



**UNIVERSIDADE FEDERAL DA PARAÍBA**  
**CENTRO DE CIÊNCIAS HUMANAS, LETRAS E ARTES**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA**  
**MESTRADO EM LINGUÍSTICA**

**RAMON BRASILEIRO GUEDES**

**ANIMACIDADE E O PROCESSAMENTO DE ORAÇÕES RELATIVAS DE**  
**SUJEITO E OBJETO POR BILÍNGUES PORTUGUÊS-INGLÊS**

**JOÃO PESSOA**

**2021**

**RAMON BRASILEIRO GUEDES**

**ANIMACIDADE E O PROCESSAMENTO DE ORAÇÕES RELATIVAS DE  
SUJEITO E OBJETO POR BILÍNGUES PORTUGUÊS-INGLÊS**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal da Paraíba (UFPB), como requisito parcial para a obtenção de título de Mestre em Linguística, sob a orientação do Professor Doutor **Márcio Martins Leitão** e coorientação da Professora Doutora **Juliana Novo Gomes**. Área de concentração: Teoria e Análise Linguística. Linha de pesquisa: Aquisição da Linguagem e Processamento Linguístico.

**JOÃO PESSOA**

**2021**

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

G924a Guedes, Ramon Brasileiro.

Animacidade e o processamento de orações relativas  
de sujeito e objeto por bilíngues português-inglês /  
Ramon Brasileiro Guedes. - João Pessoa, 2021.  
102 f. : il.

Orientação: Márcio Martins Leitão.

Coorientação: Juliana Novo Gomes.

Dissertação (Mestrado) - UFPB/CCHLA.

1. Processamento linguístico. 2. Bilinguismo. 3.  
Orações relativas. 4. Animacidade. I. Leitão, Márcio  
Martins. II. Gomes, Juliana Novo. III. Título.

UFPB/BC

CDU 81(043)



**ATA DE DEFESA DE DISSERTAÇÃO DE  
RAMON BRASILEIRO GUEDES**

Aos treze dias do mês de dezembro de dois mil e vinte e um (13/12/2021), às catorze horas, realizou-se, via Plataforma Zoom, a sessão pública de defesa de Dissertação intitulada “Animacidade e o processamento de orações relativas de sujeito e objeto por bilíngues português-inglês”, apresentada pelo(a) mestrando(a) **RAMON BRASILEIRO GUEDES**, Licenciado(a) em Letras pelo(a) Universidade Federal da Paraíba - UFPB, que concluiu os créditos para obtenção do título de MESTRE(A) EM LINGÜÍSTICA, área de concentração **Teoria e Linguística**, segundo encaminhamento do(a) Prof(a). Dr(a). José Ferrari Neto, Coordenador(a) do Programa de Pós-Graduação em Linguística da UFPB e segundo registros constantes nos arquivos da Secretaria da Coordenação do Programa. O(A) Prof(a). Dr(a). Márcio Martins Leitão (PROLING - UFPB), na qualidade de orientador(a), presidiu a Banca Examinadora da qual fizeram parte os(a)s Professores(as) Doutores(as) Juliana Novo Gomes (Coorientadora /PROLING-UFPB), Eduardo Kenedy Nunes Areas (Examinador/UFPB) e Elena Ortiz Preuss (Examinadora/UFG). Dando início aos trabalhos, o(a) senhor(a) Presidente Prof(a). Dr(a). Márcio Martins Leitão convidou os membros da Banca Examinadora para compor a mesa. Em seguida, foi concedida a palavra ao(à) Mestrando(a) para apresentar uma síntese de sua Dissertação, após o que foi arguido(a) pelos membros da banca Examinadora. Encerrando os trabalhos de arguição os examinadores deram o parecer final sobre a Dissertação, ao qual foi atribuído o conceito **APROVADO**. Proclamados os resultados pelo(a) professor(a) Dr(a). Márcio Martins Leitão, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar a presente ata foi lavrada e assinada por todos os membros da Banca Examinadora. João Pessoa, 13 de dezembro de 2021.

Observações

Prof(a). Dr(a). Márcio Martins Leitão  
(Presidente da Banca Examinadora)

Prof(a). Dr(a). Juliana Novo Gomes  
(Coorientadora)

Prof(a). Dr(a). Eduardo Kenedy Nunes  
Areas  
(Examinador)

Prof(a). Dr(a). Elena Ortiz Preuss  
(Examinadora)

## AGRADECIMENTOS

Agradeço, primeiramente, ao meu orientador Márcio Martins Leitão, cujo farto conhecimento, e, acima de tudo, generosidade, tornaram esta dissertação possível. É um privilégio ser orientado por alguém que evidentemente tem gosto pelo que faz e que se empolga tanto com as descobertas que surgem pelo caminho. Sempre diligente, por um lado, em me guiar quando necessário, Márcio também me deu liberdade para seguir minha própria curiosidade durante esta pesquisa. As virtudes deste trabalho devem-se, em imensa medida, a ele, mas as limitações que porventura houver são minhas. Espero fazer jus ao esforço investido e à confiança depositada em mim.

Além do meu orientador, tive a sorte de contar com o inestimável suporte da professora Juliana Novo Gomes, cujos conhecimentos de estatística e programação foram fundamentais na execução do nosso experimento e na compreensão dos resultados obtidos com ele. Juliana nos guiou na adaptação dos instrumentos de pesquisa ao contexto de isolamento social que vivenciamos durante estes anos de pandemia, transformando este momento de crise em um período de descoberta de possibilidades para o nosso laboratório, o Laprol.

Não posso deixar de agradecer, aliás, aos membros do Laprol (Laboratório de Processamento Linguístico) da UFPB, valiosos incentivadores e críticos no decorrer desta pesquisa. Menciono, em especial, o professor José Ferrari Neto, que tanto me auxiliou nos trâmites burocráticos do mestrado; e os amigos (em ordem alfabética) Matheus, Maylton e Renata, companheiros nos estudos sobre bilinguismo. Juntos, ajudamos uns aos outros na elaboração do teste de proficiência em língua inglesa, na coleta de participantes, no ensaio dos experimentos, nas leituras, no aprofundamento na teoria e, não menos importante, no apoio moral mútuo. Desejo imenso sucesso aos três.

Destaco o papel fundamental dos participantes desta pesquisa, que dedicaram tempo e esforço na realização tanto do teste de proficiência quanto do experimento propriamente dito, sem esperar nenhuma retribuição de qualquer espécie. Este trabalho é resultado de sua generosidade.

Sou grato à minha família, que sempre me forneceu o apoio material e emocional necessário à caminhada que trilhei até chegar à universidade. Professores de formação, meus pais nunca mediram esforços para me prover da melhor educação possível, e foi

deles que herdei a curiosidade e o gosto pela leitura. Sei que, em razão do meu ambiente familiar, comecei neste mundo com o pé direito.

Agradeço, por fim, aos amigos que me acompanham há anos, os primeiros a celebrar minha entrada no mestrado e que tanto torceram por mim até aqui. Reconheço e aprecio a paciência dos que ouviram minhas divagações sobre processamento preditivo, orações relativas e a Hipótese da Estrutura Rasa mesmo sem nenhuma experiência – ou mesmo interesse – em Psicolinguística. Agora que esta pesquisa deu frutos, prometo falar ainda mais.

*Para Luna, que tantas vezes ronronou  
entre os meus pés ou montou guarda  
sob a mesa enquanto eu escrevia.  
Despediu-se com um eclipse lunar e  
deixou para trás muita saudade.*

## RESUMO

Devido a sua complexidade estrutural, orações relativas têm sido um frequente objeto de interesse de psicolinguistas. Tradicionalmente, orações relativas de sujeito são apontadas como mais fáceis de processar que orações relativas de objeto. Essa diferença se explicaria por meio da maior complexidade sintática das últimas frente às primeiras. No entanto, quando se controla a animacidade do referente dessas estruturas, diferenças de custo de processamento entre os dois tipos de oração relativa desaparecem ou são atenuadas, o que põe em dúvida a ideia de que apenas fatores sintáticos estão acessíveis ao *parser* durante os momentos iniciais do processamento linguístico. Ainda não está claro se esses resultados são igualmente válidos no caso de processamento de segunda língua. Decidimos investigar se a animacidade do referente interfere no processamento de orações relativas em inglês por falantes não nativos, neste caso brasileiros que têm o português como idioma materno e o inglês como segunda língua. Além disso, buscamos descobrir se o nível de proficiência desses sujeitos teria efeito sobre o processamento dessas estruturas. Para tanto, utilizamos o método de leitura automonitorada com 32 bilíngues português (L1)/inglês (L2), divididos igualmente em dois grupos de acordo com o nível de proficiência na L2 (intermediário ou avançado). Os estímulos utilizados foram orações relativas de sujeito e objeto, com referente animado e inanimado, o que gerou quatro condições experimentais. Os resultados obtidos foram comparados com o *corpus* de orações relativas em língua inglesa que criamos. Observamos o desaparecimento da assimetria de processamento entre orações relativas de sujeito e objeto no caso dos falantes de nível avançado, bem como uma relação estreita entre as condições mais facilmente processadas por esse grupo e os tipos de oração relativa mais comuns no nosso *corpus*. Dados diferentes foram obtidos com os sujeitos de nível intermediário, o que sugeriria uma menor acessibilidade de fatores não sintáticos no processamento de L2 para esses falantes.

**Palavras-chave:** processamento linguístico; bilinguismo; orações relativas; animacidade.



## ABSTRACT

Due to their structural complexity, relative clauses have often been an object of study in psycholinguistics. Traditionally, subject relative clauses have been deemed easier to process than object relative clauses. This difference was supposedly due to the greater syntactic complexity of the latter. However, when the animacy in such structures is controlled, differences in processing cost between both types of relative clause disappear or are attenuated, which brings into question the idea that only syntactic factors are accessible to the parser during the initial stages of language processing. It is still not clear whether these results hold true in the case of second language processing. We decided to investigate if animacy interferes in the processing of English relative clauses by non-native speakers, in this case Brazilians who speak Portuguese as their mother tongue and English as their second language. Furthermore, we explored whether these participants' English proficiency level would have any effect on the processing of relative clauses. With this goal in mind, we used a self-paced reading experiment with 32 Portuguese (L1)/English (L2) bilinguals, equally divided into two groups according to their proficiency level in the L2 (intermediate or advanced). The stimuli that we used were subject and object relative clauses, with either animate or inanimate referents, which generated four experimental conditions. The results that we obtained were compared with our corpus of English relative clauses. We observed the disappearance of the processing asymmetry between subject and object relative clauses in the case of highly proficient speakers, as well as a clear correlation between the sentence types most easily processed by this group and the most common types of relative clauses in our corpus. Different data were gathered from the intermediate-level participants, which could suggest non-syntactic factors are not as readily available to them during the processing of their L2.

**Keywords:** language processing; bilingualism; relative clauses; animacy.

## GRÁFICOS

	PÁG.
<b>Gráfico 1:</b> <i>Corpora</i> escrito e falado do português brasileiro e do português europeu retirados de Cabral (2016)	38
<b>Gráfico 2:</b> Orações relativas com divergência de animacidade entre sujeito e objeto, classificadas de acordo com o conectivo/pronome relativo utilizado	67
<b>Gráfico 3:</b> Orações relativas com divergência de animacidade entre sujeito e objeto, classificadas por condição	68
<b>Gráfico 4:</b> Orações relativas com divergência de animacidade entre sujeito e objeto, com o uso do conectivo <i>that</i> , classificadas por condição	69
<b>Gráfico 5:</b> Tempo de leitura do segmento crítico por nível de proficiência e tipo de oração relativa	75
<b>Gráfico 6:</b> Assimetria entre o processamento de ORS e de ORO na junção dos resultados de falantes de níveis intermediário e avançado	76
<b>Gráfico 7:</b> Tempo de leitura do segmento crítico dos diferentes tipos de oração relativa por falantes de nível avançado	77
<b>Gráfico 8:</b> Tempo de leitura do segmento crítico dos diferentes tipos de oração relativa por falantes de nível intermediário	77

## FIGURAS

	PÁG.
<b>Figura 1:</b> Possibilidades de divisão lexical em marcação de caso	30
<b>Figura 2:</b> Efeito da Hierarquia de Animacidade sobre o número gramatical nas línguas humanas	32
<b>Figura 3:</b> Exemplo de cena utilizada em um dos experimentos de Altmann e Kamide (1999)	43

## SIGLAS

- AFS:** Active Filler Strategy (Estratégia do Antecedente Ativo)
- ASL:** American Sign Language (Língua de Sinais Americana)
- DLT:** Dependency Locality Theory (Teoria da Dependência Local)
- ERP:** event-related potentials (potenciais relacionados a eventos)
- FT/FA:** Full Transfer/Full Access Hypothesis (Hipótese da Transferência Total/Acesso Total)
- GU:** Gramática Universal
- L1:** língua materna
- L2:** segunda língua
- OR:** oração relativa
- ORO:** oração relativa de objeto
- OROa:** oração relativa de objeto animado
- OROi:** oração relativa de objeto inanimado
- ORS:** oração relativa de sujeito
- ORSa:** oração relativa de sujeito animado
- ORSi:** oração relativa de sujeito inanimado
- RAGE:** Reduced Ability to Generate Expectations (Habilidade Reduzida de Geração de Expectativas)
- SPLT:** Syntactic Prediction Locality Theory (Teoria da Predição Sintática Local)
- SSH:** Shallow Structure Hypothesis (Hipótese da Estrutura Rasa)
- TGP:** Teoria do *Garden Path*
- VLT:** Vocabulary Levels Test

## SUMÁRIO

	PÁG.
<b>CAPÍTULO 1: INTRODUÇÃO</b>	<b>15</b>
<b>CAPÍTULO 2: ORAÇÕES RELATIVAS E ANIMACIDADE</b>	<b>19</b>
2.1 Pronomes relativos da língua inglesa	19
2.2 Estrutura da oração relativa	21
2.3 Assimetria no processamento de ORS e ORO	23
2.4 Animacidade	26
2.4.1 Animacidade real, animacidade no sistema conceitual humano e animacidade como traço formal das gramáticas humanas	27
2.4.2 Animacidade e cognição humana	28
2.4.3 Hierarquias de animacidade	29
2.4.4 Efeitos da animacidade nas línguas	33
2.4.5 Animacidade e o processamento de orações relativas	34
<b>CAPÍTULO 3: PROCESSAMENTO DE SENTENÇAS</b>	<b>41</b>
3.1 Modelos seriais x modelos paralelos	44
3.2 Modelos modulares x modelos interativos	48
3.3 Frequência e processamento linguístico	50
3.4 Processamento de L2	54
3.5 Frequência e processamento preditivo em L2	56
3.6 Modelos de processamento de L2	59
3.6.1 Hipótese da Estrutura Rasa ( <i>Shallow Structure Hypothesis</i> )	59
3.6.2 Habilidade Reduzida de Geração de Expectativas ( <i>Reduced Ability to Generate Expectations – RAGE</i> )	60
3.6.3 Perspectivas alternativas	61
3.7 Hipótese da Transferência Total/Acesso Total ( <i>FT/FA</i> )	64
<b>CAPÍTULO 4: ESTUDO ATUAL</b>	<b>66</b>
4.1 Estudo de <i>corpus</i>	66
4.2 Teste de leitura automonitorada	71
4.2.1 Participantes	72
4.2.2 Materiais	73

<b>4.2.3 Procedimentos</b>	<b>73</b>
<b>4.2.4 Resultados</b>	<b>74</b>
<b>4.2.5 Discussão dos resultados</b>	<b>78</b>
<b>CAPÍTULO 5: CONSIDERAÇÕES FINAIS</b>	<b>82</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>85</b>
<b>ANEXO</b>	<b>93</b>

## 1. INTRODUÇÃO

Estudos na área da psicolinguística partem do pressuposto de que a linguagem humana evoluiu de modo a refletir a estrutura da nossa mente, sendo interesse do psicolinguista investigar como se dá essa relação. Especificamente, o psicolinguista se debruça sobre os processos cognitivos envolvidos no planejamento, produção, percepção e compreensão da linguagem. A aproximação aos princípios e adoção dos métodos de pesquisa da psicologia cognitiva distinguem a psicolinguística da linguística como um todo, uma vez que a última é geralmente dedicada à compreensão e descrição da estrutura das línguas e/ou da linguagem enquanto faculdade humana. No entanto, deve-se salientar a multidisciplinaridade inerente aos estudos psicolinguísticos, que se servem também dos achados da fonética, da neurolinguística, da modelagem computacional, da análise do discurso e da semântica (FIELD, 2011; TAVAKOLI, 2012).

Atualmente, psicolinguistas se dedicam à investigação de diversos fenômenos linguísticos, como a aquisição da linguagem por crianças, a aquisição de segunda língua, a complexidade linguística, o processamento linguístico, a leitura, patologias da linguagem e a relação entre universais linguísticos e cognitivos. É comum, além disso, que se adote uma abordagem baseada na coleta de evidências que possam corroborar pontos específicos da teoria linguística (TAVAKOLI, 2012). Para esse fim, podem-se utilizar diversos métodos experimentais, como leitura automonitorada, rastreamento ocular, o uso de potenciais relacionados a eventos (*event-related potentials*, ou *ERP*) em estudos de eletroencefalografia, técnicas de neuroimagem, etc. (MITCHELL, 2004).

Um dos temas que vêm ocupando os psicolinguistas é a natureza das informações que são utilizadas pelos falantes na compreensão de um determinado estímulo linguístico: seriam considerados apenas fatores sintáticos num primeiro momento, ou seria possível que informações de caráter semântico, por exemplo, também exercessem influência no processamento linguístico desde o seu início? Para estudar esse fenômeno, alguns pesquisadores se debruçam sobre o papel da animacidade, um fator de natureza semântica, no processamento de orações relativas de sujeito (doravante ORS) e de objeto (ORO).

A animacidade é a característica definidora dos entes vivos. Um substantivo dito animado denota seres humanos ou animais, mas a aceção exata de animacidade pode

diferir, até certo ponto, de acordo com a cultura sob análise. Neste trabalho, limitaremos os substantivos animados àqueles que denotam seres humanos e animais, excluindo entes sobrenaturais e forças da natureza, cujo pertencimento ao rol dos seres animados é mais dependente de especificidades culturais (BROWN & MILLER, 2013).

Por oração relativa, entendemos uma estrutura linguística que se encaixa em outra por meio de adjunção, modificando um sintagma nominal ao adjungir-lhe tantas propriedades quanto forem possíveis dentro das limitações da memória de trabalho de um falante ou ouvinte (CABRAL, 2016). É farta a literatura que indica uma maior dificuldade de processamento de ORO frente às ORS (cf. CARAMAZZA & ZURIF, 1976; WANNER & MARATSOS, 1978; FRAUENFELDER, SEGUI & MEHLER, 1980; HOLMES & O'REGAN, 1981; FORD, 1983; KING & JUST, 1991; SCHRIEFERS, FRIEDERICI & KÜHN, 1995; JUST et al. 1996; CAPLAN, ALPERT & WATERS, 1998; GORDON, HENDRICK & JOHNSON, 2001; TRAXLER, MORRIS & SEELY, 2002; OLIVEIRA, 2013). Essa assimetria de processamento se deveria a uma diferença de complexidade sintática entre os dois tipos de estrutura linguística, sendo as ORO significativamente mais complexas que as ORS (cf. FRAZIER, 1987; FRAZIER & FLORES D'ARCAIS, 1989; GIBSON, 1998).

Entretanto, outros autores encontraram resultados diferentes ao controlarem a animacidade do referente nas orações relativas (cf. CABRAL, 2016; TRAXLER, MORRIS & SEELY, 2002; MAK, VONK & SCHRIEFERS, 2002). A animacidade é um fator semântico relevante, visto que a distinção entre entes vivos e não vivos é parte fundamental da cognição humana (cf. OPFER & GELMAN, 2011; NAIRNE et al., 2017), uma vez que os eventos no mundo externo que demandam nossa atenção tendem a ser iniciados por entidades animadas, e a atividade humana, em particular, constitui o principal alvo de nosso interesse (VIHMAN & NELSON, 2019). Por essa razão, a animacidade influi não só na gramática de algumas línguas, como na frequência de estruturas sintáticas em muitas outras (BAYANATI & TOIVONEN, 2019).

Esses estudos geralmente investigam o processamento de língua materna (L1), sem considerar eventuais diferenças no caso do processamento de segunda língua (L2). Seria possível estender os resultados encontrados até aqui à compreensão de estímulos em um idioma não materno?



Tendo essa questão em mente, e com base em pesquisas que parecem indicar uma resposta afirmativa a ela (TRAXLER et al., 2014; PIÑAR et al., 2017), acreditamos que falantes bilíngues, quando se deparam com orações relativas de sujeito e de objeto em que o traço de animacidade foi rigorosamente controlado, fazem uso dessa informação de caráter semântico como instrumento facilitador no processamento de tais estruturas.

Para investigar o tema, escolhemos a técnica de leitura automonitorada, em que os estímulos linguísticos são divididos em segmentos exibidos um de cada vez, seguindo os comandos do leitor, numa tela de computador (MITCHELL, 2004). Medindo em milissegundos o tempo de leitura de cada segmento da sentença, somos capazes de determinar a diferença no tempo de leitura dos segmentos de interesse nas ORS e ORO. Dessa maneira, determinamos quais estruturas demandam mais tempo de leitura e, portanto, oferecem maior dificuldade de processamento ao leitor, levando-se em consideração não apenas a complexidade sintática dessas estruturas, mas também a possível influência da animacidade nessa assimetria. Nosso estudo compara falantes de português como L1 e inglês como L2, em dois níveis de proficiência na última língua: intermediário e avançado.

Esta dissertação está dividida da seguinte maneira: após o primeiro capítulo, introdutório, o Capítulo 2 é dedicado à revisão teórica dos conceitos que são caros à nossa pesquisa, como os pronomes relativos da língua inglesa; a estrutura e a assimetria no processamento dos dois tipos de oração relativa; a animacidade no mundo real, no sistema conceitual humano e como traço formal nas línguas humanas; as diferentes formas de se classificar a animacidade como sistema binário ou como *continuum*; a influência que a animacidade exerce sobre os idiomas e sobre o processamento linguístico.

No Capítulo 3, continuamos abordando alguns temas importantes, dessa vez com foco no processamento de L1 e L2: a princípio, revisamos os modelos teóricos que tentam explicar o modo como seres humanos compreendem os estímulos linguísticos a que são expostos; depois, o papel da frequência na compreensão desses estímulos; apresentamos duas hipóteses propostas para descrever de quais maneiras o processamento bilíngue poderia diferir do processamento de L1, bem como perspectivas alternativas a esses modelos; e, por fim, discutimos sobre a possibilidade de, no processamento de L2, ocorrer interferência advinda do conhecimento que os falantes têm de sua L1.

No Capítulo 4, apresentamos o nosso *corpus* de orações relativas em língua inglesa; expomos, mais detalhadamente, os objetivos (principal e específicos) desta pesquisa, assim como a justificativa e a hipótese que a norteiam. Além disso, descrevemos o *design* experimental que utilizamos e os sujeitos bilíngues, brasileiros que têm o inglês como L2 em nível avançado ou intermediário, que participaram do experimento proposto.

O Capítulo 5, por sua vez, apresenta os resultados de nosso experimento relacionados ao tempo de leitura do segmento crítico das orações relativas nas quatro condições experimentais, por falantes de inglês como L2 em nível intermediário e avançado. Além disso, discutimos esses resultados à luz dos dados obtidos com o nosso *corpus*, das pesquisas feitas na área do processamento linguístico de maneira geral e do processamento bilíngue em particular, contrastando-os com alguns modelos teóricos que buscam explicar esses processos.

Por fim, o Capítulo 6 expõe nossas considerações finais, incluindo um apanhado das questões levantadas nesta dissertação, bem como de alguns estudos e teorias que buscam esclarecê-las, comparando-os com os resultados que obtivemos com nosso experimento. No caso dos falantes de inglês como L2 em nível avançado, esses resultados confirmaram nossa hipótese inicial de que o processamento de L2 seria semelhante ao de L1 no que concerne ao uso de fatores além dos sintáticos – nomeadamente semânticos e probabilísticos – no processamento de orações relativas. Ensaíamos uma breve discussão sobre as implicações de nosso estudo na área do processamento bilíngue, sem deixar de reconhecer suas limitações e a necessidade de se empreenderem mais pesquisas sobre essa questão.

## 2. ORAÇÕES RELATIVAS E ANIMACIDADE

Orações relativas têm há muito tempo despertado o interesse dos linguistas, tanto por sua complexidade estrutural quanto pelo papel fundamental que desempenham na recursividade característica da linguagem humana. De acordo com Cheon et al. (2020), a construção de orações relativas é um dos principais processos na formação de sentenças complexas. Por consequência, o estudo dessas orações vem sendo guiado pela esperança de se chegar, um dia, a uma explicação universal da diversidade estrutural que apresentam nas várias línguas naturais.

O que define, então, a oração relativa? Nas palavras de Cabral (2016),

Cláusulas relativas<sup>1</sup> são estruturas linguísticas encaixadas em outras estruturas por meio de adjunção. Sintagmas complexos que, recursivamente, podem modificar um sintagma nominal, SN, adjungindo-lhe tantas propriedades quanto o discurso corrente e a memória de trabalho do falante/ouvinte permitirem (Cabral, 2016, p. 22).

Para compreendermos a razão por trás das controvérsias relacionadas às orações relativas, precisamos, primeiramente, estar a par das discussões feitas sobre sua estrutura. Uma breve descrição da constituição das orações relativas é apresentada na seção abaixo.

### 2.1 Pronomes relativos da língua inglesa

Em inglês, orações relativas fazem uso de uma série de pronomes relativos (*which*, *who*, *whom*) e do conectivo *that*, que correspondem ao português *que*. Exemplos de uso desses elementos podem ser observados abaixo.

(01) The TV set **that** I bought isn't working properly.

(02) The kid **that** stole Ben's bike has been arrested.

(03) The problem, **which** none of us could solve, put the whole project on hold.

---

<sup>1</sup> Na literatura especializada em língua portuguesa, orações relativas são por vezes denominadas "cláusulas relativas".

- (04) The man **who** showed her the way vanished in front of her.  
(05) The students **whom** Susan brought to the museum wanted to see the T-Rex.

O conectivo *that* pode ser substituído por *who(m)* ou *which*, embora cada um deles tenha usos específicos. *Who* refere-se somente a antecedentes humanos (e, em registro menos formal, a animais de estimação), enquanto que *which* requer um antecedente não humano. Em linguagem formal, *whom* substitui *who* ao ocupar a posição de objeto (DECLERCK, 1991).

- (06) The TV set **which** I bought isn't working properly.  
(07) The kid **who** stole Ben's bike has been arrested.  
(08) The heroes **whom** we admire tend to die young.

Orações relativas podem ser encaixadas ou hipotáticas. O encaixamento é caracterizado por uma relação hierárquica evidente entre as orações, sendo a oração subordinada relegada ao papel de constituinte dentro da oração principal (LEHMANN, 1988). Em inglês, *that*, *who* e *which* são utilizados nesse tipo de oração (DECLERCK, 1991).

- (09) The food **that** we cooked was too spicy.

Em orações relativas de objeto, também é possível dispensar o conectivo.

- (10) The food (**Ø**) we cooked was too spicy.

A hipotaxe, por sua vez, é determinada por um grau menor de subordinação entre orações propriamente ditas (SOUZA, 2010). No inglês escrito, a separação entre a oração relativa hipotática e a oração principal é marcada por vírgulas. O conectivo *that* não pode ser utilizado nesse tipo de estrutura (DECLERCK, 1991).

- (11) Noah's restaurant, **which** opened last year, attracts many tourists.

Com o objetivo de repetir as condições utilizadas em trabalhos como Cabral (2016) e Traxler (2002), optamos pelo uso exclusivo de orações relativas encaixadas

em nosso experimento e no nosso *corpus*. Uma vez que o conectivo *that*, nessas orações, é o que mais se assemelha ao português *que* – e, acima de tudo, é o único que não oferece pistas sobre a animacidade do referente – apenas orações relativas com *that* constituirão nossas frases experimentais.

## 2.2 Estrutura da oração relativa

Do ponto de vista semântico, orações relativas podem ser descritas de modo relativamente simples. Como exposto acima, uma relativa modifica uma expressão nominal N qualquer. No exemplo a seguir (12), a expressão N é [raven] e a oração relativa é [which Edward saw].

(12) The [raven] [which Edward saw]

Conforme explicado por Chierchia (2003), tanto a expressão N quanto a oração relativa denotam, cada uma, uma classe de objetos quaisquer. É por meio da interseção entre ambas as classes que, semanticamente, ocorre a relativização: por terem igual valor, a classe *raven* é restrita à classe *which Edward saw*. O desafio, no que se refere à descrição sintática dessas estruturas, é determinar de que maneira ocorre a interseção entre a expressão nominal N e a oração relativa.

Na tradição gerativista, o interesse pelas orações relativas remonta, pelo menos, à década de 60, ocorrendo, a partir de então, a formulação de uma série de modelos teóricos na tentativa de explicar o funcionamento das estruturas relativas. Foi nesse contexto que Chomsky (1977) passou a argumentar que as orações relativas seriam um exemplo de aplicação das chamadas *regras de movimento de qu-*, ou *wh-movement*.

Entende-se por *wh-movement* o processo pelo qual os pronomes interrogativos (*who*, *whose*, *which*, *how*) são movidos para a esquerda de orações interrogativas (BROWN & MILLER, 2013). Nas línguas indo-europeias, muitos pronomes relativos são morfologicamente idênticos aos pronomes interrogativos. Por isso, é plausível defender a ideia de que ocorre *wh-movement* também na construção de orações

relativas. O mesmo se aplicaria a relativas com *that*, embora, nesses casos, não haja nenhum pronome interrogativo correspondente (DE VRIES, 2002).<sup>2</sup>

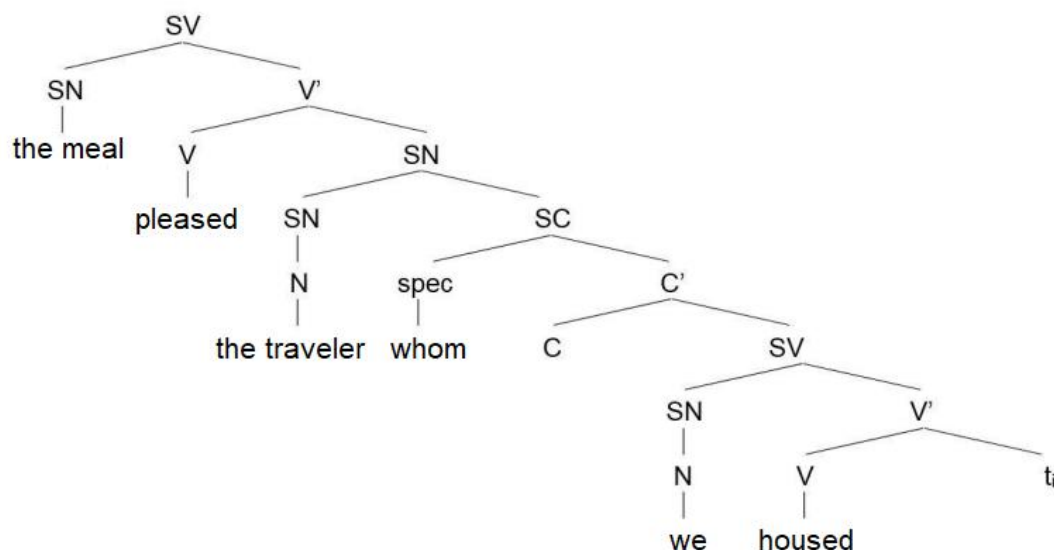
Entre as propriedades associadas por Chomsky (1977) ao *wh- movement*, podemos citar a presença de uma categoria vazia (*gap*) na estrutura resultante. Essa propriedade pode ser facilmente observada tanto em orações interrogativas quanto em relativas, como no exemplo (13), retirado de DE VRIES, 2002, p. 117.

(13) the meal which<sub>i</sub> you ate t<sub>i</sub>

Em (13), o pronome *which* é objeto de *ate*, não sendo possível a existência de outro objeto. Temos, portanto, uma categoria vazia depois do verbo, criada pelo movimento do pronome relativo para a esquerda da oração relativa.

Para compreender a proposta de Chomsky, precisamos recordar que orações relativas são adjungidas a outras orações. Isso significa que, nesse tipo de estrutura, o complementizador e o sintagma nominal estabelecem, em suas projeções máximas, uma relação de adjunção, conforme vemos no exemplo abaixo:

(14) The meal pleased the traveler whom we housed.



Num processo essencial para a gramaticalidade da sentença, o pronome relativo se move para a esquerda da oração relativa que introduz. Essa operação, chamada de

<sup>2</sup> Seguindo o exemplo de Chomsky (1977), De Vries (2002) considera que uma oração como *I fear the man that carries the gun* apresenta um operador vazio que corresponde a um pronome relativo.

*move*, gera o apagamento do constituinte movido, o que significa que o movimento desempenhado pelo pronome relativo deixa como resultado uma categoria vazia após o verbo. Por essa razão, o constituinte passa a ser interpretado em uma posição diferente daquela em que é produzido foneticamente. O sintagma verbal – representado em (14) como SV – tem seus argumentos interno e externo saturados (CABRAL, 2016).

Orações relativas, entretanto, apresentam estrutura significativamente diferente a depender do seu tipo: com extração de sujeito ou de objeto. Tais diferenças gerariam, por sua vez, uma assimetria no processamento dos dois tipos de oração relativa, sendo as de objeto consideradas, tradicionalmente, mais difíceis de se processar que as de sujeito. A natureza dessa assimetria e o histórico de estudos sobre ela são os temas da seção a seguir.

### 2.3 Assimetria no processamento de ORS e ORO

Sujeito e objeto são elementos sintaticamente distintos. O objeto, no papel de argumento interno, mantém uma relação estreita com o verbo transitivo direto. O sujeito, embora também seja necessário à saturação do verbo enquanto argumento externo, mantém com este uma relação menos próxima. Tais diferenças transparecem nos resultados de diversas pesquisas psicolinguísticas, com os mais variados procedimentos metodológicos e em diferentes línguas, que apontam uma assimetria na dificuldade de processamento de orações relativas de sujeito e de objeto (cf. CARAMAZZA & ZURIF, 1976; WANNER & MARATSOS, 1978; FRAUENFELDER, SEGUI & MEHLER, 1980; HOLMES & O'REGAN, 1981; FORD, 1983; KING & JUST, 1991; SCHRIEFERS, FRIEDERICI & KÜHN, 1995; JUST et al. 1996; CAPLAN, ALPERT & WATERS, 1998; GORDON, HENDRICK & JOHNSON, 2001; TRAXLER, MORRIS & SEELY, 2002; OLIVEIRA, 2013).

Diversos modelos teóricos têm sido propostos na tentativa de explicar o maior custo de processamento das ORO em relação às ORS. Dentre eles, um dos mais conhecidos é a Teoria da Predição Sintática Local (*Syntactic Prediction Locality Theory*, ou *SPLT*), de Gibson (1998). Segundo o autor, a teoria é constituída de dois componentes principais: o primeiro seria um componente de custo de memória que determinaria a quantidade de recursos computacionais requeridos no armazenamento de uma sentença recebida como *input*; e o segundo componente, por sua vez, estaria

relacionado ao custo de integração, de acordo com o qual se estabeleceria a quantidade necessária de recursos na integração de novas palavras às estruturas sintáticas até então construídas. Central a ambos os componentes é a ideia de localidade, pois quanto maior for o tempo que uma categoria prevista precisa ser armazenada na memória antes de a previsão ser satisfeita, maior o custo de processamento.

Para a *SPLT*, quando o ouvinte ou leitor se depara com uma oração relativa, a análise favorecida é determinada pelo custo de memorização das alternativas. Sendo assim, ao encontrar no *input* um pronome relativo, esse ouvinte/leitor preferirá interpretar a sentença em processamento como uma ORS, visto que, nesse caso, apenas dois constituintes precisarão ser mantidos na memória: o traço referente ao sintagma nominal como sujeito e um verbo. Em ORO, por outro lado, três constituintes precisam ser retidos: um sintagma nominal como sujeito, o traço referente ao sintagma nominal como objeto e o verbo. Essa estratégia de economia de recursos cognitivos seria, evidentemente, vantajosa quando a sentença que está sendo processada constitui, de fato, uma ORS. Entretanto, ao interpretar erroneamente uma ORO como ORS, a revisão da expectativa inicial geraria maior custo de processamento (MAK, VONK & SCHRIEFERS, 2002).

Uma outra teoria que busca explicar o fenômeno é a Estratégia do Antecedente Ativo (*Active Filler Strategy*, ou *AFS*). Descrita em Frazier (1987) e Frazier & Flores d'Arcais (1989), a *AFS* toma como ponto de partida as diferenças estruturais entre as ORS e as ORO. Nas primeiras, o pronome relativo se origina na posição de sujeito, sendo então movido para o início da oração. Esse movimento cria uma posição vazia, ou *gap*, no lugar de onde o pronome relativo veio, sendo este o *filler* do *gap*. Ao se deparar com um *filler*, o *parser* o atribui à posição de *gap* mais próxima. Essa estratégia se mostra apropriada para o processamento de ORS, visto que o *filler* é atribuído à posição de sujeito. Nas ORO, todavia, a posição de sujeito já se encontra ocupada, o que significa que a estratégia em questão induz o *parser* ao erro (MAK, VONK & SCHRIEFERS, 2002).

Como se pode perceber, tanto a *SPLT* quanto a *AFS* postulam fatores exclusivamente sintáticos na assimetria de processamento entre ORS e ORO, não havendo recurso a explicações de ordem semântica ou relativas à frequência das estruturas em questão. Desse modo, nenhuma das duas teorias consegue justificar o



papel que a animacidade, um fator semântico, parece exercer sobre o processamento dissimétrico dos dois tipos de oração relativa.

Diante da insuficiência desses modelos na explicação de alguns fenômenos psicolinguísticos, Gibson (2000) retoma uma variante de sua *SPLT*, chamada Teoria da Dependência Local, ou, no original em inglês, *Dependency Locality Theory (DLT)*. De acordo com esse modelo teórico, que, a princípio, tem muito em comum com a *SPLT*, o processamento de uma sentença também envolveria dois componentes distintos, ambos alimentados por recursos cognitivos. Em primeiro lugar, o ouvinte ou leitor, ao se deparar com uma sentença, precisa manter armazenada na memória sua estrutura sintática, ao mesmo tempo em que prevê quais categorias sintáticas são possíveis na continuação dessa sentença. Por exemplo, quando da aparição de um substantivo, espera-se a ocorrência de um verbo que complete a oração. O segundo componente envolvido no processamento linguístico seria a integração estrutural, isto é, a integração da palavra lida ou ouvida no momento presente a uma palavra encontrada previamente (PATEL, 2003).

É importante destacar que, segundo a *DLT*, o custo de processamento aumenta de acordo com a distância entre um novo elemento e o local de sua integração. Como exemplo, podemos analisar a relação entre as palavras *reporter* e *sent* nas sentenças (15-16). Em (15), o custo de integração entre as duas palavras é baixo, uma vez que se encontram quase adjacentes. Em (16), no entanto, a maior distância entre os dois termos geraria maior custo de integração, e portanto, de processamento. Perceba-se que (15) é uma ORS: *reporter* é o agente do verbo *sent*. Por sua vez, (16) é uma ORO, sendo *reporter* (ou, tecnicamente, o pronome relativo *who*) o objeto de *sent* (ibid.).

(15) The reporter who sent the photographer to the editor hoped for a story.

(16) The reporter who the photographer sent to the editor hoped for a story.

Além disso, Gibson (2000) defende que a velocidade de acesso a um item lexical depende de sua frequência de ocorrência. Dito de outro modo, para a *DLT*, um fator não sintático exerce influência sobre o processamento sentencial. Esse aspecto da teoria a afasta claramente de modelos teóricos como a *SPLT* e *AFS*, os quais recorrem unicamente a fatores sintáticos na explicação de diferenças de processamento entre estruturas diversas.

Gennari e MacDonald (2008), num experimento utilizando *gated sentence completion task*, descobriram que os participantes completavam orações relativas do tipo *The xxxxx that* tomando como pista a animacidade do seu referente: em caso de referente animado, a sentença era completada como ORS 85% das vezes, enquanto que os inanimados tinham uma chance de 65% de gerar ORO. Tais dados poderiam ser uma pista sobre a frequência dos diferentes tipos de orações relativas, e de como essa frequência possivelmente agiria na facilidade ou dificuldade de processamento dessas estruturas.

Dessa forma, a *DLT* pode ser um modelo apropriado na interpretação dos resultados de pesquisas que indicam a animacidade como condição relevante na assimetria de processamento de ORS e ORO, uma vez que a distribuição desigual dos quatro tipos de oração relativa (ORSa, ORSi, OROa e OROi) é um achado de alguns estudos de *corpus* em diferentes línguas (cf. ZUBIN, 1979, em alemão; MAK, VONK & SCHRIEFERS, 2002, em alemão e holandês; CABRAL, 2016, em português europeu e brasileiro). Encontramos resultado semelhante no nosso *corpus* da língua inglesa (cf. p. 66).

A compreensão da animacidade enquanto fenômeno cognitivo humano e sua atuação nas línguas naturais são os temas tratados nas seções que seguem.

## 2.4 Animacidade

O dicionário Michaelis (2020) da língua portuguesa define o adjetivo *animado* como aquilo ou aquele que “tem vida ou movimento”; ou, numa definição própria à linguística, como uma “categoria semântica de substantivos que indicam seres vivos (animais e pessoas) ou outros seres que podem ser assim considerados, como, por exemplo, anjo, demônio, ciclope, medusa, pégaso, etc.” Embora essa definição esteja correta, é necessário ir além dela se quisermos compreender o papel da animacidade na cognição humana de modo geral, e nas línguas naturais em particular.

De acordo com Vihman, Nelson e Kirby (2018), a maioria das línguas humanas – senão todas elas – são sensíveis à influência da animacidade dos referentes nominais. Distinções de animacidade podem aparecer de modo explícito na estrutura sintática de determinados idiomas, ou simplesmente motivar as preferências dos falantes no que se refere à ordem das palavras em uma frase ou ao uso de sinônimos em determinados contextos de fala. Os autores ainda afirmam que, embora, em sua essência, a

animacidade seja um construto baseado em uma realidade biológica, seres humanos o interpretam através de filtros culturais e cognitivos. Para tanto, podem ser levados em consideração aspectos como conspecificidade, empatia, textura, comestibilidade, agência e intencionalidade dos referentes.

#### **2.4.1 Animacidade real, animacidade no sistema conceitual humano e animacidade como traço formal das gramáticas humanas**

Bayanati e Toivonen (2019) propõem que a animacidade seja definida em três níveis distintos: *animacidade real* (ou *biológica*), *animacidade no sistema conceitual humano* e *animacidade como traço formal das gramáticas humanas*. No primeiro caso, diz-se que uma determinada entidade é animada se estiver viva, ou inanimada se não o estiver. A distinção biológica entre entidades animadas e inanimadas é, portanto, binária. Entidades inanimadas são destituídas de habilidades como a autopropulsão, a senciência e o livre arbítrio. Seres animados, ao contrário, definem-se justamente por essas capacidades, embora em níveis diversos.

A *animacidade no sistema conceitual humano*, ainda de acordo com as autoras, nem sempre corresponde com exatidão à realidade biológica. Como exemplo de desencontro entre os dois níveis de animacidade, pode-se citar o caso da personificação de entidades inanimadas. Além disso, enquanto no mundo externo a animacidade é uma categorização binária, no sistema conceitual humano ela se apresenta como um *continuum*, com seres mais ou menos animados entre os dois extremos. Animais de estimação são considerados mais animados que insetos, que, por sua vez, são mais animados que plantas. Dahl (2008), citado pelas autoras, defende que a característica básica de uma entidade animada é a capacidade de perceber e agir sobre o ambiente, capacidade existente em graus distintos entre as várias espécies vivas.

Por fim, ainda de acordo com Bayanati e Toivonen (ibid.), a *animacidade como traço formal das gramáticas humanas* constitui tão-somente um traço técnico com efeitos observáveis na estrutura da gramática, não possuindo um significado explícito. Devido à complexidade inerente ao conceito de animacidade, muitas línguas humanas diferem no modo como a gramaticalizam, por exemplo, em matéria de concordância e morfologia de caso. Nessas línguas, o traço de animacidade influencia o *output* linguístico de modo sistemático e previsível. Diversos idiomas, por outro lado, não

têm a animacidade como parte de seu sistema formal. Nesses casos, a animacidade dos referentes cria efeitos estatísticos significativos na distribuição de enunciados possíveis.

### **2.4.2 Animacidade e cognição humana**

A existência do sistema conceitual humano acima descrito revela a importância, do ponto de vista evolutivo, de se diferenciar as entidades do mundo externo como animadas ou inanimadas. Como explicam Nairne, VanArsdall e Cogdill (2017), um ser animado pode ser um predador, uma presa, um competidor ou um parceiro sexual. Sendo assim, é natural que a capacidade de distinguir entidades ao longo de um *continuum* de animacidade surja muito cedo no processo de desenvolvimento cognitivo humano. Revisando a literatura da área, Opfer e Gelman (2011) sugerem que bebês seriam capazes de discriminar seres animados e inanimados já no primeiro ano de vida, mostrando inclusive alguma sensibilidade a pistas de animacidade imediatamente após o nascimento. Mecanismos neurais especializados no processamento de informações de natureza física (como rostos e sons) e dinâmica (movimentos motivados por um objetivo específico) estariam em funcionamento já aos oito meses de idade. Essa compreensão inicial da animacidade seria aprimorada no decorrer do desenvolvimento cognitivo do indivíduo em contato com o seu ambiente.

O chamado *efeito de animacidade*, segundo o qual entes animados são lembrados mais facilmente que entes inanimados, apresenta-se como um fenômeno bem documentado nas pesquisas sobre memória. VanArsdall et al. (2013), por exemplo, utilizaram dois experimentos em que os participantes eram apresentados a pseudopalavras e propriedades associadas a entidades vivas (humanas) ou não vivas. Os autores descobriram que as pseudopalavras relacionadas a características animadas são memorizadas mais facilmente, o que condiz com o fato de que detectar e lembrar-se de entes animados – especialmente outros seres humanos – é essencial para a sobrevivência. Além disso, Félix, Pandeirada e Nairne (2019) empreenderam dois experimentos que envolviam a exposição dos sujeitos de pesquisa a listas de palavras que denotavam entidades animadas ou inanimadas. Em um dos experimentos, os participantes eram instruídos a memorizar a lista (*aprendizado intencional*). No outro, deveriam classificar a agradabilidade dos entes presentes na lista (*aprendizado*

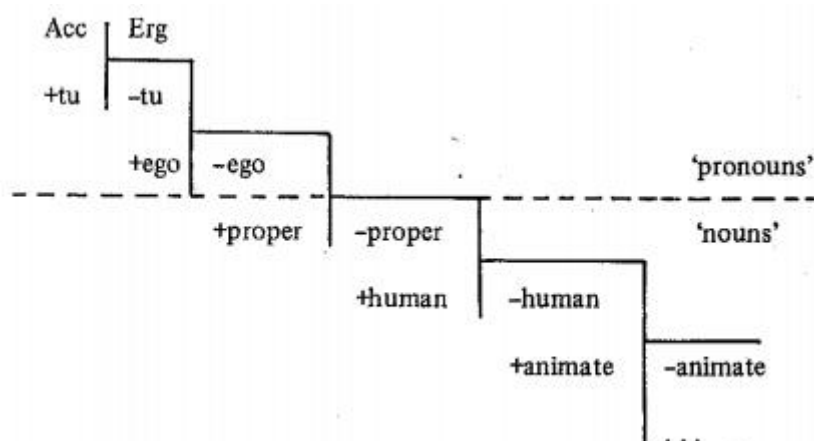
*incidental*). A memória dos sujeitos da pesquisa foi testada imediatamente após e 48h depois da realização dos experimentos; em ambos os casos, e em ambas as condições de aprendizagem, o efeito de animacidade foi detectado.

Até o momento, temos tratado a animacidade como um binário: as entidades em discussão são ora animadas, ora inanimadas. Entretanto, outras formas de compreender e classificar o fenômeno são possíveis.

### **2.4.3 Hierarquias de animacidade**

Embora a distinção entre seres vivos e não vivos seja teoricamente simples, muitos autores preferem tratar a animacidade não como um sistema binário, mas como um *continuum* que se estende das entidades mais animadas às menos. Essa percepção tem um longo histórico: já na Grécia antiga, o filósofo Aristóteles descreveu sua hierarquia da animacidade em *De Anima* (2006). Segundo ele, dentre os seres dotados de “alma”, as plantas, por sua capacidade de se alimentarem e se reproduzirem, apresentam as características básicas de qualquer ser vivo, constituindo, portanto, a base da escala; os animais inferiores, além disso, apresentam as habilidades de percepção e de mobilidade, o que os posiciona acima das plantas; os seres humanos, por sua vez, possuem todas essas capacidades, além do intelecto, que lhes é particular. Dessa forma, ocupam o topo da escala de animacidade.

Se na filosofia o conceito de animacidade tem uma longa tradição, é na linguística que ele ocupa posição privilegiada. Isso se reflete no grande número de estudos sobre categorias gramaticais (como número, classificadores, pessoa, etc.), em diversas línguas, que se utilizam do conceito (SORLIN & GARDELLE, 2018). A primeira formulação de uma hierarquia de animacidade na linguística moderna surgiria na década de 1970. Com base em um artigo originalmente apresentado em um encontro da Chicago Linguistic Society, em 1973, Silverstein (1976) descreveu uma hierarquia que expressa quão natural é para um sintagma nominal lexicalmente determinado funcionar como agente ou paciente de um verbo transitivo. Segundo ele, os sintagmas nominais no topo da hierarquia apresentam marcação de caso nominativo-acusativa, enquanto que os da base exibem marcação de caso ergativo-absolutiva.



**Figura 1:** Possibilidades de divisão lexical em marcação de caso (SILVERSTEIN, 1976, p. 176).

Matthews (2007) descreve o conceito de hierarquia de animacidade de tal modo que se aproxima bastante da abordagem proposta por Silverstein:

Uma proposta de ordenamento hierárquico de sintagmas nominais etc. que vão desde pronomes pessoais como *eu*, animados em nível máximo, até formas referentes a objetos sem vida, minimamente animados. As formas de um extremo da escala podem diferir sintaticamente daquelas do outro extremo: por exemplo, uma estrutura característica de uma língua acusativa pode ser encontrada com formas maximamente animadas, e as de uma língua ergativa com o resto. (Matthews, 2007, p. 20, tradução nossa)<sup>3</sup>

Com base no trabalho de Silverstein, Langacker (1991) formula sua **hierarquia da empatia** (*empathy hierarchy*). O autor parte do princípio de que os falantes categorizam as diversas entidades que compõem o mundo externo de acordo com um ponto de vista egocêntrico. Essas entidades são classificadas segundo sua capacidade de gerar empatia nos falantes: aquelas entidades que se assemelhem mais aos seres humanos figuram nas posições mais altas da hierarquia da empatia. Naturalmente, o topo dessa hierarquia é ocupado pelo próprio falante:

(17) Falante > interlocutor > humano > animal > objeto físico > entidade abstrata

<sup>3</sup> "A proposed hierarchical ordering of noun phrases etc. ranging from personal pronouns such as *I* as maximally 'animate' to forms referring to lifeless objects as minimally 'animate'. Those at one end of the scale may differ in syntax from those at the other: e.g. the construction characteristic of an accusative language may be found with those that are maximally animate, that of an ergative language with the remainder."

O autor explica sua esquematização argumentando que, como a coparticipação do ouvinte no evento discursivo iniciado pelo falante estabelece um “interesse comum imediato” entre os dois, é o interlocutor que vem em segundo lugar na hierarquia da empatia. Na terceira posição, temos uma pessoa além do falante e do ouvinte, um animal não humano em quarto lugar e assim por diante (LANGACKER, 1991). Para ilustrar a importância da empatia como fator determinante na gramática, o autor oferece os seguintes exemplos:

- |                             |                                  |
|-----------------------------|----------------------------------|
| (18) (a) The dog chased me. | (a') I was chased by the dog.    |
| (b) I chased the dog.       | (b') ??The dog was chased by me. |

Enquanto as frases (18a), (18b) e (18a') soam naturais, o mesmo não pode ser dito de (18b'), que subverte a escala de animacidade ao marcar como sujeito a entidade hierarquicamente inferior. Em algumas línguas, frases com (18b') não apenas soam pouco naturais, como são gramaticalmente impossíveis.

Em relação à interação entre animacidade e número gramatical, Smith-Stark (apud CORBETT, 2004) cita o caso do georgiano, em que um sujeito plural que denote um ser animado levará o verbo também para o plural, o que não ocorre com sujeitos plurais referentes a entidades inanimadas. Isso significa que os substantivos em georgiano são divididos em duas categorias simples: animados ou inanimados. Vários outros idiomas, entretanto, fazem essa divisão em pontos diferentes da escala de animacidade. Baseando-se em Smith-Stark, Corbett (ibidem, p. 56) apresenta a seguinte esquematização de uma hierarquia de animacidade:

- (19) Falante > ouvinte > 3ª pessoa > familiar > humano > animado > inanimado

Tendo em mente essa hierarquia, o autor propõe a chamada *restrição da Hierarquia de Animacidade na distinção singular-plural* (*constraint of the Animacy Hierarchy on the singular-plural distinction*), segundo a qual a distinção de número em uma língua deve afetar um segmento no topo da Hierarquia de Animacidade. A imagem abaixo, produzida pelo próprio autor, ilustra bem sua posição:





o segundo a mulheres e animais, e os dois últimos a seres inanimados. Apenas alguns dos substantivos dos dois primeiros gêneros, quase todos denotando humanos, têm uma marca específica de plural. Artigos e adjetivos podem concordar em número (plural) somente com substantivos desses dois gêneros. Em mundari, de modo semelhante, verbos concordam em número apenas com substantivos animados.

Por fim, é necessário esclarecer que a ideia de uma Hierarquia da Animacidade tem, na Linguística, uma base geralmente funcionalista. Linguistas com uma abordagem mais formal tendem a rechaçar o uso de hierarquias semânticas na descrição de padrões linguísticos (DE SWART & DE HOOP, 2018).

#### **2.4.4 Efeitos da animacidade nas línguas**

Uma importante questão a ser debatida é se as hierarquias de animacidade vistas acima classificam entidades do mundo externo ou unidades linguísticas. Segundo Sorlin e Gardelle (2018, n.p.), na linguística, a Hierarquia de Animacidade é uma classificação de substantivos, pronomes e sintagmas nominais. No entanto, as autoras destacam que, por trás desse ranqueamento de unidades linguísticas, existe uma “hierarquização de entidades que resultam em restrições gramaticais em algumas áreas das línguas”. Elas ainda pontuam que “a existência de consequências gramaticais reflete a importância da hierarquia: a gramática codifica os domínios da experiência que são importantes para a cultura” (tradução nossa).<sup>5</sup>

Exemplos de línguas que gramaticalizaram a animacidade foram dadas na seção anterior. Todavia, em muitas outras línguas, a animacidade não é explicitamente codificada na gramática, influenciando, na verdade, a distribuição estatística no uso de determinadas estruturas linguísticas. Um exemplo de efeito estatístico da animacidade nos usos linguísticos é a tendência de entidades animadas a serem expressas como sujeitos de uma oração (BAYANATI & TOIVONEN, 2019). Isso é natural quando se leva em conta que os eventos sobre os quais falamos costumam ser iniciados por seres animados – capazes de se mover de forma autônoma e de modificar seu ambiente. Ademais, nossa atenção geralmente é direcionada às atividades humanas, devido ao

---

<sup>5</sup>“But underlying this ranking of linguistic units is a hierarchisation of entities, which causes grammatical constraints in some areas of languages. The existence of grammatical consequences reflects the importance of hierarchy: grammar encodes those domains of experience which are important to cultures.”

ponto de vista egocêntrico a partir do qual nossa espécie interpreta os fatos do mundo (VIHMAN & NELSON, 2019).

Podemos citar alguns estudos que corroboram as afirmações acima. Analisando um *corpus* composto de 3.100 enunciados escritos em sueco, Dahl e Fraurud (1996) observaram que em mais de 97% das orações transitivas o sujeito é hierarquicamente superior ao objeto no quesito animacidade. Dito de outro modo, apenas 2,6% dos enunciados apresentam um sujeito que não denota um ser humano e um objeto que o denota.

Mak, Vonk e Schriefers (2002) realizaram um estudo de dois *corpora* linguísticos: um em holandês e outro em alemão. No primeiro, os autores selecionaram 286 orações relativas com verbos transitivos e que não incluíssem pronomes. Das orações com um antecedente animado, 140 eram orações relativas de sujeito (ORS) e apenas 2 de objeto (ORO). Em alemão, das 84 orações relativas com um antecedente animado, 82 eram ORS, sendo apenas 2, mais uma vez, ORO.

Quando consideramos o modo como a distinção entre seres animados e inanimados influencia explícita ou implicitamente as línguas humanas, ao ponto de determinar não apenas o que pode ou não pode ser dito, mas também a frequência de ocorrência das estruturas gramaticalmente possíveis, podemos nos perguntar se a animacidade influenciaria, ademais, a forma como compreendemos a linguagem.

#### **2.4.5 Animacidade e o processamento de orações relativas**

Embora a assimetria de processamento de ORS e ORO – especificamente, a maior facilidade no processamento das primeiras em relação às últimas – seja um fenômeno bem documentado, o papel da animacidade nesse processo nem sempre foi levado em consideração. Em um dos artigos mais influentes da área, Mak, Vonk e Schriefers (2002) descrevem um estudo de *corpus* constituído de textos jornalísticos em alemão e holandês que indica a importância da animacidade na distribuição de ORS e ORO nas duas línguas. Tal constatação evidenciou a necessidade de se controlar a animacidade na elaboração das frases experimentais utilizadas em pesquisas sobre processamento de orações relativas, uma vez que as diferentes frequências das estruturas investigadas poderiam estar enviesando os resultados obtidos até então.

Tendo isso em mente, os autores desenvolveram um experimento de leitura automonitorada em holandês, realizado com 40 participantes universitários. O material era composto de 32 conjuntos de quatro sentenças, cada uma representando um tipo diferente de oração relativa (de sujeito com objeto animado, de objeto com objeto animado, de sujeito com objeto inanimado e de objeto com objeto inanimado) conforme os exemplos de (20) a (23).

(20) Oração relativa de sujeito, objeto animado

Vanwege het onderzoek moeten de inbrekers, die de bewoner beroofd hebben, nog een tijdje op het politiebureau blijven.<sup>6</sup>

(21) Oração relativa de objeto, objeto animado

Vanwege het onderzoek moet de bewoner, die de inbrekers beroofd hebben, nog een tijdje op het politiebureau blijven.<sup>7</sup>

(22) Oração relativa de sujeito, objeto inanimado

Vanwege het onderzoek moeten de inbrekers, die de computer gestolen hebben, nog een tijdje op het politiebureau blijven.<sup>8</sup>

(23) Oração relativa de objeto, objeto inanimado

Vanwege het onderzoek moet de computer, die de inbrekers gestolen hebben, nog een tijdje op het politiebureau blijven.<sup>9</sup>

Cada participante sentava-se diante de um monitor de computador com um painel de três botões. Ao pressionar o botão do meio, o teste tinha início. Uma sentença inteira era exibida na tela, porém com todas as letras e vírgulas substituídas por traços. Apertando o mesmo botão novamente, a primeira palavra aparecia; ao repetir essa operação, a segunda palavra era exibida, enquanto que a primeira voltava a ser ocultada por traços, e assim por diante. Após lerem a última palavra e pressionarem o botão do meio mais uma vez, o próximo estímulo era exibido em 3/4 dos casos; 1/4 das sentenças eram seguidas de uma afirmação

---

<sup>6</sup>“Por causa da investigação, os invasores, que roubaram o morador, tiveram que passar um tempo na delegacia.” (Tradução nossa)

<sup>7</sup>“Por causa da investigação, o morador, que os invasores roubaram, teve que passar um tempo na delegacia.” (Tradução nossa)

<sup>8</sup>“Por causa da investigação, os invasores, que roubaram o computador, tiveram que passar um tempo na delegacia.” (Tradução nossa)

<sup>9</sup>“Por causa da investigação, o computador, que os invasores roubaram, teve que passar um tempo na delegacia.” (Tradução nossa)

sobre seu conteúdo cuja veracidade o participante devia julgar. Ao todo, cada sujeito era exposto a 32 sentenças experimentais (retiradas dos conjuntos supramencionados) distribuídas entre 40 frases distratoras. O tempo de leitura de cada frase foi aferido, partindo-se do princípio de que um estímulo que causasse mais dificuldade de processamento acarretaria um tempo maior de leitura.

Ao analisarem os dados desse experimento, os autores descobriram que não havia diferença nos tempos de leitura de orações relativas de sujeito e de objeto quando o objeto da oração relativa era inanimado. Esse resultado contrasta com os encontrados na maioria dos estudos anteriores, que não controlavam o fator da animacidade. Segundo esses estudos, orações relativas com extração de objeto geravam mais dificuldade de processamento devido a uma maior complexidade sintática.

Um outro estudo importante na área foi o empreendido por Traxler, Morris e Seely (2002). Para investigar a possível assimetria entre o processamento de ORS e ORO, os autores elaboraram três experimentos. Nos dois primeiros, sem controle de fatores semânticos, foi confirmada a maior dificuldade dos participantes na compreensão de ORO. No entanto, o terceiro experimento demonstrou que, uma vez que a animacidade do referente das orações relativas é levada em consideração, a assimetria de processamento de ORS e ORO é grandemente reduzida.

Para esse experimento, foram preparados 28 conjuntos de quatro tipos de orações relativas (ORSa, ORSi, OROa e OROi) em inglês, como exemplificadas em (24)-(27).

(24) The director that watched the movie received a prize at the film festival.

(ORSa)<sup>10</sup>

(25) The director that the movie pleased received a prize at the film festival.

(OROa)<sup>11</sup>

(26) The movie that pleased the director received a prize at the film festival.

(ORSi)<sup>12</sup>

---

<sup>10</sup>“O diretor que assistiu ao filme recebeu um prêmio no festival de cinema.” (Tradução nossa)

<sup>11</sup>“O diretor que o filme agradou recebeu um prêmio no festival de cinema.” (Tradução nossa)

<sup>12</sup>“O filme que agradou o diretor recebeu um prêmio no festival de cinema.” (Tradução nossa)

(27) The movie that the director watched received a prize at the film festival.  
(ORO<sub>i</sub>)<sup>13</sup>

Os itens foram randomizados, sendo uma sentença de cada conjunto atribuída a uma de quatro listas de itens experimentais, de forma que nenhum participante seria exposto a mais de uma sentença de cada conjunto. Além dessas frases experimentais, foram adicionadas 62 frases distratoras a cada lista. Enquanto os participantes liam os estímulos, seus movimentos oculares foram rastreados a fim de registrar a quais segmentos das frases eles dedicavam mais atenção, ou a quais retornavam. Após algumas frases distratoras, uma pergunta de compreensão deveria ser respondida, o que visava garantir que os sujeitos estariam atentos ao experimento.

Feita a análise dos dados angariados, os autores indicaram um efeito importante da animacidade no processamento de orações relativas: houve uma redução na dificuldade associada às ORO quando o sujeito da sentença era inanimado. Esse resultado não pode ser explicado se considerarmos que apenas fatores de natureza sintática estão envolvidos no processamento de orações relativas.

Igualmente notável foi o estudo empreendido por Cabral (2016), baseado em cinco experimentos de leitura automonitorada com participantes brasileiros e portugueses. Nos dois primeiros experimentos (um realizado com brasileiros e o outro com portugueses), mediu-se o tempo de processamento de orações relativas de sujeito e objeto, todas com traço animado, como nos exemplos abaixo.

(28) João / **que encontrou José** / naquele dia / não quis / pedir perdão.

(29) João / **que José encontrou** / naquele dia / não quis / pedir perdão.

A relativa encaixada – o segmento crítico – aparece em negrito. Assim como em pesquisas semelhantes, o autor encontrou resultados que corroboravam a maior dificuldade de processamento das relativas de objeto.

Nos experimentos 3 e 4, também feitos com voluntários dos dois países, testaram-se quatro condições diferentes: orações relativas com sujeito animado, sujeito inanimado, objeto animado e objeto inanimado, exemplificadas a seguir.

---

<sup>13</sup>“O filme que o diretor assistiu recebeu um prêmio no festival de cinema.” (Tradução nossa)

(30) O gato / **que encantou a menina** / de tarde / estava brincando / na grama do jardim.

(31) O livro / **que encantou a menina** / de tarde / era sobre a vida / das aves do cerrado.

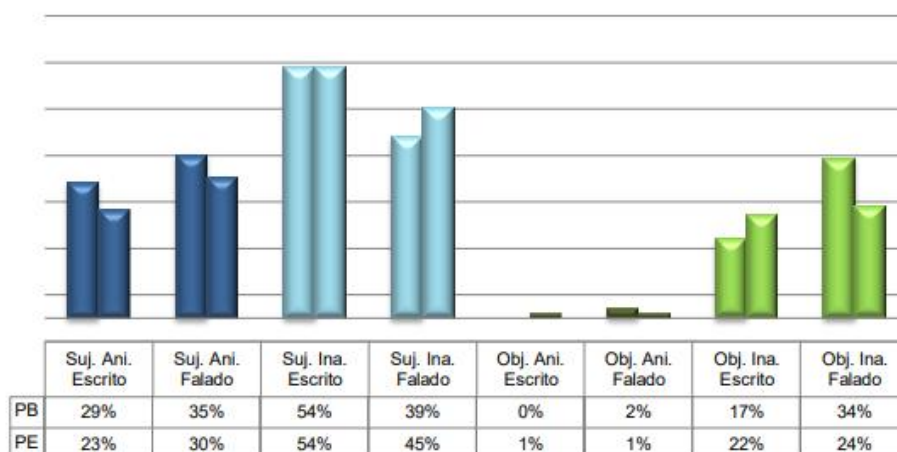
(32) O livro / **que a menina encontrou** / de tarde / na minha casa / era muito interessante.

(33) O gato / **que a menina encontrou** / de tarde / naquela praça / comeu a ração dele.

Esses experimentos indicaram um efeito principal do traço de animacidade, sendo as relativas de objeto inanimado lidas mais rapidamente que as de sujeito. Por outro lado, as OROa foram a condição experimental que mais dificuldade causou aos participantes da pesquisa.

Por fim, o autor empreendeu um último teste, dessa vez de leitura de substantivos animados e inanimados isolados (isto é, fora de um contexto sintático), apenas com falantes brasileiros. Esse experimento não gerou resultados significativos, o que apontaria que o efeito de animacidade encontrado nas quatro primeiras análises era referente às estruturas frasais como um todo, e não aos substantivos isolados.

Além dos experimentos propriamente ditos, Cabral também criou *corpora* de orações relativas em português nas variedades brasileira e europeia por meio da coleta de textos escritos e falados. Os resultados desses *corpora*, que podem ser verificados no gráfico abaixo, retirado de Cabral (2016, p. 123), demonstram que o tipo de relativa com maior custo de processamento (OROa) é também o menos frequente na língua portuguesa, registrando ocorrências baixíssimas, que vão de 0 a 2%.



**Gráfico 1:** *Corpora* escrito e falado do português brasileiro e do português europeu retirados de Cabral (2016, p. 123).

Os estudos mencionados acima investigaram o processamento por parte de falantes nativos. É possível que resultados semelhantes sejam encontrados quando se avalia o processamento de L2? Pesquisas desse tipo têm sido menos frequentes, mas destacamos o estudo realizado por Piñar et al. (2017), que também fizeram uso da técnica de rastreamento ocular. O diferencial desse trabalho é a participação de sujeitos bilíngues em inglês e Língua de Sinais Americana (*American Sign Language*, ou *ASL*), trazendo para o debate sobre processamento linguístico não apenas o bilinguismo, mas também a bimodalidade.

Foram 39 os participantes desse experimento, todos estudantes universitários surdos com idades que variavam de 18 a 33 anos. Apesar de serem bilíngues em inglês e ASL, nem todos eram usuários nativos desta última, além de variarem no nível de proficiência leitora em língua inglesa. Por outro lado, todos eles viviam num ambiente de imersão diária em língua de sinais.

Para o experimento, utilizaram-se 44 conjuntos de orações relativas, cada um contendo as quatro condições já mencionadas (ORSa, OROa, ORSi, OROi), além de 78 frases distratoras. Foram controladas a extensão e frequência dos substantivos animados e inanimados. Apenas uma sentença de cada conjunto foi atribuída a cada uma das quatro listas a serem lidas pelos participantes. Após cada sentença, o sujeito deveria responder a uma pergunta simples do tipo sim/não.

Em consonância com estudos anteriores com ouvintes que tinham o inglês como L1, os resultados desse experimento indicaram que os participantes surdos tiravam proveito de pistas tanto de natureza sintática quanto semântica ao processarem orações relativas: ORS eram lidas mais rapidamente, o que demonstrava a importância de fatores sintáticos no processo; entretanto, também ficou claro que, no caso das sentenças estruturalmente mais complexas – as ORO – informações de caráter semântico exerceram influência sobre seu processamento (OROa tiveram tempos de leitura mais longos que OROi, enquanto que nenhuma diferença significativa foi encontrada entre OROi e ORS, ou entre ORSa e ORSi).

As pesquisas vistas nesta seção demonstram como fatores distintos (sintáticos e semânticos, por exemplo, mas possivelmente também relacionados à frequência de determinadas estruturas) influenciam o processamento de orações relativas. Para

compreender o quanto esses resultados são significativos, precisamos, no entanto, considerar os modelos teóricos que há décadas buscam explicar os mecanismos cognitivos por trás do processamento de sentenças.



### 3. PROCESSAMENTO DE SENTENÇAS

A sentença tem sido tradicionalmente tomada como o domínio em que relações sintáticas são definidas. Constituindo uma entidade sintática independente, ela é o principal meio de expressão de uma atitude proposicional, e, dessa forma, de um ato de fala (BURTON-ROBERTS, 2011). Diversos psicolinguistas têm estudado os mecanismos cognitivos por trás do processamento – i.e., a compreensão – de sentenças. De acordo com Tanenhaus (2003):

[o] objetivo das pesquisas em processamento de sentenças é entender as representações que as pessoas formam enquanto compreendem uma sentença, bem como os processos envolvidos no desenvolvimento de tais representações. O processamento de sentenças envolve o reconhecimento das palavras em uma frase, a determinação das relações sintáticas e semânticas entre essas palavras e a interpretação da sentença no que diz respeito ao contexto linguístico e extralinguístico que lhe é relevante.<sup>14</sup> (Tanenhaus, 2003, p. 1; tradução nossa).

Tipos diferentes de informação guiam esse processo, nomeadamente fatores sintáticos, lexicais, prosódicos, discursivos e contextuais. O momento em que esses tipos de informação tornam-se disponíveis para o ouvinte/leitor e o modo como interagem são, todavia, motivo de intenso debate (FERREIRA & ÇOKAL, 2016; FRANÇA & GOMES, 2015). Tradicionalmente, ao menos no que se refere ao processamento linguístico em L1, defende-se o papel primordial de princípios de *parsing* sintático, sendo o resultado final do processo modulado por pistas semânticas, senso-semânticas e contextuais (PIÑAR et al, 2017).

Entretanto, de acordo com Traxler (2014) observa-se um crescente distanciamento, por parte dos pesquisadores, de descrições de um processamento linguístico estritamente *bottom up*, serial e encapsulado, privilegiando-se, em seu lugar, modelos mais interativos. O autor acrescenta:

Três tipos de desdobramentos inter-relacionados estão mudando a forma como psicolinguistas entendem a interpretação linguística de um modo geral e a natureza dos processos de *parsing* sintático em particular. Incluem-se aí a ênfase relativamente recente em

---

<sup>14</sup> “The goal of research in sentence processing is to understand the representations people form as they understand a sentence and the processes involved in developing these representations. Sentence processing involves recognizing the words in a sentence, determining the syntactic and semantic relationships among these words, and interpreting the sentence with respect to the relevant linguistic and nonlinguistic context.”

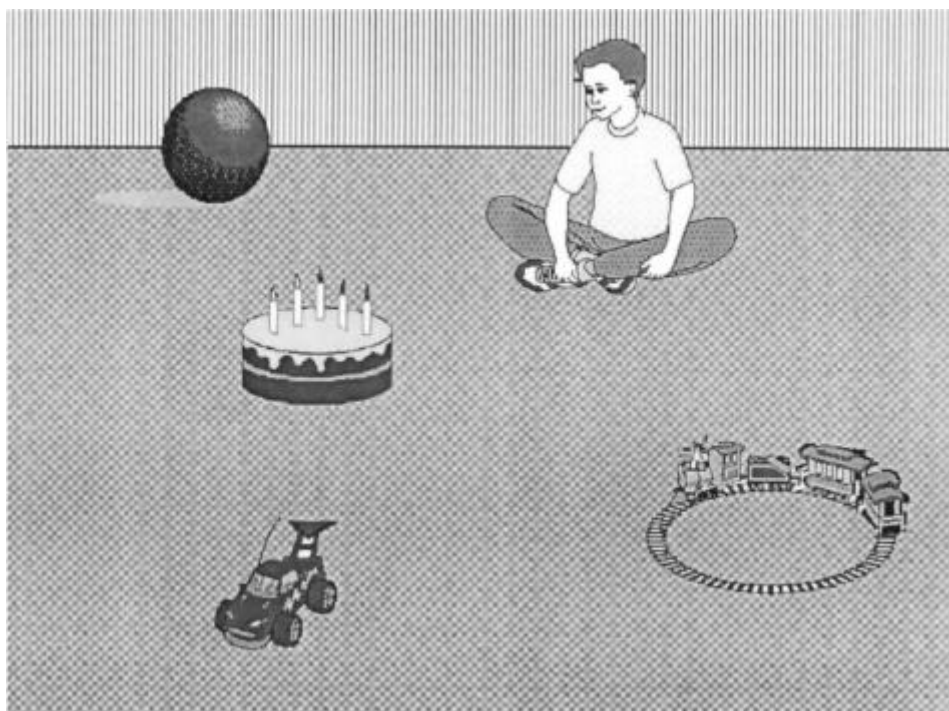
processos antecipatórios ou preditivos, a aplicação de estimativas de probabilidade bayesiana à compreensão linguística e as mudanças na concepção de compreensão linguística sob a ótica do processamento *good-enough* ou satisfatório.<sup>15</sup> (Traxler, 2014, p. 606, tradução nossa)

O processamento linguístico ocorre de forma notavelmente rápida. Uma das explicações possíveis para tal rapidez seria a de que o ouvinte ou leitor não espera passivamente pelo aparecimento da continuação de um *input* linguístico, mas cria hipóteses sobre como uma determinada sentença pode prosseguir. Um trabalho influente na área é o de Altmann e Kamide (1999), que analisaram os movimentos oculares de participantes enquanto estes eram expostos a determinados estímulos visuais e linguísticos.

A cena abaixo é um exemplo de estímulo visual utilizado na pesquisa. Enquanto visualizava a cena, o participante ouvia uma de duas frases relacionadas: “*the boy will move the cake*” ou “*the boy will eat the cake*”. No primeiro caso, o verbo *move* pode ser adequadamente associado a qualquer um dos objetos presentes na cena (à bola, ao bolo, ao trem em miniatura ou ao carrinho de brinquedo), pois todos eles são passíveis de serem deslocados pelo menino. No segundo caso, porém, o verbo *eat* leva o participante a antecipar *the cake* como complemento lógico da frase, já que o bolo é o único objeto comestível na cena. Os pesquisadores observaram que o *onset* dos movimentos oculares sacádicos dos participantes em direção ao objeto-alvo (o bolo) era significativamente mais demorado na primeira condição (“*The boy will move the cake.*”). Esse resultado sugere que os ouvintes eram capazes de antecipar o objeto direto da oração antes mesmo de seu aparecimento no enunciado.

---

<sup>15</sup> “Three sets of related developments are changing the way that psycholinguists view language interpretation in general and the nature of syntactic parsing processes in particular. These include the relatively recent emphasis on predictive or anticipatory processes, the application of Bayesian probability estimation to language comprehension, and the changing view of language comprehension through the lens of satisfying or good-enough processing.”



**Figura 3:** Exemplo de cena utilizada em um dos experimentos de Altmann e Kamide (1999, p. 250).

Em relação às abordagens de processamento linguístico paralelo, Traxler (2014) afirma que partem da hipótese de que a interpretação poderia se basear não no enunciado produzido, mas numa aproximação que o ouvinte ou leitor faz desse enunciado. Tal aproximação dependeria tanto das pistas extraídas do enunciado quanto das crenças que o sujeito pode ter sobre determinados eventos, além do que os falantes tendem a dizer sobre esses eventos e como costumam estruturar o modo como os expressam. Essas abordagens são compatíveis com a hipótese do *good-enough parsing*, proposta por Ferreira, Bailey e Ferraro (2002).

Tomando como base Sachs (1967), de acordo com quem a estrutura construída pelo processador linguístico seria frágil e tenderia a decair rapidamente, Ferreira, Bailey e Ferraro (2002) argumentam que a representação mental gerada por um enunciado precisa de suporte quase imediato, advindo do contexto ou de esquemas mentais baseados na experiência. Os autores apontam que sentenças como “*The dog was bitten by the man.*” (“O cachorro foi mordido pelo homem.”), são frequentemente (mais de 25% das vezes) mal interpretadas pelos leitores. Os mecanismos sintáticos proporcionariam uma interpretação correta da sentença em questão. Todavia, essa estrutura sintática, por sua fragilidade, necessitaria de reforço por meio do contexto,

uma vez que a interpretação correta a ela não poderia contar com a corroboração de esquemas da memória de longo prazo.

A exposição sobre processamento de sentenças até aqui empreendida teve como objetivo oferecer um panorama geral sobre a área, citando os diferentes tipos de pistas que podem desempenhar um papel importante nesse processo. A partir deste ponto, podemos discriminar melhor os diferentes modelos de processamento linguístico, classificando-os segundo quatro princípios, a saber: serialidade, paralelismo, modularidade e interatividade. Esses princípios são explanados abaixo, acompanhados de teorias que os exemplificam.

### 3.1 Modelos seriais x modelos paralelos

De acordo com os modelos seriais de processamento linguístico, o *parsing* de uma sentença envolve “uma série linear de operações mentais, em um processo em que a saída da operação anterior fornece a entrada da próxima operação” (MAIA, 2015, p. 16). O *parser*, portanto, seria capaz de lidar com apenas uma análise sintática por vez. Ainda segundo Maia, em modelos seriais, “a estrutura vai sendo construída à medida em que cada item vai sendo acessado, sem atraso, sem esperar o fim da frase”, o que caracterizaria o aspecto *incrementacional* do processamento linguístico.

Como exemplo de modelo serial, tem-se a Teoria do Labirinto ou do *Garden Path* (TGP), como proposta por Frazier (1979). Segundo essa teoria, assim como cada trajeto feito em um labirinto é formado por uma sequência única de direções percorridas, sendo necessário abandonar um percurso antes de tentar seguir outro, também o processamento linguístico seguiria uma estrutura sintática única, o que confere à Teoria do Labirinto seu caráter serial. O processamento ocorreria dessa maneira por seguir princípios baseados na economia de recursos cognitivos, o que nos levaria a escolher a interpretação advinda da estrutura mais simples do ponto de vista sintático (LEITÃO, 2011). Essa parece uma estratégia razoável, mas por vezes nos faz incorrer em interpretações equivocadas de um enunciado.

Em inglês, a expressão *lead somebody up the garden path* (literalmente: “conduzir alguém pelo trajeto do jardim”)<sup>16</sup> significa ludibriar, enganar alguém (CAMBRIDGE DICTIONARY, 2020). Desse modo, sentenças do tipo *garden path*

---

<sup>16</sup> Uma referência provável aos jardins em forma de labirinto.

são aquelas em que “pistas estruturais, ambiguidade lexical ou uma combinação de ambas induzem o leitor ou ouvinte a uma interpretação incorreta até que uma pista desambiguadora apareça mais adiante na sentença” (PAM, 2013; tradução nossa).<sup>17</sup>

São inúmeros os exemplos de sentenças do tipo *garden path*, algumas das quais se tornaram emblemáticas na área da psicolinguística. O exemplo aqui citado foi retirado de Ferreira e Henderson (1991, p. 276):

(34) Because Bill drinks wine is never kept in the house.

Uma leitura inicial de (34) levaria à interpretação de *wine* como objeto direto do verbo *drink* (“Porque Bill bebe vinho...”). No entanto, ao chegar ao verbo *is*, o leitor se dá conta de que é necessário realizar uma reanálise sintática da sentença. Ocorreu aí o efeito *garden path*. A reanálise leva o leitor à interpretação correta (“Porque Bill bebe, nunca se guarda vinho na casa.”).<sup>18</sup>

Frazier e Fodor (1978) presumem que o processamento sintático ocorre em dois passos. No primeiro, nós lexicais e sintagmáticos são associados a conjuntos de palavras dentro do enunciado em processamento. Esse estágio é chamado de *Preliminary Phrase Packager*, ou Empacotador Preliminar de Sintagmas, numa tradução livre. Também é conhecido como *Sausage Machine* (Máquina de Salsicha). O segundo passo consiste na combinação desses sintagmas devidamente estruturados, formando um marcador sintagmático completo.

Dois princípios regem o processamento de sentenças como descrito pela Teoria do *Garden Path*. O primeiro deles, chamado Princípio da Aposição Local (*Late Closure*) postula que, sempre que possível, itens lexicais que vão sendo encontrados pelo ouvinte ou leitor devem ser apostos à oração ou ao sintagma que está sendo processado no momento. Por sua vez, o Princípio da Aposição Mínima (*Minimal Attachment*) pressupõe a aposição de um item à estrutura sentencial de modo a utilizar o mínimo possível de nós sintáticos (FRAZIER, 1979, apud MAIA, 2015).

Embora inicialmente formulados para explicar fenômenos concernentes à língua inglesa, Frazier supunha que ambos os princípios seriam universalmente aplicáveis. Os resultados obtidos com pesquisas posteriores, em que se investigaram esses

---

<sup>17</sup> “Sentence where structural cues, lexical ambiguity or a combination of both mislead the reader or listener to an incorrect interpretation until a disambiguating cue appears later in the sentence.”

<sup>18</sup> Uma tradução mais literal seria “Porque Bill bebe, vinho nunca é mantido na casa”.

princípios em diferentes línguas e por meio do uso de métodos diversos, confirmaram o poder explicativo da Aposição Mínima, mas não do Princípio da Aposição Local. Cabral (2016) aponta que testes feitos em algumas línguas utilizando frases ambíguas (como a exposta abaixo) têm gerado resultados que não se justificam pela Teoria do *Garden Path*.

(35) Alguém atirou no empregado da atriz que estava na varanda.

Observe-se que uma aposição local significaria que a oração relativa *que estava na varanda* seria aposta ao sintagma nominal *a atriz*. Todavia, se a mesma relativa estiver aposta ao sintagma nominal *o empregado*, teríamos uma aposição não local. Enquanto a preferência por uma aposição local tem sido observada em inglês, trabalhos em línguas como o espanhol (CUETOS & MITCHELL, 1988) e o português (RIBEIRO, 1998) levantam sérios questionamentos quanto à sua aplicabilidade em outros idiomas. Tais resultados põem em cheque os postulados originais da Teoria do *Garden Path*, e, por extensão, de uma visão modular do processamento linguístico. Isso levou alguns pesquisadores a dedicar sua atenção a modelos de processamento que assumam um *parser* capaz de produzir simultaneamente mais de uma representação sintática da sentença a ser processada.

Parte-se do princípio de que *parsers* de arquitetura paralela, quando confrontados com uma sentença ambígua, computam paralelamente todas as análises sintáticas possíveis dessa sentença. Com o auxílio de informações não sintáticas, análises incompatíveis podem então ser descartadas (HOPF et al., 2003). Um exemplo de modelo de processamento linguístico paralelo é o proposto por MacDonald, Pearlmutter e Seidenberg (1994), que descrevem um mecanismo de processamento central baseado nas Teorias de Satisfação de Condições. Para esses autores, a resolução de ambiguidade estrutural – isto é, sintática – está diretamente relacionada à resolução de ambiguidade lexical, sendo ambas alicerçadas sobre representações cognitivas comuns. Eles consideram que o processamento lexical abarca “a ativação de diferentes tipos de informação associados com uma forma lexical, e, então, o uso dessa informação na computação de representações nos outros níveis [sintático e discursivo]” (p. 682; tradução nossa).<sup>19</sup> Além de apresentar associações entre

---

<sup>19</sup> “Lexical processing involves activating different types of information associated with a word form and then using this information to compute representations at the other levels.”

informações de ordem fonológica, ortográfica e semântica, o léxico mental de um falante comportaria dados relativos à frequência de cada um dos seus itens.

Essa frequência de itens lexicais é relevante para a presente discussão, pois sabe-se que a animacidade tem efeito sobre a distribuição de orações relativas nas línguas (CABRAL, 2016). Tanenhaus (2003), por exemplo, afirma que sentenças formadas por palavras e estruturas mais frequentes são processadas mais facilmente. Desse modo, se o falante fizer uso de informações concernentes à frequência de palavras no processamento linguístico, é possível que a animacidade afete esse processo.

Entretanto, Levy e Gibson (2013) oferecem uma crítica pertinente a modelos teóricos que tratam a frequência de uso como único fator explicativo na assimetria de processamento entre ORS e ORO, deixando de lado aspectos relacionados à complexidade sintática de tais estruturas. Os autores utilizam as duas orações relativas abaixo para explicar seu ponto.

(36) The reporter that attacked the senator admitted the error.

(37) The reporter that the senator attacked admitted the error.<sup>20</sup>

Como se sabe, a primeira sentença – uma ORS – oferece menos dificuldade de processamento que a segunda – uma ORO. Argumentos convincentes para explicar esse fenômeno são apresentados tanto por teorias baseadas na frequência de uso quanto por aquelas que apontam as limitações de nossos recursos cognitivos – como a memória – como fator mais relevante. No entanto, o *local* nas orações em que a assimetria se torna mais evidente não pode ser adequadamente explicado pelas teorias do primeiro grupo. Estas apontariam o sintagma nominal *the senator* como gerador de uma previsão equivocada por parte do ouvinte/leitor que se deparasse com a oração (37), pois uma análise de frequência o levaria a esperar que uma oração relativa com um antecedente animado seria de sujeito. A maior parte da dificuldade, no entanto, está associada ao verbo *attacked*. Como os autores explicam:

Essa dificuldade ocorre a despeito do fato de que o verbo da oração relativa em ORO deveria ser mais, não menos esperado, do que em ORS: há menos eventos sintáticos possíveis após *The reporter that*

---

<sup>20</sup>“O repórter que atacou o senador admitiu seu erro.” e “O repórter que o senador atacou admitiu seu erro.”, respectivamente (tradução nossa).

*the senator...* do que depois de *The reporter that...*, e há menos coisas que um senador possa fazer a um repórter do que um repórter possa fazer (...). (ibid., p. 2; tradução nossa).<sup>21</sup>

Continuamos, portanto, em busca de uma teoria de processamento linguístico que abarque tanto fatores de ordem sintática quanto aqueles relacionados à semântica e à frequência de uso de determinadas estruturas.

### 3.2 Modelos modulares x modelos interativos

Na área da psicolinguística, os estudos sobre processamento linguístico têm gerado inúmeras discussões entre, por um lado, aqueles que defendem um sistema em que se consideram tão-somente informações de cunho linguístico – e possivelmente apenas de natureza sintática; e, por outro, os que assumem um sistema alimentado por qualquer tipo de informação que lhe esteja disponível (FERREIRA & ÇOKAL, 2016). Os modelos defendidos pelos primeiros são chamados modulares, enquanto que o segundo grupo é responsável pela formulação de modelos ditos interativos.

Para Maia (2015, p. 15), “[o] conceito de modularidade da mente se refere à ideia de que existem subcomponentes relativamente independentes e de função específica no processamento da linguagem”. De acordo com modelos de processamento linguístico modulares, o *parser* constituiria, então, um dos módulos mentais: o módulo sintático, capaz de utilizar informações de natureza estritamente sintática no momento inicial do *parsing*. Dados relativos à frequência ou semântica de certos componentes linguísticos são, portanto, desconsiderados até que o *parser* tenha acesso à estrutura sintática completa.

Segundo a TGP, o *parser* “usa o seu conhecimento gramatical (estrutura sintática) isoladamente do conhecimento de mundo e de outras informações de caráter semântico para a identificação inicial das relações entre as palavras e os sintagmas” (LEITÃO, 2008, p. 225). Essa teoria, então, com sua proposta de encapsulamento sintático, parte de uma perspectiva modular de processamento da linguagem. Por negar a possibilidade do uso de pistas semânticas ou de frequência nos momentos iniciais do processamento, a TGP não consegue explicar o fenômeno da animacidade enquanto fator relevante no processamento de orações relativas. Isso nos leva a

---

<sup>21</sup> “This difficulty occurs despite the fact that the RC verb in ORCs should be more expected, not less expected, than in SRCs: there are fewer syntactic events that can occur after *The reporter that the senator ...* than after *The reporter that ...*, and there are fewer things that a senator can do to a reporter than that a reporter can do.”



considerar as contribuições que os modelos denominados interativos podem oferecer às pesquisas de processamento de orações relativas.

Ao contrário dos modelos modulares, os modelos interativos defendem que o *parser* tem acesso a informações que vão além das meramente morfossintáticas desde o início do processamento de um determinado *input* linguístico, sem que se privilegiem alguns tipos de informação em detrimento de outros (KENEDY, 2015). Tais modelos estão em consonância com os resultados de pesquisas que apontam a influência da animacidade dos referentes no processamento de orações relativas.

Embora, tradicionalmente, psicolinguistas tenham associado, de um lado, serialidade a modularidade e, de outro, paralelismo a não modularidade, essas vinculações não aparentam ter base empírica. De fato, a chamada Teoria da Dependência Local (*Dependency Locality Theory*, ou DLT), proposta por Gibson (2000), une interatividade a serialidade. Segundo Kenedy (ibid., p. 153), a DLT

identifica-se como um modelo interativo porque, tal como a TGP, assume a serialidade e incrementalidade na construção de representações linguísticas, mas diferentemente da TGP, sustenta que certos tipos de informações não estruturais podem ser vistos pelo *parser*.

Essa teoria é baseada em duas premissas (Gibson, 2000):

- 1) A compreensão da linguagem demanda recursos cognitivos para o *armazenamento* da estrutura sintática construída até o momento (como se pretendia em teorias anteriores) e para a *integração* a essa estrutura da palavra ouvida/lida.
- 2) A complexidade da integração estrutural está ligada à distância entre os dois elementos que são integrados.

O autor considera, ainda, que muitos efeitos de complexidade no uso de recursos cognitivos podem ser explicados pelo custo de integração, sem a necessidade de se utilizar outros fatores explicativos. Ele presume a construção de projeções máximas correspondentes às entradas lexicais para uma palavra processada em dado momento, sendo a velocidade de acesso a uma entrada lexical dependente de sua frequência de uso.

Por sua inclusão de dados relativos à semântica e à frequência de determinados itens lexicais no processamento de sentenças, uma teoria interativa como a DLT poderia explicar os resultados encontrados por estudos psicolinguísticos em que falantes parecem sensíveis ao traço de animacidade durante os primeiros estágios na compreensão de orações relativas.

Considerando-se a importância desses estudos na compreensão de quais informações estão acessíveis ao indivíduo durante o processamento linguístico, podemos nos perguntar até que ponto tais resultados seriam aplicáveis ao processamento de L2. Essa é uma questão relevante, primeiramente, porque ainda não está claro se os mecanismos utilizados no processamento de L1 são os mesmos que um bilíngue teria à disposição na compreensão de sua L2. Em segundo lugar, o bilinguismo não é uma situação anômala ou excepcional; é, pelo contrário, a condição majoritária da humanidade (FINGER, 2015). Se nosso objetivo é entender os mecanismos que subjazem o processamento linguístico, incluir o bilíngue como objeto de estudo é essencial.

De modo, portanto, a contribuir com as investigações sobre processamento de L2, elaboramos o presente estudo. Utilizamos orações relativas como estímulos linguísticos, porque, devido à sua complexidade estrutural, encontram amplo uso na literatura psicolinguística. Assim, nossos resultados poderão ser comparados aos colhidos em investigações com participantes monolíngues.

### **3.3 Frequência e processamento linguístico**

Uma aparente influência da frequência no processamento e no desenvolvimento de estruturas linguísticas é perceptível em diversos fenômenos da linguagem: elementos linguísticos que repetidamente ocorrem em conjunto tendem à redução fonética e à coalescência; expressões frequentes são mais resistentes ao mapeamento estrutural e ao nivelamento analógico; a produtividade de regras gramaticais seria determinada pelo efeito conjunto da frequência e da semelhança; durante o desenvolvimento linguístico da criança, a segmentação da fala e a aquisição de categorias sintáticas seriam guiadas pela frequência de determinadas estruturas; a escolha do falante entre construções alternativas durante o processo de produção linguística estaria relacionada à sua experiência com formas e significados particulares; e as ambiguidades estruturais no processamento linguístico parecem,

muitas vezes, ser resolvidas por meio de informação referente à frequência de estruturas e itens lexicais (DIESSEL & HILPERT, 2016).

Linguistas que estudam a linguagem centrando-a no uso defendem que a frequência é “um dos principais determinantes para a emergência de estruturas linguísticas e da organização do nosso conhecimento gramatical”<sup>22</sup>, sendo a gramática um sistema dinâmico constituído de categorias fluidas e restrições variáveis determinadas pela frequência. Sob essa perspectiva, *construções* – “sequências convencionalizadas de elementos linguísticos que combinam uma forma particular com um significado ou função específica”<sup>23</sup> – são associadas a expressões linguísticas mais frequentes em determinadas posições estruturais. (ibid., pp. 2, 6-7).

No que diz respeito ao processamento linguístico, um dos primeiros estudos a relacionar esse fenômeno à frequência foi o de Bever (1970), que, usando como base os resultados de uma série de experimentos psicolinguísticos, argumenta que uma sequência substantivo-verbo-substantivo é geralmente interpretada como indicando um evento do tipo agente-ação-objeto devido a uma possível preponderância estatística de estruturas desse tipo. Índícios de aprendizado estatístico também foram encontrados nos estudos de Green (1979), sobre o uso de palavras funcionais no delineamento de sintagmas; de Morgan, Meier e Newport (1987), sobre a integração de pistas prosódicas e morfológicas na aquisição de língua; de Brooks et al. (1993), sobre a aquisição de morfologia relacionada a gênero em uma língua artificial; e de Misyak e Christiansen (2007), cuja pesquisa sugere que a habilidade de aprendizado estatístico era um fator mais relevante na compreensão de orações relativas do que os resultados de um teste de *reading span*.

Para compreendermos de que forma a frequência poderia interferir no processamento linguístico, seria útil nos atermos, primeiro, à relação entre regularidade e frequência no reconhecimento de palavras. De acordo com Wells et al. (2009), esse fenômeno se referiria à ambiguidade característica à ortografia inglesa, em que uma sequência de letras como *int*, por exemplo, pode ser pronunciada como /ɪnt/ (como em *mint*) ou /aɪnt/ (como em *pint*). A acurácia e a velocidade de leitura de palavras com uma mesma sequência de letras dependeriam das *vizinhas* que elas possuem, i.e., de quantas outras palavras partilham com elas a relação entre ortografia

---

<sup>22</sup> “In the usage-based approach, however, frequency is one of the main determinants for the emergence of linguistic structure and the organization of our grammatical knowledge.” (Tradução nossa)

<sup>23</sup> “Most usage-based linguists conceive of syntactic constituents as constructions—conventionalized sequences of linguistic elements combining a particular form with a particular meaning or function.” (Tradução nossa)

e pronúncia. Uma palavra regular como *mint* – com uma série de vizinhas, tais quais *hint*, *lint* e *dint* – seria lida mais rapidamente e de maneira mais precisa do que um vocábulo irregular como *pint*. Ademais, esse efeito de regularidade seria mais importante para o processamento das palavras menos frequentes, uma vez que estas, se forem regulares, podem se beneficiar da experiência que o leitor tem com suas vizinhas. Por outro lado, o processamento de uma palavra irregular – e, portanto, com poucas vizinhas – depende muito mais da experiência que o leitor tem com a própria palavra. Desse modo, palavras irregulares como *have* são processadas tão facilmente quanto vocábulos regulares de frequência semelhante, mas palavras irregulares raras são lidas com mais dificuldade que as regulares de mesma frequência.

É possível que a resolução de ambiguidade no processamento de sentenças também sofra influência da interação entre regularidade e a frequência de determinadas estruturas. Sob essa perspectiva, certos tipos de sentença seriam mais regulares que outros, havendo uma relação mais clara entre a forma (a ordem das palavras) e o significado de sentenças mais regulares, enquanto que as menos regulares apresentariam uma associação sintaxe-significado idiossincrática. Desse modo, o processamento de orações menos regulares seria mais dependente da experiência direta que o ouvinte ou leitor tem com elas (Pearlmutter & MacDonald, 1995).

Segundo MacDonald e Christiansen (2002), relativas de sujeito seriam relativamente regulares por apresentarem uma ordem de palavras semelhante à de orações simples na voz ativa, bastante comuns na língua inglesa. A farta experiência que um ouvinte ou leitor tem com essas orações simples o auxiliaria no processamento de ORS, uma vez que a compreensão de estruturas linguísticas similares ocorreria por meio de processos semelhantes. Orações relativas de objeto, por outro lado, seriam caracterizadas por sintaxe mais irregular: nelas, o objeto direto precede o verbo. A experiência que o ouvinte ou leitor tem com orações simples seria, portanto, de pouca valia no processamento de ORO. Para aprimorar seu desempenho nessa tarefa, seria necessário familiarizar-se com as próprias relativas de objeto, o que dependeria da frequência com que o ouvinte/leitor se depara com elas.

Embora os estudos psicolinguísticos sobre processamento de relativas geralmente utilizem estímulos com antecedentes animados, a maioria das relativas de objeto na verdade modifica antecedentes inanimados em línguas como o português (CABRAL, 2016), alemão (ZUBIN, 1979), holandês (MAK, VONK & SCHRIEFERS, 2002) e

inglês (cf. p. 66). Levando-se em consideração esse fato e a possibilidade de haver influência da frequência no processamento de estruturas linguísticas, poderíamos explicar por que ORO com antecedentes animados causam dificuldade de processamento. O maior esforço despendido na compreensão de tais estruturas não se daria, necessariamente, por questões de ordem sintática. É possível que sua menor frequência também explique esse fenômeno. De fato, quando se utilizam como estímulos experimentais relativas de objeto com antecedentes inanimados, a assimetria no tempo de processamento de ORS e ORO some ou é grandemente reduzida (MAK, VONK & SCHRIEFERS, 2002; TRAXLER, MORRIS & SEELY, 2002; WELLS et al., 2009; CABRAL, 2016). Os resultados desses estudos sugerem que o processamento de relativas de objeto é influenciado pela experiência que o leitor ou ouvinte tem com essas estruturas.

Dentre os trabalhos supracitados, o de Wells et al. (2009) é particularmente relevante a esta discussão. Os autores partiram da hipótese de que a experiência seria um fator importante no processamento linguístico, mas que a compreensão de relativas de objeto – por apresentarem uma ordem de palavras idiossincrática na língua inglesa – seria mais beneficiada por um acréscimo na experiência direta que os participantes têm com essas estruturas do que o processamento de relativas de sujeito – caracterizadas por um grande número de “vizinhas” sintáticas.

Por meio de um estudo de grande escala, os pesquisadores manipularam a experiência de leitura de 97 adultos, falantes nativos de inglês, durante semanas (quatro sessões com um intervalo de quatro a oito dias entre cada uma delas), utilizando experimentos de *reading span* e de leitura automonitorada. Os participantes foram divididos em dois grupos: metade constituiu o grupo sujeito a experiências de leitura de orações relativas, divididas igualmente entre as de objeto e as de sujeito; e a outra metade formou o grupo controle. O objetivo da pesquisa era descobrir se os tempos de leitura dos participantes do primeiro grupo diminuiriam entre o pré-teste e o pós-teste.

Os resultados obtidos pelos autores demonstraram um efeito claro da experiência no processamento de orações relativas. Todavia, esse efeito foi assimétrico, sendo mais perceptível no que se refere à compreensão de ORO que de ORS. Os pesquisadores especulam que, da primeira à última sessão, os participantes expostos a orações relativas familiarizaram-se com a estrutura incomum inerente às relativas de objeto, mas rara em outras estruturas da língua inglesa.

É importante ressaltar, no entanto, que o uso de informações relativas a frequência na compreensão de um enunciado não significa que dados de natureza sintática seriam dispensados desse processo. Como afirmamos anteriormente, um dos modelos vigentes de processamento linguístico, a DLT (Gibson, 2000), parte do princípio de que o ouvinte/leitor armazena e integra estruturas sintáticas enquanto tenta prever de que modo o enunciado que ele ouve/lê poderia prosseguir. O armazenamento e integração das estruturas se dariam por meio de informações de natureza sintática, mas a previsão seria possível devido à experiência que o sujeito tem com enunciados semelhantes.

Antes de nos perguntarmos até que ponto resultados como os vistos acima podem ser estendidos a bilíngues, precisamos definir o que entendemos exatamente por bilinguismo, que diferenças poderia haver entre falantes nativos e não nativos, e quais modelos de processamento linguístico foram propostos para explicar tais diferenças. Essas questões são abordadas na seção a seguir.

### 3.4 Processamento de L2

Bilíngues não são um grupo homogêneo e definir o bilinguismo não é tarefa simples. Falantes de mais de um idioma apresentam níveis de proficiência variados nas línguas que utilizam, adquirem-nas em etapas diferentes de suas vidas, por razões diversas e em contextos distintos (FINGER, 2015). Pesquisadores diferentes definem bilinguismo de múltiplas maneiras em seus estudos. De modo a estabelecer um esquema resumido do que se pretende tratar quando se fala em bilinguismo, Bhatia afirma que:

[É] natural que não exista nenhuma definição ou medida amplamente aceita do bilinguismo. Em vez disso, uma rica variedade de escalas, dicotomias e categorias é empregada na descrição dos bilíngues. Um bilíngue que fala e compreende duas línguas é chamado bilíngue *produtivo*, enquanto um bilíngue *receptivo* é o indivíduo que entende mas não fala uma segunda língua. Uma criança que tenha adquirido duas línguas em seu lar antes dos cinco anos de idade (*ambiente natural*) é denominada bilíngue *simultânea* ou *precoc*, enquanto aqueles que aprenderam uma segunda língua depois dos cinco anos, seja em casa ou no

ambiente escolar, são descritos como bilíngues *tardios* ou *sequenciais* (Bhatia, 2011, p. 126, tradução nossa).<sup>24</sup>

Para os propósitos desta pesquisa, interessa-nos investigar o processamento de sentenças por bilíngues tardios, uma vez que as circunstâncias de aquisição de sua segunda língua diferem significativamente daquelas que caracterizam a aquisição de uma língua materna. Nossos participantes têm o português brasileiro como L1 e o inglês como L2 (em níveis intermediário e avançado).

Além das discussões sobre o exato significado de *bilinguismo*, também o conceito de *língua* apresenta complicações. A maioria das pessoas provavelmente concordaria que português e japonês são idiomas diferentes, mas as distinções se tornam menos claras quando tratamos de variedades muito próximas, como o português e o galego. Esse problema, embora relevante em estudos de bilinguismo enquanto fenômeno mais amplo, não se apresenta como pertinente para a pesquisa aqui descrita: as línguas faladas pelos sujeitos de nossa pesquisa – a portuguesa e a inglesa – pertencem a ramos distintos do tronco indo-europeu (KRESSLER, 2011). Isso as torna, portanto, suficientemente distantes uma da outra para que denominemos bilinguismo – e não bidialetalismo – o fenômeno aqui tratado. O ponto que se precisa frisar, todavia, é que consideramos o inglês falado por nossos participantes como sua L2 por ter sido aprendido após a L1.

Tradicionalmente, pesquisadores na área do processamento linguístico tratavam bilíngues como um grupo especial de falantes, comparáveis a indivíduos afásicos ou com déficit linguístico, e em claro contraste com os monolíngues, que constituíam os falantes ideais. Em outras palavras, o bilinguismo era visto como uma situação excepcional, pouco representativa dos fenômenos da linguagem que os estudiosos tinham interesse em investigar. Ou, paralelamente, pensava-se que o bilíngue equivalesse a dois monolíngues em um único indivíduo, uma acepção do bilinguismo denominada *fracional*. Seu estudo, desse modo, seria dispensável, uma vez que o falante monolíngue ofereceria todas as respostas necessárias às questões que se pudessem levantar sobre o processamento da linguagem, seja por monolíngues ou bilíngues (BHATIA, 2011).

---

<sup>24</sup> “[I]t is natural that no widely accepted definition or measure of bilingualism exists. Instead, a rich range of scales, dichotomies, and categories are employed to describe bilinguals. A bilingual who can speak and understand two languages is called a *productive* bilingual, whereas a receptive bilingual is an individual who can understand but cannot speak a second language. A child who has acquired two languages before the age of five at home (*natural setting*) is called a *simultaneous* or *early* bilingual, whereas those who learn a second language after the age of five, either at home or in school setting, are described as *late* or *sequential* bilinguals.”

Bilíngues não são excepcionais: é o monolinguismo que constitui a situação minoritária no mundo atual (FINGER, 2015). Tampouco parece ser verdade que o indivíduo bilíngue guarde na mente duas representações completamente distintas de suas línguas, sem sobreposição alguma entre elas. Grosjean (1989, p. 6), defendendo uma concepção *holística* do bilinguismo, afirma que o bilíngue é um todo coeso que não pode ser dividido em duas partes, não sendo, portanto, a soma de dois monolíngues (in)completos; segundo o autor, “[a] coexistência e constante interação entre as duas línguas produziu no bilíngue uma entidade linguística diferente, mas completa” (tradução nossa).<sup>25</sup>

O desempenho linguístico do bilíngue é caracterizado pelo domínio (iniciante, intermediário ou avançado) que possui de cada uma de suas línguas, pelo modo como estas se influenciam mutuamente (COOK, 2003), bem como pelas representações mentais que o bilíngue constrói das duas línguas, que as fazem diferir da maneira como estão representadas nas mentes dos monolíngues que as dominam (SELINKER, 1972; COOK, 2014). Essas especificidades tornam o bilinguismo uma área de pesquisa rica e pertinente, que tem atraído diversos psicolinguistas interessados na compreensão e descrição não apenas dos processos que caracterizam a aquisição e o processamento de L2, como também das “interações entre os sistemas cognitivos e os mecanismos neurocognitivos subjacentes à aquisição e ao processamento da linguagem” de modo geral (FINGER, 2015, p. 158).

### 3.5 Frequência e processamento preditivo em L2

Um número crescente de pesquisas na área da psicolinguística tem dado suporte à ideia de que falantes nativos são capazes de prever a informação que está por vir num enunciado, em detalhes e em vários níveis de representação linguística, tais como gênero gramatical, categoria sintática, forma fonológica e visual de uma palavra, ou a posição de *gap* de um objeto numa dependência do tipo *wh*-. Essas previsões são baseadas no conhecimento de mundo do falante, inclusive aquele relacionado à frequência de uso de determinados itens linguísticos; informação sintática como caso,

---

<sup>25</sup> “The coexistence and constant interaction of the two languages in the bilingual has produced a different but complete linguistic entity.”



gênero ou estrutura sintática precedente; propriedades do verbo; ou prosódia (FINGER, 2015).

O modo como tais informações poderiam ser aproveitadas no processamento de sentenças é algo que pode ser verificado na literatura relevante. Utilizando-se da técnica de rastreamento ocular para investigar o processamento preditivo em L2, Hopp (2013) descobriu que os participantes de sua pesquisa, falantes nativos de inglês que tinham o alemão como L2 em nível avançado ou quase nativo, eram capazes de se guiar pela concordância de gênero entre determinante e substantivo para realizar previsões durante o processamento linguístico. Esse resultado se torna ainda mais significativo quando nos lembramos de que a L1 dos sujeitos, o inglês, não possui gênero gramatical. Em estudo semelhante, Dussias et al. (2013) relataram que bilíngues inglês L1/espanhol L2 faziam uso do gênero gramatical de um artigo em espanhol para antecipar o substantivo a ser ouvido, enquanto bilíngues italiano L1/espanhol L2 eram capazes do mesmo, mas apenas quando se deparavam com um artigo feminino. Os autores sugerem que esses resultados demonstram a importância tanto da proficiência em L2 dos falantes quanto da semelhança tipológica entre a L1 e a L2.

De acordo com Kaan (2014), a partir do conhecimento adquirido durante toda a vida,

falantes nativos de uma língua aprenderam a associar certas palavras, categorias gramaticais, regras e estruturas sintáticas, formando uma base de dados sobre frequências relativas e associações. Essas frequências (de co-ocorrência) não são estáticas, mas mudam com o tempo e por meio da interação com outros usuários da língua. Enquanto escutam ou leem, falantes constroem uma representação sintática, semântica e fonológica/ortográfica do *input*, atribuindo diferentes probabilidades às várias maneiras como o enunciado corrente pode prosseguir. Se uma continuação em particular é ativamente antecipada (em outras palavras, se ela é muito provável), mas é desmentida pelo *input* ou ausente dele, gera-se dificuldade de processamento.<sup>26</sup> (Kaan, 2014, pp. 260-261, tradução nossa).

---

<sup>26</sup> “Based on a lifetime of exposure to their native language, native speakers of a language have learned to associate certain words, word categories, syntactic frames and rules, building a database of relative frequencies of occurrences and associations. These (co-occurrence) frequencies are not static but instead change over time when interacting with other users of the language. While listening or reading, language users construct a syntactic, semantic, and phonological/orthographical representation of the input, assigning likelihoods to the various ways in which the current utterance can be continued. If a particular continuation is strongly predicted (i.e., is very likely) but is contradicted by the input or not found in it, processing difficulty ensues.”

Os fatores que a autora aponta como sendo relevantes para o processamento preditivo são: a informação relativa à frequência que se tem armazenada na mente; informações em competição; a precisão e a consistência da informação utilizada; processos e estratégias sob influência do tipo de tarefa experimental; e outros fatores como motivação, recursos e controle cognitivos. Todos esses fatores estariam inter-relacionados, o que torna improvável que apenas um deles seja responsável pelas diferenças entre falantes nativos e não nativos. Aqui nos deteremos sobre o primeiro e o terceiro tipos de fatores, que consideramos de fundamental importância para compreender as diferenças no processamento linguístico desempenhado pelos dois grupos de falantes.

As informações relativas à frequência de diferentes estruturas linguísticas que um indivíduo armazena em sua mente diferem de acordo com as experiências linguísticas pelas quais ele passou. Uma segunda língua costuma ser aprendida e utilizada em contextos particulares, sofrendo muitas vezes a influência de uma linguagem mais formal e didática, bem como do uso comunicativo entre falantes não nativos. Esse tipo de linguagem pode adquirir características originadas da L1 dos falantes, ou mesmo de uma interlíngua (ibid.; SELINKER, 1972). Apesar disso, estudos experimentais têm demonstrado que falantes não nativos são tão capazes de utilizar a predição no processamento linguístico quanto nativos, diferindo destes apenas no tipo de informação sobre frequência que possuem (DUSSIAS & CRAMER SCALTZ, 2008; LEE, LU & GARNSEY, 2013). Outras pesquisas têm inclusive sugerido que falantes não nativos seriam capazes de adquirir informações sobre frequência similares às utilizadas pelos nativos, mesmo que sua L2 não tenha sido adquirida num contexto de imersão (DUSSIAS et al., 2010; MARTINEZ-GARCIA & WULFF, 2012).

A precisão e a consistência das informações linguísticas armazenadas na mente do falante diferem de acordo com a exposição deste à língua em questão. Bilíngues geralmente são menos expostos à L2 do que os falantes nativos desta, o que leva a uma representação mental mais fraca de palavras de baixa frequência. Além disso, itens lexicais são mais difíceis de distinguir em um léxico mental bilíngue, uma vez que ambas as línguas – a L1 e a L2 – são ativadas durante o processamento linguístico (KAAN, 2014). Todavia, com um aumento no nível de proficiência, um falante não nativo pode desenvolver representações lexicais de maior qualidade, podendo

utilizá-las para formar previsões sobre um *input* linguístico (DUSSIAS et al., 2013; HOPP, 2013).

### 3.6 Modelos de processamento de L2

#### 3.6.1 Hipótese da Estrutura Rasa (*Shallow Structure Hypothesis*)

Um dos modelos teóricos mais referenciados na área de processamento de L2 é a chamada Hipótese da Estrutura Rasa (ou *Shallow Structure Hypothesis* – *SSH*, em inglês), de Clahsen e Felser (2006; 2017). Segundo essa hipótese, bilíngues seriam capazes de utilizar informações de cunho lexical, semântico, pragmático e probabilístico no processamento de frases em sua L2, assim como ocorre normalmente no processamento de L1. Por outro lado, ao contrário de falantes nativos, mesmo bilíngues com alta proficiência em sua L2 apresentariam dificuldades para gerar e manipular representações sintáticas abstratas em tempo real. Os autores frisam, no entanto, que isso não significa que bilíngues não conseguiriam se tornar usuários eficientes de sua L2, mas que os mecanismos que esses falantes utilizam *preferencialmente* no processamento linguístico diferem daqueles normalmente empregados na compreensão de uma L1.

Observe-se que o que se quer dizer quando se fala em uso preferencial é que tanto nativos quanto não nativos tiram proveito de informações de cunho estrutural e não estrutural, apenas com a ressalva de que as primeiras seriam mais facilmente utilizadas por nativos, enquanto que não nativos privilegiariam as últimas. Os autores também rejeitam uma interpretação da Hipótese da Estrutura Rasa como uma “hipótese do déficit”, ressaltando que não há razão para acreditar que um processamento que priorize a estrutura sintática e morfológica de um *input* seja superior àquele baseado em informações não gramaticais desse mesmo *input*.

De acordo com Kaan (2014), a Hipótese da Estrutura Rasa poderia incorporar um componente preditivo, sendo compatível com estudos que apontam a importância da previsão no processamento linguístico, uma vez que pressupõe que falantes de L2 utilizam informações relacionadas ao contexto e ao léxico. No entanto, para que isso seja possível, as dependências em questão devem ser adjacentes. Leal, Slabakova e

Farmer (2016) entendem que, sob a perspectiva da *SSH*, falantes não nativos não deveriam ter a capacidade de gerar expectativas sobre material linguístico vindouro a partir de relações sintáticas de longa distância, devido às representações sintáticas menos detalhadas da L2 na mente desses falantes. Se esses autores estiverem corretos, a *SSH* não poderia explicar resultados que demonstrassem o uso da predição em orações relativas de objeto, por exemplo.

### **3.6.2 Habilidade Reduzida de Geração de Expectativas (*Reduced Ability to Generate Expectations – RAGE*)**

Num estudo com aprendizes de L2 de alta proficiência, Grüter, Lew-Williams e Fernald (2012) observaram que, em comparação com falantes nativos, os primeiros faziam uso menos eficiente de pistas de gênero gramatical durante o processamento de substantivos. Os autores propõem que essas diferenças de performance são originadas no nível da representação lexical, sendo fruto de divergências fundamentais no modo como crianças e adultos adquirem o léxico de uma língua.

Outro exemplo nesse sentido é o estudo de Martin et al. (2013), em que falantes monolíngues de inglês e bilíngues espanhol (L1)/inglês (L2) liam sentenças terminadas em substantivos esperados ou inesperados, enquanto seus potenciais relacionados a eventos (*event-related potentials*, ou *ERP*) eram aferidos. De acordo com os dados obtidos, os pesquisadores concluíram que falantes não nativos não preveem palavras vindouras durante a leitura de uma sentença com a mesma eficiência que nativos. Os autores atribuem essa diferença a estágios de processamento linguístico em L2 comparativamente mais lentos e menos precisos.

Buscando explicar resultados desse tipo, Grüter e Rohde (2013) propuseram uma hipótese denominada *RAGE*, ou *Reduced Ability to Generate Expectations* (Habilidade Reduzida de Geração de Expectativas). Como explicitado em seu nome, essa hipótese sugere que aprendizes de L2 seriam menos capazes, em comparação com falantes nativos, de gerar previsões sobre como um enunciado lido ou ouvido poderia prosseguir. Em artigo posterior, Grüter, Rohde e Schafer (2014, sem paginação, tradução nossa) argumentam que

[a] ideia que sustenta a hipótese *RAGE* é a de que um processador de capacidade limitada (L2) é levado ao limite por processos imediatamente requeridos para lidar com informações incrementais, como acesso lexical e integração estrutural. Em outras palavras, o processador se dedica completamente ao processamento reativo, com nenhum ou poucos recursos disponíveis para a interpretação de pistas não essenciais na atualização de expectativas, ou limitando drasticamente o escopo de previsões para alternativas de ocorrência imediata, como a próxima palavra em uma sentença.<sup>27</sup>

É importante não interpretar a *RAGE* como uma hipótese da “deficiência” do falante de L2 frente ao falante nativo. Grüter, Lau e Ling (2019) acreditam que o processamento linguístico típico dos nativos não deve ser tomado como objetivo ou ideal sobre o qual se pautaria o sucesso no processamento de uma L2. Em vez disso, as autoras reconhecem a possibilidade de se atingir máxima eficiência no processamento a partir de “estratégias diferentes, em diferentes contextos e em diferentes circunstâncias” (tradução nossa)<sup>28</sup>, e que se falantes não nativos fizessem uso dos mesmos meios que os monolíngues, essa eficiência poderia ser prejudicada.

### 3.6.3 Perspectivas alternativas

Está claro que a *SSH* e a *RAGE* não podem condizer com a realidade se considerarmos o bilíngue capaz não apenas de fazer previsões sobre a continuidade de um enunciado, mas também de fazê-lo em sentenças em que as dependências não são adjacentes uma à outra, como no caso de ORO.

Com base nos resultados encontrados em estudos anteriores (cf. TRAXLER, 2014; PIÑAR et al, 2017), pressupomos que os participantes bilíngues desta pesquisa terão a habilidade de utilizar pistas de cunho semântico e probabilístico no processamento de orações relativas. Dito de outro modo, espera-se que esses sujeitos sejam sensíveis ao traço de animacidade dos referentes das orações relativas, o que levará a uma atuação perceptível da frequência de cada variável (ORSa, ORSi, OROA,

---

<sup>27</sup> “The intuition underlying the RAGE hypothesis is that a limited capacity (L2) processor is stretched to its limits by processes that are immediately required for dealing with incrementally incoming information, such as lexical access and structural integration. In other words, it is fully occupied by reactive processing, with little or no resources left for taking up non-essential cues to update expectations, or severely limiting the scope of expectations to immediately upcoming choices, such as the next word in a sentence.” (Tradução nossa)

<sup>28</sup> “[W]e come to recognise that maximal processing efficiency may be achieved through different means in different contexts and under different circumstances.” (Tradução nossa)

OROi) no processamento de tais estruturas. Nossa hipótese não está, portanto, em consonância com a *RAGE* nem com a Hipótese da Estrutura Rasa, a menos que esta inclua a possibilidade de um componente preditivo no processamento de L2.

Em contrapartida aos modelos teóricos supracitados, alguns autores defendem que os mecanismos utilizados no processamento de L2 são semelhantes aos empregados na compreensão de uma L1, sendo as diferenças entre os dois casos atribuídas a fatores independentes. As razões mencionadas por Cherepovskaia, Reutova e Slioussar (2021) para explicar essas diferenças são: o maior custo cognitivo do processamento de L2 em relação ao de L1, a interferência da L1 no compreensão da L2 e o nível de proficiência na L2.

Uma possível explicação para *o maior custo cognitivo no processamento de L2* seria a redução da automaticidade e da velocidade desse processo. Por meio de um teste de leitura automonitorada, Kaan, Ballantyne e Wijnen (2015) investigaram o processamento da concordância entre sujeito e verbo em ORO por falantes nativos e não nativos de inglês. Os autores descobriram que diferenças entre os dois grupos podiam ser explicadas basicamente por diferenças de tempo de leitura: quando esse dado era controlado, o grupo não nativo demonstrou padrões de leitura semelhantes aos dos nativos. Observe-se que, ainda assim, as respostas às perguntas de compreensão feitas após a leitura das sentenças demonstraram uma diferença importante entre os dois grupos de participantes: nesse quesito, apenas os leitores mais rápidos entre os não nativos apresentaram padrões similares aos dos nativos.

Uma outra razão apontada para o maior custo cognitivo do processamento de L2 seria o acesso lexical mais limitado nessa condição. Segundo Gollan et al. (2005), a conexão entre a representação fonológica e a representação semântica nos idiomas de um bilíngue seria mais tênue que no idioma falado por um monolíngue, uma vez que um bilíngue desenvolve duas representações fonológicas para cada conceito semântico, tendo menos oportunidades de praticar e sedimentar a conexão entre essas duas representações do que um monolíngue. Também é possível que haja interferência entre as duas representações fonológicas, da L1 e da L2, durante o acesso lexical.

Ainda em relação à diferença de custo de processamento, podemos citar a integração sintática dos componentes de um enunciado. Hopp (2013) dividiu 75

falantes não nativos de inglês em três grupos de igual tamanho, de acordo com os resultados de integração sintática que esses participantes obtiveram num experimento anterior envolvendo monitoramento de palavras (*word monitoring task*): integração baixa, intermediária ou avançada. Foram computados, então, os tempos de leitura de cada grupo. De modo geral, todos os bilíngues demonstraram precisão na compreensão dos estímulos linguísticos, mas os três grupos exibiram padrões de comportamento distintos na reanálise das ambiguidades de sujeito/objeto a que foram expostos. Nas comparações de transitividade e de caso, os bilíngues com alta capacidade de integração sintática assemelharam-se aos participantes nativos que também haviam realizado o experimento.

Quanto à *interferência da L1 na compreensão da L2*, Foucart e Frenck-Mestre (2010) compararam a performance de alemães que tinham o francês como L2 e a de falantes nativos de francês, utilizando-se, para tanto, de três experimentos com coleta de dados eletrofisiológicos. Todos os experimentos envolviam violações de concordância de gênero, seja entre determinante e substantivo (Experimento 1), entre o adjetivo posposto e o substantivo (Experimento 2) ou entre o adjetivo preposto e o substantivo (Experimento 3). Os resultados obtidos indicaram o mesmo efeito P600 para os dois grupos de participantes quando a violação era similar em francês e em alemão, mas nenhum efeito foi detectado entre os não nativos nos casos em que havia variação nas regras de concordância da L1 e da L2, o que sugeriria interferência da primeira sobre o processamento da segunda.

Por fim, o *nível de proficiência na L2* também é um fator relevante no que se refere ao processamento linguístico realizado por falantes não nativos. Hopp (2006) elaborou dois experimentos – de leitura automonitorada e de julgamento de aceitabilidade – com dois grupos: falantes de inglês ou holandês como L1 e alemão como L2, bem como falantes de alemão como L1. Ambos os experimentos envolviam ambiguidades de sujeito/objeto em língua alemã. O autor destaca que os sujeitos bilíngues não apresentaram diferenças de acordo com sua L1, e sim com seu nível de proficiência na L2. Muito embora todos eles tivessem longa experiência com o idioma alemão – sendo todos de nível avançado ou quase nativo – os falantes de nível avançado demonstraram menor precisão no processamento da L2. Por sua vez, aqueles com proficiência quase nativa geraram resultados que convergem com os de

nativos na acurácia do processamento e nos padrões de reanálise dos estímulos linguísticos.

Nesta seção, mencionamos a influência que o conhecimento da L1 pode exercer sobre a compreensão da L2. É importante, portanto, que tenhamos ciência do modo como tal interferência poderia confundir os resultados obtidos no nosso experimento: seria possível que as informações de frequência utilizadas pelos nossos bilíngues no processamento de orações relativas viriam não da experiência que têm com a L2, mas com a sua L1? Abordamos esse tema na seção que segue.

### 3.7 Hipótese da Transferência Total/Acesso Total

Dentre as diversas hipóteses que buscam explicar de que modo as representações mentais de falantes bilíngues diferem daquelas desenvolvidas por nativos, podemos citar também a Hipótese da Transferência Total/Acesso Total (*Full Transfer/Full Access Hypothesis*, ou *FT/FA*). De acordo com esse modelo, a totalidade da gramática da L1 – com exceção das matrizes fonéticas de itens lexicais e morfológicos – corresponde ao estado inicial da L2 a ser adquirida. Provém daí o termo “Transferência Total”. No entanto, esse estado inicial da L2 sofre mudanças à medida que o aprendiz se expõe a *input* na língua-alvo que contradiga essa gramática provisória, o que força uma reestruturação desse sistema com base nas opções possíveis da Gramática Universal (GU), à qual o aprendiz tem “Acesso Total” (SCHWARTZ & SPROUSE, 1996).

Isso significa que, para o modelo *FT/FA*, os processos cognitivos que sustentam a aquisição de L1 e L2 são os mesmos, muito embora tanto os pontos de partida (o estágio inicial na estruturação da gramática da L2 corresponde ao último estágio da L1) quanto os pontos de chegada (o domínio das duas línguas pelo falante raramente é equivalente) desses processos sejam distintos. Em outras palavras, os mecanismos que induzem o desenvolvimento da gramática de uma L2 são os mesmos que possibilitam a aquisição de uma L1 (a GU e a capacidade de aprendizagem do falante) (*ibid.*).

Para os autores supracitados, aprendizes de uma L2 com línguas maternas tipologicamente distintas passam por etapas diferentes na aquisição da língua-alvo,



especialmente no que se refere ao desenvolvimento sintático. Desse modo, “a causa de quaisquer diferenças de evolução na aquisição de uma língua em particular que sejam covariáveis com a língua materna devem estar presentes desde o princípio (isto é, devem resultar do estado inicial da L2), precisamente porque a causa de tal covariação não pode ser o *input*, uma vez que este permanece constante.” (ibid., p. 67, tradução nossa).<sup>29</sup>

A hipótese da *FT/FA* é importante para a nossa discussão porque um dos tipos de conhecimento sobre a L1 que podem ser transferidos para a representação mental que o bilíngue tem de sua L2 é aquele que diz respeito à frequência de determinadas estruturas sintáticas. Isso significa que, se a frequência dos quatro tipos de orações relativas investigados neste estudo for semelhante tanto em português quanto em inglês, será difícil, senão impossível, determinar se os resultados encontrados para o processamento de relativas em inglês se devem aos conhecimentos que os participantes adquiriram sobre o uso dessas estruturas em sua L2, ou se apenas transferiram os dados sobre a ocorrência dos tipos de relativas em português para o inglês.

Todavia, Skalicky e Chen (2020) ressaltam que, à medida que aumenta a proficiência de um falante em sua L2, diminui nesta a influência exercida pela L1 (e, por outro lado, maior conhecimento e proficiência na L2 podem passar a interferir na performance em L1 do bilíngue). Sendo assim, se a razão para um processamento diferenciado dos quatro tipos de relativas em inglês estiver relacionada a uma transferência dos conhecimentos estatísticos do português, essa transferência deve ter impacto desigual nos resultados de bilíngues de nível intermediário e avançado. E se esses dois grupos se comportarem de modo distinto, é possível supor que os resultados dos bilíngues de proficiência superior sejam causados por mais experiência e um maior conhecimento sobre a distribuição relativa das estruturas sintáticas da língua inglesa.

---

<sup>29</sup> “In short, the cause of any (developmental) differences in the L2 acquisition of a particular language that covary with native language must be present from the beginning (namely, must be the result of the L2 initial state), precisely because the cause of such covariation cannot be in the input, since this remains constant.”

## 4. ESTUDO ATUAL

### 4.1 Estudo de *corpus*

Apresentamos, neste capítulo, o *corpus* de orações relativas em língua inglesa que produzimos, bem como uma análise dos resultados que obtivemos com ele. Essas orações foram classificadas de acordo com a animacidade de seus referentes e com o conectivo ou pronome relativo utilizado (*that*, *which* ou *who/whom*<sup>30</sup>). Desse modo, adicionamos nosso *corpus* à lista de *corpora* semelhantes, como os de Zubin (1979), em alemão; o de Mak, Vonk e Schriefers (2002), nas línguas holandesa e alemã; e, em particular, o *corpus* em português brasileiro e europeu produzido por Cabral (2016), com o qual comparamos o nosso, uma vez que as condições analisadas em seu estudo se aproximam das que investigamos.

Nosso *corpus* é composto de textos escritos retirados de dois tipos de fonte. O primeiro é o periódico *online* britânico *The Guardian* ([theguardian.com/international](http://theguardian.com/international)), do qual foram selecionados 20 artigos publicados entre 5 de janeiro de 2017 e 15 de julho de 2021, somando 70.732 palavras. Esses artigos, veiculados nas seções *News*, *Lifestyle*, *Culture*, *World* e *Opinion* lidavam com temas diversos, sendo representativos da variedade padrão do inglês britânico escrito.

Nossa segunda fonte é constituída das versões digitais dos livros didáticos *Diálogo* (NIHONMATSU, 2020), *New Alive High* (BRAGA, RACILAN & GOMES, 2020), *English and More!* (WEIGEL & RESCHKE, 2020) e *Take Action!* (RICHTER & LARRÉ, 2020). Essas obras didáticas de língua inglesa são disponibilizadas a alunos da rede pública de todo o Brasil por meio do Plano Nacional do Livro Didático, ciclo 2021, e foram escolhidas por representarem usos do inglês típicos da sala de aula, com os quais presume-se que brasileiros falantes de inglês como L2 tenham familiaridade. Os textos retirados desses livros são de gêneros diversos, sobre temas variados, com registros que vão da formalidade às variedades mais coloquiais do idioma, de autores nativos ou não, incluindo transcrições de diálogos falados. São, no total, 83 textos, contabilizando 29.936 palavras.

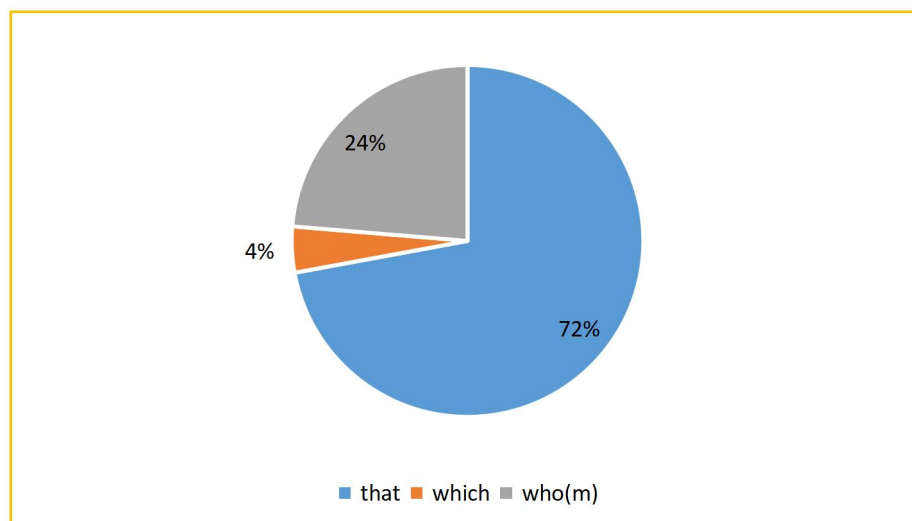
---

<sup>30</sup> Apenas orações relativas com o conectivo *that* e com os pronomes relativos *which* e *who(m)* foram consideradas.

Juntando todo esse material, constatamos a presença de 327 orações relativas restritivas em que se tem o uso do conectivo (*that*) ou de um dos pronomes relativos já mencionados (*which* ou *who/whom*). É importante frisar que nosso *corpus* não conta com orações relativas ditas reduzidas (*reduced relative clauses*), em que não se utiliza conectivo ou pronome relativo (ex.: *The shirt Ø I wore was comfortable.*). Embora sejam estruturas comuns na língua inglesa, elas são de detecção automática mais difícil na análise de um *corpus* digital; são estruturalmente distintas das orações relativas realizadas em português, dificultando a comparação entre os resultados do processamento nas duas línguas; e não foram utilizadas nas pesquisas de processamento de orações relativas em inglês que fazem parte da bibliografia deste estudo (TRAXLER, MORRIS & SEELY, 2012; TRAXLER et al., 2013; PILAR et al., 2017), com cujos resultados com falantes nativos e não nativos seria interessante comparar nossos achados.

Além disso, e mais uma vez em consonância com os materiais utilizados nas pesquisas supracitadas, bem como com as frases de que fizemos uso em nosso experimento, selecionamos apenas aquelas orações relativas em que a animacidade de sujeito e objeto diferia. Desse modo, das 327 orações relativas coletadas, consideramos para o nosso estudo apenas 190.

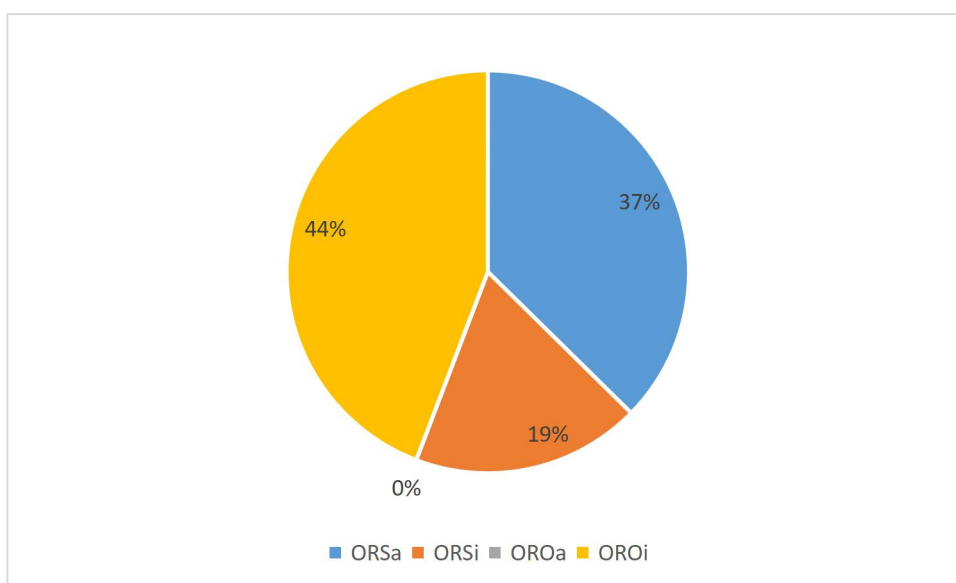
À parte, portanto, das orações relativas explicativas, das orações relativas reduzidas e daquelas em que sujeito e objeto não diferem em animacidade, obtivemos em nosso *corpus* a seguinte distribuição de orações relativas de acordo com o pronome relativo ou conectivo utilizado:



**Gráfico 2:** Orações relativas com divergência de animacidade entre sujeito e objeto, classificadas de acordo com o conectivo/pronome relativo utilizado.

Podemos observar a clara preponderância de relativas com o conectivo *that*, constituindo mais de 2/3 do total. Orações nessa condição são de nosso interesse não apenas por sua alta frequência, mas também porque *that* não denuncia a animacidade do referente da oração, podendo aludir tanto a entidades animadas quanto inanimadas. Nesse sentido, ele difere tanto de *who(m)* – que indica um referente obrigatoriamente humano ou antropomorfizado – quanto de *which* – que, por sua vez, aponta para entidades inanimadas. O conectivo *that*, no entanto, se assemelha bastante ao português *que*, igualmente ambíguo no que diz respeito ao fator da animacidade.

Se classificarmos essas orações de acordo com as quatro condições utilizadas em nosso experimento (oração relativa de sujeito animado, de sujeito inanimado, de objeto animado e de objeto inanimado), obteremos a seguinte distribuição:

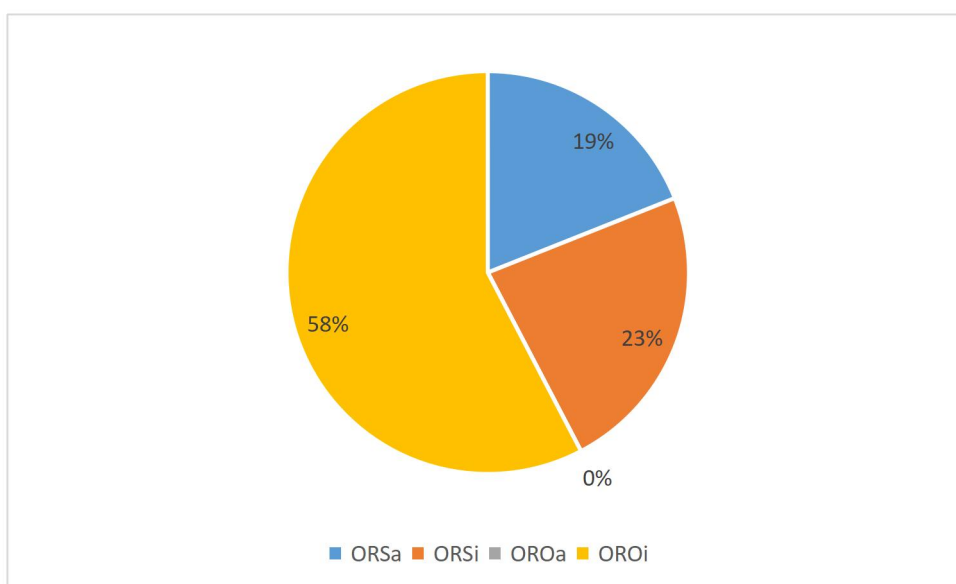


**Gráfico 3:** Orações relativas com divergência de animacidade entre sujeito e objeto, classificadas por condição.

O aspecto mais notável desses números é a completa ausência de orações relativas de objeto animado. Esse fato é importante porque, na literatura sobre processamento de relativas, as OROa têm sido tomadas como únicas representativas

das relativas de objeto de modo geral, sendo contrastadas com as ORSa, que, em nosso *corpus*, constituem mais de 1/3 do número total de relativas. Essa grande assimetria na frequência das duas condições poderia explicar a maior dificuldade de processamento das ORO frente às ORS nas pesquisas mais tradicionais em que a animacidade não foi controlada, não havendo necessariamente (apenas) um fator de origem sintática nessa discrepância. Também significativa é a alta ocorrência de OROi, conspicuamente ausentes desses mesmos estudos (cf. bibliografia citada na p. 16).

Por fim, podemos direcionar nossa atenção às orações relativas com o conectivo *that*, o mesmo tipo utilizado no nosso experimento de leitura automonitorada. Com esse recorte, observamos a distribuição abaixo:



**Gráfico 4:** Orações relativas com divergência de animacidade entre sujeito e objeto, com o uso do conectivo *that*, classificadas por condição.

Uma vez mais, evidentemente, percebemos a ausência de relativas na condição OROa. Devemos salientar, entretanto, que quando se consideram apenas as relativas com *that*, a condição OROi não apenas é a mais comum, como constitui mais da metade de todo o conjunto de relativas. Em contraste, as duas condições de relativas com referente animado somam apenas 42% de todas as ocorrências. Após a realização de um teste qui-quadrado, podemos afirmar que tais resultados são estatisticamente

significativos:  $X^2 = 40,84$ ;  $p < 0,01$ . Repetimos que essa configuração das taxas de frequência de orações relativas foi durante muito tempo desconsiderada nos estudos que indicavam uma assimetria de processamento entre ORS e ORO, ignorando as estruturas na condição majoritária (OROi).

Partimos, então, para a comparação dos nossos resultados com aqueles coletados por Cabral (2016) nos seus *corpora* do português escrito e falado nas variedades brasileira e europeia (cf. pp. 38-39). Ressaltamos que essa comparação não pode ser perfeita por três razões: primeiramente, porque as fontes em língua inglesa que selecionamos eram todas textos escritos (embora uma minoria representasse transcrições de textos falados); em segundo lugar, devido ao fato de termos selecionado apenas aquelas relativas em que havia distinção de animacidade entre sujeito e objeto, algo que Cabral não menciona ter feito; mas, principalmente, porque o português *que* corresponde a *that*, *which*, *who* e *whom* em inglês. Uma vez que os três últimos denunciam a animacidade do referente, a análise do nosso *corpus* teve como foco orações relativas em que se utilizava *that*, semelhante em sua ambiguidade a *que*.

Percebemos de imediato um achado comum no nosso *corpus* e nos *corpora* de Cabral: a raridade da condição OROa, que, em Cabral, atinge a frequência máxima de 2% no português brasileiro falado e, no nosso *corpus*, sequer chega a ser registrada. Por outro lado, a distribuição das outras condições difere entre as duas línguas: se em português a ordem decrescente de frequência é ORSi > ORSa > OROi > OROa, em inglês encontramos as seguintes ordens: OROi > ORSa > ORSi > OROa, quando incluímos orações com qualquer um dos pronomes relativos considerados mais o conectivo *that*; e OROi > ORSi > ORSa > OROa ao selecionarmos apenas orações relativas com *that*. Reconhecendo as limitações inerentes a qualquer *corpus*, podemos supor a existência de diferenças não apenas na maneira como as orações relativas são formadas em cada língua, mas na distribuição estatística das quatro condições investigadas.

Tendo em mãos os resultados obtidos na construção desse *corpus*, perguntamo-nos se tais dados de frequência em inglês estão disponíveis de alguma forma para os falantes não nativos dessa língua, e se esse conhecimento poderia exercer influência no processamento de orações relativas por parte desses falantes. Além disso, seria possível que dados semelhantes, mas referentes à distribuição

dessas estruturas na L1, interfiram no processamento de L2? É justamente a possibilidade de que as diferentes frequências de determinadas estruturas linguísticas, intermediadas por um fator de caráter semântico, influenciariam o processamento bilíngue que nos levou a desenvolver o estudo atual.

## 4.2 Teste de leitura automonitorada

Diante dos resultados encontrados por pesquisas sobre o processamento nativo de orações relativas e o papel da animacidade nesse processo, e cientes da escassez de estudos semelhantes com participantes bilíngues, decidimos averiguar se produziríamos resultados comparáveis com um grupo não nativo. Pretendemos, com esta pesquisa, contribuir para a discussão em torno dos modelos de processamento linguístico vigentes e sua aplicabilidade ao processamento de L2, especificamente no que se refere à relevância de informações não sintáticas nos estágios iniciais do processamento – neste caso, de orações relativas.

Nosso objetivo geral é analisar o efeito da animacidade do antecedente no processamento de orações relativas de sujeito e de objeto por falantes bilíngues português (L1)/inglês (L2), de níveis intermediário e avançado. Como objetivos específicos, procuramos: criar um *corpus* de ORS e ORO em língua inglesa para determinar a frequência de seus diferentes tipos, de acordo com o traço de animacidade do antecedente (ORSa, ORSi, OROa, OROi); comparar as diferenças de frequência dos quatro tipos de oração relativa em inglês com sua frequência em língua portuguesa, tomando como parâmetro o *corpus* coletado por Cabral (2016); por fim, queremos contrastar os resultados do processamento de orações relativas por parte de falantes bilíngues português-inglês de acordo com seu nível de proficiência (intermediário ou avançado).

Para o estudo atual, temos como variáveis independentes a animacidade e o tipo de oração relativa (de sujeito ou de objeto), bem como o domínio da língua inglesa por parte dos participantes (intermediário ou avançado). Como variável dependente, citamos o tempo de processamento do segmento crítico (oração relativa). Para garantir que os sujeitos estavam atentos às sentenças lidas, verificamos o tempo de resposta à pergunta feita após a leitura dos estímulos e o índice de respostas corretas.

Tomando como base os resultados obtidos em estudos anteriores (cf. TRAXLER, 2013; PINÃR et al, 2017), parte-se da hipótese de que os participantes bilíngues desta

pesquisa serão capazes de utilizar pistas semânticas no processamento de orações relativas, e que essa habilidade variará de acordo com o nível de proficiência dos sujeitos da pesquisa. Em outras palavras, acredita-se que falantes bilíngues, especialmente os de nível avançado de proficiência na L2, serão sensíveis ao traço de animacidade em orações relativas nessa L2, de modo que a assimetria no tempo de leitura de ORS e ORO desaparecerá ou será muito atenuada, como ocorre nos experimentos com falantes nativos. Isso ocorreria porque, ao aprender uma outra língua, os falantes, especialmente os de maior experiência na L2, armazenariam e fariam uso de dados relativos à frequência dos diferentes tipos de relativa, de acordo com a animacidade dos referentes nessas orações.

#### 4.2.1 Participantes

Como amostra, reunimos um total de 32 participantes divididos em dois grupos: 16 falantes de inglês em nível intermediário e 16 em nível avançado<sup>31</sup>. Para classificá-los desse modo, utilizamos o *Vocabulary Levels Test* ou *VLT* (NATION, 1990) em uma versão *online* que disponibilizamos na plataforma *ClassMarker* (<https://www.classmarker.com>). Voluntários que não atingissem pelo menos o nível intermediário não puderam participar do experimento.

Embora o *VLT*, em sua configuração original, pudesse ser respondido na velocidade desejada pelo participante, decidimos estabelecer um limite de tempo de 10 minutos para a resolução do teste. Tomamos como base os estudos de Laufer e Nation (2001), que encontraram uma associação positiva entre velocidade de resposta e dimensão de vocabulário do participante; e de Souza, Duarte e Berg (2015), que indicaram 10 minutos, ou 20 segundos por item, como tempo suficiente para a associação de cada palavra à sua definição.

Em matéria de escolaridade, todos os sujeitos da pesquisa tinham pelo menos nível superior completo, com exceção de dois, que haviam concluído apenas o Ensino Médio. Suas idades variavam entre 22 e 60 anos, sendo 24 mulheres e 10 homens. Todos eles são falantes nativos de português brasileiro, tendo o inglês como segunda língua. Foram recrutados principalmente entre os professores de inglês da Rede Federal de Educação Profissional, Científica e Tecnológica (apenas 7 participantes

---

<sup>31</sup> Havia, inicialmente, 18 participantes de nível intermediário, mas dois deles foram excluídos da análise dos dados por se comportarem como *outliers*.



não se encaixam nesse perfil, por não serem professores ou por trabalharem em outras instituições), sendo provenientes de distintas regiões do Brasil.

#### 4.2.2 Materiais

Os estímulos utilizados foram 16 conjuntos de quatro sentenças cada, referentes às quatro condições investigadas (ORSa, OROa, ORSi, OROi). Retirando-se uma sentença de cada conjunto, criaram-se quatro listas diferentes com 16 frases experimentais cada, além de 32 frases distratoras cujo propósito era impedir que os participantes percebessem qual estrutura linguística era o foco do experimento. Esses estímulos foram tirados e/ou adaptados com base em Traxler (2013) e Piñar et al. (2017). Controlamos a extensão do segmento crítico (contendo a oração relativa encaixada) e do pós-crítico (o verbo principal da sentença). A distinção de animacidade era clara e binária: entidades animadas eram animais ou seres humanos, enquanto que as inanimadas eram todas objetos sem vida. Um exemplo de conjunto experimental, demonstrando as quatro condições analisadas, pode ser visto abaixo.

(38) The cowboy that concealed the pistol was known to be unreliable. (ORSa)<sup>32</sup>

(39) The cowboy that the pistol injured was known to be unreliable. (OROa)<sup>33</sup>

(40) The pistol that injured the cowboy was known to be unreliable. (ORSi)<sup>34</sup>

(41) The pistol that the cowboy concealed was known to be unreliable. (OROi)<sup>35</sup>

Assim como ocorreu com a aplicação do *VLT*, o experimento foi rodado nos computadores pessoais dos próprios participantes, por meio da Plataforma de Experimentação Psicolinguística (Zehr & Schwarz, 2018), PCIBex Farm (<https://expt.pcibex.net/>).

#### 4.2.3 Procedimentos

Para o experimento em questão, usamos o método de leitura automonitorada do tipo *moving window*. A tarefa do participante era ler sentenças na tela de um

---

<sup>32</sup> “O caubói que escondeu a pistola era conhecido por ser pouco confiável.” (Tradução nossa)

<sup>33</sup> “O caubói que a pistola feriu era conhecido por ser pouco confiável.” (Tradução nossa)

<sup>34</sup> “A pistola que feriu o caubói era conhecida por ser pouco confiável.” (Tradução nossa)

<sup>35</sup> “A pistola que o caubói escondeu era conhecida por ser pouco confiável.” (Tradução nossa)

computador. Inicialmente as sentenças estavam ocultas e os segmentos marcados por quatro traços. Ao apertar o botão de espaço, o primeiro traço era substituído pelo primeiro segmento da sentença; apertando-se mais uma vez o mesmo botão, o segundo segmento aparecia no lugar do segundo traço, enquanto o segmento anterior era novamente ocultado, e assim por diante até que a sentença inteira pudesse ser lida. O tempo de leitura de cada segmento era aferido em milissegundos. Partimos do pressuposto de que um tempo de leitura mais longo está relacionado a uma maior dificuldade de processamento do segmento em questão. Após isso, uma pergunta de compreensão simples, do tipo sim/não, era feita. O participante respondia utilizando comandos do teclado do computador (C para “sim”, M para “não”).

Antes de iniciar o experimento, dentro do próprio programa rodado na plataforma *online*, os sujeitos deviam concordar com um termo de consentimento livre e esclarecido, responder a um questionário sociolinguístico e passar por uma seção de prática com dez sentenças que não faziam parte dos estímulos utilizados na pesquisa. A escolha pelo uso de plataformas *online* para a aplicação do teste de proficiência e do experimento em si se deveu às dificuldades impostas pela pandemia de Covid-19, que levou ao fechamento temporário das universidades federais, bem como ao distanciamento social dos sujeitos envolvidos na pesquisa.

O experimento nos proporcionou resultados significativos, que serão expostos, analisados e interpretados na seção seguinte.

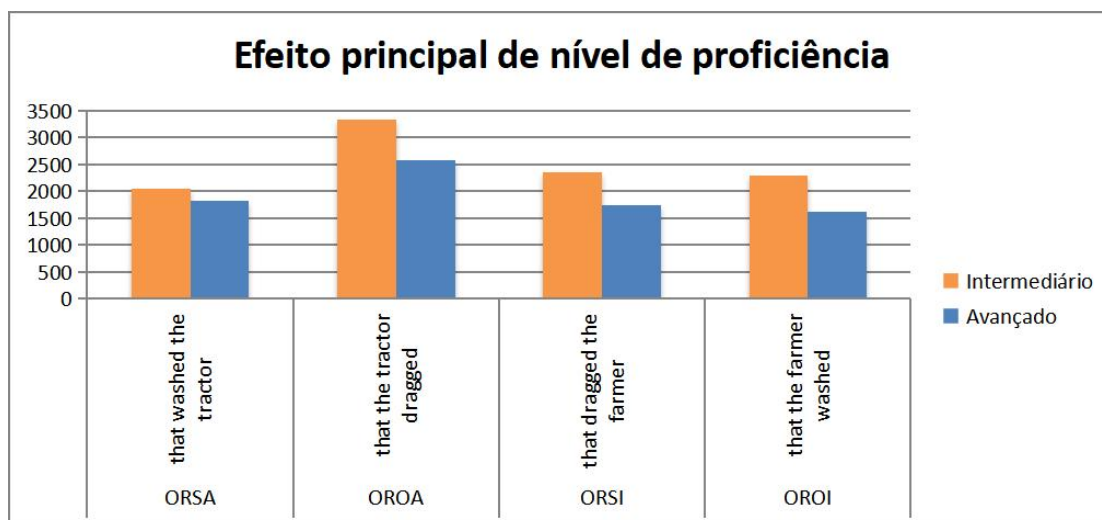
#### 4.2.4 Resultados

Os dados obtidos com o experimento geraram uma medida *online*: o tempo de leitura dos segmentos (sujeito, oração relativa encaixada, verbo e segmento final), conforme o exemplo (42). Além disso, registraram-se o índice de acerto e o tempo de resposta às perguntas feitas após cada sentença lida.

(42) The farmer / that washed the tractor / remained / next to the barn.

Reportamos aqui os efeitos para o segmento crítico (oração relativa encaixada) – no caso do exemplo (42), uma ORSa, *washed the tractor*.

Observando o Gráfico 5, destaquemos, primeiramente, a existência de efeito principal de nível de proficiência na leitura do segmento crítico, OR; ANOVA  $F(1,47) = 15.2$ ;  $p < 0,000305$ .<sup>36</sup>

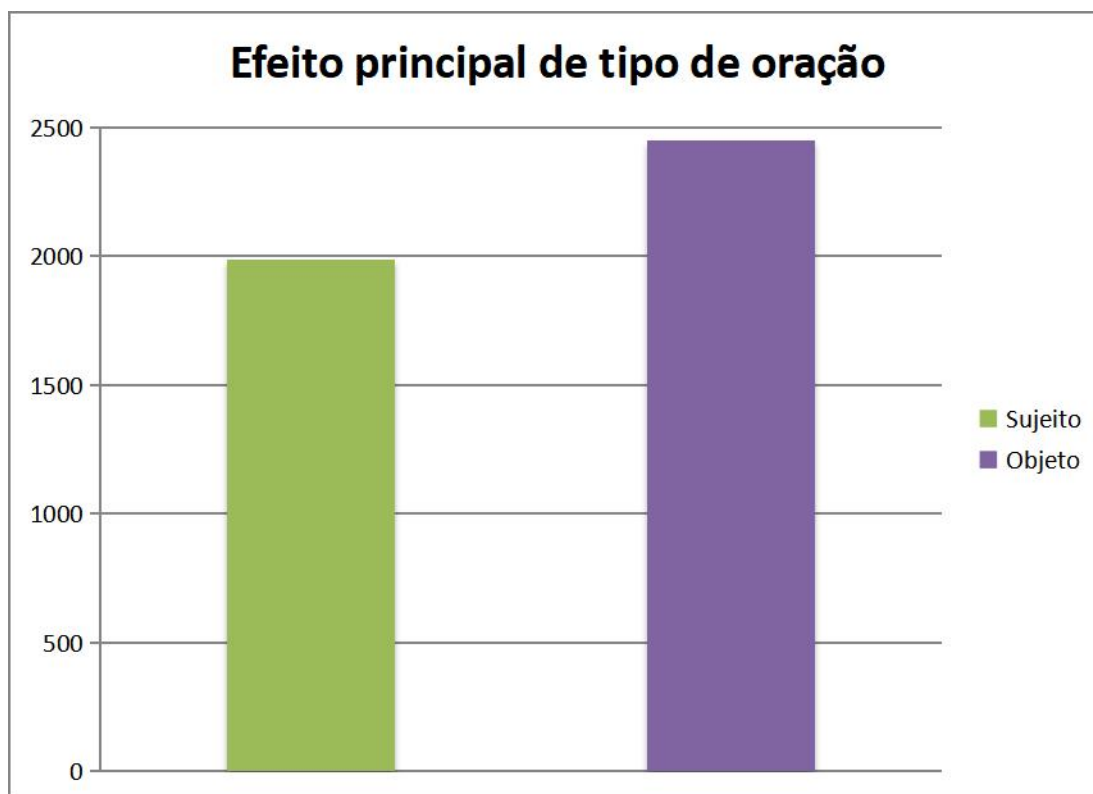


**Gráfico 5:** Tempo de leitura do segmento crítico por nível de proficiência e tipo de oração relativa.

Podemos visualizar, no gráfico acima, o impacto do nível de proficiência no processamento desse segmento, uma vez que em todas as condições experimentais – independentemente de tipo de relativa ou da animacidade do referente – falantes de nível avançado apresentaram tempos de leitura inferiores àqueles de nível intermediário. Também observamos efeito de interação entre nível de proficiência e tipo de oração relativa: ANOVA  $F(1,47) = 1,40$ ;  $p < 0,002419$ ; bem como entre nível de proficiência e animacidade do referente: ANOVA  $F(1,47) = 8,50$ ;  $p < 0,005419$ ; tipo de relativa e animacidade: ANOVA  $F(1,47) = 53,7$ ;  $p < 0,000001$ ; e, finalmente, nível de proficiência, tipo de relativa e animacidade: ANOVA  $F(1,47) = 2,41$ ;  $p < 0,001274$ . Todos esses resultados são estatisticamente significativos:  $p < 0,05$ .

Também verificamos um efeito principal de tipo de oração: ANOVA  $F(3,141) = 10,8$ ;  $p < 0,000002$ , que pode ser observado no Gráfico 6, abaixo.

<sup>36</sup> Utilizamos a análise de variância (ANOVA) a fim de manter a comparabilidade com estudos prévios na área (CABRAL, 2016; TRAXLER, MORRIS & SEELY, 2002; MAK, VONK & SCHRIEFERS, 2002).

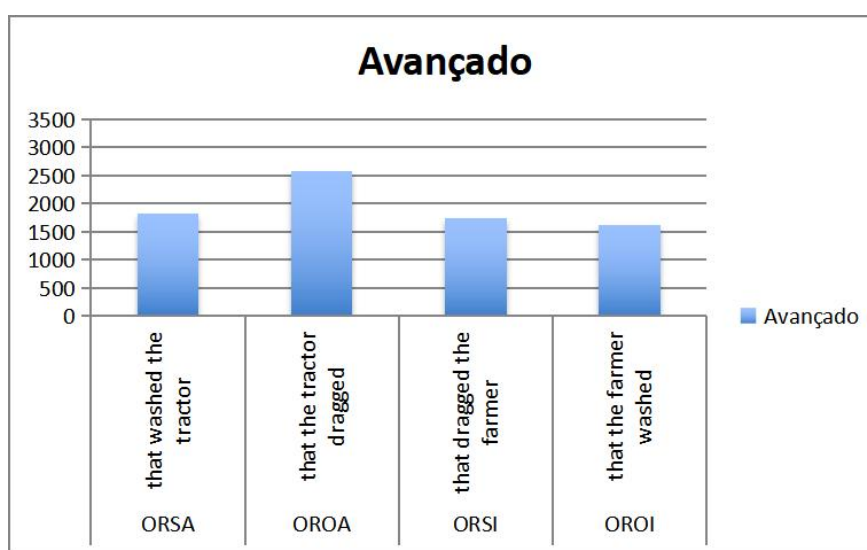


**Gráfico 6:** Assimetria entre o processamento de ORS e de ORO na junção dos resultados de falantes de nível intermediário e avançado.

Devemos destacar que, exclusivamente para fins ilustrativos, o Gráfico 6 mostra as médias obtidas a partir da junção dos resultados de falantes de nível tanto intermediário quanto avançado. Essa fusão, apesar de servir como um índice de tendência, pode mascarar a análise dos dados. Observe que, enquanto que no primeiro grupo permanece a assimetria de processamento entre ORS e ORO apontada pelos primeiros estudos na área, no segundo grupo essa assimetria desaparece. Na união dos resultados de ambos os grupos, a dissimetria pode ser constatada, como no gráfico acima. Percebe-se a maior facilidade de processamento das ORS frente às ORO.

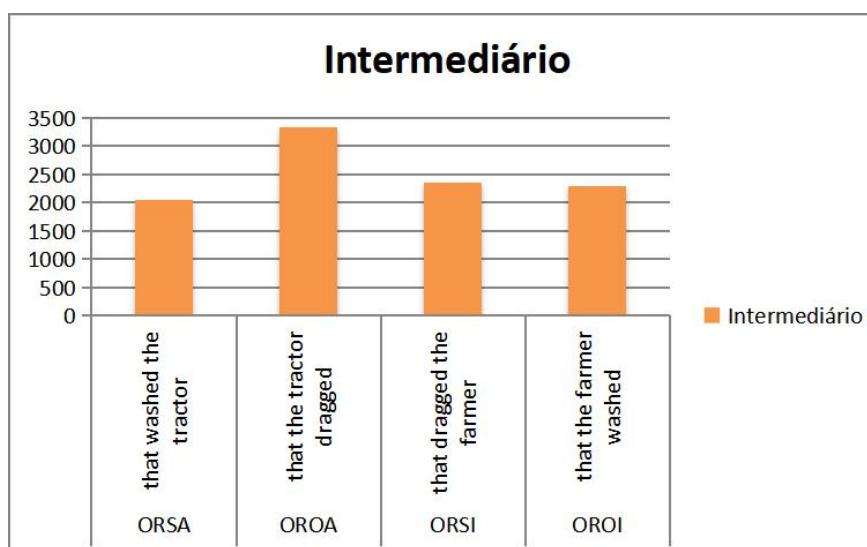
Debruçando-nos sobre os resultados que dizem respeito aos falantes de inglês como L2 em nível avançado, percebemos a discrepância entre a condição experimental que exigiu maior tempo de leitura (OROa) e a que demandou o menor tempo (OROi):  $t(47)=13,18$ ;  $p < 0,0001$ . Nota-se também a diferença entre ORSa e OROa:  $t(47)=31,38$ ;  $p < 0,0001$ . É importante lembrar que eram essas duas condições que geralmente figuravam nas pesquisas tradicionais sobre processamento de orações relativas. Pela exclusão das outras duas condições aqui investigadas (ORSi e OROi), criava-se a ilusão de uma assimetria entre ORS e ORO que não se explicava senão por

fatores sintáticos. Ainda de acordo com os nossos dados, disponíveis no Gráfico 7, os participantes com nível avançado processaram mais rapidamente as OR de sujeito e objeto inanimados, do que aquelas OR com sujeito e objeto animados, sendo estatisticamente relevante o contraste entre ORS e ORO ( $ORSa \times OROa$   $t(47)=7.90$ ;  $p < 0,0001$ ;  $ORSi \times OROi$   $t(47)=2.89$ ;  $p < 0.0058$ ). Entretanto, essa assimetria não se mantém nos falantes de nível intermediário, como se poderá ver nas análises subsequentes.



**Gráfico 7:** Tempo de leitura do segmento crítico dos diferentes tipos de oração relativa por falantes de nível avançado.

Observemos o Gráfico 8 abaixo, referente aos tempos de leitura dos segmentos críticos (OR) dos leitores com nível intermediário de inglês.



**Gráfico 8:** Tempo de leitura do segmento crítico dos diferentes tipos de oração relativa por falantes de nível intermediário.

O Gráfico 8 mostra os dados de leitura do segmento OR do grupo de falantes de inglês como L2 de nível intermediário. Também neste caso as OROa exigiram maior tempo de leitura que as outras condições. Todavia, foram as ORSa e não as OROi – como entre os sujeitos de nível avançado – que se mostraram o tipo de relativa mais facilmente processado. A diferença entre os tempos de leitura de ambas as condições é estatisticamente significativa:  $t(47)=2,34$ ;  $p < 0,0236$ . Os dados referentes ao processamento empreendido por este grupo contrastam com os obtidos com falantes nativos e bilíngues de nível avançado. No caso dos falantes de nível intermediário a assimetria sujeito-objeto se mantém mesmo quando a animacidade é controlada. Ou seja, quando comparamos ORSa x OROa há diferença estatística entre ORS e ORO ( $t(47)=2,34$ ;  $p < 0,0236$ ), porém na comparação entre ORS inanimado e ORO inanimado, os falantes parecem processá-las com o mesmo nível de dificuldade ( $t(47)=1,61$ ;  $p < 0,1134$ ).

Tendo em mãos esses resultados, passemos para as discussões que eles suscitam.

#### 4.2.5 Discussão dos resultados

A primeira constatação que podemos fazer em relação aos resultados obtidos é que, no que se refere aos tempos de leitura do segmento crítico nas quatro condições experimentais (ORSa, ORSi, OROa, OROi), os falantes de nível avançado foram mais rápidos que os de nível intermediário. Isso nos parece mais um indício da validade do *Vocabulary Levels Test (VLT)*, de Nation (1990), uma vez que os participantes que esse teste indicou como tendo maior proficiência em língua inglesa de fato demonstraram mais facilidade no processamento dos estímulos linguísticos utilizados em nosso experimento.<sup>37</sup>

Também podemos perceber que, à semelhança dos resultados obtidos em estudos de processamento em L1 (cf. CABRAL, 2016; TRAXLER, 2002; MAK, VONK & SCHRIEFERS, 2002), nossos sujeitos bilíngues demonstraram ser capazes de utilizar o traço semântico da animacidade no processamento de orações relativas em sua L2,

---

<sup>37</sup> Para um estudo sobre a capacidade do *VLT* de prever os conhecimentos gramaticais de bilíngues português (L1)/inglês (L2), cf. Souza & Soares-Silva (2015).

em particular os falantes de inglês em nível avançado. Para estes, a condição que demandou menos custo de processamento foi a OROi, enquanto que a condição OROa exigiu maior tempo de leitura. Um resultado não muito diferente foi encontrado para os bilíngues de nível intermediário; no entanto, esse grupo não apresentou assimetria entre o processamento de ORS e ORO, quando inanimados. Ainda assim, o padrão de leitura se manteve: ORS lidas mais rapidamente que as ORO. As ORSa, para esses sujeitos, parece configurar a condição de processamento mais fácil.

A diferença encontrada entre o tempo de processamento das ORSi e OROi entre os grupos de bilíngues intermediários e avançados é um dado interessante, pois sugere que, enquanto falantes de nível avançado poderiam utilizar-se de informações de cunho semântico assim como nativos, os de nível intermediário precisariam fazer uso de estratégias distintas durante esse processo, talvez estritamente sintáticas. Uma questão a ser levantada é se bilíngues de nível intermediário se guiam menos pela animacidade durante o processamento de relativas por ainda não terem desenvolvido conhecimento extenso sobre a frequência dos diferentes tipos de ORO e ORS.

Os resultados também condizem com os dados encontrados em nosso *corpus* de orações relativas em inglês, em especial no que se refere ao processamento empreendido por falantes de nível avançado. No caso deles, a ordem de condições experimentais de acordo com seu custo de processamento, da mais facilmente à mais dificilmente processada, é idêntica à ordem decrescente de frequência dessas condições no *corpus*: OROi > ORSi > ORSa > OROa.

Apenas no que se refere aos falantes de nível intermediário há alguma discrepância entre os dados do *corpus* e os resultados do nosso experimento. No caso desses participantes, a ordem de condições experimentais, da mais fácil à mais difícil de ser processada, é ORSa < OROi = ORSi < OROa. Podemos ver que a condição mais comum no *corpus*, OROi, não ocupa a primeira posição no *ranking* de facilidade de processamento. A de maior dificuldade, no entanto, é de fato causada pela menos frequente: OROa.

Também podemos comparar esses resultados com os já citados *corpora* nas línguas portuguesa, alemã e holandesa (cf. CABRAL, 2016; MAK, VONK & SCHRIEFERS, 2002). Todos esses estudos apontaram a condição OROa como a estrutura menos frequente no que se refere às orações relativas, justamente a que demanda maior tempo de leitura em nosso experimento. É importante ressaltar que,

no que se refere à frequência das outras condições, nossos resultados não condizem com os achados do *corpus* de Cabral, o que poderia indicar que os participantes do nosso experimento não fizeram uso de dados de frequência da L1 no processamento da L2. Por outro lado, os sujeitos de nossa pesquisa que falam inglês em nível avançado geraram resultados semelhantes aos brasileiros que participaram do experimento de Cabral: nos dois casos, o *ranking* de tipos de relativas, da mais rapidamente processada à mais lenta, foi o mesmo:  $ORO_i < ORS_i < ORS_a < ORO_a$ . Não consideramos provável que esse seja um indício de transferência dos conhecimentos de frequência de relativas na L1, já que nossos participantes de nível intermediário – que deveriam mostrar indícios ainda mais claros de transferência (SKALICKY & CHEN, 2020) – geraram um *ranking* diferente:  $ORS_a < ORO_i = ORS_i < ORO_a$ .

A ideia de que falantes não nativos seriam capazes de fazer previsões quanto à continuação de um *input* linguístico não pode ser explicada por um modelo teórico como a Habilidade Reduzida de Geração de Expectativas (*RAGE*). Como sabemos, essa hipótese postula que bilíngues teriam um *parser* de capacidade limitada, que se ocuparia exclusivamente com o processamento reativo, sobrando-lhe poucos recursos para a interpretação de pistas não essenciais que poderiam ser utilizadas para atualizar expectativas (GRÜTER, ROHDE & SCHAFER, 2014). Ora, os sujeitos de nossa pesquisa aparentam ter demonstrado sensibilidade ao traço semântico de animacidade, utilizando-o a seu favor no processamento das estruturas mais frequentes (isto é, que seguiam suas expectativas). Para esses bilíngues, o *parser* parece ter dado conta não apenas de informações incrementais, como acesso lexical e integração estrutural, mas também da criação e revisão de expectativas feitas quanto à continuação das sentenças lidas.

O outro modelo de processamento de L2 que abordamos foi a Hipótese da Estrutura Rasa (*SSH*) (CLAHSEN & FELSNER, 2006; 2017), segundo a qual bilíngues, por um lado, fariam uso normal de informações lexicais, semânticas, pragmáticas e probabilísticas ao processarem estímulos em sua L2, assim como ocorre no processamento de L1, mas apenas se as dependências processadas forem adjacentes na sentença (KAAN, 2014). Além disso, não nativos – mesmo os de alta proficiência na L2 – não conseguiriam gerar e manipular representações sintáticas abstratas em tempo real em um segundo idioma com a mesma desenvoltura que falantes nativos dessa língua. Os dados que obtivemos com nosso experimento não se encaixam com o



primeiro postulado da *SSH*, uma vez que as ORO, que envolvem relações sintáticas de longa distância, parecem ter tido seu processamento facilitado com o uso do dado de animacidade pelos nossos participantes. A segunda afirmação da hipótese não pode ser desmentida nem confirmada por nossos resultados. Mais estudos serão necessários para jogar luz sobre essa questão.

Poucos pesquisadores se debruçaram sobre a questão do processamento não nativo de orações relativas. Duas das pesquisas na área foram citadas aqui, nomeadamente as de Traxler et al. (2013) e Piñar et al. (2017). Por meio, respectivamente, dos métodos de leitura automonitorada e de rastreamento ocular, e com a participação de sujeitos surdos que tinham a língua inglesa como L2, esses estudos geraram dados que apontam para a capacidade de falantes não nativos de se utilizarem da animacidade como informação relevante durante o processamento de orações relativas em sua L2. Nosso experimento confirma os achados dessas pesquisas e aponta para a possibilidade de que fatores individuais, e não o uso de mecanismos diferentes, explicariam as eventuais diferenças de processamento de L1 e L2.

Indícios do uso de informações de cunho semântico e probabilístico, além de sintático, no processamento linguístico, tanto por falantes nativos quanto não nativos, reacendem a discussão sobre a natureza desse processamento. Modelos téoricos modulares, como a *TGP* e a *SPLT*, mostram-se inadequados para explicar tais resultados, uma vez que postulam a primazia de fatores sintáticos no processamento linguístico. Por outro lado, uma teoria interativa como a *DLT*, em que se considera o uso de dados probabilísticos como elemento facilitador das estruturas mais frequentes, permanece viável como explicação dos processos que caracterizam a compreensão de orações relativas em particular, e da linguagem de maneira mais ampla. Para isso, no entanto, precisaríamos considerar a *DLT* como um modelo de processamento linguístico de modo geral, e não apenas de L1.

## 5. CONSIDERAÇÕES FINAIS

Para o experimento de leitura automonitorada que empreendemos, utilizamos orações relativas de sujeito e de objeto, controlando a animacidade do referente em todos os casos. Estudos anteriores em que esse traço semântico era desconsiderado, privilegiando o uso de orações relativas cujos referentes eram sempre animados, encontravam como resultado uma maior dificuldade no processamento de ORO que de ORS. Isso se deveria a uma diferença de complexidade sintática entre as duas estruturas, sendo as ORO mais complexas que as ORS.

Algumas teorias baseadas na diferença de complexidade estrutural foram propostas na tentativa de explicar essa assimetria de processamento. A Teoria da Predição Sintática Local (*Syntactic Prediction Locality Theory*, ou *SPLT*) de Gibson (1998), por exemplo, postula que a análise favorecida no processamento de uma oração relativa é escolhida de acordo com o custo de memorização de todas as alternativas. Ao se deparar com um pronome relativo, o ouvinte ou leitor preferiria interpretar o *input* como uma ORS porque esse tipo de estrutura demanda a memorização de apenas dois constituintes, em vez dos três exigidos no caso de uma ORO. Tal explicação seria satisfatória para os resultados que se encontravam nas pesquisas psicolinguísticas de então.

No entanto, autores como Mak, Vonk e Schriefers (2002) perceberam, por meio de análise de *corpus*, que em línguas como o holandês e o alemão a distribuição de ORS e ORO sofre influência da animacidade dos referentes de tais estruturas. Os pesquisadores se perguntaram, então, se as diferenças de frequência entre os tipos de oração relativa poderiam estar enviesando os resultados apontados na literatura da época. Por meio de um teste de leitura automonitorada, os autores confirmaram sua suspeita: quando o objeto da oração relativa era inanimado, a assimetria na dificuldade de processamento de ORS e ORO desaparecia.

Um resultado como esse não pode ser explicado por modelos de processamento linguístico ditos modulares, como a *SPLT*, segundo os quais apenas fatores de ordem sintática são utilizados pelo *parser* nos momentos iniciais do processamento. Frente a resultados semelhantes, Gibson (2000) propôs uma reformulação da *SPLT*, que ele batizou de Teoria da Dependência Local, ou *Dependency Locality Theory (DLT)*. De acordo com essa teoria, além de fatores sintáticos, o *parser* também teria acesso a

informações concernentes à frequência de determinadas estruturas, sendo facilitado o processamento daquelas mais comuns.

Todavia, os estudos que indicaram a importância da animacidade na compreensão de orações relativas foram geralmente empreendidos com participantes lendo ou ouvindo estímulos em sua língua materna. Resultados semelhantes seriam encontrados se utilizássemos sujeitos expostos a orações relativas em sua L2? Nosso estudo sugere que sim, especialmente no que se refere aos bilíngues de nível avançado.

Ao dividirmos as orações relativas em quatro tipos distintos, de acordo não apenas com sua classificação em ORO e ORS, mas também segundo a animacidade de seus referentes, descobrimos que, ao menos no que concerne ao processamento linguístico empreendido por bilíngues de nível avançado, o tipo mais facilmente compreendido era o de OROi, justamente aquele que não figurava nas pesquisas tradicionais sobre processamento de orações relativas. Nestas, apenas as OROa eram utilizadas, embora aparentem ser o tipo menos frequente de oração relativa na língua inglesa. Tal raridade, e não sua complexidade estrutural, pode ser a explicação para sua leitura mais lenta por parte de falantes nativos e não nativos de nível avançado. Notavelmente, falantes não nativos de nível intermediário apresentaram resultados ligeiramente diferentes: para eles, as ORSa eram a estrutura mais rapidamente processada, embora, deva-se destacar, não tenha havido diferença significativa entre os tempos de leitura de ORSi e OROi. Mais uma vez, é posta em dúvida a ideia de que as ORO sempre acarretam maior dificuldade de compreensão.

Nossa pesquisa aponta para uma grande semelhança entre o processamento linguístico empreendido por falantes nativos e não nativos, uma vez que ambos os grupos demonstram fazer uso de informações de naturezas distintas durante a compreensão de orações relativas. Além de fatores sintáticos, também o traço semântico de animacidade aparenta exercer influência sobre esse processo. Isso indica que falantes não nativos seriam capazes de prever o tipo de oração relativa estatisticamente mais provável após um referente animado ou inanimado, o que põe em xeque a ideia, defendida pela hipótese da Habilidade Reduzida de Geração de Expectativas (*Reduced Ability to Generate Expectations*, ou *RAGE*) (GRÜTER & ROHDE, 2013) de que previsões não são feitas durante o processamento de L2, bem como o postulado da Hipótese da Estrutura Rasa (*SSH*) (CLAHSEN & FELSER,

2006; 2017) de que tais previsões só seriam possíveis no processamento de dependências adjacentes.

Os dados obtidos com este estudo são relevantes e contribuem para a importante discussão sobre o papel de fatores não sintáticos no processamento de segunda língua, ajudando na compreensão da natureza da representação linguística que um falante não nativo tem de sua L2. Dito isso, nossa contribuição está longe de encerrar definitivamente as questões levantadas nessa área da psicolinguística. Para caminharmos em direção a um entendimento mais aprofundado do processamento linguístico de modo geral, e do processamento bilíngue em particular, são necessárias pesquisas que se utilizem de outras configurações de participantes, em línguas diversas, por meio de métodos experimentais variados e analisando diferentes estruturas linguísticas. Dessa forma, poderemos testar a aplicabilidade e universalidade dos resultados que temos encontrado. De nossa parte, pretendemos, no futuro, aplicar novamente o nosso experimento, dessa vez com falantes de inglês como L1. Assim, confirmaremos ou contextualizaremos os resultados que descrevemos aqui.

Embora com um foco específico, buscamos dar nossa contribuição para as pesquisas de processamento bilíngue. Acreditamos firmemente na importância de se tomar o falante não nativo como objeto de estudo legítimo na psicolinguística. Num contexto global em que monolíngues constituem uma minoria, é imprescindível que modelos de processamento linguístico levem em consideração as representações mentais que indivíduos bilíngues têm de suas línguas, sem enxergá-los como falantes excepcionais ou defeituosos. Se hipóteses como *SSH* e *RAGE* estiverem corretas quando postulam que o processamento de L2, mesmo por parte de falantes com nível avançado de proficiência, seria essencialmente diferente do empreendido por monolíngues, é necessário que obtenhamos conhecimento dessas diferenças. Apenas assim poderemos nos aproximar de um retrato completo dos mecanismos que subjazem a capacidade humana para a linguagem.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALTMANN, Gerry T. M.; KAMIDE, Yuki. Incremental Interpretation at Verbs: Restricting the Domain of Subsequent Reference. **Cognition**, vol. 73, pp. 247-264, 1999.
- ARISTÓTELES. **De Anima**. Tradução de Maria Cecília Gomes dos Reis. São Paulo: Editora 34, 2006.
- BAYANATI, Shiva; TOIVONEN, Ida. Humans, Animals, Things and Animacy. **Open Linguistics**, v. 5, pp. 156-170, 2019.
- BEVER, Thomas. The Cognitive Basis for Linguistic Structure. In: HAYES, John R. (ed.). **Cognition and the Development of Language**. Nova Iorque: Wiley & Sons, 1970, pp. 279-363.
- BHATIA, Tej K. Bilingualism and multilingualism. In: HOGAN, Patrick Colm (org.). **The Cambridge Encyclopedia of the Language Sciences**. Cambridge: Cambridge University Press, 2011, pp. 125-127.
- BRAGA, Junia; RACILAN, Marcos; GOMES, Ronaldo. **New Alive High**. São Paulo: SM, 2020.
- BROOKS, Patricia J.; BRAINE, Martin D. S.; CATALANO, Lisa; BRODY, Ruth E.; SUDHALTER, Vicki. Acquisition of Gender-like Noun Subclasses in an Artificial Language: The Contribution of Phonological Markers to Learning. **Journal of Memory and Language**, v. 32, n. 1, pp. 76-95, 1993.
- BROWN, Keith; MILLER, Jim. **The Cambridge Dictionary of Linguistics**. Cambridge: Cambridge University Press, 2013.
- BURTON-ROBERTS, Noel. Sentence. In: HOGAN, Patrick Colm (org.). **The Cambridge Encyclopedia of the Language Sciences**. Cambridge: Cambridge University Press, 2011.
- CABRAL, Althiere F. V. **A influência da animacidade no processamento de relativas de sujeito e de objeto no português brasileiro e europeu**. 2016. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal da Paraíba, João Pessoa.
- CAMBRIDGE DICTIONARY. **Lead sb up the garden path**. Disponível em: <<https://dictionary.cambridge.org/dictionary/english/lead-sb-up-the-garden-path>>. Acesso em: 25 de abril de 2020.
- CAPLAN, David; ALPERT, Nathaniel; WATERS, Gloria. Effects of Syntactic Structure and Propositional Number on Patterns of Regional Cerebral Blood Flow. **Journal of Cognitive Neuroscience**, v. 10, n. 4, pp. 541-552, 1998.
- CARAMAZZA, Alfonso; ZURIF, Edgar B. Dissociation of Algorithmic and Heuristic Processes in Language Comprehension: Evidence from Aphasia. **Brain and Language**, v. 3, n. 4, pp. 572-582, 1976.
- CHEON, Kyung-Hwan; KIM, Youngjoo; YOON, Hee-Dong; NAM, Ki-Chun; LEE, Sun-Young; JEON, Hyeon-Ae. Syntactic Comprehension of Relative Clauses and Center Embedding Using Pseudowords. **Brain Sciences**, v. 10, n. 4, n.p., 2020.
- CHEREPOVSKAIA, Natalia; REUTOVA, Elizaveta; SLIOUSSAR, Natalia. Becoming Native-Like for Good or Ill: Online and Offline Processing of Case Forms in L2 Russian. **Frontiers in Psychology**, v. 12, n.p., 2021.
- CHIERCHIA, Gennaro. **Semântica**. Campinas: Unicamp, 2003.

- CHOMSKY, Noam. On Wh- Movement. In: CULICOVER, Peter W.; WASOW, Thomas; AKMAJIAN, Adrian (Eds.). **Formal Syntax**. Nova York: Academic Press, pp. 71-132, 1977.
- CLAHSEN, Harald; FELSER, Claudia. Grammatical Processing in Language Learners. **Applied Psycholinguistics**, vol. 27, pp. 3-42, 2006.
- \_\_\_\_\_. Some Notes on the Shallow Structure Hypothesis. **Studies in Second Language Acquisition**, vol. 40, n. 3, pp. 693-706, 2017.
- COOK, Vivian. Introduction: The Changing L1 in the L2 User's Mind. In: COOK, Vivian (Org.). **Effects of the Second Language on the First**. Clevedon: Multilingual Matters, 2003, pp. 11-38
- \_\_\_\_\_. **Linguistics and Second Language Acquisition**. Nova York: Macmillan, 1993.
- \_\_\_\_\_. How Do Different Languages Connect in Our Minds? In: COOK, Vivian; SINGLETON, David (org.). **Key Topics in Second Language Acquisition**. Bristol: Multilingual Matters, 2014, pp. 20-51.
- CORBETT, Greville G. **Number**. 2. ed. Cambridge: Cambridge University Press, 2004.
- CUETOS, Fernando; MITCHELL, Don C. Cross-linguistic Differences in Parsing: Restrictions on the Use of the Late Closure Strategy in Spanish. **Cognition**, v. 30, p. 73-105, 1988.
- DAHL, Östen. Animacy and Egophoricity: Grammar, Ontology and Philogeny. **Lingua**, v. 118, pp. 141-150, 2008.
- DAHL, Östen; FRAURUD, Kari. Animacy in Grammar and Discourse. In: FRETHERM, Thorstein; GUNDEL, Jeanette K. **Reference and Referent Accessibility**. Amsterdã: John Benjamins Publishing Company, 1996.
- DECLERCK, Renaat. **A Comprehensive Descriptive Grammar of English**. Tóquio: Kaitakusha, 1991.
- DE SWART, Peter; DE HOOP, Helen. Shifting Animacy. In: KRIFKA, Manfred; GÄRTNER, Hans-Martin (Org.). **Theoretical Linguistics**, vol. 44, n. 1-2, 2018.
- DE VRIES, Mark. **The Syntax of Relativization**. Amsterdã: LOT, 2002.
- DIESSEL, Holger; HILPERT, Martin. **Frequency effects in grammar**, 2016. Disponível em: <[https://www.researchgate.net/publication/303540830\\_Frequency\\_effects\\_in\\_grammar](https://www.researchgate.net/publication/303540830_Frequency_effects_in_grammar)>. Acesso em: 12 nov. 2021.
- DUSSIAS, Paola E.; CRAMER SCALTZ, Tracy R. Spanish-English L2 Speakers' Use of Subcategorization Bias Information in the Resolution of Temporary Ambiguity During Second Language Reading. **Acta Psychologica**, vol 128, n. 3, pp. 501-513, 2008.
- DUSSIAS, Paola E.; KROFF, Jorge R. Valdés; TAMARGO, Rosa E. Guzzardo; GERFEN, Chip. When Gender and Looking Go Hand in Hand: Grammatical Gender Processing in L2 Spanish. **Studies in Second Language Acquisition**, vol. 35, n. 2, pp. 353-387, 2013.
- DUSSIAS, Paola E.; MARFUL, Alejandra; GERFEN, Chip; MOLINA, María Teresa Bajo. Usage Frequencies of Complement-Taking Verbs in Spanish and English: data from Spanish Monolinguals and Spanish-English Bilinguals. **Behavior Research Methods**, vol. 42, n. 4, pp. 1004-1011, 2010.

- FÉLIX, Sara B.; PANDEIRADA, Josefa N. S.; NAIRNE, James S. Adaptive memory: longevity and learning intentionality of the animacy effect. **Journal of Cognitive Psychology**, vol. 31, n. 3, pp. 251-260, 2019.
- FERREIRA, Fernanda; BAILEY, Karl G. D.; FERRARO, Vittoria. Good-Enough Representations in Language Comprehension. **Current Directions in Psychological Science**, vol. 11, pp. 11-15, 2002.
- FERREIRA, Fernanda; ÇOKAL, Derya. Sentence Processing. In: HICKOK, Gregory; SMALL, Steven L. (org.). **Neurobiology of Language**. Amsterdã: Elsevier, 2016.
- FERREIRA, Fernanda; HENDERSON, John M. Recovery from Misanalyses of Garden-Path Sentences. **Journal of Memory and Language**, v. 30, pp. 725-745, 1991.
- FIELD, John. Psycholinguistics. In: HOGAN, Patrick Colm (org.). **The Cambridge Encyclopedia of the Language Sciences**. Nova York: Cambridge University Press, 2011, pp. 686-689.
- FINGER, Ingrid. Processamento de segunda língua. In: MAIA, Marcus (org.). **Psicolinguística, psicolinguísticas: uma introdução**. São Paulo: Contexto, 2015, pp. 157-170.
- FORD, Marilyn. A Method for Obtaining Measures of Local Parsing Complexity through Sentences. **Journal of Verbal Learning and Verbal Behavior**, v. 22, n. 2, pp. 203-218, 1983.
- FOUCARD, Alice; FRENCK-MESTRE, C. Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1-L2 syntactic similarity. **Bilingualism: Language and Cognition**, v. 14, pp. 379-399, 2010.
- FRANÇA, Aniela Improta; GOMES, Juliana Novo. A técnica de ERP: investigando a assimetria sujeito-objeto na interface sintaxe-semântica com EEG. **Letras de Hoje**, v. 50, n. 3, pp. 360-370, 2015.
- FRAUENFELDER, Ulrich; SEGUI, Juan; MEHLER, Jacques. Monitoring around the Relative Clause. **Journal of Verbal Learning and Verbal Behavior**, v. 19, n. 3, pp. 328-337, 1980.
- FRAZIER, Lyn. **On comprehending sentences: Syntactic parsing strategies**. Bloomington: Indiana University Linguistics Club, 1979.
- \_\_\_\_\_. Syntactic Processing: Evidence from Dutch. **Natural Language & Linguistic Theory**, v. 5, n. 4, pp. 519-559, 1987.
- FRAZIER, Lyn; FLORES D'ARCAIS, Giovanni B. Filler Driven Parsing: A Study of Gap Filling in Dutch. **Journal of Memory and Language**, v. 28, n. 3, pp. 331-344, 1989.
- FRAZIER, Lyn; FODOR, Janet Dean. The sausage-machine: A new two-stage parsing model. **Cognition**, Lausanne, v. 6, pp. 291-325, 1978.
- GENNARI, Silvia P.; MACDONALD, Maryellen C. Semantic Indeterminacy in Object Relative Clauses. **Journal of Memory and Language**, v. 58, pp. 161-187, 2008.
- GIBSON, Edward. Linguistic Complexity: Locality of Syntactic Dependencies. **Cognition**, v. 68, n. 1, pp. 1-76, 1998.
- \_\_\_\_\_. The Dependency Locality Theory: A Distance-Based Theory of Linguistic Complexity. In: MARANTZ, Alec; MIYASHITA, Yasushi; O'NEIL, Wayne (org.). **Image, Language, Brain: Papers from the First Mind Articulation Project Symposium**. Cambridge, Massachusetts: MIT Press, 2000, pp. 95-126.

- GOLLAN, Tamar H.; MONTOYA, Rosa I.; FENNEMA-NOTESTINE, Christine; MORRIS, Shauna K. Bilingualism affects picture naming but not picture classification. **Memory & Cognition**, v. 33, n. 7, pp. 1220-1234, 2005.
- GORDON, Peter C.; HENDRICK, Randall; JOHNSON, Marcus. Memory Interference During Language Processing. **Journal of Experimental Psychology: Learning, Memory, and Cognition**, v. 27, n. 6, pp. 1411-1423, 2001.
- GREEN, T. R. G. The Necessity of Syntax Markers: Two Experiments with Artificial Languages. **Journal of Verbal Learning and Verbal Behavior**, vol. 18, pp. 481-496, 1979.
- GROSJEAN, François. Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person. **Brain and Language**, v. 36, pp. 3-15, 1989.
- GRÜTER, Theres; LAU, Elaine; LING, WENYI. How classifiers facilitate predictive processing in L1 and L2 Chinese: The role of semantic and grammatical cues. **Language, Cognition and Neuroscience**, v. 34, pp. 1-14, 2019.
- GRÜTER, Theres; LEW-WILLIAMS, Casey; FERNALD, Anne. Grammatical Gender in L2: A Production or a Real-Time Processing Problem? **Second Language Research**, vol. 28, n. 2, pp. 191-215, 2012.
- GRÜTER, Theres; ROHDE, Hannah. L2 Processing Is Affected by Rage: Evidence from Reference Resolution. In: CONFERENCE ON GENERATIVE APPROACHES TO SECOND LANGUAGE ACQUISITION (GASLA), 12, 2013, Gainesville.
- GRÜTER, Theres; ROHDE, Hannah; SCHAFFER, Amy J. The Role of Discourse-Level Expectations in Non-Native Speakers' Referential Choices. In: BOSTON UNIVERSITY CONFERENCE ON LANGUAGE DEVELOPMENT, 38, 2014, Boston. **Anais...** Boston: Cascadilla Press, 2014, pp. 179-191.
- HOLMES, Virginia M.; O'REGAN, J. Kevin. Eye Fixation Patterns during the Reading of Relative-Clause Sentences. **Journal of Verbal Learning and Verbal Behavior**, v. 20, n. 4, pp. 417-430, 1981.
- HOPF, Jens-Max; BADER, Markus; MENG, Michael; BAYER, Josef. Is human sentence parsing serial or parallel?: Evidence from event-related brain potentials. **Cognitive Brain Research**, v. 15, n. 2, pp. 165-177, 2003.
- HOPP, Holger. Grammatical Gender in Adult L2 Acquisition: Relations Between Lexical and Syntactic Variability. **Second Language Research**, vol. 29, n. 1, pp. 33-56, 2013.
- \_\_\_\_\_. Individual differences in the second language processing of object-subject ambiguities. **Applied Psycholinguistics**, pp. 1-45, 2013.
- \_\_\_\_\_. Syntactic features and reanalysis in near-native processing. **Second Language Research**, v. 22, n. 3, pp. 369-397, 2006.
- JUST, Marcel Adam; CARPENTER, Patricia A.; KELLER, Timothy A.; EDDY, William F.; THULBORN, Keith R. Brain Activation Modulated by Sentence Comprehension. **Science**, v. 274, n. 5284, pp. 114-116, 1996.
- KAAN, Edith; BALLANTYNE, Jocelyn; WIJNEN, Frank. Effects of reading speed on second-language sentence processing. **Applied Psycholinguistics**, v. 36, n. 4, pp. 799, 830, 2015.
- KAAN, Edith. Predictive Sentence Processing in L2 and L1: What is Different? **Linguistic Approaches to Bilingualism**, vol. 4, n. 2, pp. 257-282, 2014.



- KENEDY, Eduardo. Estruturas sintáticas de orações relativas. In: BISPO, Edvaldo Balduino; OLIVEIRA, Mariangela Rios de. (Org.). **Orações relativas do português brasileiro: diferentes perspectivas**. Niterói: EdUFF, 2014; pp. 11-46.
- \_\_\_\_\_. Modelos interativos no processamento de frases: a Teoria da Dependência Local. **Revista Prolíngua**, v. 10, n. 1, p. 150-162, 2015.
- KING, Jonathan; JUST, Marcel Adam. Individual Differences in Syntactic Processing: The Role of Working Memory. **Journal of Memory and Language**, v. 30, n. 5, pp. 580-602, 1991.
- KRESSLER, Brett. Language Families. In: HOGAN, Patrick Colm (org.). **The Cambridge Encyclopedia of the Language Sciences**. Cambridge: Cambridge University Press, pp. 414-417, 2011.
- LANGACKER, Ronald W. **Foundations of Cognitive Grammar – Volume 2: Descriptive Application**. Stanford: Stanford University Press, 1991.
- LAUFER, Batia; NATION, Paul. Passive vocabulary size and speed of meaning recognition: Are they related? **EUROSLA Yearbook**, v. 1, pp. 7-28, 2001.
- LEAL, Tania; SLABAKOVA, Roumyana; FARMER, Thomas A. The Fine-Tuning of Linguistic Expectations over the Course of L2 Learning. **Studies in Second Language Acquisition**, 2016.
- LEE, Eun-Kyung; LU, Dora Hsin-Yi; GARNSEY, Susan M. L1 Word Order and Sensitivity to Verb Bias in L2 Processing. **Bilingualism: Language and Cognition**, vol. 16, n. 4, pp. 761-775, 2013.
- LEHMANN, Christian. Towards a typology of clause linkage. In: HAIMAN, John; THOMPSON, Sandra (eds.). **Clause combining in grammar and discourse**. Amsterdã, J. Benjamins, 1988, pp. 181-225.
- LEITÃO, Márcio Martins. Psicolinguística experimental: focalizando o processamento da linguagem. In: MARTELOTTA, Mário Eduardo (Org.). **Manual de Linguística**. São Paulo: Contexto, 2008.
- LEVY, Roger; GIBSON, Edward. Surprisal, the PDC, and the Primary Locus of Processing Difficulty in Relative Clauses. **Frontiers in Psychology**, v. 4, p. 1-3, 2013.
- MACDONALD, Maryellen C.; CHRISTIANSEN, Morten H. Reassessing Working Memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). **Psychological Review**, v. 109, n. 1, 35-54, 2002.
- MACDONALD, Maryellen C.; PEARLMUTTER, Neal J.; SEIDENBERG, Mark S. Lexical Nature of Syntactic Ambiguity Resolution. **Psychological Review**, v. 101, n. 4, 676-703, 1994.
- MAIA, Marcus. **Psicolinguística, psicolinguísticas: uma introdução**. São Paulo: Contexto, 2015.
- MAK, Willem M.; VONK, Wietske; SCHRIEFERS, Herbert. The Influence of Animacy on Relative Clause Processing. **Journal of Memory and Language**, v. 47, n. 1, pp. 50-68, 2002.
- MARTIN, Clara D.; THIERRY, Guillaume; KUIPERS, Jan-Rouke; BOUTONNET, Bastien; FOUCART, Alice; COSTA, Albert. Bilinguals Reading in Their Second Language Do Not Predict Upcoming Words As Native Readers Do. **Journal of Memory and Language**, vol. 69, n. 4, pp. 574-588, 2013.

MARTINEZ-GARCIA, María Teresa; WULFF, Stefanie. Not Wrong, Yet Not Quite Right: Spanish ESL Students' Use of Gerundial and Infinitival Complementation. **International Journal of Applied Linguistics**, vol. 22, n. 2, pp. 225-244, 2012.

MATTHEWS, P. H. **The Concise Oxford Dictionary of Linguistics**. 2 ed. Oxford: Oxford University Press, 2007.

MICHAELIS. **Animado**. Disponível em: <<https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/animado/>>. Acesso em: 9 de abril de 2020.

MISYAK, Jennifer B.; CHRISTIANSEN, Morten H. Extending Statistical Learning Farther and Further: Long-Distance Dependencies, and Individual Differences in Statistical Learning and Language. In: McNAMARA, D. S.; TRAFTON, J. Gregory. (eds.). **Proceedings of the 29<sup>th</sup> Annual Cognitive Science Society Conference**, 2007, pp. 1307-1312.

MITCHELL, Don C. On-line methods in language processing: Introduction and historical review. In: CARREIRAS, Manuel; CLIFTON, Jr., Charles. **The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond**. Nova York: Psychology Press, 2004, pp. 15-32.

MORGAN, James L.; MEIER, Richard P.; NEWPORT, Elissa L. Structural Packaging in the Input to Language Learning: Contributions of Prosodic and Morphological Marking of Phrases to the Acquisition of Language. **Cognitive Psychology**, v. 19, pp. 498-550, 1987.

NATION, Paul. **Teaching and Learning Vocabulary**. Boston, MA: Heinle & Heinle, 1990.

NAIRNE, James S.; VANARSDALL, Joshua E.; COGDILL, Mindi. Remembering the Living: Episodic Memory is Tuned to Animacy. **Current Directions in Psychological Science**, v. 26, n. 1, pp. 22-27, 2017.

NIHONMATSU, Karino Otsuka. **Diálogo**. São Paulo: Moderna, 2020.

OLIVEIRA, Fernando Lúcio de. **O processamento da assimetria sujeito-objeto em construções do tipo QU no português brasileiro: interrogativas e relativas**. Dissertação de Mestrado – Universidade Federal do Rio de Janeiro, 2013.

OPFER, John E.; GELMAN, Susan A. Development of the Animate-Inanimate distinction. In: GOSWAMI, Usha (org.). **The Wiley-Blackwell Handbook of Childhood Cognitive Development**. 2. ed. Malden: Wiley-Blackwell, 2011, p. 213-238.

PAM, N. **Garden-Path Sentence**, 2013. Disponível em: <<https://psychologydictionary.org/garden-path-sentence/>>. Acesso em: 25 de abril de 2020.

PATEL, Aniruddh D. Language, music, syntax and the brain. **Nature Neuroscience**, v. 6, n. 7, pp. 674-681, 2003.

PEARLMUTTER, Neal J.; MACDONALD, Maryellen C. Individual differences and probabilistic constraints in syntactic ambiguity resolution. **Journal of Memory and Language**, v. 34, pp. 521-542, 1995.

PIÑAR, Pilar; CARLSON, Matthew T.; MORFORD, Jill P.; DUSSIAS, Paola E. Bilingual deaf readers' use of semantic and syntactic cues in the processing of English relative clauses. **Bilingualism: Language and Cognition**, v. 20, n. 5, pp. 980-998, 2017.

- RIBEIRO, Antônio João Carvalho. **Um caso de não aplicação de Late Closure no português do Brasil**. Manuscrito. Rio de Janeiro: UFRJ/FL, 1998.
- RICHTER, Carla; LARRÉ, Julia. **Take Action!** São Paulo: Ática, 2020.
- SACHS, Jacqueline Strunk. Recognition Memory for Syntactic and Semantic Aspects of Connected Discourse. **Perception & Psychophysics**, v. 2, n. 9, pp.437-442, 1967.
- SCHRIEFERS, Herbert; FRIEDERICI, Angela D.; KÜHN, Katja. The Processing of Locally Ambiguous Relative Clauses in German. **Journal of Memory and Language**, v. 34, n. 4, pp. 499-520, 1995.
- SCHWARTZ, Bonnie D.; SPROUSE, Rex A. L2 cognitive states and the Full Transfer/Full Access model. **Second Language Research**, v. 12, n. 1, pp. 40-72, 1996.
- SELINKER, Larry. Interlanguage. **International Review of Applied Linguistics in Language Teaching**, v. 10, n. 3, pp. 209-231, 1972.
- SHA, Long; HAXBY, James V.; ABDI, Herve; GUNTUPALLI, J. Swaroop; OOSTERHOF, Nikolaas N.; HALCHENKO, Yaroslav O.; CONNOLY, Andrew C. The Animacy Continuum in the Human Ventral Vision Pathway. **Journal of Cognitive Neuroscience**, v. 27, n. 4, pp. 665-678, 2015.
- SILVERSTEIN, Michael. Hierarchy of Features and Ergativity. In: DIXON, R. M. W. **Grammatical Categories in Australian Languages**, v. 22, pp. 112-171, 1976.
- SKALICKY, Stephen; CHEN, Victoria. Forward and backward transfer of sentence processing cues in English and Mandarin Chinese: A call for approximate replication of Liu, Bates, and Li (1992) and Su (2001). **Language Teaching**, pp. 1-9, 2020.
- SORLIN, Sandrine; GARDELLE, Laure. Anthropocentrism, egocentrism and the notion of Animacy Hierarchy. **International Journal of Language and Culture**, v. 5, n. 2, n.p., 2018.
- SOUZA, Elenice Santos de Assis Costa de. Cláusulas relativas: um caso de interface entre sintaxe e prosódia. **ReVEL**, v. 8, n. 15, 2010.
- SOUZA, Ricardo Augusto de; SOARES-SILVA, Jesiel. Exploring the measurement of vocabulary size to differentiate Brazilian Portuguese-English Bilinguals' access to grammatical knowledge in the L2. **Revista Linguística**, v. 11, n. 1, 2015.
- SOUZA, Ricardo Augusto de; DUARTE, Denise; BERG, Isadora Barreto. A avaliação de habilidades linguísticas em L2: uma questão metodológica em estudos de influências translinguísticas. In: BUCHWEITZ, A; MOTA, M. B. (Eds.) **Linguagem e Cognição: Processamento, Aquisição e Cérebro**. Porto Alegre: EdIPUCRS, 2015.
- TANENHAUS, Michael K. Sentence Processing. In: NADEL, Lynn (org.). **Encyclopedia of Cognitive Science**. Hoboken: Wiley, 2003.
- TAVAKOLI, Hossein. **A Dictionary of Language Acquisition: A Comprehensive Overview of Key Terms in First and Second Language Acquisition**. Teerã: Rahnama Press, 2012.
- TRAXLER, Matthew J.; CORINA, David P.; MORFORD, Jill P.; HAFER, Sarah; HOVERSTEN, Liv J. Deaf readers' response to syntactic complexity: Evidence from self-paced reading. **Memory & Cognition**, v. 42, n. 1, pp. 97-111, 2013.
- TRAXLER, Matthew J.; MORRIS, Robin K.; SEELY, Rachel E. Processing Subject and Object Relative Clauses: Evidence from Eye Movements. **Journal of Memory and Language**, v. 47, n. 1, pp. 69-90, 2002.

- TRAXLER, Matthew J. Trends in Syntactic Parsing: Anticipation, Bayesian Estimation, and Good-Enough Parsing. **Trends in Cognitive Science**, v. 18, n. 11, pp. 605-611, 2014.
- VANARSDALL, Joshua E.; NAIRNE, James S.; PANDEIRADA, Josefa N. S.; BLUNT, Janell R. Adaptive Memory: Animacy Processing Produces Mnemonic Advantages. **Experimental Psychology**, v. 60, n. 3, pp. 172-178, 2013.
- VIHMAN, Virve-Anneli; NELSON, Diane. Effects of Animacy in Grammar and Cognition: Introduction to Special Issue. **Open Linguistics**, v. 5, pp. 260-267, 2019.
- VIHMAN, Virve-Anneli; NELSON, Diane; KIRBY, Simon. Animacy Distinctions Arise from Iterated Learning. **Open Linguistics**, v. 4, pp.552-565, 2018.
- WANNER, Eric; MARATSOS, Michael. An ATN Approach in Comprehension. In: HALLE, Morris; BRESNAN, Joan; MILLER, George A. (Org.). **Linguistic Theory and Psychological Reality**. Cambridge, Mass.: MIT Press, 1978.
- WEIGEL, Adriana; RESCHKE, Tatiana. **English and More!** São Paulo: Richmond, 2020.
- WELLS, Justine B.; CHRISTIANSEN, Morten H.; RACE, David S.; ACHESON, Daniel J.; MACDONALD, Maryellen C. Experience and Sentence Processing: Statistical Learning and Relative Clause Comprehension. **Cognitive Psychology**, v. 58, n. 2, pp. 250-271, 2009.
- Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). <https://doi.org/10.17605/OSF.IO/MD832>
- ZUBIN, David A. Discourse Function of Morphology: The Focus System in German. In: GIVON, Thomas (Ed.). **Syntax and Semantics: Discourse and Syntax**. Nova York: Academic Press, 1979.

## ANEXO

### FRASES EXPERIMENTAIS

The farmer / that washed the tractor / remained / next to the barn. Did the farmer remain next to the barn?

The farmer / that the tractor dragged / remained / next to the barn. Did the tractor remain next to the barn?

The tractor / that dragged the farmer / remained / next to the barn. Did the tractor remain next to the barn?

The tractor / that the farmer washed / remained / next to the barn. Did the farmer remain next to the barn?

The plumber / that brought the hammer / was found / near the back door. Was the hammer found near the back door?

The plumber / that the hammer bruised / was found / near the back door. Was the plumber found near the back door?

The hammer / that bruised the plumber / was found / near the back door. Was the plumber found near the back door?

The hammer / that the plumber brought / was found / near the back door. Was the hammer found near the back door?

The burglar / that found the revolver / stayed / in the bedroom. Did the burglar stay in the bedroom?

The burglar / that the revolver shot / stayed / in the bedroom. Did the revolver stay in the bedroom?

The revolver / that shot the burglar / stayed / in the bedroom. Did the revolver stay in the bedroom?

The revolver / that the burglar found / stayed / in the bedroom. Did the burglar stay in the bedroom?

The senator / that reviewed the article / mentioned / the Constitution. Did the article mention the Constitution?

The senator / that the article accused / mentioned / the Constitution. Did the senator mention the Constitution?

The article / that accused the senator / mentioned / the Constitution. Did the senator mention the Constitution?

The article / that the senator reviewed / mentioned / the Constitution. Did the article mention the Constitution?

The actor / that bought the razor / appeared / in the horror movie. Did the actor appear in the horror movie?

The actor / that the razor shaved / appeared / in the horror movie. Did the razor appear in the horror movie?

The razor / that shaved the actor / appeared / in the horror movie. Did the razor appear in the horror movie?

The razor / that the actor bought / appeared / in the horror movie. Did the actor appear in the horror movie?

The woman / that started the accident / was shown / on TV last night. Was the accident shown on TV last night?

The woman / that the accident killed / was shown / on TV last night. Was the woman shown on TV last night?

The accident that killed the woman / was shown / on TV last night. Was the woman shown on TV last night?

The accident / that the woman started / was shown / on TV last night. Was the accident shown on TV last night?

The student / that attended the school / took part / in the competition. Did the student take part in the competition?

The student / that the school educated / took part / in the competition. Did the school take part in the competition?

The school / that educated the student / took part / in the competition. Did the school take part in the competition?

The school / that the student attended / took part / in the competition. Did the student take part in the competition?

The people / that boarded the train / attracted / a good deal of attention. Did the train attract attention?

The people / that the train carried / attracted / a good deal of attention. Did the people attract attention?

The train / that carried the people / attracted / a good deal of attention. Did the people attract attention?

The train / that the people boarded / attracted / a good deal of attention. Did the train attract attention?

The cattle / that destroyed the grass / vanished / in the blink of an eye. Did the cattle vanish in the blink of an eye?

The cattle / that the grass nourished / vanished / in the blink of an eye. Did the grass vanish in the blink of an eye?

The grass / that nourished the cattle / vanished / in the blink of an eye. Did the grass vanish in the blink of an eye?

The grass / that the cattle destroyed / vanished / in the blink of an eye. Did the cattle vanish in the blink of an eye?

The woman / that prepared the water / stayed / still in the bath tub. Did the water stay still in the bath tub?

The woman / that the water burned / stayed / still in the bath tub. Did the woman stay still in the bath tub?

The water / that burned the woman / stayed / still in the bath tub. Did the woman stay still in the bath tub?

The water / that the woman prepared / stayed / still in the bath tub. Did the water stay still in the bath tub?

The spies / that printed the notes / got lost / in the basement. Did the spies get lost in the basement?

The spies / that the notes disturbed / got lost / in the basement. Did the notes get lost in the basement?

The notes / that disturbed the spies / got lost / in the basement. Did the notes get lost in the basement?

The notes / that the spies printed / got lost / in the basement. Did the spies get lost in the basement?

The pilot / that displayed the tattoo / looked / uglier than sin. Did the tattoo look ugly?

The pilot / that the tattoo covered / looked / uglier than sin. Did the pilot look ugly?

The tattoo / that covered the pilot / looked / uglier than sin. Did the pilot look ugly?

The tattoo / that the pilot displayed / looked / uglier than sin. Did the tattoo look ugly?

The children / that found the fossil / became / very famous in France. Did the children become famous?

The children / that the fossil scared / became / very famous in France. Did the fossil become famous?

The fossil / that scared the children / became / very famous in France. Did the fossil become famous?

The fossil / that the children found / became / very famous in France. Did the children become famous?

The girls / that trashed the church / looked / very shabby at night. Did the church look shabby?

The girls / that the church housed / looked / very shabby at night. Did the girls look shabby?

The church / that housed the girls / looked / very shabby at night. Did the girls look shabby?

The church / that the girls trashed / looked / very shabby at night. Did the church look shabby?

The cowboy / that concealed the pistol / was known / to be unreliable. Was the cowboy known to be unreliable?

The cowboy / that the pistol injured / was known / to be unreliable. Was the pistol known to be unreliable?

The pistol / that injured the cowboy / was known / to be unreliable. Was the pistol known to be unreliable?

The pistol / that the cowboy concealed / was known / to be unreliable. Was the cowboy known to be unreliable?

The director / that watched the movie / received / a prize at the festival. Did the movie receive a prize at the festival?

The director / that the movie pleased / received / a prize at the festival. Did the director receive a prize at the festival?

The movie / that pleased the director / received / a prize at the festival. Did the director receive a prize at the festival?

The movie / that the director watched / received / a prize at the festival. Did the movie receive a prize at the festival?

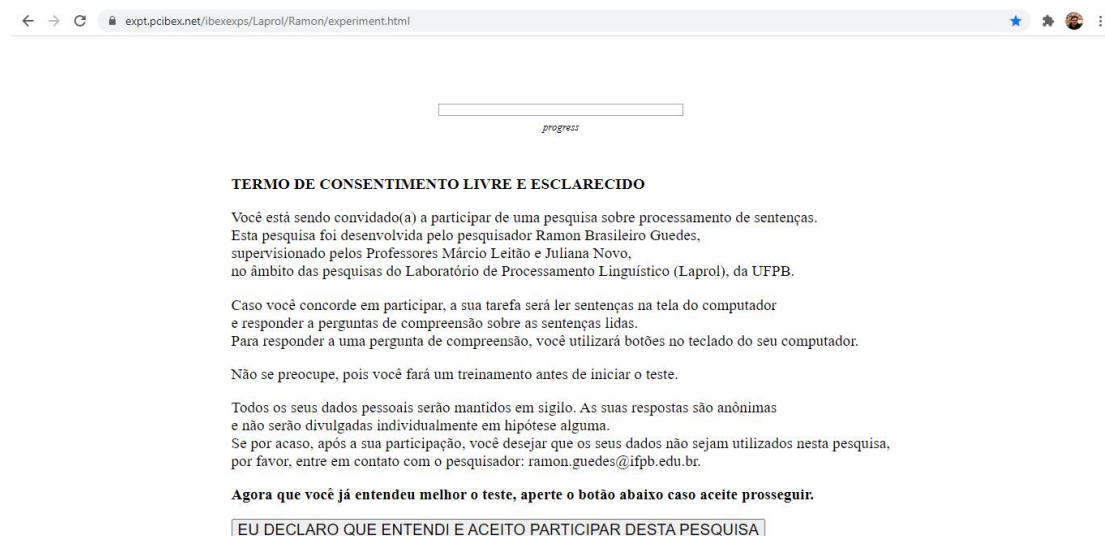
### FRASES DISTRATORAS

1. Monica's daughters / told me / that they enjoy / animated movies. Do Monica's daughters hate animated movies?
2. Morgan / assured me / that his wife is / a talented interior designer. Does Morgan think his wife is a good interior designer?
3. French people / who live abroad / always long / for their homeland. Do French people who live abroad want to go back home?
4. The lawyer / who lied in court / was arrested / for drunk driving. Was the lawyer fined?
5. The student / whose backpack was stolen / wanted to talk / to the principal. Did the student want to talk to the mayor?
6. Religious people / pray a lot / and go to church / every week. Do religious people go to church every day?
7. Robert / played football / with his friends / near the school. Did Robert play football with his friends?
8. Emily and Clara / never accepted / the divorce / of their parents. Were Emily and Clara unhappy about their parents' divorce?
9. The gardener / quit his job / because he decided / to become a chef. Does the gardener want to become a chef?
10. The teacher / whose students I met / moved / to another country. Did the students move to another country?
11. The old women / who showed up at noon / demanded / to talk to Claudia. Did the old women want to talk to Claudia?
12. Harry's friends / disappeared / before the police / could interrogate them. Were Harry's friends interrogated by the police?
13. Brandon / remained in the kitchen / in case / he was needed. Did Brandon remain in the courtyard?
14. The books / were criticized / for dealing with / a controversial subject. Were the books deemed controversial?
15. The birds / flew over / the villages / and the dense forest. Did the birds fly over the open sea?



16. All the politicians / promised / way more / than they could deliver. Did the politicians break their promises?
17. The snake / that slithered away / was venomous / and brightly coloured. Did the snake go away?
18. Lucas and Dylan / fought over / the box of crayons / that Lucas stole. Did Dylan steal a box of crayons?
19. Marta / often writes about / her childhood / in the south of Spain. Did Marta grow up in Spain?
20. Russians / think that people / shouldn't smile / for no reason. Do Russians admire people who smile for no reason?
21. All the boys / are very excited / about Tim's birthday party / next week. Are the boys looking forward to Tim's birthday party?
22. Cell phones / are extraordinary tools / but people / should use them less. Should people use cell phones more often?
23. Majestic tigers / roam free / among the mountains / of eastern Russia. Are there tigers in Russia?
24. Cats / often use their tongues / to lick themselves / clean. Do cats take showers?
25. People / must exercise / and eat healthily / to live longer. Should people eat more to live longer?
26. Matthew / must hide / his true self / from his family and friends. Is Matthew keeping a secret?
27. Romantic men / write poems / and read Goethe / all night long. Do romantic men read Nietzsche?
28. European castles / are visited / by multitudes of tourists / in the summer. Do tourists like European castles?
29. NASA engineers / want to send / an Ecuadorian iguana / to outer space. Do NASA engineers want to send a bird to space?
30. Housewives / often dream / of dangerous adventures / in the Wild West. Do housewives dream of adventures?
31. Rosa's boyfriend / showed her / how to swim / against the current. Can Rosa's boyfriend swim?
32. Argentinians / went to the streets / to demand / the president's resignation. Are Argentinians happy about their president?

## EXPERIMENTO



A screenshot of a web browser displaying a consent form. The browser's address bar shows the URL "expt.pcibex.net/ibexexps/Laprol/Ramon/experiment.html". At the top, there is a progress bar. The main heading is "TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO". The text explains that the participant is being invited to participate in a research project on sentence processing, supervised by Professors Márcio Leitão and Juliana Novo at the Linguistic Processing Laboratory (Laprol) of UFPB. It details the tasks: reading sentences on a computer and answering comprehension questions. It also states that personal data will be kept confidential and that the participant can contact the researcher at ramon.guedes@ifpb.edu.br. At the bottom, there is a button labeled "EU DECLARO QUE ENTENDI E ACEITO PARTICIPAR DESTA PESQUISA".

progress

### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado(a) a participar de uma pesquisa sobre processamento de sentenças. Esta pesquisa foi desenvolvida pelo pesquisador Ramon Brasileiro Guedes, supervisionado pelos Professores Márcio Leitão e Juliana Novo, no âmbito das pesquisas do Laboratório de Processamento Linguístico (Laprol), da UFPB.

Caso você concorde em participar, a sua tarefa será ler sentenças na tela do computador e responder a perguntas de compreensão sobre as sentenças lidas. Para responder a uma pergunta de compreensão, você utilizará botões no teclado do seu computador.

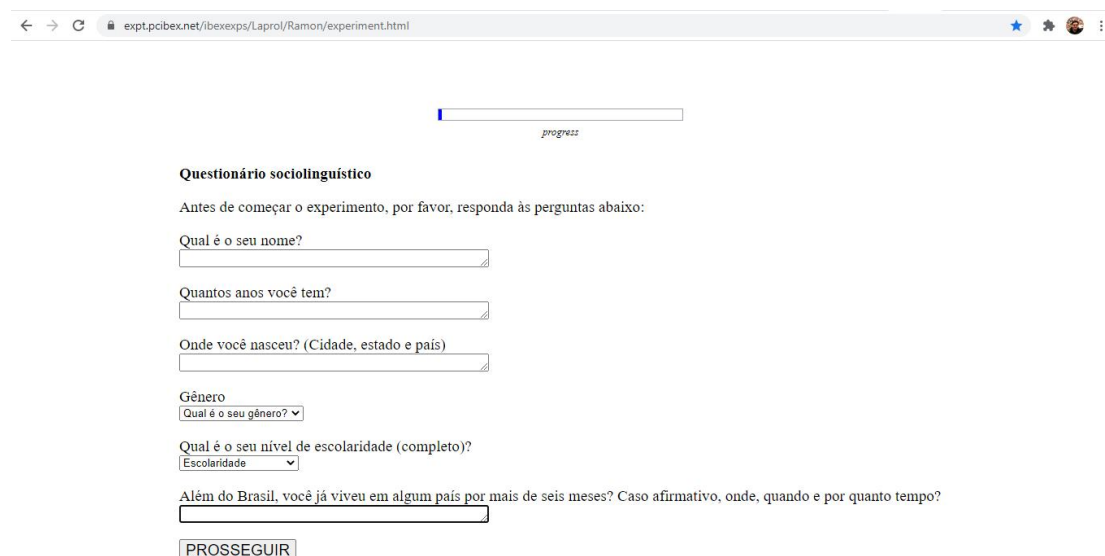
Não se preocupe, pois você fará um treinamento antes de iniciar o teste.

Todos os seus dados pessoais serão mantidos em sigilo. As suas respostas são anônimas e não serão divulgadas individualmente em hipótese alguma. Se por acaso, após a sua participação, você desejar que os seus dados não sejam utilizados nesta pesquisa, por favor, entre em contato com o pesquisador: ramon.guedes@ifpb.edu.br.

**Agora que você já entendeu melhor o teste, aperte o botão abaixo caso aceite prosseguir.**

EU DECLARO QUE ENTENDI E ACEITO PARTICIPAR DESTA PESQUISA

Termo de Consentimento Livre e Esclarecido apresentado ao participante antes do questionário sociolinguístico que, por sua vez, antecedia as instruções sobre o experimento. A plataforma utilizada é a PCibex Farm.



A screenshot of a web browser displaying a sociolinguistic questionnaire. The browser's address bar shows the URL "expt.pcibex.net/ibexexps/Laprol/Ramon/experiment.html". At the top, there is a progress bar. The heading is "Questionário sociolinguístico". The text asks the participant to answer the following questions before starting the experiment: "Qual é o seu nome?", "Quantos anos você tem?", "Onde você nasceu? (Cidade, estado e país)", "Gênero" (with a dropdown menu showing "Qual é o seu gênero?"), "Qual é o seu nível de escolaridade (completo)?" (with a dropdown menu showing "Escolaridade"), and "Além do Brasil, você já viveu em algum país por mais de seis meses? Caso afirmativo, onde, quando e por quanto tempo?". At the bottom, there is a button labeled "PROSSEGUIR".

progress

### Questionário sociolinguístico

Antes de começar o experimento, por favor, responda às perguntas abaixo:

Qual é o seu nome?

Quantos anos você tem?

Onde você nasceu? (Cidade, estado e país)

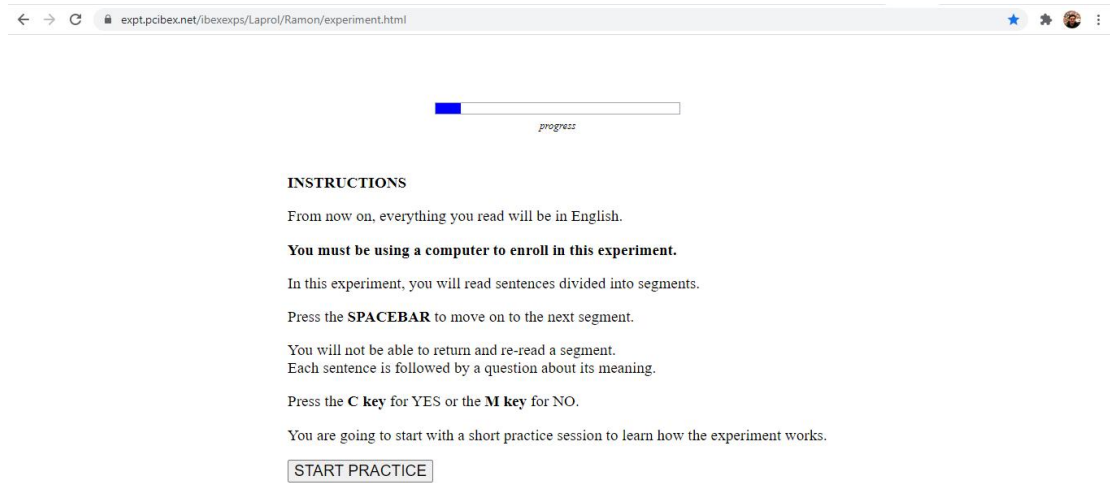
Gênero  
Qual é o seu gênero?

Qual é o seu nível de escolaridade (completo)?  
Escolaridade

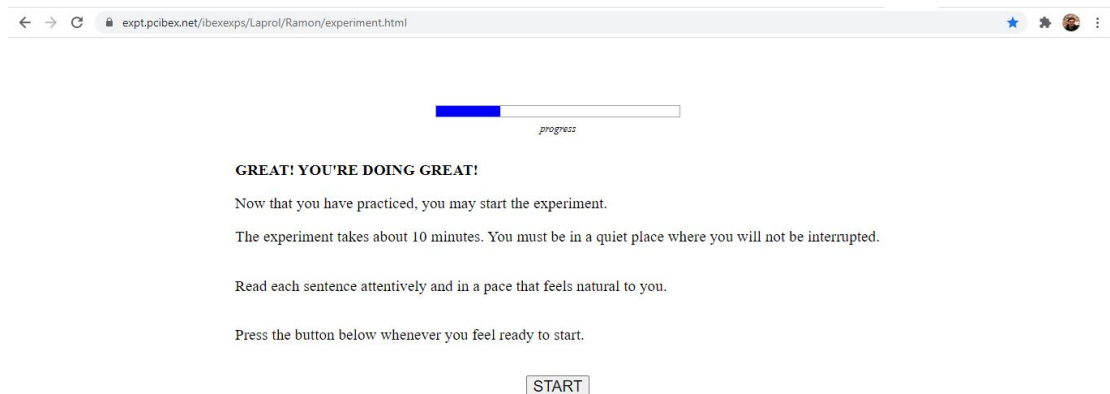
Além do Brasil, você já viveu em algum país por mais de seis meses? Caso afirmativo, onde, quando e por quanto tempo?

PROSSEGUIR

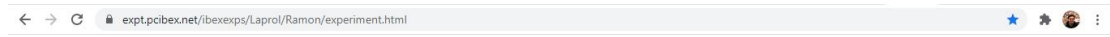
Primeira página do questionário sociolinguístico, respondido pelos participantes antes de lerem as instruções sobre o experimento. A íntegra do questionário pode ser conferida ao fim deste anexo.



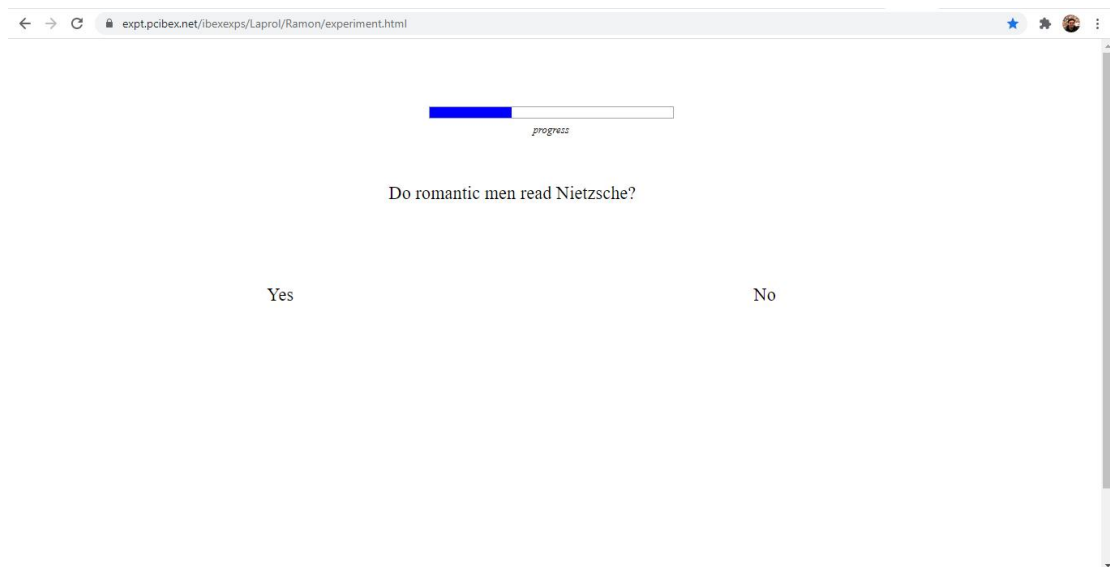
Página de instruções anterior à seção de prática que os participantes realizavam para se familiarizarem com o método de leitura automonitorada. A partir deste ponto, tudo que os sujeitos liam estava em inglês, para evitar que a troca entre línguas interferisse em sua performance no experimento.



Página de instruções posterior à seção prática e anterior ao experimento propriamente dito.



Experimento de leitura automonitorada, com exemplo de sentença escondida por traços. Para exibir cada segmento da sentença, o participante deveria aperta a tecla barra de espaço.



Exemplo de pergunta de compreensão feita após a leitura de uma sentença, neste caso uma frase distratora. Para responder “sim”, o participante deveria pressionar a tecla C. Em caso de resposta negativa, a tecla acionada era a M.

## QUESTIONÁRIO SOCIOLINGUÍSTICO

Antes de começar o experimento, por favor, responda às perguntas abaixo.

Qual é o seu nome?

Quantos anos você tem?

Onde você nasceu? (Cidade, estado e país)

Qual é o seu gênero?

Masculino

Feminino

Prefiro não responder

Outro

Qual é o seu nível de escolaridade (completo)?

Ensino Fundamental

Ensino Médio

Graduação

Especialização

Mestrado

Doutorado

Além do Brasil, você já viveu em algum país por mais de seis meses? Em caso afirmativo, onde, quando e por quanto tempo?

Qual(is) língua(s) você adquiriu antes de entrar na escola (com menos de seis anos)?

Quantos anos você tinha quando começou a adquiri-la(s)?

Em qual(is) língua(s) você foi escolarizado(a) (Ensino Fundamental e Médio)?

Qual(is) língua(s) você aprendeu em um ambiente educacional formal?

Qual(is) língua(s) você aprendeu fora de um ambiente educacional formal?

Em geral, como você classificaria sua proficiência em inglês?

Fala (produção oral)

Nenhuma

Péssima

Ruim

Suficiente

Boa

Excelente

Escuta (compreensão oral)

Nenhuma

Péssima

Ruim

Suficiente

Boa  
Excelente  
Proficiência geral  
Nenhuma  
Péssima  
Ruim  
Suficiente  
Boa  
Excelente

Com que frequência você fala inglês?

Nunca  
Raramente  
Poucas vezes ao ano  
Mensalmente  
Semanalmente  
Diariamente

Você se sente mais confortável falando inglês ou português?

Português  
Inglês  
Sem preferência

Você poderia explicar melhor a sua resposta anterior? Por que você se sente mais confortável usando o português ou o inglês? Ou, se for o seu caso, por que você não tem nenhuma preferência?

Indique na grade abaixo a frequência com que você usa inglês em cada contexto:

Com a família

Nunca  
Muito raramente  
Raramente  
Às vezes  
Frequentemente  
O tempo todo  
N/A

Com amigos

Nunca  
Muito raramente  
Raramente  
Às vezes  
Frequentemente  
O tempo todo  
N/A

Com animais de estimação

Nunca  
Muito raramente  
Raramente  
Às vezes  
Frequentemente

O tempo todo  
 N/A  
 No trabalho  
 Nunca  
 Muito raramente  
 Raramente  
 Às vezes  
 Frequentemente  
 O tempo todo  
 N/A  
 Na escola  
 Nunca  
 Muito raramente  
 Raramente  
 Às vezes  
 Frequentemente  
 O tempo todo  
 N/A  
 Na igreja  
 Nunca  
 Muito raramente  
 Raramente  
 Às vezes  
 Frequentemente  
 O tempo todo  
 N/A  
 Em lojas/supermercados  
 Nunca  
 Muito raramente  
 Raramente  
 Às vezes  
 Frequentemente  
 O tempo todo  
 N/A  
 Em clubes/organizações  
 Nunca  
 Muito raramente  
 Raramente  
 Às vezes  
 Frequentemente  
 O tempo todo  
 N/A  
 Você escuta músicas em inglês?  
 Sim  
 Não  
 Você assiste a filmes ou vídeos em inglês?  
 Sim  
 Não

Você escuta programas de rádio ou *podcasts* em inglês?

Sim

Não

Você lê livros, jornais, revistas ou blogs em inglês?

Sim

Não

Você se sente desconfortável falando em inglês com um falante nativo dessa língua?

Sim, sempre.

Sim, às vezes.

Não, nunca.

Você chegou ao fim do questionário. Há algo que gostaria de mencionar? Escreva qualquer comentário que você ache pertinente ou relacionado à pesquisa.