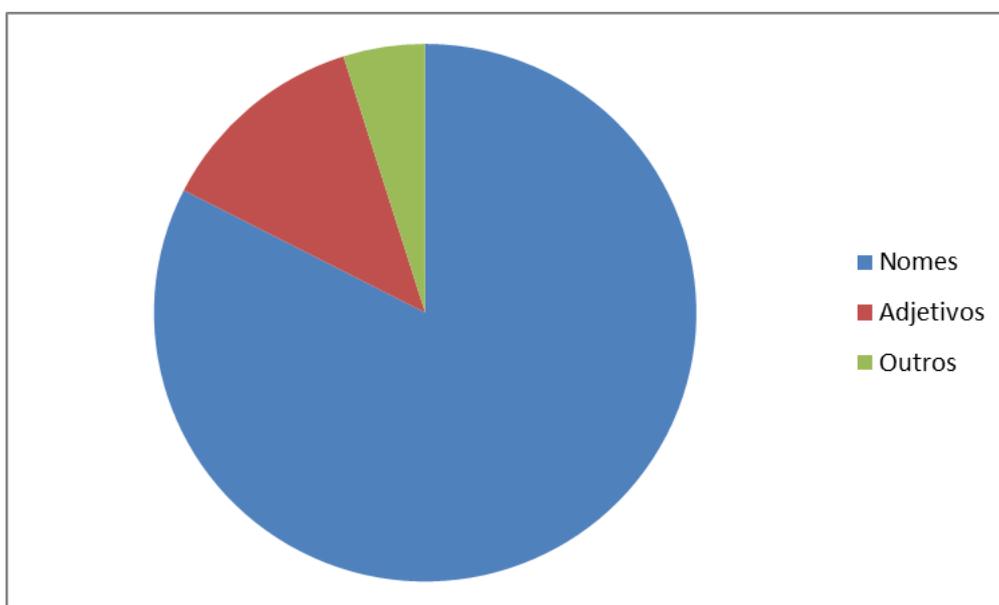
No Gráfico 1, destaca-se, de forma bem marcada, a palavra *Pérola*. De facto, encontramos, no *corpus*, diversos estabelecimentos que elegeram esta palavra na sua designação, isolada ou expandida por modificador preposicional ou adjetival, como por exemplo: *Bazar Pérola do Oceano*, *Boutique Pérola*, *Churrascaria Pérola de Gaia*, *Dancing Pérola Negra*, *Papelaria Pérola Branca*. A conotação positiva do lexema – nome de algo de precioso, perfeito, belo e raro – não será alheia ao seu sucesso, sendo escolhido para referir empresas dos mais diversos ramos de atividade e não necessariamente joalherias, como seria expectável.

Também frequentes, mas apesar de tudo distantes deste primeiro lugar, vamos encontrar os seguintes nomes, que apresentam entre 32 e 59 ocorrências: *Arcada*, *Modelo*, *Zero*, *Alegre* e *Rotunda*. A popularidade dos lexemas *Arcada* e *Rotunda* explica-se pela utilização da estratégia de designar a localização do estabelecimento no seu próprio nome (fica junto ou numa rotunda, fica perto ou dentro de arcadas), o que facilita a deslocação do público que o procura. Já os nomes *Modelo* e *Alegre* serão certamente escolhidos pela conotação positiva a eles associada e que se pretende transmitir como imagem da loja. Por sua vez, o uso de *Zero* na designação comercial origina, por vezes, interessantes e originais jogos semânticos, como nos seguintes exemplos do *corpus*: *Zerocáries Clínica Médica Dentária*, *Ano Zero Fotografia*,

*Cabeleireiros Pente Zero, Café Mais Que Zero, Administração de Condomínios Chatice Zero, Escola de Condução Tolerância Zero, Zero ao Infinito Organização de Eventos.*

Os restantes exemplos podem ser agrupados em áreas vocabulares predominantes. Uma delas é a da música, de que são exemplo os seguintes lexemas: *piano, concerto, adágio, sinfonia, sonata, flauta, ária, arpejo, cadência, canção, madrigal, ópera*. Destacamos ainda outras áreas lexicais agregadoras destes italianismos, tais como: a arquitetura (*terraço, pórtico, balcão, mezanino, parapeito, pérgula*), a navegação (*fragata, bergantim, piloto, lancha, galera, regata, gôndola*) e a alimentação (*cantina, banquete, biscoito, sarrabulho, ravióli, salame, talharim, merengue, esparguete, capuchino, cebola*).

Observemos agora os resultados obtidos para a análise morfológica do corpus. Também aqui encontramos a presença de uma desproporção, tal como foi observada para o léxico. A distribuição dos nossos italianismos por classe de palavras apresenta a configuração que está patente no Gráfico 2.



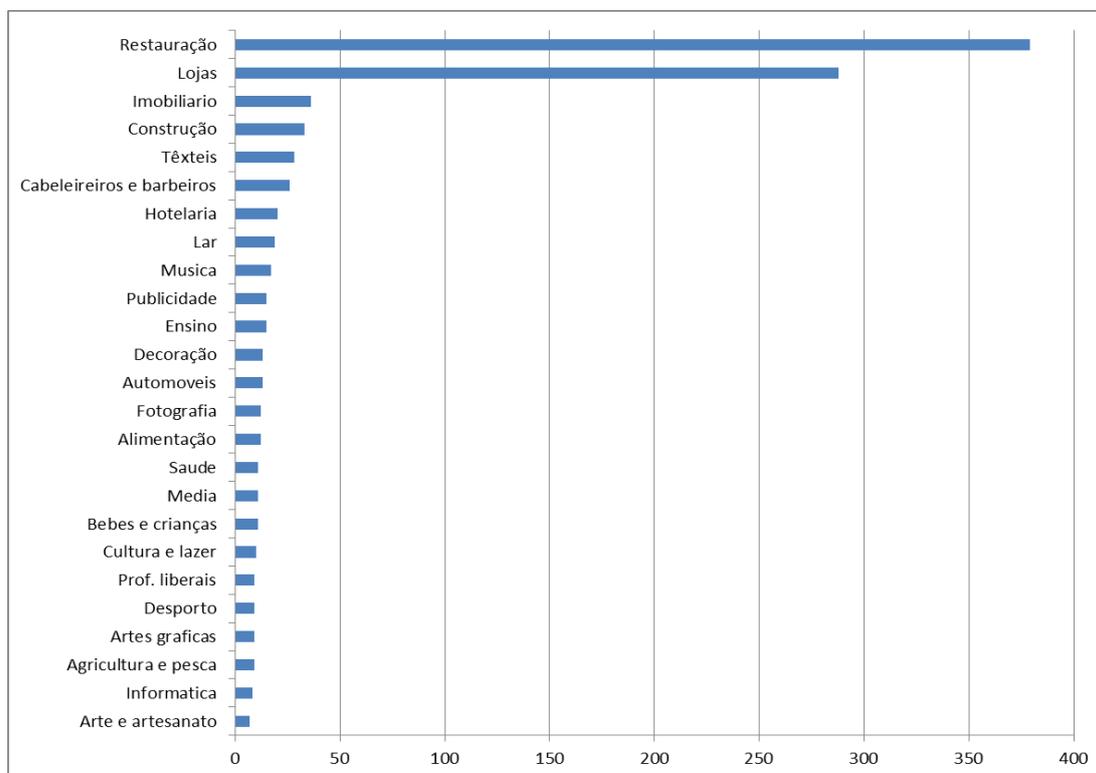
**Gráfico 2. Frequências absolutas dos italianismos nos nomes de empresas por classe morfológica.**

Verificamos, pela observação do Gráfico 2, que a maior parte dos italianismos presentes nos nomes empresariais em Portugal é constituída por substantivos. Em número significativamente menor encontramos os adjetivos. Duas outras categorias morfológicas, pouco representadas, surgem no corpus: a interjeição (*Tchau*) e o numeral (*Zero*). Não foram encontrados verbos, certamente por não constituírem uma categoria morfológica adequada à designação de realidades, mas antes de ações, processos e estados.

Estes resultados são consistentes com estudos que demonstram que a maior parte dos neologismos por empréstimo nas línguas são, de facto, os nomes e, mais raramente adjetivos e verbos (Alves 1984, p. 122; Alves & Maraneze 2004, p. 33).

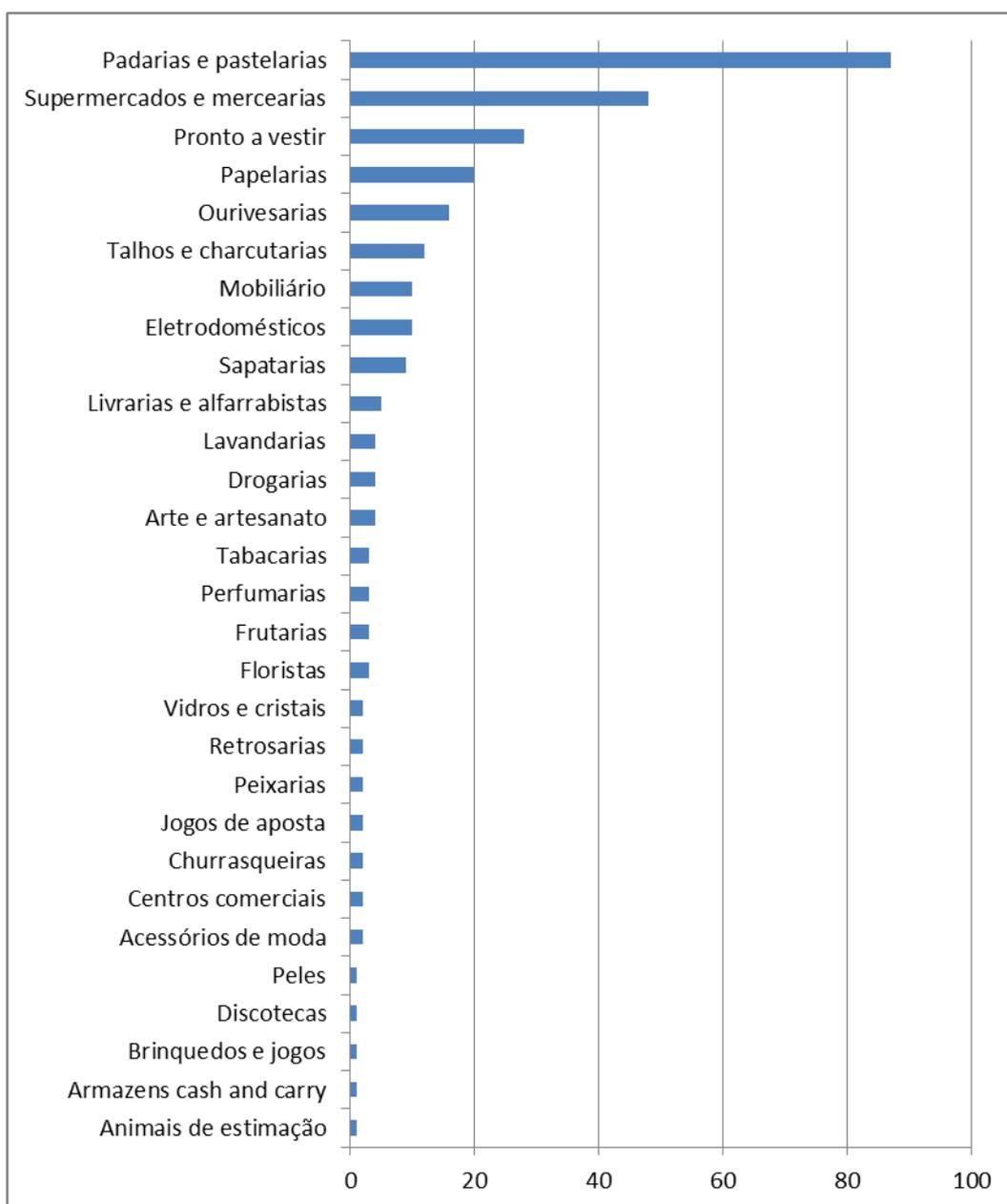
Seguidamente, fizemos a distribuição dos italianismos do *corpus* pelas categorias referentes ao ramo de atividade das empresas por eles designadas. Para tal, socorremo-

nos das categorias fornecidas pela base de dados (ver Figura 2 acima) e obtivemos a distribuição que está patente no Gráfico 3. Neste gráfico, listamos apenas as categorias com mais de 6 ocorrências.



**Gráfico 3. Frequências absolutas dos italianismos nos nomes de empresas por ramo de atividade (ver tipos de Lojas no Gráfico 4).**

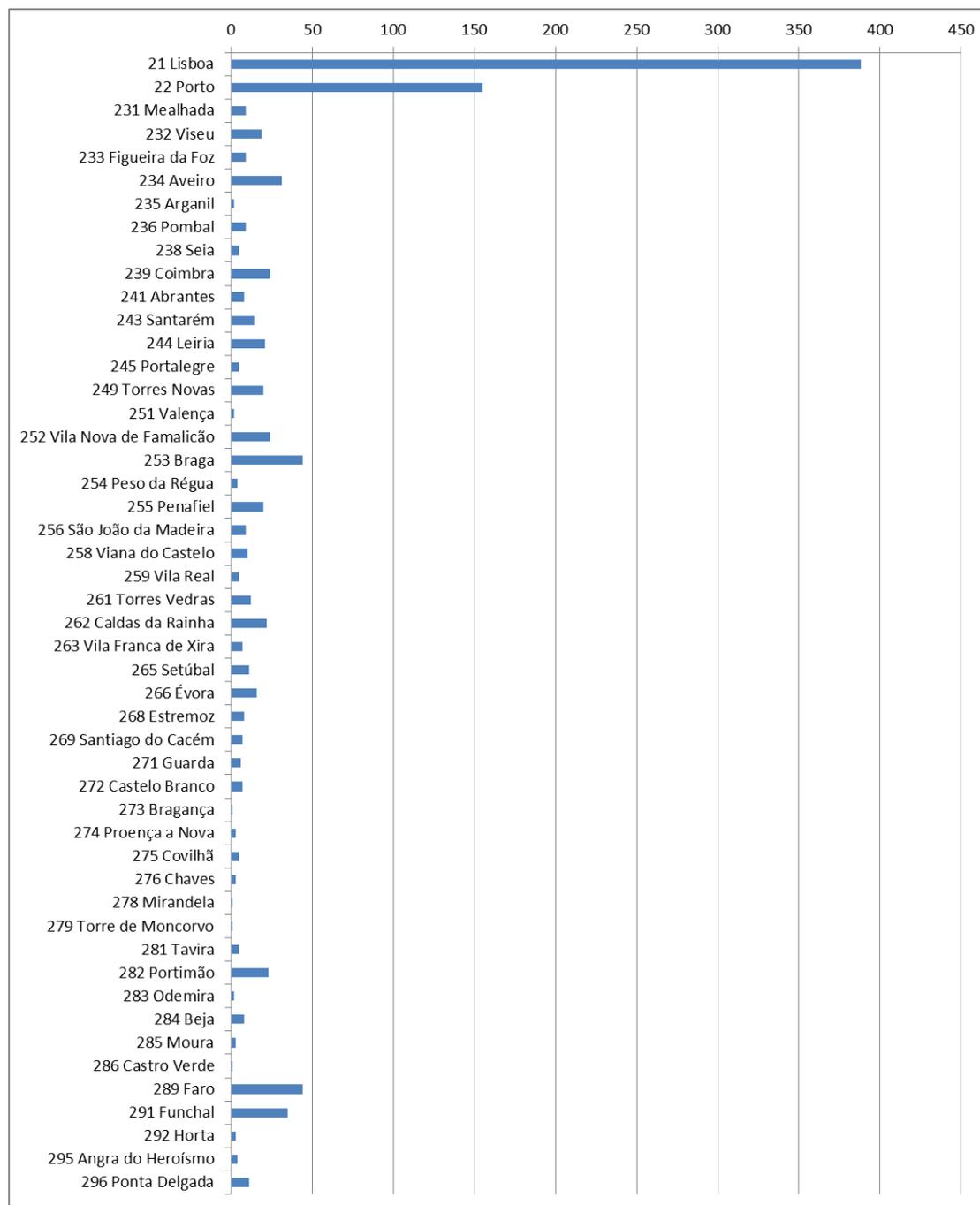
Na elaboração do Gráfico 3, tivemos em conta as categorias superordenadas das fichas da base de dados (Figura 2). Note-se que as duas primeiras categorias – restauração e lojas – atingem valores que se destacam completamente de todos os outros, atingindo, juntos, 62% da totalidade de ocorrências, ou seja, a maioria. Na categoria Restauração encontramos subcategorias como Bar, Café, Restaurante, Taberna, Snack-bar, Cervejaria, Pronto-a-comer, todas elas relacionadas com o setor da alimentação, daí não termos elaborado a sua repartição nestes subgrupos. No entanto, já em relação a Lojas, por considerarmos ser um termo muito genérico e abarcando, simultaneamente, ramos de atividade muito diversificados tivemos em conta as subcategorias da base de dados. Deste modo, fizemos o seu levantamento, obtendo-se, assim, o Gráfico 4.



**Gráfico 4. Italianismos nos nomes de lojas por tipo de loja.**

Se atentarmos agora no Gráfico 4, verificamos que uma grande parte dos estabelecimentos colocados na categoria Lojas estão relacionados com o setor alimentar, apesar de não estarem incluídos no item Restauração. É o caso das seguintes cinco subcategorias: Peixarias, Frutarias, Talhos e charcutarias, Supermercados e mercearias, Padarias e pastelarias. Estas subcategorias de Loja representam 54% de todas as lojas. Se somarmos estas ocorrências à categoria Restauração, obtemos um total de 533 estabelecimentos, ou seja, praticamente metade de todo o *corpus*.

Decidimos também apurar a distribuição geográfica das empresas em questão, fazendo o levantamento da sua localização com base na pesquisa por indicativo telefónico, fornecido pela base de dados consultada e que colámos no Gráfico 5 à esquerda do nome da localidade. Constatamos que se concentram em Lisboa e Porto (ver Gráfico 5).



**Gráfico 5 – Italianismos nos nomes de lojas por região (contagem com base nos indicativos telefónicos).**

Os resultados patentes no Gráfico 5 têm de ser lidos sem grandes especulações visto que, sendo as duas maiores cidades portuguesas, é natural que reúnam a maior quantidade de nomes desta natureza. Um estudo mais apurado poderia levar-nos a melhor entender se esta distribuição tem a ver apenas com a densidade populacional ou se poderia também haver uma hipotética relação com movimentos de populações, tanto mais que não encontramos uma correlação entre a quantidade de italianismos e a densidade populacional das outras localidades.

#### 4. Observações finais

Após a realização desta pesquisa, seria de salientar alguns aspetos que mereceriam vir a ser aprofundados. De entre os vários aspetos, destacamos o facto de os italianismos que surgem na designação de empresas em Portugal não terem uma distribuição regular.

No que se refere às classes morfológicas, apenas encontramos nomes e alguns adjetivos, verificando-se uma ausência total de verbos nas categorias procuradas. Tal predomínio seria, à partida, espectável, já que as formas nominais são claramente mais adequadas à função de referir e designar uma realidade extralinguística.

No que diz respeito à categoria semântica, esta mesma irregularidade de distribuição também se verifica, com um claro predomínio da área alimentar, embora também tenha expressividade as áreas da música, navegação, artes e arquitetura. A forte presença de empréstimos na área alimentar pode ser interpretada pela boa adesão e percepção de valor de marca (Guerrero 2008) que a cozinha italiana tem em todo o mundo, nomeadamente no que se refere a pratos de massas e pizzas. Escolher italianismos nos nomes empresariais ligados à restauração e alimentação poderá ser visto como uma forma de salientar a genuinidade dos referidos produtos.

Quanto às empresas que utilizam na sua designação italianismos, elas pertencem, como referido, predominantemente à área alimentar e situam-se, na sua maioria, nas áreas metropolitanas de Lisboa e Porto. O espaço urbano em geral, e as grandes cidades em particular, são pontos de confluência de línguas e culturas, uma vez que “ a cidade, como instância pós-moderna, é a arena privilegiada dos encontros e dos desencontros das línguas e das culturas do homem urbano que nela se instala e se diversifica – em cada momento diferenciado – através das comunhões simbólicas” (Corrêa-Cardoso 2016, p. 80), daí que o espaço urbano se revele um espaço propício ao aparecimento destas mesclas linguísticas, enriquecendo a língua com novos vocábulos.

Tendo em conta a mobilidade característica da sociedade portuguesa, e a diversidade dos fluxos migratórios, seria interessante, na continuidade deste estudo, pesquisar outro tipo de estrangeirismos e respetivas distribuições geográficas em Portugal.

**Financiamento:** Este trabalho é financiado por fundos nacionais através da FCT, Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto UIDB/04188/2020.

#### Referências

- Alves, I. M. (1984). A integração dos neologismos por empréstimo ao léxico português. *Alfa*, 28 (supl.), 119–126. Consultado em <http://hdl.handle.net/11449/107590>.
- Alves, I. M. & Maraneze, B. O. (2004). Italianismos na língua portuguesa contemporânea. *Revista de Italianistica*, 9, 29–36. <https://doi.org/10.11606/issn.2238-8281.v0i9p29-36>

- Assis-Peterson (2008). Como ser feliz no meio de anglicismos: processos translógicos e transculturais. *Trabalhos em Linguística Aplicada*, 47(2), 323–340. <https://doi.org/10.11606/issn.2238-8281.v0i9p29-36>
- Coimbra, R. L. & Moutinho, L. C. (2009). É uma empresa portuguesa com certeza. Um estudo onomástico. In A. M. Ferreira (Ed.), *Lusofilias* (pp. 235–254). Aveiro: UA Editora.
- Coimbra, R. L. & Moutinho, L. C. (2012a). Memória real em empresas e instituições portuguesas. Um estudo onomástico. In P. Petrov, P. Q. Sousa, R. L. Samartim, & E. J. Torres Feijó (Eds.), *Avanços em ciências da linguagem* (pp. 288–310). Santiago de Compostela/Faro: Associação Internacional de Lusitanistas / Através Editora.
- Coimbra, R. L. & Moutinho, L. C. (2012b). Perfumaria Paris e Pizaria Romana: Topónimos europeus em nomes de empresas em Portugal. *Revista da Universidade de Aveiro – Letras*, 1, 349–359. <https://doi.org/10.34624/rual.v0i1.9587>
- Corrêa-Cardoso, J. N. (2016). Estudo do contacto de línguas em ambiente urbano. In J. Corrêa-Cardoso & M. C. Fialho (Eds.), *A linguagem na pólis* (pp. 77–97). Coimbra: Imprensa da Universidade de Coimbra.
- Dicionário Priberam da Língua Portuguesa. (2008-2016). Editora Priberam, Portugal. Disponível em <http://www.priberam.pt/dlpo/dlpo.aspx>
- Guerrero, N. E. C. (2008). Percepción del valor de la marca en restaurantes a traves de tres modelos de evaluación. *Revista de Administração da UNIMEP*, 6(2). <https://doi.org/10.15600/1679-5350/rau.v6n2p1-25>
- Moutinho, L. C. & Coimbra, R. L. (1997). VEJAKY sobre desvios linguísticos nos nomes das lojas. *Revista da Universidade de Aveiro – Letras*, 14, 127–141.
- Moutinho, L. C. & Coimbra, R. L. (1998). O nome é a alma do negócio: Estudo linguístico dos nomes das lojas em Portugal. In *Atas do XIII Encontro da Associação Portuguesa de Linguística* (Vol. II, pp. 93–104) Lisboa: Colibri.
- Moutinho, L. & Coimbra, R. L. (2019). Sociolinguística urbana: um estudo de caso na cidade do Porto. In A. M. Ferreira, C. Morais, M. F. Brasete & R. L. Coimbra (Eds.), *Pelos mares da língua portuguesa 4* (pp. 471–481). Aveiro: UA Editora.
- Piel, J.-M. (1989). Origens e estruturação histórica do léxico português. In J.-M. Piel (Ed.), *Estudos de linguística histórica galego-portuguesa* (pp. 9–16). Lisboa: IN-CM.
- Raposo, E. B. P. & Nascimento, M. F. B. (2013). Nomes próprios. In E. B. P. Raposo, M. F. B. Nascimento, M. A. C. Mota, L. Segura & A. Mendes (Eds.), *Gramática do português* (Vol. I, pp. 993–1041). Lisboa: Fundação Calouste Gulbenkian
- Sabbatini, R. M. E. (2007). *As contribuições do idioma italiano ao português: Estrangeirismos que ficaram*. Disponível em <http://www.renato.sabbatini.com/papers/italianismos.htm>
- Valadares, F. B. & Bertozzi, A. L. G. (2015). Estrangeirismos em propagandas de revistas brasileiras: Usos/abusos?!. *Dominios de Lingu@gem*, 9(59), 247–265. <https://doi.org/10.14393/DLE-v9n5a2015-13>
- Vilela, M. (1995). *Léxico e gramática*. Coimbra: Almedina.
- Villalva, A. (2008). *Morfologia do português*. Lisboa: Universidade Aberta.

[recebido em 24 de janeiro de 2021 e aceite para publicação em 27 de fevereiro de 2022]

## INSTITUIÇÕES DE FORMAÇÃO DE PROFESSORES COMO COMUNIDADES DE ESCRITA: ESTUDO COM BASE NOS TÍTULOS DOS RELATÓRIOS DE MESTRADO

HIGHER EDUCATION INSTITUTIONS AS ACADEMIC WRITING  
COMMUNITIES:  
A STUDY BASED ON THE ANALYSIS OF MASTER REPORT TITLES

Luís Filipe Barbeiro\*

luis.barbeiro@ipleiria.pt

José António Brandão Carvalho \*\*

jabrandao@ie.uminho.pt

Célia Barbeiro \*\*\*

celiapb7@gmail.com

As escolhas e decisões na escrita académica incluem o próprio título. As escolhas podem envolver estruturas linguísticas, marcas de pontuação, estratégias comunicativas, aspetos gráficos. Elas refletem-se na adequação aos objetivos sociocomunicativos do género e na projeção alcançada na comunidade académica. O presente estudo incide sobre os títulos dos relatórios de mestrado em quatro instituições de ensino superior portuguesas de formação de educadores de infância e professores. O objetivo do estudo é observar se as características dos títulos diferenciam as instituições, apontando para a existência de comunidades próprias de implementação destes mestrados e de escrita académica. Para a realização do estudo, reuniu-se um *corpus* de 800 relatórios de mestrado e efetuou-se a análise comparativa dos títulos quanto ao número de elementos, estrutura, estratégias comunicativas, presença no título das dimensões de prática profissional e de investigação e outros indicadores facultados pela Linguística de *Corpus*, como frequência de palavras e palavras-chave. Os resultados revelam diferenças entre as instituições quanto à predominância das escolhas em relação aos títulos, designadamente quanto ao número de elementos, estruturas sintático-semânticas, estratégias, palavras-chave e saliência dada às dimensões pedagógica e investigativa. Estes resultados indicam que as instituições se constituem como comunidades locais, nas quais predominam determinadas escolhas. As diferenças remetem para o seu funcionamento enquanto comunidade científica de formação, comunidade discursiva e também comunidade de escrita académica. Nesta última dimensão, o desafio é que a comunidade de cada instituição, sem anular a sua identidade, potencie as funções que quer atribuir aos títulos.

---

\* Escola Superior de Educação e Ciências Sociais, Politécnico de Leiria, Leiria, Portugal / CELGA-ILTEC, Universidade de Coimbra, Coimbra, Portugal. ORCID: 0000-0001-5798-2904

\*\* Centro de Investigação em Educação - CIEd, Instituto de Educação, Universidade do Minho, Braga, Portugal. ORCID: 0000-0003-0876-3101

\*\*\* Escola Dr. Correia Mateus, Leiria, Portugal / CELGA-ILTEC, Universidade de Coimbra, Coimbra, Portugal. ORCID: 0000-0002-7277-9573

**Palavras-chave:** Títulos. Escrita. Escrita Académica. Relatórios de mestrado. Formação de Professores.

Choices and decisions made when writing academic genres involve several aspects, including those concerning the title, such as linguistic structures, punctuation marks, communicative strategies, length. These aspects impact the achievement of the socio-communicative objectives of the genre and the projection the text may attain among the academic community. This study focuses on the titles of master reports produced by students of pre-service teacher training master programmes in four Portuguese Higher Education institutions. In addition to characterizing the titles and identifying their predominant features, it aims at verifying whether such features can be considered as identity and distinctive marks of each institution, viewed as an academic writing community. The study, based on a corpus of 800 master reports, consisted of a comparative analysis of the titles, considering the following aspects: number of elements; structure; communicative strategies; length; the reference to pedagogical and/or research dimensions; other indicators provided by Corpus Linguistics, such as word frequency, and keywords. According to the results, there are differences between the HE institutions regarding some aspects of the titles: number of elements, syntactic-semantic structures, strategies, keywords and the salience of pedagogical and research dimensions. The predominance of certain choices, according to the HE institution, may suggest that each institution constitutes itself as a local academic writing community with its identity characteristics. The differences point to their functioning as scientific communities, discursive communities and also academic writing communities. Concerning this last dimension, by being aware of these characteristics, each community may enhance the functions assigned to the titles.

**Keywords:** Title. Writing. Academic writing. Master reports. Teacher training.

•

## 1. Introdução

Na sequência das alterações trazidas pelo designado “Processo de Bolonha”, a formação inicial de educadores de infância e de professores do ensino básico e secundário passou a incluir o segundo ciclo de estudos de ensino superior, ou seja, o mestrado (conforme estabelece o regime jurídico de habilitação profissional para a docência na educação pré-escolar e nos ensinos básico e secundário). Este ciclo de estudos inclui, para a respetiva conclusão e, conseqüentemente, para a obtenção da habilitação profissional como educador ou professor, a redação e defesa de um relatório relativo à componente de prática de ensino desenvolvida durante a formação.

A elaboração do documento implica escolhas e decisões em variados níveis, nas vertentes de conteúdo e também de construção linguística desse conteúdo (Martin, Maton, & Doran 2020; O’Donnell 2013). Essas escolhas e decisões incluem o próprio título.

O contexto para a realização das escolhas no processo de elaboração do relatório ou outro documento final do mestrado, como a dissertação, no âmbito da escrita académica, não se circunscreve ao estudante, mas envolve o(s) orientador(es) e está ligado à própria instituição, pela coexistência de uma pluralidade de estudantes e de

orientadores a encarar a tarefa e pelas práticas instituídas num determinado momento ou ao longo do tempo.

Considerando o enquadramento global do regime jurídico e do processo de Bolonha, mas também a margem existente para cada instituição concretizar a prática de ensino supervisionada e respetivo relatório, ao cabo de uma dezena de anos de elaboração de relatórios de mestrado neste campo, é relevante investigar as suas características. O presente estudo incide sobre as escolhas que emergem nos títulos. Para além do objetivo geral de contribuir para a sua caracterização, pretende observar se as características dos títulos diferenciam as instituições, apontando para a existência de comunidades próprias de concretização destes mestrados e de escrita académica.

## **2. Enquadramento**

### **2.1. Enquadramento contextual: O relatório de Prática de Ensino Supervisionada na formação de professores**

Na sequência das alterações trazidas pelo “Processo de Bolonha”, o sistema de formação de professores em Portugal passou a estar organizado em dois ciclos de estudo: um primeiro ciclo, correspondente à licenciatura, e um segundo ciclo, correspondente ao mestrado (Decreto-Lei n.º 43/2007, de 22 de fevereiro; Decreto-Lei n.º 79/2014 de 14 de maio).

Conforme é explicitado na fundamentação do Decreto-Lei n.º 79/2014, “ao primeiro ciclo, a licenciatura, cabe assegurar a formação de base na área da docência.”, salientando-se, para o segundo ciclo, as funções de “assegurar um complemento dessa formação que reforce e aprofunde a formação académica, incidindo sobre os conhecimentos necessários à docência nas áreas de conteúdo e nas disciplinas abrangidas (...)” e de “assegurar a formação educacional geral, a formação nas didáticas específicas da área da docência, a formação nas áreas cultural, social e ética e a iniciação à prática profissional, que culmina com a prática supervisionada.”

Entre os princípios que regem a componente de iniciação à prática profissional, incluem-se os de facultar aos formandos “experiências de planificação, ensino e avaliação”, de articular “o conhecimento e a forma de o transmitir visando a aprendizagem” e de ser “concebida numa perspetiva de desenvolvimento profissional dos formandos” e de promover “uma atitude orientada para a permanente melhoria da aprendizagem” (Decreto-Lei n.º 79/2014). A prática supervisionada é objeto de relatório final exigido para a conclusão do mestrado. Esse relatório deverá evidenciar a capacidade, associada ao ciclo de estudos de mestrado, de “comunicar as suas conclusões e os conhecimentos e raciocínios a elas subjacentes, quer a especialistas e não especialistas, de uma forma clara e sem ambiguidades” (Quadro Nacional de Qualificações do Ensino Superior 2011, p. 59). As escolhas relativas ao título do relatório podem ser consideradas no âmbito da estratégia de comunicação.

A elaboração do relatório tem, por conseguinte, como referência os princípios referidos, orientadores da iniciação à prática profissional, e os próprios princípios estabelecidos para o ciclo de estudos correspondente ao mestrado (Decreto-Lei, 74/2006).

Estes princípios, baseados nos descritores de Dublin (Joint Quality Initiative Informal Group 2004) estabelecem a relação com a atividade de investigação, pelo que esta componente deverá emergir no relatório, em ligação ao processo de ensino-aprendizagem, como é salientado por Ponte (2006).

A inclusão da dimensão de investigação na iniciação à prática profissional e a sua articulação com as outras dimensões, no âmbito da concretização do processo de Bolonha, tem sido analisada e discutida (Arcadinho, Folque, & Costa 2020; Flores 2018; Vieira, Flores, Silva, Almeida, & Vilaça 2020a; Vieira, Flores, & Almeida 2020b; Vieira, Silva, & Vilaça 2021). De um modo geral, embora apontando limitações, as conclusões dos estudos continuam a evidenciar as potencialidades dessa inclusão e articulação, associando-a à condição de professor-investigador (Vieira *et al.* 2020a), ou seja, à capacidade de fundar a ação pedagógica na investigação das próprias práticas (Vieira *et al.* 2021).

Sendo o título um lugar de denominação e saliência, no quadro da questão da presença e articulação das dimensões de prática profissional e de investigação nestes mestrados (Alves, Silva, & Silva 2017; Arcadinho *et al.* 2020; Leal & Leite 2017), o presente estudo, para além de evidenciar os recursos linguístico-estruturais utilizados, coloca também o foco no relevo que estas dimensões adquirem nos títulos atribuídos aos relatórios de mestrado.

## 2.2. Enquadramento concetual: os títulos como domínio de escolhas e de estratégia

A função básica do título é permitir referir a obra, identificá-la. Numa perspetiva estrita, para cumprir essa função, bastaria a aposição de uma marca externa, como qualquer número arbitrário, à obra em si, sem qualquer relação com o conteúdo. Contudo, de forma geral, não foi esta a via que foi sendo construída, para a inscrição na pequena etiqueta, designada *titulus*, presa à extremidade do rolo de madeira em torno do qual se encontrava enrolado o papiro, constituindo o *volumen*. Embora exterior ao *volumen*, a inscrição no *titulus*, étimo de título, estabelecia a relação com a obra, dando informações sobre ela. Essa inscrição permitia conhecer, designadamente, o autor ou o assunto, sem necessidade de desenrolar o *volumen* (Hoek 1981). A autonomização dos elementos inscritos fez também emergir concetualmente o título e estabeleceu-o como domínio de escolhas linguísticas e estratégicas em relação às funções que é chamado a desempenhar (Barbeiro, Pereira, Coimbra, & Calil 2020; Hoek 1981; Vicea 2003), em conjugação com a saliência que adquiriu. Essa saliência é concetual pelos elementos selecionados para o título e é também visual, pela autonomização gráfica na obra e na página (Sharples 1999).

Por sobre a dimensão correspondente à função primordial do título de referir ou designar a obra, em si, de lhe dar um nome (Camprubi 2002), as escolhas quanto ao nome que é atribuído ativam outras funções, como a de informar globalmente sobre o conteúdo ou de atrair o leitor (Hartley 2005, 2007; Hoek 1981). Estas duas dimensões não se manifestam apenas no domínio das obras literárias (Genette & Crampé 1988; Paquin 2008; Roy 2008), mas emergem também no domínio da escrita académica (Hartley 2005, 2007; Garcia, Gattaz, & Gattaz 2019; Lamanauskas 2019; Whissel 2012). Em cada um destes domínios, em associação a essas funções, abre-se uma amplitude de possibilidades.

No domínio da escrita académica, Hartley (2007, 2008) enumera mais de uma dezena de tipos de títulos, orientados predominantemente para uma ou para outra destas funções. Em relação à orientação para informar, considera as seguintes possibilidades: podem anunciar o assunto geral; particularizar um aspeto colocado como subtítulo; apresentar a pergunta a que o texto dá a resposta; indicar os resultados; anunciar que a resposta a uma pergunta implícita ou explícita vai ser dada no texto; indicar o sentido da argumentação apresentada pelo autor; enfatizar a metodologia usada na pesquisa; anunciar a apresentação de categorias em contraste. A orientação para atrair a atenção inclui também diferentes possibilidades, tendo por base as estratégias de: surpreender por meio de aparentes incoerências; atrair por meio de aliterações; alusões literárias; trocadilhos; sugestões ou evocações (Hartley 2007, pp. 96–98; Hartley 2008, pp. 23–25).

As múltiplas possibilidades para a construção dos títulos, no campo da escrita académica, são também evidenciadas por Sahragard e Meihami (2016). Estes autores analisam as componentes de informação que emergem nos títulos de artigos de revistas da área da linguística aplicada. A saliência no título pode ser dada a diferentes componentes da investigação: metodologia, tópico, resultados, dados analisados e conclusões. Jamali e Nikzad (2011) estabelecem três categorias para a sua análise: os descritivos ou neutros (que apenas indicam o tópico ou assunto); os declarativos (que, para além do assunto incluem a conclusão do estudo); e os interrogativos (que apresentam o tópico sob a forma de uma pergunta, para captar a curiosidade do leitor).

Para além da análise funcional, com relevo para as funções de informar e atrair, as propriedades linguístico-estruturais dos títulos têm também estado em foco, em textos de diversos campos (Coimbra 1999, para o campo jornalístico; Barbeiro *et al.* 2020, Calil 2020; Pereira, Coimbra, & Calil 2017, para as histórias das crianças), incluindo o campo da escrita académica (Hudson 2016; Soler 2007; Whissel 2012). Neste campo, a descrição estrutural tem sido cruzada com a procura de contrastes entre títulos de diferentes áreas científicas, como as ciências biológicas e as ciências sociais, e de géneros textuais, como o artigo de investigação e o de revisão (Soler 2007). Na sua análise, Soler (2007) encontrou peculiaridades ligadas ao cruzamento destes eixos, com destaque para o facto de os títulos correspondentes a frases emergirem de forma saliente nos artigos de investigação na área das ciências biológicas. Outro eixo de pesquisa tem-se orientado para a procura de relação entre as características dos títulos e a saliência alcançada perante os destinatários/leitores, medida por indicadores da métrica relativa aos artigos publicados, como o número de leituras, de descarregamentos, de citações, ou a repercussão nos media (Fox & Burns 2015; Guo, Ma, Shi, & Zhong 2018; Jamali & Nikzad 2011; Zagorova, Weller, Janosov, Wagner, & Peters 2018) ou por questionários de apreciação da sua efetividade (Hallock & Bennett 2021). Entre as características linguístico-estruturais que têm sido consideradas, encontram-se: a extensão (número de palavras); o número de elementos (título principal, subtítulos); a natureza sintática (nominal, verbal, etc.); a natureza discursiva (afirmativa, interrogativa, exclamativa); o recurso a estratégias particulares (citações, metáforas, etc.).

Os resultados destes estudos evidenciam tendências em relação a alguns indicadores, como o tamanho, sem que a evolução seja unidirecional (Fox & Burns 2015; Guo *et al.* 2018). Jamali e Nikzad (2011) encontram diferenças entre o tipo de títulos e o

número de descarregamentos e citações, mas não existe convergência quanto a estes dois indicadores, constatando-se, por exemplo, que títulos sob a forma de perguntas tendem a ser mais descarregados, mas menos citados.

Hallock e Bennett (2021) não partiram dos produtos publicados, mas focaram-se na perspectiva dos destinatários e potenciais autores, os estudantes universitários de licenciaturas procurando apreender as suas preferências. Estas privilegiaram os títulos longos em relação aos curtos, os que contêm mais do que um elemento separados pelo sinal de pontuação “dois pontos” e os títulos sob a forma de perguntas.

Embora os resultados encontrados pelos diferentes estudos não sejam generalizáveis, pois podem ter por base apenas uma área ou mesmo uma revista (Fox & Burns 2015), e os efeitos não sejam absolutos ou simples, mas complexos, os estudos têm a virtualidade de mostrar a amplitude de possibilidades de recursos e de vias linguístico-discursivas para a construção dos títulos.

O presente estudo não tem como propósito prosseguir ou aprofundar a análise em relação aos eixos considerados pelos artigos referidos, mas efetuar o cruzamento entre as características dos títulos e um novo eixo, o correspondente às instituições de formação, observando se, também pela via das decisões e escolhas relativas aos títulos, essas instituições podem constituir-se como comunidades de escrita.

Por comunidade de escrita, podemos entender a projeção para a construção de textos escritos da comunidade discursiva. Esta, tendo como referência Swales (1990), assenta em objetivos comuns, mecanismos de participação e de troca de informações, uso de géneros textuais e de termos específicos, interação que envolve elementos experientes e aprendizes. As comunidades discursivas desenvolvem modos específicos de se referirem ao mundo que observam e investigam (Creme & Lea 2000; Vásquez 2001). Em relação à comunidade de escrita, Graham (2018) propõe o modelo *Writer(s)-Within-Community*. Este modelo defende que a escrita é simultaneamente moldada e vinculada pelas características, capacidade e variabilidade das comunidades nas quais é produzida e pelas características cognitivas, capacidades e diferenças individuais de quem a produz. Na base do reconhecimento do papel da comunidade, encontra-se a assunção de que a escrita é uma atividade social que se realiza dentro de contextos específicos, contextos que Graham (2018) faz corresponder a comunidades de escrita. Nesses contextos, podem emergir como traços caracterizadores da comunidade orientados para a escrita, objetivos, instrumentos, atividades desenvolvidas, tipos de participantes, papéis por estes desempenhados, funcionamento em ambientes físicos e sociais específicos, e também produtos. No presente estudo, desde logo, existem alguns elementos comuns às instituições cujos títulos de relatórios de mestrado são analisados, por exemplo, os objetivos destes relatórios, o enquadramento jurídico e profissional em que são produzidos, os papéis de participantes como estagiários e orientadores. O estudo coloca o foco num dos elementos do produto, o título, procurando apreender se são preponderantes as características comuns entre as instituições ou se este elemento se torna diferenciador entre elas.

Os conceitos de comunidade discursiva e de comunidade de escrita colocam o foco na vertente da língua. Considerámos que a comunidade de escrita recebia a projeção para a construção dos textos escritos de traços da comunidade discursiva. Por sua vez, a

comunidade discursiva pode receber a projeção de traços ou escolhas com um alcance mais alargado, que vão para além da língua. Estes traços incluem as correntes ou perspectivas pedagógicas e de investigação dos docentes (designadamente no seu papel de supervisores de prática pedagógica e/ou de orientadores dos relatórios), os campos em que realizam investigação, os projetos e rede de relações em que a instituição e os docentes estão envolvidos. Esta dimensão de comunidade alargada a vertentes não linguísticas projeta-se na comunidade discursiva e na comunidade de escrita. Consideramo-la aqui sob a designação de comunidade científica de formação, que assume áreas e características próprias que a especificam no âmbito da comunidade académica mais vasta que se dedica à formação de professores.

### 3. Metodologia

#### 3.1. *Corpus*

O *corpus* sobre o qual recai a análise é constituído por 800 títulos de relatórios de mestrado, repartidos por quatro instituições de formação de professores. Foram recolhidos 200 títulos em cada instituição. A instituição A e a instituição B são duas Escolas Superiores de Educação, ou seja, integram-se no subsistema de ensino politécnico. As instituições C e D estão integradas em universidades.

Os cursos de mestrado em causa, destinados à formação de educadores de infância e de professores do ensino básico e secundário, são oferecidos pelas instituições na sequência da licenciatura em Educação Básica, no caso da formação de educadores de infância e professores dos 1.º e 2.º ciclos do ensino básico, ou de licenciaturas nas áreas das respetivas especialidades, no que respeita aos professores do 3.º ciclo do ensino básico e do ensino secundário. A formação de educadores e professores até ao segundo ciclo do ensino básico tem lugar no ensino politécnico; no ensino universitário, para além dessa, realiza-se também a dos professores do terceiro ciclo do ensino básico e do ensino secundário. As instituições universitárias integradas no estudo oferecem cursos destinados à formação desde a educação pré-escolar até ao ensino secundário.

Os títulos que constituem o *corpus* repartem-se pelos seguintes níveis e cursos: educação pré-escolar, de forma autónoma (205 títulos); ensino do primeiro ciclo do ensino básico, de forma autónoma (4); educação pré-escolar e ensino do primeiro ciclo do ensino básico, em conjunto (329); ensinos do primeiro e do segundo ciclos do ensino básico, nas quatro áreas do segundo ciclo: Português, História e Geografia de Portugal, Matemática e Ciências da Natureza, em conjunto (100); ensino do primeiro e do segundo ciclos do ensino básico nas áreas Português, História e Geografia de Portugal do segundo (36); ensino do primeiro e do segundo ciclos do ensino básico nas áreas de Matemática e Ciências da Natureza do segundo ciclo (65); ensino de Inglês e de outra língua estrangeira no ensino básico (14); ensino em diferentes áreas do terceiro ciclo do ensino básico e do ensino secundário (47).

O acesso aos títulos e respetiva recolha foram realizados através dos repositórios institucionais na Internet. Efetuou-se a recolha a partir dos resultados das consultas nos repositórios, recorrendo-se a palavras-chave ligadas às designações dos cursos, quando

não se encontravam organizados por cursos. A constituição foi sendo feita, recuando na data de defesa do relatório até se perfazer uma amostra de 200 por cada instituição. Os títulos recolhidos vão desde 2021, até ao ano de 2013, como data mais recuada; o *corpus* cobre, por conseguinte, em larga medida, o período de vigência do novo regime jurídico da habilitação profissional para a docência na educação pré-escolar e nos ensinos básico e secundário.

Além da consulta dos títulos nos repositórios, acedeu-se também aos resumos e aos próprios documentos. Esta consulta foi realizada para confirmação do tópico que foi objeto de investigação e da coocorrência do título, na capa e/ou na folha de rosto, com outros elementos suscetíveis de contextualizarem o próprio título. Assim, os documentos de cada instituição seguem modelos adotados no seu âmbito, que, nos casos das instituições A e B, incluem na capa a indicação do mestrado e do tipo de documento elaborado no seu âmbito. Nos modelos seguidos nas instituições C e D, essa informação é apresentada na folha de rosto, limitando-se a capa à indicação do autor e do título (além dos elementos relativos à identidade visual da instituição e ao ano). A indicação do tipo de documento tem por base a palavra “relatório”, apresentando variantes quanto aos elementos que o complementam, designadamente: “... de mestrado”, “... de prática de ensino supervisionada”, “... de estágio”. A contextualização, no conjunto da formulação em que surge esta indicação, pode ainda incluir outros elementos como a denominação da instituição e o orientador ou orientadores.

### 3.2. Análise

Uma vez constituído o *corpus*, os títulos foram analisados quanto aos seguintes parâmetros: número de elementos do título, extensão, categoria sintática, funções e estratégias proeminentes e dimensões projetadas nos títulos. Em relação aos elementos do título, tomaram-se como marcadores de fronteira os sinais de pontuação dois pontos, ponto final, ponto de exclamação, ponto de interrogação e travessão e a configuração gráfica de mudança de linha. O primeiro elemento foi tomado como o título principal e os restantes como subtítulos. Alguns repositórios já apresentam a separação entre títulos e subtítulos. A extensão foi determinada com base no número de palavras. A categoria sintática foi estabelecida ao nível dos sintagmas ou grupos sintáticos e ao nível da frase.

A determinação das funções e estratégias tomou como base as funções correspondentes a informar e a captar a atenção do leitor. No âmbito da função de informar, pode ser adotada uma perspetiva referencial (focada em diversos elementos ou componentes do estudo), a perspetiva processual (focada num processo associado ao trabalho desenvolvido, que pode incluir a elaboração do relatório) e a perspetiva declarativa (focada na conclusão decorrente desse trabalho).

Em relação à função de captar a atenção do leitor, o aprofundamento da análise procurou apreender as estratégias adotadas (Hartley 2007), como o recurso a citação, pergunta, jogos de linguagem, etc.

Considerando a relevância para estes relatórios das dimensões relativas à prática profissional e à investigação, conforme expresso no enquadramento, fez-se incidir a análise também sobre a emergência destas dimensões nos títulos. Para tal, quando a

formulação do título não se revelava suficientemente elucidativa (por exemplo, quanto ao objeto de estudo da investigação), procedeu-se à confirmação por meio da consulta do resumo e do documento, a partir do acesso facultado pelo repositório.

A análise procurou caracterizar os títulos tendo por base os indicadores referidos, de um modo global, e colocando em contraste as diversas instituições, a fim de verificar se existem diferenças entre elas. Para determinar a existência de diferenças significativas em relação ao número de elementos e ao número de palavras, procedeu-se à análise estatística, por meio de testes de análise de variância unifatorial e de testes *post-hoc* (calculados com recurso ao programa SPSS, versão 28).

A análise baseou-se ainda em instrumentos da Linguística de *Corpus*, designadamente a frequência das palavras e a identificação de palavras-chave (para o que se recorreu aos programas *AntConc* (Anthony 2021). A identificação das palavras-chave permite encontrar as palavras que são significativamente mais frequentes num *corpus* particular, em comparação com um *corpus* de referência (“we can therefore say that keywords are words that are typical of the corpus of interest when compared to another corpus”, Brezina 2018, pp. 79–80).

Para encontrar as palavras-chave correspondentes ao *subcorpus* de cada instituição, tomámos o respetivo *subcorpus* por si, enquanto *corpus-alvo* ou *corpus de interesse*, e comparámo-lo com o conjunto dos *subcorpora* das outras instituições, tomados no seu conjunto como *corpus* de referência). O facto de os *corpora* em comparação serem da mesma natureza (títulos dos relatórios de mestrado) e de o *corpus* de referência ser maior que o *corpus-alvo* reforça a solidez da comparação (“Generally speaking, the larger and the more similar the reference corpus is to the corpus of interest the more reliable and focused the comparison is.” (Brezina 2018, p. 81).

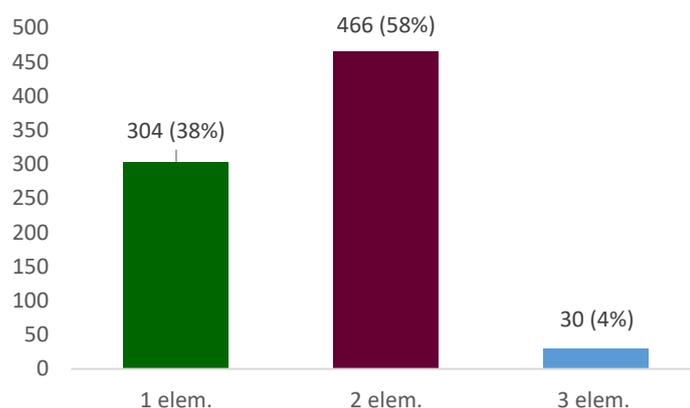
## 4. Resultados

Apresentam-se os resultados repartidos por duas secções. A primeira é dedicada aos constituintes linguístico-estruturais dos títulos. A segunda incide sobre as funções e estratégias que evidenciam.

Os resultados são apresentados articulando uma perspetiva global com a discriminação em relação a cada instituição, a fim de permitir a observação de diferenças entre elas.

### 4.1. Constituintes linguístico-estruturais

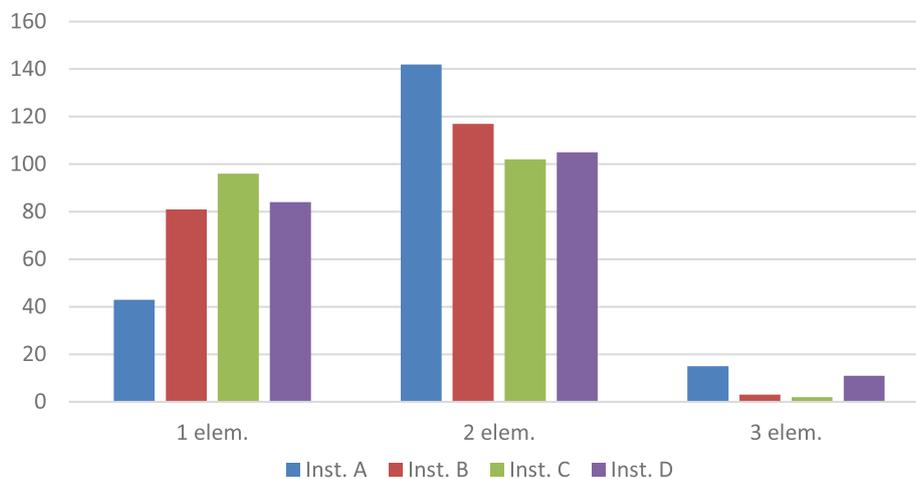
O primeiro indicador cujos resultados são apresentados é relativo ao número de elementos dos títulos, ou seja, à sua subdivisão título principal e subtítulo(s). O Gráfico 1 apresenta os resultados globais, quanto ao número de elementos.



**Gráfico 1. Número de elementos constituintes do título: valores globais**

Como se observa no Gráfico 1, predomina a estrutura constituída por dois elementos: título principal e subtítulo. Esta estrutura representa mais de metade dos casos (58%). Para além desta, a estrutura constituída apenas por um elemento ocupa também uma posição relevante (38%). Os títulos com três elementos, ou seja, com dois subtítulos, representam uma proporção muito mais reduzida (4%).

Tendo como pano de fundo este panorama global, o Gráfico 2 permite observar se as suas tendências se mantêm quando se considera cada instituição por si.



**Gráfico 2. Número de elementos constituintes do título: valores por instituição**

Nas quatro instituições, predomina a categoria de títulos com dois elementos, seguido da categoria apenas com um elemento. Contudo, apesar desta tendência geral quanto à ordenação das categorias, os valores atingidos por cada categoria nas instituições são diferentes, o que se reflete no número médio de elementos: 1,9 para a instituição A; 1,6 para a instituição B; 1,5 para a instituição C; e 1,6 para a instituição D. A análise de variância revela a existência de diferenças significativas quanto ao valor médio de elementos ( $F_{3,796} = 14,084$ ,  $p = 0,000$ ,  $\eta^2 = 0,050$ ). Essa diferença encontra-se entre a

instituição A e as restantes, tal como é revelado pelo teste *post-hoc* de comparações múltiplas (teste de Tamhane T2), para amostras sem igualdade de variâncias.

O indicador seguinte consiste no número de palavras, cujos resultados são apresentados na Tabela 1. Como se pode observar, os valores médios variam entre 20,3 para a instituição A e 12,6 para a instituição C.

**Tabela 1. Número de palavras dos títulos (valor médio)**

Instituição	N.º de palavras
A	20,3
B	14,6
C	12,6
D	16,4
Global	16,0

A análise de variância revela diferenças significativas ( $F_{3,796} = 68,436$ ,  $p = 0,000$ ,  $\eta^2 = 0,250$ ). Essas diferenças surgem entre os valores de todas as instituições (teste de Tamhane T2). Ainda que a variável independente não seja ordenada, se tivermos em conta o tamanho do efeito, observa-se que é reduzido no caso do número de elementos, passando a poder ser considerado médio no que diz respeito ao número de palavras (Amoroso 2018).

No que diz respeito à natureza sintática dos elementos, os resultados são apresentados na Tabela 2, considerando os valores globais e discriminados por elemento e instituição.

Globalmente, a natureza nominal é predominante, correspondendo a quase três quartos dos elementos que constituem os títulos (74%). Complementam os resultados os elementos que correspondem a Grupos Verbais (18%), a Grupos Preposicionais (5%), a Frases (3%), e a Grupos Adverbiais (0,2%). Os elementos inseridos na categoria de GV apresentam uma predicação enunciada de forma genérica, sem explicitar a função de sujeito e iniciando-se pela forma verbal, que surge no gerúndio (por exemplo: *Refletindo sobre...*, *Construindo...*, *Revelando...*) ou no infinitivo (*Refletir...*, *Aprender...*). Por sua vez, os elementos que se apresentam sob a forma de Frases, para além da predicação verbal, incluem a função de sujeito (por exemplo: *Isso é coisa de miúdas!*, *Todos temos uma história:*). Nos elementos que apresentam a forma de Grupos Preposicionais, emerge sobretudo a construção que estabelece um percurso entre a origem e um destino ou ponto de chegada, em relação à escolaridade (por exemplo, *Da Creche ao Jardim de Infância...*, que apresenta um número elevado de ocorrências na categoria) ou em relação à construção do conhecimento (por exemplo: *Da Reflexão à Investigação*). Nas construções com base em Grupos Adverbiais, emerge o advérbio *além*, por exemplo em: *Além das quatro paredes*.

Tabela 2. Natureza sintática dos elementos do título

Inst.	Elem.	Estruturas					Total
		GN	GV	F	GP	GADV	
A	1.º	31	150		19		200
	2.º	120	20	2	15		157
	3.º	13			2		15
	<i>Subtotal</i>	<i>164</i>	<i>170</i>	<i>2</i>	<i>36</i>		<i>372</i>
	<i>(%)</i>	<i>(44%)</i>	<i>(46%)</i>	<i>(0,5%)</i>	<i>(10%)</i>		
B	1.º	147	13	29	8	3	200
	2.º	109	7		3		119
	3.º	3					3
	<i>Subtotal</i>	<i>259</i>	<i>20</i>	<i>29</i>	<i>11</i>	<i>3</i>	<i>322</i>
	<i>(%)</i>	<i>(80%)</i>	<i>(6%)</i>	<i>(9%)</i>	<i>(3%)</i>	<i>(1%)</i>	
C	1.º	173	16	7	4		200
	2.º	100	2	1	1		104
	3.º	2					2
	<i>Subtotal</i>	<i>275</i>	<i>18</i>	<i>8</i>	<i>5</i>		<i>306</i>
	<i>(%)</i>	<i>(90%)</i>	<i>(6%)</i>	<i>(3%)</i>	<i>(2%)</i>		
D	1.º	160	27	6	7		200
	2.º	108	3	1	5		117
	3.º	11					11
	<i>Subtotal</i>	<i>279</i>	<i>30</i>	<i>7</i>	<i>12</i>		<i>328</i>
	<i>(%)</i>	<i>(85%)</i>	<i>(9%)</i>	<i>(2%)</i>	<i>(4%)</i>		
<b>Global</b>		<b>977</b>	<b>238</b>	<b>46</b>	<b>64</b>	<b>3</b>	<b>1328</b>
<b>(%)</b>		<b>(74%)</b>	<b>(18%)</b>	<b>(3%)</b>	<b>(5%)</b>	<b>(0,2%)</b>	

Nota. GN = Grupo Nominal; GV = Grupo Verbal; F = Frase; GP = Grupo Preposicional; GADV = Grupo Adverbial.

Traçado o quadro global, coloca-se como anteriormente a questão de saber se ele se projeta nas quatro instituições ou se surgem quadros específicos. A distribuição dos elementos dos títulos pelas principais categorias sintáticas (GN, GV, F e GP), em cada uma das instituições, encontra-se também na Tabela 2. Os valores apresentados evidenciam um claro contraste entre a instituição A e as restantes. Enquanto nas outras são os elementos de natureza nominal que predominam, na instituição A é a categoria de Grupo Verbal que ocupa uma posição predominante, com relevo para o primeiro ou único elemento do título. Neste elemento, a categoria GV apresenta 150 ocorrências (75%, no conjunto dos 200 títulos do *subcorpus* desta instituição). Esta frequência contrasta fortemente com os valores das outras instituições (13, 16 e 27, no primeiro elemento, para as instituições B, C e D, respetivamente). Embora o alcance dos valores não seja tão acentuado, outros contrastes podem ser estabelecidos, com base na natureza sintática dos elementos: é o caso respeitante à ocorrência da categoria GP, que apresenta uma frequência mais elevada na instituição A, com mais de metade dos casos; é também o caso da categoria F que ocorre com maior frequência na instituição B, ocorrendo de forma diminuta nas instituições A, C e D.

A constituição do *corpus* permitiu integrar na análise indicadores habituais da Linguística de *Corpus*, como a frequência de palavras e as palavras-chave. Na Tabela 3,

são apresentadas as dez palavras mais frequentes nos *subcorpora* de cada uma das instituições (A, B, C e D).

Tabela 3. Palavras mais frequentes

Ord.	Instituição A		Instituição B		Instituição C		Instituição D	
	Palavra	Freq.	Palavra	Freq.	Palavra	Freq.	Palavra	Freq.
1	de	327	de	258	de	159	de	235
2	a	182	do	117	e	119	do	189
3	e	177	a	111	o	117	e	161
4	do	159	e	108	no	100	o	154
5	sobre	159	o	76	do	99	<i>ensino</i>	121
6	o	156	o	74	a	73	a	119
7	<i>refletindo</i>	140	da	73	<i>educação</i>	65	no	90
8	<i>prática</i>	125	no	58	da	64	<i>básico</i>	86
9	<i>pedagógica</i>	118	<i>infância</i>	56	o	57	da	73
10	em	109	<i>crianças</i>	51	<i>escolar</i>	55	<i>ciclo</i>	69
...								

Como é habitual, ocupam os lugares cimeiros da frequência as palavras funcionais como as preposições (a preposição *de* ocupa o lugar cimeiro em todos os *subcorpora*), os artigos, as conjunções (designadamente a conjunção *e*). Uma posição relevante é também ocupada pelo elemento correspondente à parte final da abreviatura do numeral ordinal presente em 1.º, 2.º, 3.º, relativos aos ciclos do ensino básico e que emergem nos títulos, de forma transversal às diferentes instituições. Contudo, a transversalidade deixa de existir, se, para além das palavras funcionais e do elemento de abreviatura do numeral ordinal, consideramos as palavras lexicais que emergem nos títulos (assinaladas em itálico). Em vez da similaridade, observa-se que as palavras lexicais que emergem como mais frequentes são diferentes em cada instituição (A: *refletindo*, *prática* e *pedagógica*; B: *infância*, *crianças*; C: *educação*, *escolar*; D: *ensino*, *básico*, *ciclo*). Evidencia-se, por conseguinte, o contraste entre as instituições quanto às palavras lexicais mais frequentes que emergem nos títulos, não constituindo escolhas comuns, apesar de o campo de formação ser largamente partilhado entre as instituições.

O contraste encontrado é confirmado pelo indicador correspondente às palavras-chave. Tomando o conjunto dos títulos das outras instituições como *corpus* de referência, o programa *AntConc* identificou as palavras apresentadas na Tabela 4 como palavras-chave positivas, ou seja, como sendo significativamente mais frequentes no *corpus* da instituição em causa, ou *corpus-alvo*.

A Tabela 4 encontra-se ordenada pelo valor do teste estatístico relativo à significância do contraste entre os *corpora*, quanto ao seu caráter-chave no *corpus-alvo* (o teste de Log Likelihood, representado na tabela por LL).

Pela observação da Tabela 4, conclui-se que o número de palavras-chave de cada instituição é diferente, variando entre quatro, no caso da instituição B, e 21, no caso da instituição C. Em termos de conteúdo, as palavras-chave presentes na tabela remetem para alguns eixos ou nichos de formação e prática pedagógica ou de investigação que assumem maior relevo nalgumas instituições: pode ser o caso da *creche* na instituição A, da *diversidade linguística*, *plurilingue* e *línguas* ou *cts*, *kamishibai*, *fish* e *peixe* na

instituição C, da *educação física e desporto* em D. A comparação com base nas palavras-chave revela, por conseguinte, áreas específicas de afirmação das instituições na comunidade académica alargada da formação de professores. Esta diferenciação corresponde, assim, à afirmação enquanto comunidade científica com facetas específicas.

Tabela 4. Palavras-chave

Ord	Instituição A		Instituição B		Instituição C		Instituição D	
	Palavra	LL	Palavra	LL	Palavra	LL	Palavra	LL
1	<b>refletindo</b>	327.1	é	20.4	diversidade	87.2	estágio	35.4
2	<b>sobre</b>	240.9	conceções	16.4	linguística	63.3	ensino	28.0
3	prática	189.0	grupo	16.1	ceb	49.4	ciclos	27.9
4	pedagógica	152.6	crianças	16.1	cultural	38.3	física	26.4
5	creche	78.2			cts	32.3	básico	25.6
6	ao	63.6			<b>estudo</b>	29.9	<b>relatório</b>	24.8
7	<b>investigando</b>	60.2			sensibilização	29.0	integrado	24.5
8	infância	52.2			no	27.1	motivação	19.4
9	<b>percurso</b>	34.4			pré	25.7	desporto	19.0
10	em	27.4			pensamento	22.3	do	18.9
11	<b>caminho</b>	18.5			escolar	20.1	secundário	16.3
12	jardim	16.9			plurilingue	19.4	inglês	16.1
13	aprendizagens	16.2			caso	19.2		
14	capacidade	16.1			um	17.7		
15					crítico	16.9		
16					didática	16.7		
17					fish	16.1		
18					kamishibai	16.1		
19					línguas	16.1		
20					peixe	16.1		
21					sustentável	16.1		

Nota. LL = Log Likelihood

Por outro lado, a comparação por meio do indicador constituído pelas palavras-chave também revela outros eixos de escolhas diferenciadoras em relação à denominação de entidades que podem considerar-se comuns às instituições, por estarem associadas ao campo de intervenção e formação que é partilhado. Estão neste caso palavras como *infância*, *jardim* (de infância), *prática*, *pedagógica*, *aprendizagens*, para a instituição A; *crianças*, *grupo* (de crianças), para a instituição B; a abreviatura *ceb*, *pré* e *escolar* (*pré-escolar*), *didática*, para a instituição C; *ensino*, *ciclos*, *básico*, *estágio*, para a instituição D. De forma diferente da vertente anterior (que reflete especificidades académicas e científicas) todos estes termos poderiam surgir em qualquer das instituições. A diferenciação refletida nestas palavras-chave corresponde, de algum modo, à afirmação da instituição enquanto comunidade discursiva, ou seja, uma comunidade que constrói, por meio do discurso, entre as possibilidades facultadas pela língua, as suas escolhas de recursos semióticos.

Numa terceira vertente, podemos considerar as opções diferenciadas quanto à construção do próprio título, o que se reflete na sua estrutura e estratégias e também nas palavras-chave. Esta vertente integra palavras-chave assinaladas a negrito na tabela, como *refletindo*, *investigando*, *sobre*, *percurso*, *caminho*, para a instituição A; *é* para a instituição B; *estudo* para a instituição C; *relatório* para a instituição D. Note-se que o

que está subjacente a estas palavras-chave é comum às instituições: a prática pedagógica e a realização de uma componente de investigação a ela associada. Em relação à vertente anterior, não está aqui em causa a construção diferenciada da referência ao campo de formação e investigação, mas a construção do objeto escrito, dando, por meio do título, saliência a alguns elementos da atividade que é objeto de relatório, ou seja, que o relatório constrói discursivamente. Esta diferenciação corresponde à afirmação da instituição como comunidade de escrita.

No caso da instituição A, as palavras *refletindo* e *sobre*, complementadas com *prática* e *pedagógica*, encontram-se também em evidência na Tabela 3, relativa à frequência de palavras. O contraste acentuado entre a instituição A e as restantes revela-se também em relação a estas palavras-chave pelo valor bastante mais elevado do resultado do teste estatístico de Log Likelihood. Estas palavras apresentam ainda a particularidade de surgirem frequentemente combinadas nos títulos nesta instituição. A formulação *Refletindo sobre a Prática Pedagógica* pode ser considerada como uma marca característica da instituição A em relação aos títulos dos relatórios de mestrado em foco no presente estudo. Embora não esgotando as ocorrências de cada um dos termos que a formam, esta formulação é largamente recorrente nesta instituição: apresenta 103 presenças em 200 títulos, ou seja, em mais de metade dos casos que constituem este *corpus*. Em contrapartida, não ocorre nos títulos das outras instituições. As palavras-chave *investigando*, *percurso* e *caminho* reforçam esta via de construção dos títulos, por meio do recurso ao gerúndio, expressando continuidade, duração, que também está presente em *percurso* e *caminho*.

No caso da instituição B, emerge a ocorrência como palavra-chave da forma verbal *é*. Esta saliência está relacionada com a formulação dos títulos sob a forma de frases, seja como perguntas, por exemplo, “*Como é que o brincar no recreio pode potenciar as aprendizagens?*”, ou sob a forma de declarações, por exemplo “*A amizade é felicidade*”. Nas instituições C e D, nesta vertente, ganha relevo a referência ao próprio objeto escrito, respetivamente como *estudo* ou *relatório*.

## 4.2. Funções, estratégias e dimensões

Em conformidade com os resultados apresentados na secção anterior, determinadas funções e estratégias surgem associadas predominantemente a determinadas instituições. Globalmente, considerando o título principal, predomina a função informativa, numa perspetiva ou vertente referencial, expressa por meio de um Grupo Nominal. A vertente referencial, pelo conteúdo que expressa, pode dar saliência a diferentes componentes ou elementos; entre eles, encontram-se: a indicação do objeto ou campo de estudo (259 ocorrências, e.g.: *A pontuação no 1º Ciclo do Ensino Básico*); a referência a uma atividade, estratégia ou dispositivo pedagógico (86 oc., e.g.: *Dinamização da hora do conto*; *O jogo na aprendizagem da geometria na educação pré-escolar e no 4.º ano de escolaridade*; *O roleplaying como estratégia para o ensino das migrações*); a expressão de um fator para a obtenção de um efeito (64 oc., e.g.: *O potencial das artes na aprendizagem e desenvolvimento da criança do 1º ciclo do ensino básico*) e valorização do fator em causa (15 oc., e.g.: *A importância das atividades experimentais no Ensino*

das Ciências no 1.º e 2.º Ciclos do Ensino Básico); a articulação de duas áreas no processo pedagógico (23 oc., e.g.: *Poesia e ciência no 1º Ciclo do Ensino Básico*); a incidência no objetivo de desenvolvimento (20 oc., e.g.: *O desenvolvimento da compreensão inferencial em alunos do 3º ano do 1º Ciclo do Ensino Básico*); a construção do título com uma designação meta-analítica (25 oc., e.g.: *Estudo sobre as diferenças interativas e comunicativas das educadoras e das mães; Tipologia de exercícios de expressão oral*, sendo adotados além de “estudo” termos como “olhar”) ou por meio de uma designação genológica (2 oc., *Relatório de prática pedagógica*). De uma forma quase generalizada, a perspetiva referencial é externa em relação ao objeto que emerge no título do relatório, no entanto, também ocorre a adoção de uma perspetiva pessoal que é expressa por meio do recurso à primeira pessoa (3 oc.; *O meu percurso da Creche ao Jardim de Infância; As minhas aprendizagens, da Creche ao Jardim de Infância*).

As possibilidades enunciadas no parágrafo anterior, em ligação à vertente referencial, repartem-se pelas diversas instituições, sem constituir marca saliente de uma delas. De modo inverso, a adoção da perspetiva processual, expressa por meio do Grupo Verbal, a que nos referimos anteriormente, emerge como uma marca preponderante nos títulos na instituição A.

Ligado à função de atrair o leitor, surge o recurso a diversos mecanismos. Entre eles, encontra-se a citação. O recurso à citação apresenta maior incidência na instituição B, que recebe 20 das 26 ocorrências, ocorrendo as restantes seis nas instituições C e D, mas não na instituição A. Frequentemente, trata-se de citações das próprias crianças (13 das 20 ocorrências, e.g.: *Quero escrever outras coisas...um texto*), o que não é encontrado nas restantes instituições.

A formulação do título como pergunta é também associada à função de despertar o interesse no leitor. Tal como a citação, é um recurso que ocorre com maior relevo na instituição B (que apresenta 13 das 16 ocorrências, e.g.: “Como é que o brincar no recreio pode potenciar aprendizagens?”), sendo as restantes nas instituições C e D).

Ainda no âmbito da função de atrair o leitor, ocorre o jogo verbal com a homofonia da palavra inglesa *fish* com a palavra portuguesa *fixe* (e.g.: *O peixe é fish do mar ao prato*). Este jogo sonoro é retomado em quatro títulos de relatórios, que têm o mesmo orientador, o que aponta para o papel potencial deste interveniente no processo, que pelo menos subscreveu a adoção deste recurso partilhado entre os relatórios.

Anteriormente, apresentámos os resultados relativos aos constituintes linguístico-estruturais (subsecção 4.1.) e à associação dos títulos às funções de informar e de atrair a atenção do leitor. Na Tabela 5, apresentam-se os resultados relativos à presença nos diversos elementos do título das dimensões respeitantes à prática profissional e à investigação.

O contraste encontrado anteriormente entre a instituição A e as restantes surge também de forma proeminente nos valores inscritos na Tabela 5, relativos à presença das dimensões de prática profissional e de investigação nos elementos do título. Enquanto na instituição A, a dimensão respeitante à prática profissional é predominante no título, nas outras instituições a presença desta dimensão é diminuta. No caso da instituição A, a dimensão de prática profissional ocupa uma larga proporção dos primeiros ou únicos elementos do título (154 ocorrências, nos 200 títulos, correspondendo a 77%).

Estabelecendo a relação com os resultados anteriores, esta presença da dimensão da prática profissional corresponde à categoria sintática GV e à função processual, de que a formulação *Refletindo sobre a Prática Pedagógica* é a mais recorrente.

**Tabela 5. Dimensões: prática profissional e investigação**

Inst.	Elem.	Dimensões			Total
		PP	I	PP+I	
A	1.º	154	18	28	200
	2.º	38	117	2	157
	3.º	2	13		15
	<i>Subtotal</i>	<i>194</i>	<i>148</i>	<i>30</i>	<i>372</i>
	(%)	(52%)	(39%)	(8%)	
B	1.º	6	194		200
	2.º	3	116		119
	3.º		3		3
	<i>Subtotal</i>	<i>9</i>	<i>313</i>		<i>322</i>
	(%)	(3%)	(97%)		
C	1.º	1	199		200
	2.º	2	102		104
	3.º		2		2
	<i>Subtotal</i>	<i>3</i>	<i>303</i>		<i>306</i>
	(%)	(1%)	(99%)		
D	1.º	1	199		200
	2.º	4	113		117
	3.º	8	3		11
	<i>Subtotal</i>	<i>13</i>	<i>315</i>		<i>328</i>
	(%)	(4%)	(96%)		
<b>Global</b>		<b>219</b>	<b>1079</b>	<b>30</b>	<b>1328</b>
(%)		(17%)	(81%)	(2%)	

Nota. PP = dimensão de Prática Profissional; I = dimensão de Investigação

A dimensão investigativa, nos títulos dos relatórios da instituição A, é predominantemente remetida para o segundo elemento do título. Além disso, alguns títulos procuram associar as duas dimensões (correspondendo na tabela à categoria PP+I), o que é feito, por exemplo, através da formulação *Refletindo e Investigando sobre a Prática Pedagógica*.

Nas restantes instituições, a dimensão investigativa é preponderante quer no primeiro, quer no segundo elemento do título. Frequentemente, após a enunciação do tema de estudo, no primeiro elemento, o segundo elemento apresenta elementos especificadores, como o objeto particular em foco ou o contexto em que se desenvolve a investigação (por exemplo: *Educação literária: o texto dramático no 1.º ciclo do ensino básico*). O terceiro elemento continua habitualmente o movimento de especificação (por exemplo, em *Alimentos: equilíbrio, hábitos e cuidados: uma intervenção pedagógica com alunos do 1.º e 2.º ciclos do Ensino Básico*), por vezes com indicação do tipo de estudo, designadamente quando se trata de estudo de caso (*Autonomia e bem estar em jardim de infância: um olhar da pedagogia Waldorf e abordagem Reggio Emilia: um estudo de caso*).

## 5. Conclusão

Os títulos constituem um campo de escolhas entre múltiplas possibilidades. Essas possibilidades tornam-se manifestas formalmente pelo número de elementos ou partes do título e pela natureza sintática dessas partes e funcionalmente pela ativação da finalidade de informar ou de atrair o interesse do leitor (Hartley 2005, 2007; Hoek 1981; Jamali & Nikzad 2011; Lamanauskas 2019; Sahragard & Meihami 2016; Whissel 2012). Na escrita acadêmica, a função de informar pode adotar uma perspectiva referencial, processual ou declarativa. Na vertente referencial, abrem-se possibilidades variadas quanto à componente do relatório e/ou da investigação a que é dada saliência no título (nomeadamente o campo ou objeto de estudo, metodologia, tipo de estudo, fator em análise, estratégia de intervenção, género textual). A vertente processual integra estes elementos num processo mais alargado. A vertente declarativa expressa a conclusão a que o estudo chegou. A função de atrair o leitor pode ativar recursos assentes no conteúdo, pelo efeito de surpresa, ou na forma, pelo jogo verbal (Hartley 2005, 2007).

Os resultados do estudo que se apresentou confirmam que nos títulos dos relatórios de mestrado de formação de professores é adotada uma grande amplitude de possibilidades. A vertente referencial, assente em expressões de natureza nominal, alcança um maior relevo. Dentro desta funcionalidade, emerge de forma mais saliente a indicação, no título, do campo ou objeto de estudo, mas outras componentes como a referência a estratégias, atividades, dispositivos pedagógicos, fatores em análise adquirem também relevo nos títulos dos relatórios neste campo.

A vertente referencial não esgota os títulos e a perspectiva processual ocupa também um lugar saliente nos títulos dos relatórios. Nesta vertente, destaca-se, numa instituição, o processo reflexivo (*Refletindo sobre...*) em que o foco é colocado no próprio formando, enquanto sujeito da reflexão que incide sobre a prática pedagógica e, dentro desta, sobre atividades ou estratégias específicas (colocadas habitualmente como subtítulos), que foram objeto de investigação. Estes títulos dão, assim, também relevo ao relatório enquanto instrumento de reflexão, no processo formativo.

Apesar de ser uma das possibilidades recorrentes no campo da escrita acadêmica, não emergem no presente estudo os títulos declarativos (Jamali & Nikzad 2011), constituídos por frases que expressam a conclusão da pesquisa. As frases enquanto constitutivas dos títulos estão presentes, mas sobretudo por meio de citações, enquanto recurso para atrair o interesse do leitor. Nesta vertente, outros recursos são ativados, designadamente a formulação de perguntas e o jogo verbal.

Os resultados confirmaram a amplitude de possibilidades a que os autores dos relatórios recorreram para a construção dos títulos. Tendo como pano de fundo essa variedade, a segunda questão em foco no presente estudo incidiu sobre a constituição das instituições de formação como comunidades de escrita cuja influência se reflete no conjunto das escolhas realizadas em relação aos títulos. A análise dos *corpora* revelou que há características dos títulos que tendem a ser mais salientes em determinadas instituições e mesmo a ser predominantes. É o que se observa em relação a uma das instituições (a instituição A) em que a vertente processual é adotada de forma ampla, o

que não acontece nas outras instituições. Também a formulação dos títulos como perguntas ou a colocação de citações, nomeadamente de enunciados das próprias crianças, constituem recursos (para a captação da atenção do leitor) que ocorrem com maior frequência em algumas instituições.

A diferenciação encontrada nos títulos, designadamente por meio do indicador correspondente às palavras-chave facultado pela Linguística de *Corpus*, pode ser associada a três dimensões das instituições enquanto comunidades: comunidade científica, comunidade discursiva e comunidade de escrita. Estas três dimensões projetam-se em elementos diferenciadores nos títulos: as áreas científicas de investigação com maior relevo na instituição que tendem a refletir-se nos trabalhos realizados pelos formandos; os recursos semióticos em uso na instituição para referir os objetos do seu campo de ação e formação; e as estratégias de utilização desses recursos semióticos para construir os próprios títulos na sua relação com o texto (relatório) e com a atividade que este expressa discursivamente.

Uma questão que decorre do enquadramento respeitante a estes mestrados de formação inicial de professores, no quadro em vigor, diz respeito à incidência do relatório na prática profissional e também à inclusão da dimensão de investigação. Os resultados mostram que o desafio de contemplar estas duas dimensões se projeta de forma diferente nos títulos dos relatórios das diferentes instituições. Globalmente, exceto numa instituição, é a dimensão investigativa que recebe saliência por meio dos títulos. O relevo atribuído no título pode estar associado ao facto de ser a dimensão em que o contributo individual para o conhecimento é mais marcante. Surge a questão de saber como, no corpo do relatório, as duas dimensões se articulam. Surge igualmente a questão de saber se a atribuição de relevo à dimensão investigativa nos títulos corresponde a modalidades específicas de concretização do estágio / prática de ensino supervisionada ou de construção do relatório (e do respetivo título). A resposta a estas questões implica o alargamento do estudo ao corpo dos relatórios, designadamente no que diz respeito à sua organização e às dimensões que são contempladas. Esse estudo, que pretendemos realizar, poderá revelar se, no caso destes relatórios, esta presença da dimensão de investigação nos títulos corresponde a modalidades de funcionamento do estágio ou de construção do relatório específicas ou se corresponde a escolhas circunscritas ao título.

No enquadramento dado pelo estágio / prática de ensino supervisionada – cujas linhas configuradoras, como ficou expresso, incluem “experiências de planificação, ensino e avaliação”, a articulação entre o “conhecimento e a forma de o transmitir visando a aprendizagem”, a “perspetiva de desenvolvimento profissional dos formandos” e a promoção de “uma atitude orientada para a permanente melhoria da aprendizagem” (Decreto-Lei n.º 79/2014) – a construção dos títulos apresenta uma dimensão estratégica de afirmação das componentes que nele estão presentes. Os resultados indicam que essa afirmação de componentes no título está associada à própria instituição em que o mestrado e o relatório é realizado. Em relação a uma das instituições (A), podemos considerar que emerge no seu seio a afirmação (e, por esse meio, a valorização) da dimensão reflexiva sobre a prática pedagógica desenvolvida. Nos títulos dos relatórios desta instituição, esta dimensão surge predominantemente como enquadradora da própria investigação. Nas outras instituições, esse enquadramento na prática pedagógica não é

dado nos títulos; ele encontra-se apenas no contexto do mestrado em que surge o relatório (contexto que o próprio documento inscreve na capa e/ou na folha de rosto), dando lugar a que, por sobre esse enquadramento, emerja nos títulos apenas a dimensão investigativa. Em qualquer dos casos, as escolhas realizadas em relação aos títulos podem ser tomadas como o recurso aos signos considerados mais aptos, no contexto em que se inserem, para construir o significado que se pretende que neles figurem (Bock 2016; Jewitt 2006; Kress & Jewitt 2003). Os resultados do presente estudo mostram que a instituição em que se realiza o mestrado constitui uma instância relevante desse contexto para as escolhas efetuadas.

O presente estudo apresenta, assim, como contributo mostrar que uma comunidade constituída por uma instituição de formação pode projetar-se e, por essa via, afirmar-se enquanto comunidade por meio do relevo que determinadas características apresentam na construção dos títulos dos textos associados a essa formação, neste caso, os títulos dos relatórios de mestrado da formação inicial de professores. Ao fazer incidir a análise nos objetos constituídos pelos títulos, emerge a limitação do estudo no que se refere às relações entre os títulos e o corpo dos próprios relatórios, o que remete para o desenvolvimento do estudo, designadamente com o objetivo de verificar se as escolhas relativas aos títulos estão também ligadas à afirmação da comunidade por meio de modalidades diferenciadas de concretização das componentes da formação. Por outro lado, no presente estudo, a análise incidiu sobre os títulos já inscritos nos relatórios de mestrado, ou seja, tomando-os como produtos. Para a captação do papel das comunidades de escrita, projeta-se tomar como perspectiva o próprio processo, tentando compreender quais as possibilidades consideradas nesse processo e respetivos elementos, quais os modelos, os fatores tidos em conta e os participantes que intervieram no processo de construção e decisão.

## Referências

- Alves, E., Silva, B., & Silva, R. (2017). Mapeamento dos estudos sobre a formação de professores no âmbito do processo de Bolonha em Portugal. *Revista Observatório*, 3(6), 248–273. <https://doi.org/10.20873/UFT.2447-4266.2017V3N6P248>
- Amoroso, L. K. (2018). Analyzing group differences. In A. Phakiti, P. De Costa, L. Plonsky & S. Starfield (Eds.). *The Palgrave handbook of applied linguistics research methodology* (pp. 501–521). Palgrave.
- Anthony, L. (2021). *AntConc* (Version 4.0.2) [Computer Software]. Tokyo, Japan: Waseda University. Disponível em <http://www.antlab.sci.waseda.ac.jp/>
- Arcadinho, A., Folque, M., & Costa, C. (2020). Dimensão investigativa, docência e formação inicial de professores: Uma revisão sistemática de literatura. *Instrumento: Revista de Estudo e Pesquisa em Educação*, 22(1), 5–23. <https://doi.org/10.34019/1984-5499.2020.v22.29173>
- Barbeiro, L., Pereira, L. Á., Coimbra, R. L., & Calil, E. (2020). Os títulos no processo de escrita de histórias por alunos do 2.º ano de escolaridade. *Revista Portuguesa de Educação*, 33(2), 71–94. <https://doi.org/10.21814/rpe.18339>
- Bock, Z. (2016). Multimodality, creativity and children's meaning-making: Drawings, writings, imaginings. *Stellenbosch Papers in Linguistics Plus*, 49, 1–21. <http://dx.doi.org/10.5842/49-0-669>

- Brezina, V. (2018). *Statistics in Corpus Linguistics: A practical guide*. Cambridge University Press.
- Calil, E. (2020). Gênese textual de um título: Estudo sobre as redes associativas e ativação de memórias em alunos recém-alfabetizados. *Bakhtiniana*, 16(1), 184–210. <https://doi.org/10.1590/2176-457346306>
- Camprubi, J. B. (2002). Les fonctions du titre. *Nouveaux Actes Sémiotiques*, 82. Limoges : PULIM - Presses Universitaires de Limoges.
- Coimbra, R. L. (1999). *Estudo linguístico dos títulos de imprensa em Portugal: A linguagem metafórica*. (Tese de doutoramento, Universidade de Aveiro). Disponível em <https://ria.ua.pt/bitstream/10773/4615/3/Tese.pdf>
- Cremer, P., & Lea, M. (2000). *Escribir en la Universidad*. Editorial Gedisa.
- Flores, M. (2018). Linking teaching and research in initial teacher education: Knowledge mobilisation and research informed practice. *Journal of Education for Teaching*, 44(5), 621–636 <https://doi.org/10.1080/02607476.2018.1516351>
- Fox, C., & Burns, C (2015). The relationship between manuscript title structure and success: Editorial decisions and citation performance for an ecological journal. *Ecology and Evolution*, 5(10), 1970–1980. <https://doi.org/10.1002/ece3.1480>
- Garcia, D., Gattaz, C, & Gattaz, N. (2019). The relevance of title, abstract and keywords for scientific paper writing. *Revista de Administração Contemporânea / Journal of Contemporary Administration*, 23(3). <http://doi.org/10.1590/1982-7849rac2019190178>
- Genette, G., & Crampé, B. (1988). Structure and functions of the title in literature. *Critical Inquiry*, 14(4), 692–720. <https://doi.org/10.1086/448462>
- Graham, S. (2018). A revised Writer(s)-Within-Community Model of Writing, *Educational Psychologist*, 53(4), 258–279. <https://doi.org/10.1080/00461520.2018.1481406>
- Guo, F., Ma, C., Shi, Q., & Zhong, Q. (2018). Succinct effect or informative effect: the relationship between title length and the number of citations. *Scientometrics*, 116, 1531–1539. <https://doi.org/10.1007/s11192-018-2805-8>
- Hallock, R., & Bennett, T. (2021). I'll read that!: What title elements attract readers to an article? *Teaching of Psychology*, 48(1), 26–31. <https://doi.org/10.1177/0098628320959948>
- Hartley, J. (2005). To attract or to inform: What are titles for? *Journal of Technical Writing and Communication*, 35(2), 203–213. <https://doi.org/10.2190/NV6E-FN3N-7NGN-TWQT>
- Hartley, J. (2007). There's more to the title than meets the eye: Exploring the possibilities. *Journal of Technical Writing and Communication*, 37(1), 95–101. <https://doi.org/10.2190/BJ16-8385-7Q73-1162>
- Hartley, J. (2008). *Academic writing and publishing: A practical handbook*. Routledge.
- Hoek, L. H. (1981). *La marque du titre*. De Gruyter.
- Hudson, J. (2016). An analysis of the titles of papers submitted to the UK REF in 2014: Authors, disciplines, and stylistic details. *Scientometrics*, 109, 871–889. <https://doi.org/10.1007/s11192-016-2081-4>.
- Jamali, H., & Nikzad, M. (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics*, 88, 653–661. <https://doi.org/10.1007/s11192-011-0412-z>
- Jewitt, C. (2006). *Technology, literacy and learning: A multimodal approach*. Routledge.
- Joint Quality Initiative Informal Group (2004). Shared Dublin descriptors for Short Cycle, First Cycle, Second Cycle and Third Cycle Awards. Disponível em [https://www.aqu.cat/doc/doc\\_24496811\\_1.pdf](https://www.aqu.cat/doc/doc_24496811_1.pdf)
- Kress, G., & Jewitt, C. (2003). Introduction. In C. Jewitt & G. Kress (Eds.), *Multimodal literacy* (pp. 1–18). Lausanne: Peter Lang.
- Lamauskas, V. (2019). Scientific article preparation: Title, abstract and keywords. *Problems of Education in the 21st Century*, 77(4), 456–462. <https://doi.org/10.33225/pec/19.77.456>
- Leal, A., & Leite, C. (2017). Políticas da formação inicial de professores, em Portugal, decorrentes de Bolonha: Posições da comunidade científica. *Interritórios: Revista da Universidade Federal de Pernambuco*, 4(3). Disponível em <https://periodicos.ufpe.br/revistas/interritorios/article/view/22523/18682>

- Martin, J. R., Maton, K., & Doran, Y. J. (2020). Academic discourse: An interdisciplinary dialogue. In J. R. Martin, K. Maton & Y. J. Doran (Eds.), *Accessing academic discourse* (pp. 1–31). London: Routledge.
- O'Donnell, M. (2013). A dynamic view of choice in writing: Composition as text evolution. In L. Fontaine, T. Bartlett & G. O'Grady (Eds.), *Systemic Functional Linguistics: Exploring choice* (pp. 247–266). Cambridge: University Press.
- Paquin, N. (2008). Le titre des oeuvres : un "titulus" polyvalent. *Protée*, 36(3), 5–10. <https://doi.org/10.7202/019629ar>
- Pereira, L. A., Coimbra, R. L., & Calil, E. (2017). Os títulos de contos que crianças (re)contam: Uma “poética” da brevidade sem a angústia da influência. *Forma Breve*, 14, 617–630. <https://doi.org/10.34624/fb.v0i14.541>
- Ponte, J. P. (2006). Os desafios do Processo de Bolonha para a formação inicial de professores. *Revista da Educação*, 14(1), 19–36.
- Quadro Nacional de Qualificações do Ensino Superior. (2011). Relatório da Comissão Internacional sobre a Verificação da Compatibilidade com o Quadro de Qualificações do Espaço Europeu do Ensino Superior. Disponível em [https://wwwcdn.dges.gov.pt/sites/default/files/relatorio\\_referenciacao\\_ensino\\_superior\\_portugal\\_qq-ees\\_0.pdf](https://wwwcdn.dges.gov.pt/sites/default/files/relatorio_referenciacao_ensino_superior_portugal_qq-ees_0.pdf)
- Roy, M. (2008). Du titre littéraire et de ses effets de lecture. *Protée*, 36(3), 47–56. <https://doi.org/10.7202/019633ar>
- Sahragard, R., & Meihami, H. (2016). A diachronic study on the information provided by the research titles of applied linguistics journals. *Scientometrics*, 108, 1315–1331. <https://doi.org/10.1007/s11192-016-2049-4>
- Sharples, M. (1999). *How we write: Writing as a creative design*. Routledge.
- Soler, V. (2007). Writing titles in science: An exploratory study. *English for Specific Purposes*, 26, 90–102. <https://doi.org/10.1016/j.esp.2006.08.001>
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Vásquez, G. (2001). *Guía didáctica del discurso académico escrito*. Editorial Edinumen.
- Vicea, M. P. (2003). Le titre est-il un désignateur rigide?. In I. I. Las Heras & M. J. Cascante (Coords.), *El texto como encrucijada. Estudios franceses y francófonos* (Voil. II, pp. 251–258). La Rioja: Universidad de La Rioja.
- Vieira, F., Flores, M., Silva, J., Almeida, M., & Vilaça, T. (2021). Inquiry-based professional learning in the practicum: Potential and shortcomings. *Teaching and Teacher Education*, 105 (103429). <https://doi.org/10.1016/j.tate.2021.103429>
- Vieira, F., Flores, M., & Almeida, M. (2020a). Avaliando o modelo de estágio dos mestrados em ensino da Universidade do Minho: Entre a qualidade desejada e a qualidade percebida. *Instrumento: Revista de Estudo e Pesquisa em Educação*, 22(2), 231–247. <https://doi.org/10.34019/1984-5499.2020.v22.29962>
- Vieira, F., Silva, J., & Vilaça, M. (2020b). Formação de professores baseada na investigação pedagógica: Um estudo sobre o estágio nos mestrados em ensino. *Educação em Perspectiva*, e020012, 1–17. <https://doi.org/10.22294/eduper/ppge/ufv.v11i.9060>
- Whissel, C. (2012). The trend toward more attractive and informative titles: American psychologist 1946–2010. *Psychological Reports*, 110(2), 427–444. <https://doi.org/10.2466/17.28.PR0.110.2.427-444>
- Zagovora, O., Weller, K., Janosov, M., Wagner, C., & Peters, I. (2018, September). What increases (social) media attention: Research impact, author prominence or title attractiveness? *Proceedings of the 23rd International Conference on Science and Technology Indicators*. Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands. <https://doi.org/10.31235/osf.io/mwxye>

[recebido em 18 de janeiro de 2022 e aceite para publicação em 29 de março de 2022

FCT - UID/ELT/00305/2019

**FCT**  
Fundação para a Ciência e a Tecnologia  
INSTITUTO DE CIÊNCIA, TECNOLOGIA E INOVAÇÃO

  
**COMPETE**  
PROGRAMA OPERACIONAL FACTORES DE COMPETITIVIDADE

  
QUADRO  
DE REFERÊNCIA  
ESTRATÉGICO  
NACIONAL  
PORTUGAL 2020-2027



**UNIÃO EUROPEIA**

Fundo Europeu de  
Desenvolvimento Regional