

Universidade Federal de São Carlos
Centro de Educação e Ciências Humanas
Programa de Pós-Graduação em Psicologia

Extração de palavras e mapeamento com estímulos visuais ao longo de tentativas ambíguas
em paralelo

Rodrigo Dal Ben de Souza

São Carlos – SP
2020

Universidade Federal de São Carlos
Centro de Educação e Ciências Humanas
Programa de Pós-Graduação em Psicologia

Extração de palavras e mapeamento com estímulos visuais ao longo de tentativas ambíguas
em paralelo

Rodrigo Dal Ben de Souza

Texto apresentado ao Programa de Pós-Graduação
em Psicologia da Universidade Federal de São
Carlos como requisito parcial para o título de
Doutor em Psicologia

Orientadora: Prof.^a Dr.^a Débora de Hollanda Souza
Coorientadora: Prof.^a Dr.^a Jessica S. F. Hay¹

São Carlos – SP
2020

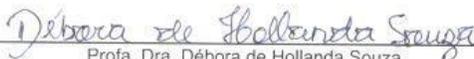
¹ The University of Tennessee, Knoxville, Estados Unidos da América.

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

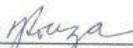
Centro de Educação e Ciências Humanas
Programa de Pós-Graduação em Psicologia

Folha de Aprovação

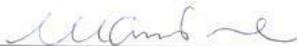
Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado do candidato Rodrigo Dal Ben de Souza, realizada em 10/03/2020:



Profa. Dra. Débora de Hollanda Souza
UFSCar



Profa. Dra. Deisy das Graças de Souza
UFSCar

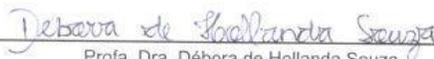


Profa. Dra. Maria Cristina Lobo Name
UFJF

Prof. Dr. Gustavo Lopes Estivalet
UFPB

Prof. Dr. João Henrique de Almeida
UEL

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Gustavo Lopes Estivalet, João Henrique de Almeida e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ao) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.



Profa. Dra. Débora de Hollanda Souza

Financiamento

O presente trabalho foi financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo por meio de uma bolsa de doutorado (processo nº 2015/26389-7) e de uma bolsa de estágio em pesquisa no exterior (processo nº 2018/04226-7).

Apoio

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), por meio do Programa de Excelência Acadêmica (PROEX), do qual faz parte o Programa de Pós-Graduação em Psicologia da Universidade Federal de São Carlos (PPGpsi/UFSCar).

O presente trabalho foi desenvolvido no âmbito do Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE; CNPq processo n° 465686/2014-1; FAPESP processo n° 2014/50909-8) que forneceu parte da infraestrutura necessária para a realização da pesquisa.

Agradecimentos

Durante a minha caminhada como estudante de Psicologia, eu sempre tive o privilégio de ter excelentes professores. Eu sou grato a todos eles. Não obstante, dois foram especialmente importantes. No primeiro ano de graduação, o professor João Juliani me mostrou a importância do pensamento sistemático, da parcimônia, e da honestidade intelectual. Na última etapa como estudante, a profa. Débora de Hollanda Souza me ensinou, principalmente pelo exemplo, que a habilidade mais valiosa de um pesquisador é estar sempre pronto a sair da zona de conforto e aprender coisas novas. Por todas as lições e pelo respeito que sempre recebi de ambos, eu sou muito grato.

Poder me dedicar aos estudos por tantos anos foi um imenso privilégio, mas o custo foi alto. Eu não teria conseguido trilhar esse caminho sem o apoio da minha família. Meus pais sempre estiveram prontos para ajudar e aconselhar. Inúmeras vezes eles abriram mão das suas atividades para me dar suporte, por isso eu sou muito grato.

Ao longo de todos esses anos, eu tive a enorme benção de ter uma companheira sábia e edificadora. Minha esposa abriu mão de tanto por mim que eu penso se um dia vou conseguir equilibrar a balança. Sem os momentos de alegria e tristeza que compartilhamos e o amor que cultivamos ao longo desses anos, todos os esforços até agora teriam sido em vão.

No princípio era o Verbo e o Verbo estava com Deus, e o Verbo era Deus. Ele estava no princípio com Deus. Todas as coisas foram feitas por intermédio dele, e, sem ele, nada do que foi feito se fez. A vida estava nele e a vida era a luz dos homens. A luz resplandece nas trevas, e as trevas não prevaleceram contra ela.

João 1:1-5

Souza, R. DB. (2020). *Extração de palavras e mapeamento com estímulos visuais ao longo de tentativas ambíguas em paralelo* (Tese de doutorado). Programa de Pós-Graduação em Psicologia, Universidade Federal de São Carlos, São Carlos, SP, Brasil. 186 pp.

Resumo

A aprendizagem de uma língua oral envolve vários desafios, entre eles: extrair palavras da fala contínua e mapeá-las a seus referentes. Uma fonte de informação importante para quem aprende uma língua são as regularidades presentes na própria língua oral (e.g., prosódia, probabilidade fonotática, probabilidade transicional). Há evidências de que a maior ou menor probabilidade de uma sílaba ser seguida por outra sílaba (probabilidade transicional) pode informar sobre os limites das palavras de uma fala contínua. Além disso, os processos de extração de palavras e mapeamento das mesmas aos seus referentes podem ocorrer sequencialmente ou simultaneamente, por meio do rastreamento da coocorrência entre as palavras e referentes. No entanto, há um número ainda restrito de estudos investigando processos simultâneos de extração de palavras com base na probabilidade transicional e o mapeamento com base em coocorrências ao longo de tentativas ambíguas (nas quais mais de uma palavra e mais de um referente são apresentados). A presente investigação é guiada pela pergunta: palavras são extraídas de falas contínuas e mapeadas a estímulos visuais simultaneamente ao longo de tentativas ambíguas? Uma sequência de três estudos foi desenvolvida para investigar essa questão. O primeiro estudo, com nove experimentos, investigou se pseudopalavras (previamente selecionadas) podiam ser extraídas de falas contínuas com base na probabilidade transicional de suas sílabas. Participaram adultos falantes do português brasileiro e do inglês norte-americano. Os resultados indicaram que as palavras podem ser extraídas com base na probabilidade transicional, porém outras dicas estatísticas, como a probabilidade fonotática das palavras, devem ser controladas. O segundo estudo, com dois experimentos, investigou se as palavras das línguas utilizadas no primeiro estudo, já extraídas, seriam mapeadas a figuras abstratas ao longo de tentativas ambíguas. Participaram adultos falantes do português brasileiro. Os resultados sugerem que as palavras podem ser mapeadas e que o mapeamento melhora quando há repetições da tarefa. Por fim, o terceiro estudo, com dois experimentos, investigou se os processos de extração e mapeamento ao longo de tentativas ambíguas podem ocorrer em paralelo. Participaram adultos falantes do português brasileiro e do inglês norte-americano. Os resultados indicam que tais processos podem ocorrer em paralelo e que a acurácia na tarefa em paralelo é superior à acurácia nas tarefas separadas (Estudo 1 e 2). Ao combinar a probabilidade transicional e a coocorrência de estímulos ao longo de tentativas ambíguas, a presente investigação contribui para o avanço do conhecimento atual sobre o papel das regularidades linguísticas na aprendizagem de linguagem. Ao mesmo tempo, o trabalho inaugurou essa linha de investigação em âmbito nacional, podendo motivar futuras investigações sobre o assunto.

Palavras-chave: processamento de linguagem, desenvolvimento de linguagem, aprendizagem de palavras, aprendizagem estatística

Souza, R. DB. (2020). *Word segmentation and cross-situational word learning in parallel* (Doctoral Dissertation). Graduate Program in Psychology, Universidade Federal de São Carlos, São Carlos, SP, Brazil. 186 pp.

Abstract

Learning an oral language presents several challenges, two of which are: extracting words from continuous speech and mapping them to referents. Statistical cues (e.g., prosody, phonotactical probability) are an important source of information to overcome such challenges. Evidence suggests that the higher or lower probability of a syllable being followed by another one (transitional probability) can be indicative of word boundaries in continuous speech. Additionally, segmentation and mapping processes can occur sequentially and simultaneously as a result of tracking the co-occurrence between words and its referents. Nonetheless, there is still a limited number of studies investigating simultaneous processes of word segmentation based on transitional probability and mapping based on cooccurrences across ambiguous trials (when more than one word and one referent is presented). The present investigation aims to answer the following question: can words be simultaneously segmented from continuous speech and mapped to referents across ambiguous trials? Three studies were conducted to investigate this question. The first study, with nine experiments, investigated whether pseudowords (previously selected) could be extracted from continuous speech based on their transitional probability. Brazilian-Portuguese and English adult speakers participated. Results indicated that words can be segmented, but only when other statistical cues, such as the phonotactical probabilities, are controlled. The second study, with two experiments, investigated whether segmented words from the languages used in the first study could be mapped to abstract pictures during cross-situational word learning tasks. Brazilian-Portuguese adult speakers participated. Results suggest that words can be mapped and that performance improves when the task is repeated. Finally, the third study, with two experiments, investigated whether extraction and mapping over ambiguous trials can occur in parallel. Brazilian-Portuguese and English adult speakers participated. Results suggest that these processes can occur together, at the same time, and that accuracy rates in the parallel task are higher than accuracy rates in individual tasks (Studies 1 and 2). By combining transitional probability and the cooccurrence of stimuli across ambiguous trials, the present investigation helps to advance the current understanding of the role played by linguistic regularities in language learning. Additionally, this work inaugurated this line of investigation in Brazil, which can motivate future investigations of the topic.

Keywords: language processing, language development, word learning, statistical learning

Sumário

Financiamento	iv
Apoio	v
Agradecimentos	vi
Resumo	viii
Abstract.....	ix
Lista de Tabelas	xiv
Lista de Figuras	xv
Apresentação	18
Capítulo 1	22
Descobrimdo palavras	22
Mapeando palavras	30
Pergunta de Pesquisa e Plano de Trabalho	37
Capítulo 2	42
Elementos comuns aos Estudos.....	42
Procedimentos Éticos	42
Questionário pós-experimental.....	43
Estímulos experimentais.....	43
Estímulos visuais	63
Análise de dados.....	67
Tamanho das amostras.....	67
Capítulo 3	70
Estudo 1 – Extração de Palavras.....	70
Contexto	70
Línguas de baixa familiaridade.....	71
Experimento 1	71

Método.....	72
Resultados e Discussão.....	75
Experimento 2	77
Método.....	78
Resultados e Discussão.....	78
Experimento 3	79
Método.....	80
Resultados e Discussão.....	80
Discussão Geral	81
Línguas de alta familiaridade	83
Experimento 4	83
Método.....	84
Resultados e Discussão.....	86
Experimento 5	89
Método.....	89
Resultados e Discussão.....	90
Experimento 6	92
Método.....	93
Resultados e Discussão.....	93
Experimento 7	95
Método.....	96
Resultados e Discussão.....	96
Experimento 8	98
Método.....	99
Resultados e Discussão.....	99

Experimento 9	101
Método.....	101
Resultados e Discussão.....	102
Discussão Geral	104
Capítulo 4	107
Estudo 2 – Mapeamento de palavras ao longo de tentativas ambíguas.....	107
Contexto	107
Experimento 1	108
Método.....	108
Resultados e Discussão.....	112
Experimento 2	114
Método.....	115
Resultados e Discussão.....	116
Discussão Geral	118
Capítulo 5	120
Estudo 3 – Extração e mapeamento ao longo de tentativas ambíguas em paralelo.....	120
Contexto	120
Experimento 1	121
Método.....	121
Resultados e Discussão.....	126
Experimento 2	129
Método.....	129
Resultados e Discussão.....	131
Capítulo 6	136
Considerações Finais	136

Referências	139
Apêndice A	164
Apêndice B	174
Apêndice C	176
Apêndice D	179
Apêndice E	180
Apêndice F	181
Apêndice G	182
Apêndice H	183
Apêndice I	184

Lista de Tabelas

Tabela 1 - <i>Língua Nativa dos participantes (LN), Nível de Familiaridade da Língua com a Língua Nativa (NF), Número da Língua (L), Probabilidade Transicional (PT), Frequência, Probabilidade Fonotática (Pbf), n de Coltheart, e Distância Ortográfica de Levenshtein (DOL20) das Palavras (P; IPA), Parte-palavras (PP; IPA), e Não-palavras (PP*; IPA)..</i>	45
Tabela 2 - <i>Número de Identificação na Base de Imagens NOUN (Id), Grau de Discriminabilidade (Disc, %), Novidade (Nov, %), e Saliência Agregada (Sal, %) para cada Estímulo Visual de cada Conjunto</i>	66
Tabela 3 - <i>Número de Identificação dos Pares (Id), Palavra, Figura, para a Língua 6 do português brasileiro</i>	122
Tabela 4 - <i>Número de Identificação dos Pares (Id), Palavra, Figura, para as Línguas 10A e 10B do inglês</i>	130

Lista de Figuras

<i>Figura 1.</i> Representação gráfica das probabilidades transicionais das sílabas intra-palavra (parte superior) e entre palavras (parte inferior). A fonte em itálico indica os limites das palavras.	24
<i>Figura 2.</i> Esquema das fases de treino e teste de Cunillera et al. (2010). No treino (A), são indicadas as figuras-alvo, as figuras não alvo, as relações com as palavras-alvo e não alvo e um exemplo da sequência da fala contínua. No teste de aprendizagem de palavras (B1), palavras-alvo eram comparadas com duas figuras-alvo ou palavras não alvo eram comparadas com figuras-alvo e figuras não alvo. No teste de segmentação (B2), uma palavra era comparada com uma parte-palavra (Cunillera et al., 2010, Fig.1, p. 297). Figura reproduzida com permissão da editora Elsevier.	32
<i>Figura 3.</i> Três tentativas seguindo o paradigma CSWL (2 x 2). As setas tracejadas indicam as possíveis relações entre os estímulos. Se o participante mapear a coocorrência entre palavras e figuras ao longo das tentativas, então, na terceira tentativa, as relações entre as palavras e as figuras podem ser estabelecidas (setas cheias).....	35
<i>Figura 4.</i> À esquerda, modelo sequencial de aprendizagem de palavras no qual as palavras são primeiro extraídas da fala contínua, formando um proto-léxico, e em seguida mapeadas com referentes. À direita, modelo unificado proposto por Räsänen e Rasilo (2015) no qual palavras são extraídas e mapeadas simultaneamente a partir da coocorrência de referentes com sons da fala (Räsänen & Rasilo, 2015, Fig. 2, p. 798). Figura reproduzida com permissão da <i>American Psychological Association</i>	39
<i>Figura 5.</i> Plano de Pesquisa (Estudos 1 a 3).	40
<i>Figura 6.</i> Processo de criação das palavras das línguas de baixa familiaridade do português brasileiro (esquerda) e alta familiaridade do português brasileiro (PB) e do inglês (direita).	48
<i>Figura 7.</i> Distribuições das probabilidades fonotáticas dos bifones CV, VC, CV na primeira, segunda e terceira posição das palavras do português brasileiro.	51
<i>Figura 8.</i> Pontuação média e intervalo de confiança (95%) para cada palavra (48) avaliada na tarefa de <i>wordlikeness</i>	56
<i>Figura 9.</i> Processo de gravação das falas naturais.	59

<i>Figura 10.</i> (A) Fase de familiarização com a fala contínua e (B) teste de escolha forçada entre palavra e parte-palavra, do Estudo 1.	74
<i>Figura 11.</i> Taxa geral de acerto para cada participante e média geral de acertos para os Experimentos 1, 2, e 3 do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, <i>bootstrap</i>) e a linha pontilhada indica o nível do acaso (0,5).	76
<i>Figura 12.</i> Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 3A, 3B, e 4 do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, <i>bootstrap</i>) e a linha pontilhada indica o nível do acaso (0,5).	87
<i>Figura 13.</i> Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 5 e 6 do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, <i>bootstrap</i>) e a linha pontilhada indica o nível do acaso (0,5).	91
<i>Figura 14.</i> Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 7A e 7B do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, <i>bootstrap</i>) e a linha pontilhada indica o nível do acaso (0,5).	94
<i>Figura 15.</i> Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 8A e 8B do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, <i>bootstrap</i>) e a linha pontilhada indica o nível do acaso (0,5).	97
<i>Figura 16.</i> Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 9A e 9B do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, <i>bootstrap</i>) e a linha pontilhada indica o nível do acaso (0,5).	100
<i>Figura 17.</i> Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 10A e 10B (Estudo 1) por item de comparação (não-palavra ou parte-palavra). As barras de erro indicam o intervalo de confiança (95%, <i>bootstrap</i>) e a linha pontilhada indica o nível do acaso (0,5).	103
<i>Figura 18.</i> (A) Fase de familiarização (2 x 2) e (B) Teste (escolha forçada entre 4 alternativas) do Estudo 2.	110
<i>Figura 19.</i> Taxa geral de acerto para cada participante e média geral de acertos para o Experimento 1 do Estudo 2. As barras de erro indicam o intervalo de confiança (95%, <i>bootstrap</i>) e a linha pontilhada indica o nível do acaso (0,25).	113
<i>Figura 20.</i> (A) Taxa geral de acerto para cada participante e média geral de acertos para cada bloco do Experimento 2 do Estudo 2; (B) Taxa geral de acerto para cada participante e	

média geral de acertos para cada bloco por local de coleta e ordem dos conjuntos. A ordem “1-2” indica que o Conjunto 1 (com estímulos com a probabilidade fonotática mais alta) era apresentado no primeiro bloco; o inverso é verdadeiro para a ordem “2-1”. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,25)..... 117

Figura 21. (A) Estrutura da fase de familiarização com a fala contínua e pares; (B) Estrutura do teste de extração de palavras; (C) Estrutura do teste de mapeamento do Estudo 3. 124

Figura 22. Taxa geral de acerto para cada participante e média geral de acertos para os testes de extração e mapeamento do Experimento 1 do Estudo 3. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e as linhas pontilhadas indicam os níveis do acaso (0,5 e 0,25)..... 126

Figura 23. Taxa geral de acerto para cada participante e média geral de acertos para o teste de extração para a Língua 6, Experimento 5 do Estudo 1 e para o Experimento 1 do Estudo 3; e taxa geral de acerto para cada participante e média geral de acertos para o teste de mapeamento para o Bloco 1 do Experimento 2 do Estudo 2 e para o Experimento 1 do Estudo 3. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e as linhas pontilhadas indicam os níveis do acaso (0,5 e 0,25). 128

Figura 24. Taxa geral de acerto para cada participante e média geral de acertos para os testes de extração e mapeamento para as Línguas 10A e 10B do Experimento 2 do Estudo 3. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e as linhas pontilhadas indicam os níveis do acaso (0,5 e 0,25). 132

Figura 25. Taxa geral de acerto para cada participante e média geral de acertos para os testes de extração das Línguas 10A e 10B no Experimento 9 (Estudo 1) e no Experimento 2 (Estudo 3). As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e as linhas pontilhadas indicam os níveis do acaso (0,5). 134

Apresentação

A aprendizagem de linguagem tem fascinado cientistas e filósofos por séculos e diversas propostas foram feitas sobre os possíveis processos e mecanismos envolvidos na superação desse desafio (Bloom, 2002; Ornat & Gallo, 2004). Nas últimas décadas, pesquisas em Psicologia do Desenvolvimento e Psicolinguística têm revelado como o conjunto da prontidão dos sistemas perceptivos logo no início da vida e a estimulação ambiental geram um desenvolvimento de linguagem rápido e encadeado (e.g., Choi, Black, & Werker, 2018; Jusczyk, 1999; Werker & Gervain, 2013; Werker & Hensch, 2015). Pesquisas demonstram, por exemplo, que antes mesmo do nascimento, entre a 38^a e 39^a semana de gestação, os batimentos cardíacos de fetos se alteram quando eles são expostos às vozes de suas mães comparadas com as vozes de estranhos ou de seus pais. Os batimentos também se alteram quando a sua língua nativa é comparada a uma língua estrangeira (e.g., inglês vs. mandarim, Kisilevsky et al. 2009). Logo após o nascimento, recém-nascidos continuam a discriminar vozes maternas de vozes estranhas (e.g., DeCasper & Fifer, 1980), e a sua língua nativa em comparação à línguas estrangeiras; medida pela alteração da taxa de sucção durante a tarefa experimental (e.g., inglês vs. espanhol: Moon, Cooper, & Fifer, 1993; inglês vs. sueco, Moon, Lagercrantz, & Kuhl, 2013). Além disso, recém-nascidos preferem qualquer língua a sons igualmente complexos (Vouloumanos, Hauser, Werker, & Martin, 2010) e, com base em diferenças rítmicas, são capazes de diferenciar línguas que nunca ouviram (e.g., francês vs. inglês vs. japonês; Nazzi, Bertoncini, & Mehler, 1998).

Somado a isso, desde o nascimento, bebês percebem contrastes fonéticos sutis (e.g., /ba/ vs. /da/, pronunciadas no inglês vs. hindi; Werker & Tees, 1984); aos 6 meses, usam dicas estatísticas para descobrir palavras em falas contínuas (Saffran, Aslin, & Newport,

1996) e começam a mapear as palavras aos seus referentes no mundo (Bergelson & Swingley, 2012; Shukla, White, & Aslin, 2011); aos 12 meses, identificam estruturas gramaticais (Gomez & Gerken, 1999).

Um dos construtos mais investigados nesse campo de estudos nos últimos anos é o de *aprendizagem estatística* (*statistical learning*). De maneira ampla, a aprendizagem estatística pode ser operacionalizada como a capacidade de rastrear e extrair regularidades presentes em estímulos sensoriais (Armstrong, Frost, & Christiansen, 2017). De maneira mais específica, e partindo da definição de aprendizagem como o processo pelo qual um comportamento novo passa a ser emitido e generalizado/abstraído (Catania, 2012), a aprendizagem estatística pode ser caracterizada como um processo cujas variáveis críticas para a emissão e generalização/abstração de comportamentos novos são regularidades ambientais. Por exemplo, ao rastrear probabilidades transicionais, aprendizes são capazes de extrair novas palavras de uma fala contínua, ou, a partir de distribuições fonéticas, eles são capazes de fazer discriminações fonéticas de novas línguas.

Após o estudo de Saffran, Aslin e Newport (1996), a aprendizagem estatística foi demonstrada com diferentes modalidades sensoriais (auditivo, visual, tátil; Bulf, Johnson, & Valenza, 2011; Conway & Christiansen, 2005; Endress & Mehler, 2009; Kirkham, Slemmer, & Johnson, 2002), com estímulos linguísticos e não linguísticos (Gebhart, Newport, & Aslin, 2009; Pelucchi, Hay, & Saffran, 2009b), com regularidades adjacentes e não adjacentes (Gebhart et al., 2009; Newport & Aslin, 2004), em tarefas que demandavam atenção explícita e implícita (Perruchet & Pacton, 2006; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997), com pessoas com desenvolvimento típico e atípico (Haebig, Saffran, & Ellis Weismer, 2017; McGregor, Rost, Arenas, Farris-Trimble, & Stiles, 2013), e com

animais não humanos de várias espécies (ratos, macacos, pássaros; Santolin & Saffran, 2018).

Esse conjunto de evidências gerou dois grandes desafios para a área. O primeiro envolve integrar a aprendizagem estatística com processos cognitivos e neurofisiológicos (e.g., percepção, categorização, generalização, maturação, aquisição de linguagem; Armstrong et al., 2017; Frost, Armstrong, Siegelman, & Christiansen, 2015; Siegelman & Frost, 2015). O segundo envolve integrar diferentes processos de aprendizagem estatística entre si (Johnson, 2016; Saffran, 2014; Smith, Suanda, & Yu, 2014).

O foco do presente trabalho é investigar a interação entre dois processos de aprendizagem estatística: 1) a extração de palavras de falas contínuas, definida como a identificação das fronteiras entre palavras (começo e fim das palavras) com base nas probabilidades sequenciais entre unidades sonoras (e.g., sílabas²) de um discurso contínuo (Cannistraci et al., 2019); e 2) o mapeamento de palavras ao longo de tentativas ambíguas, definido como a aprendizagem de relações simbólicas entre palavras e referentes (e.g., estímulos visuais ou objetos) com base em suas coocorrências ao longo de tentativas nas quais mais de uma palavra e mais de um referente são apresentados (Smith et al., 2014).

A pergunta de pesquisa que guia a presente investigação é: *Palavras são extraídas de falas contínuas e mapeadas a estímulos visuais simultaneamente ao longo de tentativas ambíguas?*

² Rigorosamente, as unidades sonoras utilizadas na presente investigação foram bifones (sequência entre dois fones). Fones se referem aos sons efetivamente produzidos na fala e se contrapõe à representação dos falantes de unidades sonoras de um sistema fonológico, fonema (para uma breve apresentação veja: <http://ceale.fae.ufmg.br/app/webroot/glossarioceale/verbetes/fone> , acessado em 20 de janeiro de 2020). Porém, no presente texto, os termos sílaba e bifone são utilizados como equivalentes, por duas razões. Primeiro porque a literatura da área utiliza os dois termos de maneira equivalente. Segundo, assume-se que os possíveis leitores desse trabalho possuirão maior familiaridade com o termo “sílabas”.

Buscar responder tal pergunta avança a investigação da integração de processos de aprendizagem estatística e contribui para a solução do segundo desafio mencionado anteriormente. Além disso, a integração desses processos é altamente relevante para o estudo e explicação dos processos de aprendizagem de linguagem nativa ou não nativa. Entretanto, vale ressaltar que, se por um lado, esforços de integração teórico-conceitual entre pesquisas empíricas de aprendizagem estatística e teorias de aprendizagem de linguagem são altamente valiosos e desejáveis (e.g., Armstrong et al., 2017; Bates & Elman, 1996; Schoneberger, 2010; Yurovsky & Frank, 2015), por outro, tal empreitada está para além do escopo do presente trabalho. Não obstante, esforços futuros de integração teórico-conceitual poderão se beneficiar dos dados aqui apresentados.

O texto se divide em cinco capítulos que contextualizam e testam a pergunta de pesquisa. No primeiro capítulo, a literatura pertinente aos processos de extração e de mapeamento por meio de regularidades estatísticas é revisada. Em seguida, são apresentados as perguntas de pesquisa e o plano de trabalho. No segundo capítulo, são apresentados os elementos comuns aos estudos (i.e., procedimentos éticos, seleção de participantes, questionário pós-experimental, estímulos experimentais, e análise de dados). Do terceiro ao quinto capítulo, os três estudos são descritos e no capítulo 6, são apresentadas as considerações finais.

Capítulo 1

Descobrimo palavras

Um dos desafios para bebês aprendendo a sua língua nativa e para adultos aprendendo uma segunda língua é descobrir quais são as palavras dessa língua.

Diferentemente da escrita, a fala não apresenta pausas consistentes que indicam o começo e o fim das palavras e poucas delas são faladas em isolamento ($\approx 10\%$, Brent & Siskind, 2001).

Felizmente, várias dicas probabilísticas podem atuar em conjunto para auxiliar o aprendiz a superar esse desafio (Johnson, 2016; Jusczyk, 2002). Por exemplo, padrões prosódicos como a entonação mais forte no começo ou no final das palavras (trocaico, iâmbico; Houston, Santelmann, & Jusczyk, 2004; Nazzi, Iakimova, Bertoncini, Fredones, & Alcantara, 2006; Shukla, White, & Aslin, 2011; Thiessen & Saffran, 2003), co-articulação entre sílabas (Johnson & Jusczyk, 2001), regularidades fonotáticas, como a não ocorrência de algumas sílabas em determinadas posições das palavras (e.g., *lha* começando palavras no português brasileiro, com exceção da palavra emprestada, *lhama*; Blanchard, Heinz, & Golinkoff, 2010; Finn & Hudson Kam, 2008; Mattys & Jusczyk, 2001; Mersad & Nazzi, 2011), variações na pronúncia das mesmas sílabas (alofonia; Jusczyk, Hohne, & Bauman, 1999), estrutura silábica (Saffran & Thiessen, 2003), final de sentenças (Johnson, Seidl, & Tyler, 2014; Jusczyk & Aslin, 1995), e a presença de palavras já aprendidas no discurso (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Cunillera, Càmarà, Laine, & Rodríguez-Fornells, 2010; Cunillera, Laine, & Antoni, 2016; Mersad & Nazzi, 2012).

No entanto, tais dicas variam para cada língua e, portanto, exigem conhecimento linguístico específico para serem utilizadas na aprendizagem de linguagem. Então, como

um aprendiz que possui pouco conhecimento linguístico específico sobre uma língua começa a descobrir suas palavras e seus padrões probabilísticos?

Uma hipótese que ganha cada vez mais suporte empírico é a de que estatísticas sequenciais entre sílabas, em níveis intra e entre palavras, possibilitam a extração de palavras de falas contínuas (Saffran et al., 1996; Aslin, Saffran, & Newport, 1998). Tais estatísticas condicionais, chamadas de probabilidade transicional (PT), podem ser calculadas para sequências silábicas (e.g., sílabas XY) independentemente de qualquer conhecimento específico de uma dada língua, seguindo as fórmulas:

$$\text{PT para frente} = P(Y | X) = \frac{\text{frequência}(XY)}{\text{frequência}(X)}$$

$$\text{PT para trás} = P(X | Y) = \frac{\text{frequência}(XY)}{\text{frequência}(Y)}$$

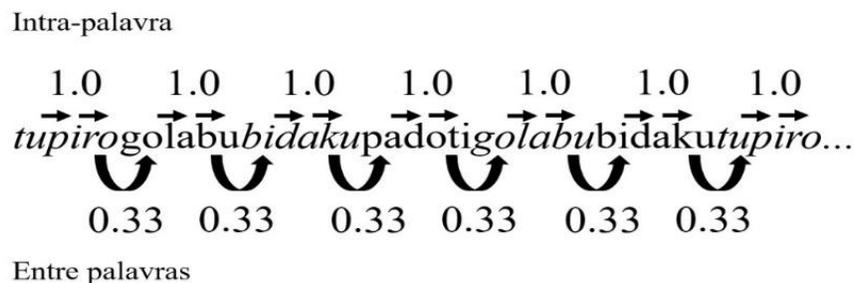
A PT para frente, primeira fórmula, indica a probabilidade da sílaba X ser seguida da sílaba Y. Por outro lado, a PT para trás, segunda fórmula, indica a probabilidade da sílaba Y ser antecedida da sílaba X. Por exemplo, na frase “o rato rolou a capa roxa do rei de Roma”, a probabilidade da sílaba “ro” ser seguida da sílaba “ma” é de 33%; já a probabilidade de “ra” ser antecedido da sílaba “o” é de 100%.

Suporte para a hipótese de que a PT entre sílabas, para frente e para trás, contribui para a aprendizagem de linguagem vem de análises de corpora³ naturais utilizando simulações computacionais (e.g., Swingley, 1999, 2005; mas veja Yang, 2004) e

³ Conjuntos de textos escritos ou registros orais em uma determinada língua. Para um exemplo detalhado das características de um corpus (singular) do Português-Brasileiro, veja Estivalet e Meunier (2015).

demonstrações empíricas. Considerando a natureza da presente investigação, as principais demonstrações empíricas serão apresentadas a seguir (para uma introdução sobre as investigações com simulações computacionais, Räsänen & Rasilo, 2015).

Em um estudo engenhoso, Saffran et al. (1996a) demonstraram que bebês com 8 meses de idade utilizam a PT para extrair palavras de falas contínuas. Quatro palavras trissílabas inventadas (e.g., tupiro, golabu) foram combinadas por 45 vezes para formar uma fala contínua de aproximadamente dois minutos. A PT, para frente, das sílabas intra-palavras era de 1 (sempre coocorriam) enquanto a PT das sílabas entre palavras era de 0,33 (Figura 1). A língua foi gravada usando um sintetizador que padronizou a duração, a frequência fundamental (F0), e a coarticulação das sílabas.



direcionaram a percepção dos bebês para a identificação e generalização de padrões que são incomuns na sua língua materna (Name, 2012).

Os achados de Saffran et al. (1996a) trouxeram uma nova perspectiva para o longo diálogo sobre a aquisição da linguagem e sua importância foi prontamente reconhecida (Bates & Elman, 1996). Nos anos que se seguiram, o papel da PT na extração de palavras de falas contínuas foi investigado em inúmeras replicações e extensões (para revisões, Krogh, Vlach, & Johnson, 2013; Romberg & Saffran, 2010; para uma meta-análise, Black & Bergmann, 2017). Por exemplo, Saffran, Newport, e Aslin (1996b) demonstraram que adultos também são capazes de extrair palavras de falas contínuas com base nas PTs entre sílabas, embora eles necessitem de um número maior de repetições de cada palavra em relação aos bebês (300 vs. 45) – provavelmente devido às extensas experiências linguísticas pré-experimentais.

Saffran et al. (1997) modificaram ligeiramente o procedimento das pesquisas anteriores e demonstraram que a extração por meio de PT independe de atenção explícita. No início do experimento, os participantes foram informados de que a pesquisa buscava medir a influência de estímulos auditivos na criatividade e foram instruídos a desenhar livremente em um software (i.e., KidPix) enquanto ouviam uma nova língua. Apesar da atividade distratora durante a familiarização com a língua, no teste, palavras foram selecionadas acima do esperado pelo acaso quando comparadas com parte-palavras e não-palavras. Pesquisas subsequentes relataram um maior nível de atividade cerebral em tarefas explícitas de extração, nas quais os participantes foram instruídos a prestar atenção na fala contínua, em comparação com tarefas implícitas, nas quais os participantes jogavam jogos distratores durante a fala contínua, porém não foram encontradas diferenças significativas na extração final, reforçando o argumento de que o rastreamento de PTs se desenvolve

implicitamente (Arciuli, Torkildsen, Stevens, & Simpson, 2014; Batterink, Reber, & Paller, 2015; Perruchet & Pacton, 2006).

Aslin et al. (1998) aumentaram a confiabilidade do papel da PT na extração da fala ao balancear a frequência de *palavras e parte-palavras* nas falas utilizadas por pesquisas anteriores. Dado que as parte-palavras são formadas pela junção de duas palavras, ao repetir as palavras um mesmo número de vezes (como em Saffran et al., 1996a e 1996b), uma fala cuja frequência das palavras é aproximadamente o dobro da frequência das parte-palavras é gerada. Tal diferença possibilita a extração com base na frequência e não na PT. Para resolver esse problema, os autores criaram uma fala balanceada cuja frequência de metade das palavras foi dobrada em relação à outra metade (90 vs. 45), gerando uma língua na qual a frequência das parte-palavras formadas a partir das palavras mais frequentes era igual à frequência das palavras menos frequentes (i.e., 45 vezes). Após serem familiarizados com a fala contínua equilibrada, bebês de 8 meses de idade preferiram palavras de baixa frequência às parte-palavras de alta frequência (ambas com 45 repetições), reforçando o papel da PT na extração das palavras.

As evidências descritas até o momento mostram como a probabilidade de uma sílaba X ser seguida da sílaba Y influencia a extração de palavras (PT para frente). No entanto, a probabilidade de uma sílaba Y ser precedida de uma sílaba X, PT para trás, também influencia a extração. Intuitivamente, a capacidade de uma sílaba prever a seguinte parece ser mais informativa para a extração de palavras do que o contrário. No entanto, análises de corpora naturais demonstraram que a combinação entre TP para frente e para trás (informação mútua) dobra o poder de extração de palavras em comparação com o uso de PT para frente apenas (de $\approx 23\%$ para $\approx 50\%$, Swingley, 1999, 2005). Nas línguas naturais, a PT para trás pode ser especialmente informativa na extração de substantivos. Por

exemplo, a probabilidade das palavras *cadeira* ou *bola* ou *caixa* serem antecidas pelo artigo *a* é maior do que o contrário (Willits, Seidenberg, & Saffran, 2009).

Seguindo essa linha, Perruchet e Desaulty (2008) demonstraram que a PT para trás é tão informativa quando a PT para frente na extração de palavras. Eles dividiram adultos universitários em dois grupos. Um dos grupos foi exposto a palavras cuja PT para frente era informativa (1 vs. 0,33, respectivamente) enquanto a PT para trás era mantida estável (0,33), não dando dicas sobre os limites das palavras. Para o outro grupo, as probabilidades foram invertidas. Os dois grupos foram capazes de extrair as palavras, sem diferença significativa entre eles. Esses resultados foram replicados com bebês de 8 meses de idade, mesmo quando a PT para trás foi apresentada em uma língua natural (Pelucchi, Hay, & Saffran, 2009a).

O uso de línguas naturais aumenta a validade ecológica dos estudos em aprendizagem estatística, com o custo de abrir mão de parte do controle experimental. Por exemplo, Pelucchi et al. (2009b) manipularam dicas estatísticas presentes em uma língua natural desconhecida dos participantes: o italiano. A língua era composta por 15 frases gramaticalmente e semanticamente corretas. Ela foi gravada por uma falante nativa do italiano, e todas as características prosódicas foram preservadas. Ao longo da língua ocorriam dois tipos de palavras-alvo. As palavras com alta PT não compartilhavam suas sílabas com nenhuma outra palavra da fala (PT = 1); enquanto as palavras com baixa PT compartilhavam suas sílabas iniciais com outras palavras, resultando numa PT para frente de 0,33. Após serem expostos à língua, bebês de 8 meses de idade preferiram palavras com alta PT (1) às palavras com baixa PT (0,33) e às não-palavras (não apresentadas na fala, PT = 0), demonstrando que a PT entre sílabas permite a extração de palavras de falas naturais.

O mesmo é verdadeiro quando a PT para trás é a única dica presente (Pelucchi et al., 2009a).

Embora a natureza geral da PT a torne uma dica eficaz para a extração inicial de palavras das falas contínuas, ela provavelmente interage com dicas específicas durante o processo de aprendizagem de uma dada língua (e.g., Mersad & Nazzi, 2011; Saksida, Langus, & Nespore, 2017; Yang, 2004). Por exemplo, a entonação típica da fala direcionada ao bebê (*motherese* ou *manhês*) facilita a extração de palavras de discursos contínuos por bebês de 7 meses de idade (Thiessen, Hill, & Saffran, 2005). Além disso, Thiessen e Saffran (2003) demonstraram que, embora bebês de 7 meses de idade utilizem mais PT do que dicas prosódicas (específicas) na extração de palavras, aos 9 meses, esse padrão se inverte (cf. Johnson & Jusczyk, 2001a; Johnson & Seidl, 2009; Thiessen & Saffran, 2007). O ritmo também interage com a PT. Saffran et al. (1996b) demonstraram que, quando a duração da última sílaba das palavras era aumentada, mimetizando o padrão da língua nativa dos participantes (inglês americano), a precisão da extração aumentava (dados similares foram relatados por Shatzman & McQueen, 2006).

A presença de palavras conhecidas junto com palavras desconhecidas em uma fala contínua é outra dica que facilita a extração de palavras, o que é conhecido como um efeito âncora. Cunillera, Càmaras, Laine, e Rodríguez-Fornells (2010) familiarizaram um grupo de adultos com duas pseudopalavras em isolamento e em seguida, utilizando um procedimento similar ao de Saffran et al. (1996), os expuseram a uma fala contínua formada pelas duas pseudopalavras aprendidas em isolamento (âncoras) e por outras seis pseudopalavras novas. Outro grupo de participantes foi exposto diretamente à fala contínua, sem familiarização prévia com as pseudopalavras em isolamento. Quando testados em tentativas de escolha forçada entre duas alternativas, o grupo âncora escolheu palavras mais acuradamente do

que o grupo sem familiarização prévia com as palavras em isolamento. Pesquisas subsequentes identificaram correlatos neurais antecipatórios relacionados ao efeito âncora (Cunillera, Laine, & Antoni, 2016) e também demonstraram que o efeito independe das características estatísticas do estímulo, ocorrendo também quando *parte-palavras* (PT = 0,5) são apresentadas em isolamento antes da fase de familiarização com a fala contínua (Poulin-Charronnat, Perruchet, Tillmann, & Peereman, 2017). Além disso, Mersad e Nazzi (2012) demonstraram que bebês extraem pseudopalavras com diferentes números de sílabas (di e trissílabas) de uma fala contínua quando palavras existentes em sua língua materna (i.e., *maman*, francês) estão presentes na fala. No entanto, quando as palavras conhecidas foram substituídas por pseudopalavras desconhecidas, a extração de pseudopalavras com diferentes números de sílabas não diferiu do esperado pelo acaso (replicando Johnson & Tyler, 2010).

Por fim, embora a frequência de palavras em isolamento em falas naturais seja baixa (Brent & Siskind, 2001), sua ocorrência ao longo de falas contínuas é outra dica que facilita a extração (Lew-Williams, Pelucchi, & Saffran, 2011). Além disso, elas oferecem suporte para a memória de longo prazo de palavras recém extraídas (Karaman & Hay, 2018).

Investigações com pessoas com desenvolvimento atípico também têm demonstrado que a capacidade de rastrear PT pode permanecer preservada em algumas populações (i.e., Síndrome de Williams e Transtorno do Espectro Autista; Cashon, Ha, Graf Estes, Saffran, & Mervis, 2016; Mayo & Eigsti, 2012; cf., Kovelman et al., 2015), mas não em outras (Distúrbio Específico da Linguagem; Evans, Saffran, & Robe-Torres, 2009; para uma revisão, veja, Lammertink, Boersma, Wijnen, & Rispen, 2017).

Extraír palavras de falas contínuas é apenas um dos desafios para quem aprende uma língua. A próxima seção abordará como dicas estatísticas possibilitam o mapeamento de palavras com referentes.

Mapeando palavras

Além de descobrir as sequências sonoras que formam palavras, outro desafio para quem aprende uma língua é mapear tais palavras com seus referentes (Saffran, 2014). Ainda antes do primeiro aniversário, entre 6 e 8 meses de idade, bebês já conhecem o significado de palavras concretas (i.e., partes do corpo e alimentos, Bergelson & Swingley, 2012; Tincoff & Jusczyk, 2012), demonstrando atenção à sequências sonoras que formam bons candidatos às palavras e aos seus possíveis referentes.

A atenção às palavras em potencial pode ser guiada tanto por dicas específicas da língua à qual o aprendiz está exposto (e.g., probabilidade fonotática, densidade de vizinhança e dicas prosódicas; Gonzalez-Gomez, Poltrock, & Nazzi, 2013; Graf Estes & Bowen, 2013; MacKenzie, Curtin, & Graham, 2012; Storkel, Bontempo, Aschenbrenner, Maekawa, & Lee, 2013), quanto por dicas gerais como a PT entre sílabas. Por exemplo, Estes, Evans, Alibali, e Saffran (2007) familiarizam bebês de 17 meses de idade com uma fala contínua formada por quatro palavras dissílabas. Em seguida, os participantes foram divididos em três grupos. O primeiro grupo foi exposto a palavras (PT = 1) pareadas com figuras geométricas ao longo de tentativas não ambíguas (apenas uma palavra pareada com uma figura). O segundo grupo foi exposto ao mesmo treino, mas com parte-palavras (PT = 0,33); e o terceiro com não-palavras (PT = 0). Na fase de teste, eram apresentadas tentativas iguais e trocadas. Nas tentativas iguais, o mesmo referente e estímulo auditivo da fase de aprendizagem das palavras eram apresentados (e.g., palavra A com referente A).

Nas tentativas trocadas, os estímulos auditivos e os referentes eram trocados entre si (e.g., palavra A com o referente B). Apesar do mesmo número de pareamentos para todos os grupos, apenas os participantes do primeiro grupo olharam por mais tempo para os referentes durante as tentativas trocadas, indicando que a PT dos estímulos auditivos influencia o mapeamento de palavras com referentes visuais. Um padrão de resultados semelhante foi encontrado com línguas naturais (Hay, Pelucchi, Estes, & Saffran, 2011).

Em um arranjo similar, Mirman, Magnuson, Estes, e Dixon (2008) demonstraram que adultos mapeiam palavras ($PT = 1$) a referentes mais rapidamente do que parte-palavras ($PT = 0,33$). Após a familiarização com a fala contínua, duas figuras geométricas e uma palavra ou parte-palavra eram apresentadas. A seleção de uma figura, inicialmente ao acaso, recebia *feedback* sobre o acerto ou erro. O número de tentativas necessário para mapear as palavras ($PT = 1$) foi menor em comparação às parte-palavras ($PT = 0,33$).

Tais demonstrações, de que palavras extraídas com base em probabilidade transicional são preferidas no estabelecimento de relações com estímulos visuais em tarefas subsequentes, levantam uma questão sobre a natureza do processo de extração e mapeamento de palavras em situações não ambíguas: se ele é necessariamente sequencial. Cunillera, Laine, Càmara e Rodríguez-Fornells (2010) familiarizaram adultos com uma fala contínua, composta por 6 palavras trissílabas ($PT = 1$), ao mesmo tempo em que figuras eram apresentadas na tela em sincronia com as palavras (Figura 2, A). Três palavras-alvo foram pareadas consistentemente com objetos familiares (e.g., *makusi* + óculos) e as outras três foram pareadas inconsistentemente (não alvos, e.g., *kiture* + formiga e cenoura). No teste (Figura 2, B), tanto as palavras-alvo quanto às palavras não alvo foram extraídas e mapeadas às figuras correspondentes, demonstrando que palavras podem ser extraídas e mapeadas simultaneamente ao longo de tentativas não ambíguas. Tais resultados foram

Na mesma linha, Shukla et al. (2011) demonstraram que bebês com 6 meses de idade extraem palavras de falas contínuas e as mapeiam a referentes visuais simultaneamente quando dicas prosódicas são combinadas com dicas estatísticas. A cada tentativa, eram apresentadas duas frases faladas compostas por cinco sílabas. Apenas duas sílabas coocorriam consistentemente ($PT = 1$), formando as palavras. As demais sílabas variavam de posição entre as frases. Para um dos grupos, tal palavra foi apresentada com entonação e ritmo típico da língua nativa dos participantes (inglês americano). Para o outro grupo, as mesmas dicas prosódicas foram distribuídas ao longo das sílabas, sem alinhamento com a palavra-alvo. Ao mesmo tempo em que as frases eram faladas, três figuras geométricas, uma alvo e duas distratoras, eram apresentadas. No entanto, as duas figuras distratoras permaneciam estáticas enquanto a figura-alvo se movimentava de um lado para o outro e crescia e diminuía em tamanho (diminuindo a ambiguidade). Apenas os participantes do grupo para o qual as dicas prosódicas foram alinhadas com palavras-alvo relacionaram-nas com as figuras-alvo. Os autores chamam a atenção para a pronta integração de dicas prosódicas e estatísticas logo no início da aprendizagem de uma língua. Além disso, eles também chamam a atenção para a natureza ambígua do contexto no qual tais processos ocorrem (Quine, 1960).

A ambiguidade na aprendizagem de palavras pode ser diminuída logo no primeiro encontro com uma palavra nova com base em dicas atencionais, sociais, linguísticas, e representacionais (*fast mapping*; Bloom, 2002; Souza, 2008; ou responder por exclusão, Dixon, 1977⁴). Além disso, pesquisas têm demonstrado que a coocorrência entre estímulos

⁴ Vale ressaltar que as relações entre estímulos auditivos e visuais são estudadas por outras linhas de pesquisa da Psicologia, usualmente com diferentes nomenclaturas (e.g., de Rose, Gil, & Souza, 2014). Embora o esforço de integração conceitual seja altamente valioso e relevante, ele não é o foco do presente trabalho e, portanto, não será abordado.

auditivos e visuais pode resolver a ambiguidade ao longo de tentativas nas quais mais de uma palavra é apresentada com mais de uma figura. A investigação desse processo por Yu e Smith (2007) inaugurou a linha de pesquisa que ficou conhecida como *cross-situational word learning*⁵ (CSWL) ou mapeamento ao longo de tentativas ambíguas⁶. O paradigma CSWL busca investigar se as relações entre palavras e referentes podem ser aprendidas por meio de coocorrências ao longo de tentativas individualmente ambíguas⁷ (Figura 3). No estudo inaugural, Yu e Smith testaram se adultos aprenderiam as relações entre 18 pares de palavras e figuras por meio da coocorrência entre palavras e referentes. Durante o treino, os participantes eram expostos a dois, três, ou quatro pares por tentativa (grupos 2 x 2, 3 x 3, 4 x 4). Cada par foi apresentado seis vezes, combinado aleatoriamente com os demais. No teste, uma palavra era tocada e os participantes escolhiam uma de quatro figuras (uma figura-alvo e três distratoras). Todos os grupos mapearam os pares acima do esperado pelo acaso, indicando que a aprendizagem das relações a partir de coocorrências, segundo o paradigma CSWL, pode ocorrer em diferentes níveis de ambiguidade.

⁵ Variações do termo são encontradas na literatura: *cross-situational learning*, *cross-situational language learning*, *cross-situational statistical learning*. No presente texto, utilizaremos *cross-situational word learning* (CSWL).

⁶ Tradução dos autores.

⁷ Os mecanismos envolvidos nesse processo são tema de constantes debates na literatura. Alguns pesquisadores defendem que as relações são aprendidas de maneira associativa e cumulativa ao longo das tentativas ambíguas e outros defendem que a cada tentativa o aprendiz elabora hipóteses sobre as possíveis relações e as testam nas tentativas seguintes, confirmando-as ou descartando-as (Trueswell, Medina, Hafri, & Gleitman, 2013). Discussões recentes apontam para um uso conjunto de tais estratégias, a depender do grau de dificuldade das tarefas (Yurovsky & Frank, 2015). Apesar de instigante, o presente trabalho não abordará essa discussão.

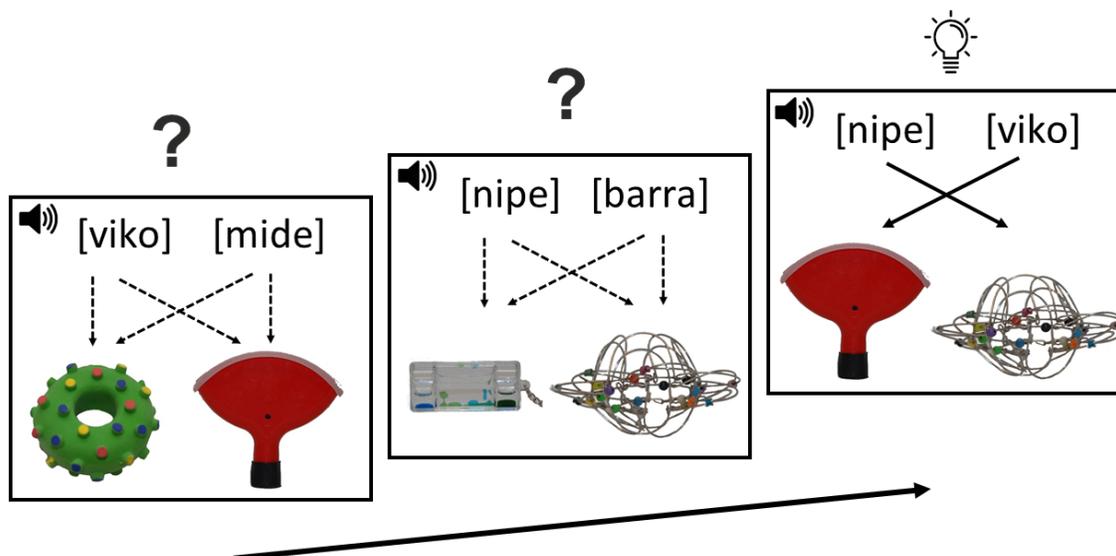


Figura 3. Três tentativas seguindo o paradigma CSWL (2 x 2). As setas tracejadas indicam as possíveis relações entre os estímulos. Se o participante mapear a coocorrência entre palavras e figuras ao longo das tentativas, então, na terceira tentativa, as relações entre as palavras e as figuras podem ser estabelecidas (setas cheias).

Desde a pesquisa inaugural de Yu e Smith (2007), mais de 50 estudos empíricos foram publicados (Dal Ben & Souza, em preparação), replicando os achados do CSWL com bebês (Smith & Yu, 2008; Yu & Smith, 2011), crianças (Fitneva & Christiansen, 2017; Suanda, Mugwanya, & Namy, 2014), com pessoas diagnosticadas com transtorno do espectro autista (McGregor et al., 2013) e com afasia (Peñaloza et al., 2017). Ademais, variações do paradigma investigaram o papel da diversidade contextual (Ramscar, Dye, & Klein, 2013), de dicas sociais (MacDonald, Yurovsky, & Frank, 2017), e de diferenças individuais (Hu, 2017; Krogh et al., 2013; Smith & Yu, 2013; Vlach & DeBrock, 2017) no desempenho em tarefas CSWL.

Entre as variações, duas pesquisas investigaram o papel de dicas estatísticas e estruturais no desempenho em tarefas CSWL. Yurovsky, Yu, e Smith (2012) demonstraram que adultos extraíram palavras (mono ou dissílabas, $PT = 1$) de frases e as mapearam com referentes em situações ambíguas quando (a) tais palavras foram precedidas consistentemente por segmentos (mono e dissílabos) que sinalizavam sua ocorrência (função comparável à de artigos) e (b) estavam localizadas no final das frases. No entanto, os participantes falharam em extraí-las quando os segmentos e as palavras foram movidos para o meio das frases, violando as dicas presentes em suas línguas naturais (inglês americano; resultados semelhantes com crianças são relatados por Filippi, Laaha, & Fitch, 2017).

Em uma direção similar, Filippi, Gingras, e Fitch (2014) realizaram uma série de estudos nos quais adultos foram expostos simultaneamente a uma imagem e uma frase falada composta por cinco sílabas. As imagens podiam ser de pessoas, ou animais ou montanhas, e uma das sílabas (palavra-alvo) sempre coocorria com uma categoria de imagens. Os participantes foram divididos em seis grupos. Um grupo foi exposto apenas à coocorrência entre palavras e figuras. Para os demais, além da coocorrência, também foi adicionada uma das seguintes propriedades: entonação consistente com a palavra-alvo, entonação inconsistente com a palavra-alvo, aumento na duração da palavra-alvo, sinal sonoro (“buzz”) sinalizando a ocorrência da palavra-alvo, ou um sinal visual sinalizando a ocorrência da palavra-alvo. Durante o teste, três figuras eram apresentadas com uma nova frase falada que continha uma das palavras-alvo e os participantes foram instruídos a escolher a figura correspondente. Todos os grupos, exceto o de entonação inconsistente, mapearam acima do esperado pelo acaso. No entanto, o desempenho do grupo exposto à

entonação consistente com a palavra-alvo foi significativamente superior aos demais grupos, chamando atenção para o papel da prosódia na extração e mapeamento.

A busca por integração entre dicas linguísticas específicas (e.g., prosódia, posição) e dicas estatísticas gerais (e.g., PT, coocorrência, diversidade contextual) tem motivado arranjos cada vez mais ecológicos que abordam a ambiguidade em contextos cotidianos (e.g., Clerkin, Hart, Rehg, Yu, & Smith, 2017; Yurovsky, Smith, & Yu, 2013). Além disso, há um esforço crescente em integrar diferentes processos de aprendizagem estatística (Estes, 2009; Saffran, 2014; Smith et al., 2014; Smith, Jayaraman, Clerkin, & Yu, 2018), de modo a gerar explicações mais completas sobre o papel de tais processos na aquisição de linguagem (Koehne & Crocker, 2015). A próxima seção descreve como o presente trabalho busca contribuir para essa integração ao investigar a extração de palavras de falas contínuas e o mapeamento ao longo de tentativas ambíguas simultaneamente.

Pergunta de Pesquisa e Plano de Trabalho

O conjunto de evidências apresentado até o momento informa que palavras podem ser extraídas de falas contínuas com base em uma série de dicas linguísticas, entre elas, a probabilidade transicional entre suas sílabas; que palavras recém extraídas são preferidas no estabelecimento de relações com estímulos visuais em situações não ambíguas, sequenciais ou simultâneas; e que palavras, em isolamento ou extraídas a partir de frases, podem ser mapeadas a referentes com base na coocorrência ao longo de tentativas ambíguas. Ao apreciar o estado atual das investigações, um questionamento possível é se os processos de extração de palavras de falas contínuas e mapeamento em situações ambíguas interagem durante a aprendizagem de uma nova língua.

Perseguindo essa questão, Räsänen e Rasilo (2015) apresentam simulações computacionais indicando que o paradigma CSWL pode explicar mais do que apenas o mapeamento em situações ambíguas a partir de palavras previamente extraídas (proto-léxico). Os autores propõem um modelo de processamento estatístico unificado no qual a extração de palavras de falas contínuas ocorreria como um subproduto do rastreamento das coocorrências entre sons da fala e referentes ao longo de tentativas ambíguas (Figura 4). Nesse contexto, a coocorrência entre estímulos auditivos e visuais funcionaria como uma *cola* que possibilita a extração de palavras e mapeamento simultâneo ao longo de tentativas ambíguas (Räsänen e Rasilo, 2015, p. 795). Os autores apresentam simulações computacionais com línguas naturais e comparações com dados empíricos⁸ que oferecem suporte ao modelo. Não obstante, no limite do nosso conhecimento, as evidências empíricas para tal proposta ainda não foram relatadas na literatura.

⁸ As comparações são feitas com os dados empíricos de Thiessen (2010), Yu e Smith (2007), e Yurovsky, Yu e Smith (2013), que investigaram a interação entre extração e dicas visuais, mapeamento por meio do paradigma CSWL, e competição ao longo de tarefas de CSWL, respectivamente.

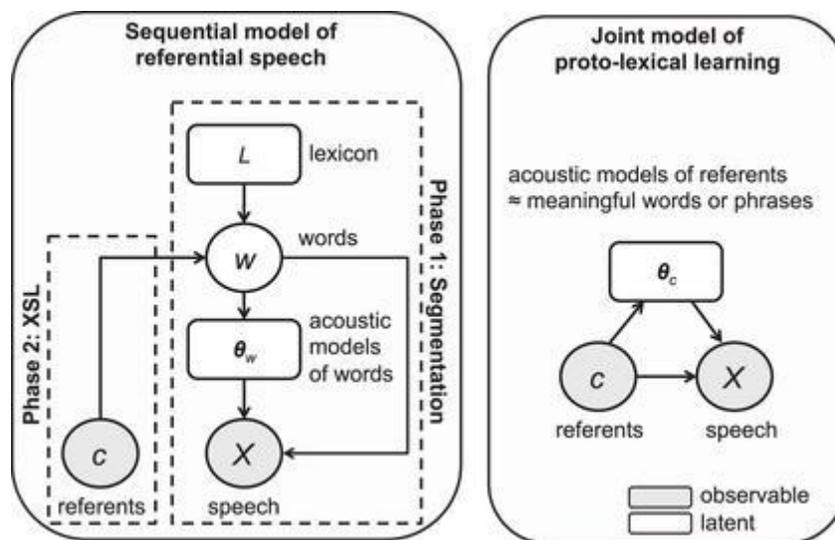


Figura 4. À esquerda, modelo sequencial de aprendizagem de palavras no qual as palavras são primeiro extraídas da fala contínua, formando um proto-léxico, e em seguida mapeadas com referentes. À direita, modelo unificado proposto por Räsänen e Rasilo (2015) no qual palavras são extraídas e mapeadas simultaneamente a partir da coocorrência de referentes com sons da fala (Räsänen & Rasilo, 2015, Fig. 2, p. 798). Figura reproduzida com permissão da *American Psychological Association*.

Diante desse cenário, a presente investigação é guiada pela pergunta de pesquisa:

Palavras são extraídas de falas contínuas e mapeadas a estímulos visuais simultaneamente ao longo de tentativas ambíguas?

Uma sequência de três estudos sequenciais e complementares foi delineada (Figura 5). O primeiro estudo é uma replicação de Saffran et al. (1997) e busca garantir que os estímulos auditivos selecionados podem ser extraídos de uma fala contínua com base na PT entre suas sílabas. O segundo estudo é uma replicação de Yu e Smith (2007) e busca garantir que as palavras e parte-palavras utilizadas no primeiro estudo podem ser mapeadas a estímulos visuais em uma tarefa de CSWL. Por fim, após garantir que as palavras selecionadas podem ser extraídas (Estudo 1) e mapeadas ao longo de tentativas ambíguas

(Estudo 2) em tarefas separadas, o terceiro estudo aborda a pergunta de pesquisa ao investigar os processos de extração e CSWL por meio de uma tarefa única, simultânea.

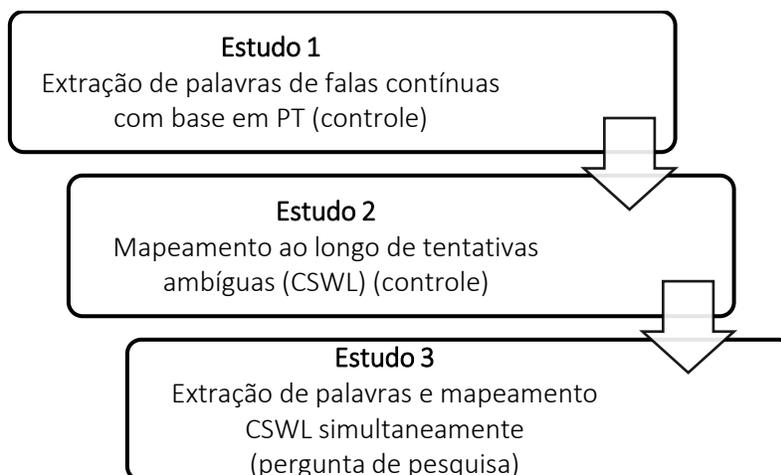


Figura 5. Plano de Pesquisa (Estudos 1 a 3).

Nessa sequência experimental, os dois primeiros estudos servem como controle para o terceiro estudo, que aborda, efetivamente, a pergunta de pesquisa. Tal sequência é especialmente importante considerando dois fatores. Primeiro, no limite do nosso conhecimento, essa é a primeira investigação que testa os efeitos da PT e coocorrência sobre a aprendizagem de palavras com falantes do Português-Brasileiro (PB)⁹. Portanto, é relevante estabelecer linhas de base para os processos de extração e mapeamento CSWL (estudos 1 e 2). Uma ampla busca na literatura nacional revelou discussões sobre a relação entre dicas estatísticas (i.e., PT) e (a) dicas linguísticas específicas na extração de palavras (e.g., prosódia, *manhês*, morfologia; Barbosa & Martins, 2014; Matsuoka, 2006; Molina, 2014; Name, 2011), (b) métodos de pesquisa e análise de dados (Name, 2012), (c) teorias

⁹ Vale notar que a abreviatura PB para português brasileiro é bastante similar a outras abreviaturas utilizadas ao longo do texto (i.e., PT para Probabilidade Transicional e Pbf para Probabilidade Fonotática). Não obstante, o uso frequente do termo justifica sua abreviatura.

de aquisição de linguagem (e.g., Corrêa, 2007). Além disso, investigações teóricas e simulações computacionais sobre o paradigma CSWL também foram desenvolvidas por Fontanari e colaboradores (e.g., Fontanari & Cangelosi, 2011). No entanto, nenhuma investigação empírica manipulando dicas estatísticas foi encontrada.

Em segundo lugar, os participantes são adultos com extensas histórias de aprendizagem linguística. Portanto, os primeiros estudos buscam controlar variáveis indesejadas (por exemplo, conhecimento linguístico, idiossincrasias dos estímulos experimentais) que possam afetar o desempenho na tarefa experimental.

É importante destacar que falantes do português brasileiro e do inglês norte-americano participaram da pesquisa. Sempre que pertinente, diferenças nos estímulos experimentais e no procedimento serão descritas. Além disso, a investigação com duas populações possibilita verificar a generalidade dos achados e comparações transculturais, quando pertinentes.

As hipóteses de cada estudo, os delineamentos experimentais, e as manipulações realizadas em cada experimento serão apresentadas nos Capítulos 3, 4, e 5. O próximo capítulo descreverá os elementos comuns aos três estudos.

Capítulo 2

Elementos comuns aos Estudos

Os três estudos que formam essa pesquisa seguem os mesmos procedimentos éticos, utilizam os mesmos questionários pós-experimentais, estímulos experimentais, estratégia de análise de dados, e critérios de tamanho da amostra descritos nesse capítulo.

Procedimentos Éticos

Todos os procedimentos adotados ao longo dos três estudos foram aprovados pelo Comitê de Ética em Pesquisa em Seres Humanos da Universidade Federal de São Carlos (parecer 1.484.847 e 3.085.914) e pela *Institutional Review Board da University of Tennessee, Knoxville* (UTK IRB-14-01842 XP) (Apêndice A).

Todos os participantes concordaram em participar e assinaram um Termo de Consentimento Livre e Esclarecido que ressaltava o caráter voluntário da participação e apresentava o objetivo geral da pesquisa, seus possíveis riscos e benefícios (Apêndice B e Apêndice C). O anonimato foi garantido a todos os participantes durante toda a pesquisa e todos os participantes receberam uma devolutiva ao final da pesquisa.

Os participantes brasileiros foram recrutados por meio de anúncios no grupo da UFSCar no Facebook e não tiveram nenhum retorno financeiro ou acadêmico pela sua participação. Os participantes norte-americanos foram recrutados na base de participantes do Departamento de Psicologia da University of Tennessee (SONA) ou por meio de cartazes espalhados pelo campus e por postagens online nos meios eletrônicos da universidade. Os participantes recrutados no banco de participantes receberam de 0,5 a 1 crédito acadêmico pela participação (a depender do estudo) e os demais participantes receberam um *gift card* no valor de 10 dólares americanos.

Questionário pós-experimental

Todos os participantes responderam a um questionário pós-experimental que buscava mensurar fatores indicados pela literatura como possíveis moderadores na aprendizagem de uma nova língua (Li, Sepanski, & Zhao, 2006). O questionário perguntava: o gênero, a idade, nível educacional, nacionalidade, se já haviam residido no exterior e tempo de residência, língua nativa, proficiência em outras línguas que não o PB ou o inglês (a depender da língua nativa), e a existência de condições que afetavam sua percepção auditiva e visual (Apêndice D e Apêndice E).

Estímulos experimentais

Estímulos auditivos e visuais foram utilizados nos estudos. Os processos de seleção e gravação dos estímulos auditivos são descritos a seguir. Por fim, o processo de seleção dos estímulos visuais será descrito.

Estímulos auditivos. Os estímulos auditivos foram selecionados respeitando as dimensões fonéticas da língua nativa dos participantes. Nesse sentido, um conjunto de línguas foi criado respeitando as características do Português Brasileiro (PB) e outro conjunto de línguas foi criado respeitando as características do inglês. Considerando a semelhança no processo de seleção, ambos os conjuntos serão descritos em paralelo nos próximos tópicos.

Cinco dimensões linguísticas foram controladas na criação das línguas: (1) tamanho (*word length*), (2) estrutura silábica, (3) probabilidade fonotática, (4) densidade de vizinhança, e (5) probabilidade transicional. Além dessas, o grau de *wordlikeness* também foi controlado para as palavras de uma das línguas do PB (ver tópico *wordlikeness*).

Ao controlar tais dimensões, doze palavras¹⁰ dissílabas, inexistentes no PB, foram selecionadas para criar seis línguas com diferentes níveis de familiaridade¹¹ e que respeitavam as características do PB (Tabela 1). Outras 47 palavras dissílabas, inexistentes do inglês, foram selecionadas para criar 4 línguas (com duas versões cada) que respeitavam as características do inglês (Tabela 1).

O tamanho, a estrutura silábica e a probabilidade transicional foram as mesmas para as palavras de todas as línguas. Os passos para o controle das demais dimensões (i.e., probabilidade fonotática, densidade de vizinhança) variaram para as línguas do PB e do inglês. Para as línguas que respeitavam as características do PB, eles variaram de acordo com a familiaridade da língua (veja o esquema na Figura 6). Já para as línguas que respeitavam as características do inglês, dois critérios foram utilizados. Primeiro, as primeiras duas línguas (Tabela 1) foram selecionadas a partir da recombinação dos estímulos utilizados por Mirman et al. (2008). As demais línguas foram selecionadas seguindo os passos esquematizados na Figura 6.

¹⁰ Vale lembrar que as *palavras* experimentais são sequências silábicas com $PT = 1$ e inexitem no PB ou no inglês.

¹¹ O termo *familiaridade* serve somente para indicar uma maior ou menor semelhança com as dimensões presentes no PB ou no inglês. Todas as línguas eram desconhecidas dos participantes. Além disso, a criação das línguas foi guiada por achados empíricos do primeiro estudo, conforme descrito no Capítulo 3.

Tabela 1

Língua Nativa dos participantes (LN), Nível de Familiaridade da Língua com a Língua Nativa (NF), Número da Língua (L), Probabilidade Transicional (PT), Frequência, Probabilidade Fonotática (Pbf), n de Coltheart, e Distância Ortográfica de Levenshtein (DOL20) das Palavras (P; IPA), Parte-palavras (PP; IPA), e Não-palavras (PP; IPA)*

LN/NF	L	P	PT	Freq	Pbf	Colt	DOL20	PP	PT	Pbf	Colt	DOL20	
Português / baixa familiaridade	1	[kire]	1	300	0,003	1	1,9	[rero]	0,5	0,0007	3	1,6	
		[ziru]	1	300	0,003	0	2	[ruki]	0,5	0,0060	1	1,9	
		[rozu]	1	300	0,003	0	2	[zuzi]	0,5	0,0007	1	1,95	
		[zote]	1	150	0,003	1	1,35						
		[reru]	1	150	0,003	1	1,85						
		[vura]	1	150	0,003		2						
	2	[zote]	1	300	0,003	1	1,35	[tevu]	0,5	0,0139	2	1,8	
		[reru]	1	300	0,003	1	1,85	[ruzo]	0,5	0,0034	5	1,45	
		[vura]	1	300	0,003	0	2	[rare]	0,5	0,0154	12	1	
		[kire]	1	150	0,003	1	1,9						
		[ziru]	1	150	0,003	0	2						
		[rozu]	1	150	0,003	0	2						
		3A	[d̥ʒini]	1	300	0,060	10	1,4	[nipe]	0,5	0,012	4	1,65
			[deta]	1	300	0,042	19	1	[tad̥ʒi]	0,5	0,015	2	1,8
[pemi]	1		300	0,030	7	1,6	[mide]	0,5	0,016	19	1		
[sute]	1		150	0,036	6	1,2							
[viko]	1		150	0,021	5	1,35							
[bara]	1		150	0,035	25	1							
3B	[teba]		1	300	0,014	12	1,15	[sute]	0,5	0,036	6	1,2	
	[kosu]	1	300	0,029	1	1,95	[viko]	0,5	0,021	5	1,35		
	[ravi]	1	300	0,013	7	1,55	[bara]	0,5	0,035	25	1		
	[nipe]	1	150	0,012	4	1,65							
	[tad̥ʒi]	1	150	0,015	2	1,8							
	[mide]	1	150	0,016	19	1							
	4	[sute]	1	300	0,036	6	1,2	[teba]	0,5	0,014	12	1,15	
[viko]		1	300	0,021	5	1,35	[kosu]	0,5	0,029	1	1,95		
[bara]		1	300	0,035	25	1	[ravi]	0,5	0,013	7	1,55		
[d̥ʒini]		1	150	0,060	10	1,4							
[deta]		1	150	0,042	19	1							
[pemi]		1	150	0,030	7	1,6							

Tabela 1, continuação												
Ln/ NF	L	P	PT	Freq	Pbf	Colt	DOL20	PP	PT	Pbf	Colt	DOL20
Português / alta familiaridade	5	[nipe]	1	300	0,012	4	0,012	[d̥ʒini]	0,5	0,060	10	1,4
		[tad̥ʒi]	1	300	0,015	2	0,015	[deta]	0,5	0,042	19	1
		[mide]	1	300	0,016	19	0,016	[pemi]	0,5	0,030	7	1,6
		[sute]	1	150	0,036	6	1,2					
		[viko]	1	150	0,021	5	1,35					
		[bara]	1	150	0,035	25	1					
		[sute]	1	300	0,036	6	1,2	[teba]	0,5	0,014	12	1,15
	6	[viko]	1	300	0,021	5	1,35	[kosu]	0,5	0,029	1	1,95
		[bara]	1	300	0,035	25	1	[ravi]	0,5	0,013	7	1,55
		[nipe]	1	150	0,012	4	0,012					
		[tad̥ʒi]	1	150	0,015	2	0,015					
		[mide]	1	150	0,016	19	0,016					
		[bædu]	1	300	0,0090	1	1,9	[dupo]	0,5	0,0071	1	1,85
		inglês / Estímulos recombinados de Mirman et al. (2008)	7A	[poti]	1	300	0,0163	2	1,8	[titæ]	0,5	0,0125
[tædi]	1			300	0,0101	3	1,8	[dibæ]	0,5	0,0060	1	1,9
[pæbo]	1			150	0,0076	1	1,9					
[pibu]	1			150	0,0073	0	2					
[dotu]	1			150	0,0046	3	1,8					
7B	[tupæ]		1	300	0,0044	4	1,65	[pæbo]	0,5	0,0076	1	1,9
	[bopi]		1	300	0,0076	4	1,75	[pibu]	0,5	0,0073	0	2
	[budo]		1	300	0,0075	5	1,75	[dotu]	0,5	0,0046	3	1,8
	[dupo]		1	150	0,0071	1	1,85					
	[titæ]		1	150	0,0125	5	1,65					
[dibæ]	1	150	0,0060	1	1,9							
inglês / Alta familiaridade	8A	[nɔbi]	1	300	0,0020	4	1,75	[kɛnɔ]	0,5	0	7	1,45
		[pɔke]	1	300	0,0029	11	1,40	[bipɔ]	0,5	0,0037	0	1,95
		[tɪkæ]	1	300	0,0045	5	1,75	[keti]	0,5	0,0080	2	1,85
		[mæsi]	1	150	0,0193	6	1,65					
		[læto]	1	150	0,0118	8	1,55					
		[fido]	1	150	0,0050	6	1,70					
		[silɐ]	1	300	0	9	1,55	[mæsi]	0,5	0,0193	6	1,65
	8B	[tofi]	1	300	0,0033	3	1,80	[læto]	0,5	0,0118	8	1,55
		[domæ]	1	300	0,0048	8	1,55	[fido]	0,5	0,0050	6	1,70
		[kɛnɔ]	1	150	0	7	1,45					
		[bipɔ]	1	150	0,0037	0	1,95					
		[keti]	1	150	0,0080	2	1,85					

Tabela 1, continuação												
Ln/ NF	L	P	PT	Freq	Pbf	Colt	DOL20	PP	PT	Pbf	Colt	DOL20
inglês / Alta familiaridade	9A	[nobi]	1	300	0,0063	4	1,75	[kæno]	0,5	0,0276	7	1,45
		[poke]	1	300	0,0087	11	1,40	[bipo]	0,5	0,0051	0	1,95
		[tikæ]	1	300	0,0045	5	1,75	[keti]	0,5	0,0080	2	1,85
		[mæsi]	1	150	0,0193	6	1,65					
		[læto]	1	150	0,0118	8	1,55					
		[fido]	1	150	0,0050	6	1,70					
	9B	[silæ]	1	300	0,0073	9	1,55	[mæsi]	0,5	0,0193	6	1,65
		[tofi]	1	300	0,0033	3	1,80	[læto]	0,5	0,0118	8	1,55
		[domæ]	1	300	0,0048	8	1,55	[fido]	0,5	0,0050	6	1,70
		[kæno]	1	150	0,0276	7	1,45					
		[bipo]	1	150	0,0051	0	1,95					
		[keti]	1	150	0,0080	2	1,85					
	10A	[siko]	1	300	0,0063	7	1,60	[mæsi]	0,5	0,0193	6	1,65
		[sumæ]	1	300	0,0056	6	1,50	[koti]	0,5	0,0135	2	1,85
		[tifæ]	1	300	0,0036	0	2,00	[fæsu]	0,5	0,0110	3	1,80
		[kæni]	1	150	0,0304	8	1,55	[kofæ] ^a	0	0,0071	4	1,80
		[potu]	1	150	0,0098	1	1,90	[mæko] ^a	0	0,0168	4	1,75
		[dæbi]	1	150	0,0089	2	1,85	[fæmæ] ^a	0	0,0095	7	1,60
	10B	[nipo]	1	300	0,0046	0	1,95	[kæni]	0,5	0,0304	8	1,55
		[tudæ]	1	300	0,0042	4	1,75	[potu]	0,5	0,0098	1	1,90
		[bikæ]	1	300	0,0055	4	1,80	[dæbi]	0,5	0,0089	2	1,85
[mæsi]		1	150	0,0193	6	1,65	[pokæ] ^a	0	0,0085	5	1,75	
[koti]		1	150	0,0135	2	1,85	[dæpo] ^a	0	0,0081	5	1,70	
[fæsu]		1	150	0,0110	3	1,80	[kædæ] ^a	0	0,0150	9	1,50	

Nota a. Não-palavras (PT = 0).

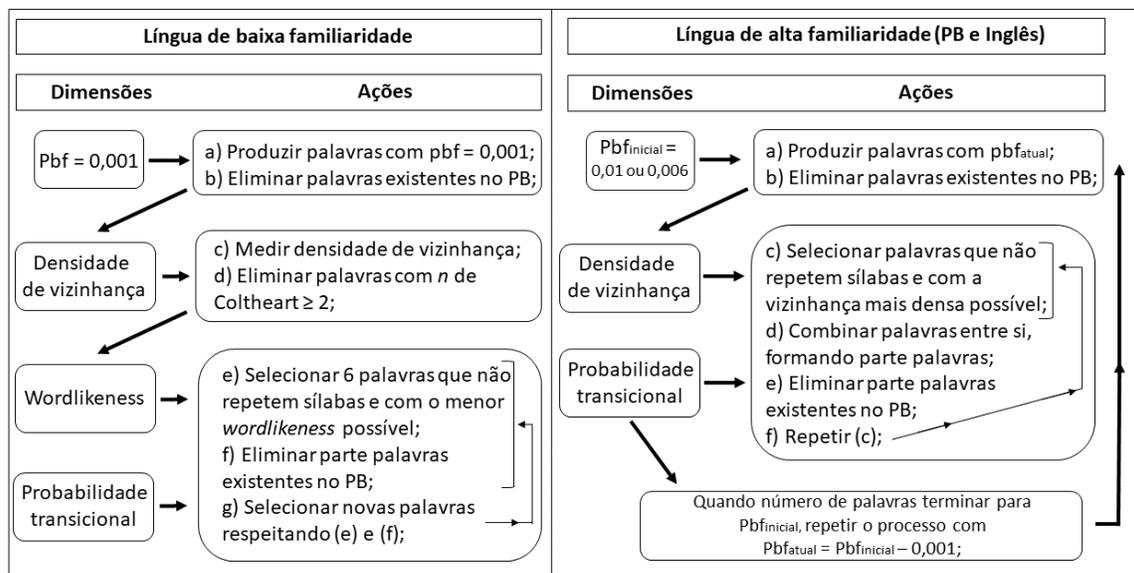


Figura 6. Processo de criação das palavras das línguas de baixa familiaridade do português brasileiro (esquerda) e alta familiaridade do português brasileiro (PB) e do inglês (direita).

Tamanho e estrutura silábica. O tamanho das palavras foi padronizado considerando evidências da influência de tal variável na extração de palavras de falas contínuas. Por exemplo, Johnson e Tyler (2010) demonstraram que crianças de 5 e 8 meses de idade não foram capazes de extrair palavras de uma fala contínua formada por palavras dissílabas e trissílabas, embora fossem capazes de extraí-las quando o tamanho das palavras era uniforme (dissílabas). Resultados semelhantes foram relatados por Mersad e Nazzi (2012).

O tamanho escolhido para os estímulos auditivos foi o dissilábico. Esse tamanho já foi usado com sucesso em pesquisas de extração e de mapeamento ao longo de tentativas ambíguas (e.g., Mirman et al., 2008; Suanda & Namy, 2012). A principal vantagem de palavras dissílabas é a maior facilidade na padronização fonotática e na criação de parte-palavras, além de resultar em uma menor duração da tarefa experimental.

A estrutura silábica utilizada foi a consoante-vogal (viz., CV.CV). Essa estrutura foi escolhida por ser uma das mais frequentes no PB (Estivalet, 2018; Marques, 2008) e por ser a estrutura mais frequentemente utilizada em pesquisas de extração e mapeamento ao longo de tentativas ambíguas (e.g., Saffran et al., 1996; Yu & Smith, 2007).

Probabilidade fonotática. O arranjo posicional de segmentos fonéticos possíveis em uma língua forma a fonotática dessa língua (Chambers, Onishi, & Fisher, 2003). A probabilidade fonotática (Pbf), por sua vez, é a probabilidade de um dado segmento fonético (e.g., fone, bifone) ocorrer em uma determinada posição de uma palavra existente na língua¹² (e.g., *cama*, *coca*, *barraca*). A Pbf pode ser calculada de várias formas. No presente estudo, ela foi calculada pela divisão da soma da frequência (*type*¹³) de um segmento (x, e.g., *ca*) em determinada posição (n, e.g., 2º bifone) de palavras pela soma da frequência das palavras que possuem qualquer segmento na mesma posição (*type*):

$$P(\text{seg } x \cup \text{posição } n) = \frac{\sum \text{frequência}(\text{seg } x \cup \text{posição } n)}{\sum \text{frequência}(\text{palavras} \cup \text{posições } \geq n)}$$

Por exemplo, a probabilidade da sílaba¹⁴ *do* na terceira posição de uma palavra (e.g., *enfado*) foi calculada pela divisão da soma do número de ocorrência (*type*) de palavras que possuem *do* na terceira posição (e.g., *enfado*, *cajado*, *estado*, *causado*) em um dado corpus

¹² A probabilidade fonotática também pode ser calculada *entre* as palavras existentes na língua (e.g., isso *ou* aquilo..., tudo *ou* nada..., agora *ou* nunca...). Para os propósitos desse estudo, apenas o cálculo das probabilidades intra-palavras foi considerado.

¹³ Para uma discussão sobre a distinção type-token, veja Wetzell (2014).

¹⁴ Para clareza do exemplo, serão utilizados grafemas e não fonemas.

pela soma da frequência de todas as palavras do mesmo corpus que possuem pelo menos três segmentos (e.g., caneta, sopapo, acato; mas não, pato, bis, tato)¹⁵.

A influência da Pbf em tarefas linguísticas tem sido documentada em investigações sobre categorização fonética (Albano, 2007), extração de palavras de falas contínuas (e.g., Estes & Bowen, 2013; Mattys, Jusczyk, Luce, & Morgan, 1999; Mersad & Nazzi, 2011), reconhecimento de palavras (e.g., Vitevitch & Luce, 2005; Vitevitch, Luce, Pisoni, & Auer, 1999), aprendizagem de palavras (e.g., Gonzalez-Gomez et al., 2013; MacKenzie et al., 2012), produção verbal (e.g., Goldrick & Larson, 2008; Zamuner, Gerken, & Hammond, 2004), avaliação de *wordlikeness* (e.g., Bailey & Hahn, 2001; Frisch, Large, & Pisoni, 2000), e em outros processos cognitivos tais como memória (e.g., Gathercole, Frankish, Pickering, & Peaker, 1999) e tomada de decisão (e.g., Vitevitch & Donoso, 2012).

Com base nessas evidências, buscou-se controlar os possíveis efeitos da Pbf das palavras utilizadas nos estudos. Considerando as inúmeras combinações fonéticas possíveis em uma língua, o cálculo das probabilidades fonotáticas é uma tarefa praticamente impossível sem ferramentas computacionais. Pesquisadores interessados na língua inglesa, espanhola, e árabe contam com calculadoras online que oferecem tais probabilidades (Aljasser & Vitevitch, 2018; Storkel, 2013; Storkel & Hoover, 2010; Vitevitch & Luce, 2004). No entanto, esse não é caso do PB. Diante da ausência desse importante recurso, uma calculadora das probabilidades fonotáticas dos fonemas e bifonemas do português brasileiro foi construída pelo pesquisador em colaboração com o Dr. Gustavo L. Estivalet,

¹⁵ O conjunto completo das palavras de uma língua e suas ocorrências dificilmente será efetivamente incluído no cálculo. Porém, grandes corpora (e.g., Estivalet & Meunier, 2015; Brysbaert & New, 2009) fornecem uma aproximação para tal conjunto.

psicolinguista computacional filiado à Universidade Federal da Paraíba (Estivalet & Dal Ben, em preparação).

O algoritmo descrito por Vitevitch e Luce (2004) foi adaptado e aplicado a um corpus do PB (Estivalet & Meunier, 2015) e todas as probabilidades de todos os fones e bifones do PB em todas as possíveis posições foram calculadas. Em seguida, considerando o tamanho e estrutura silábica escolhida (CVCV), as distribuições das probabilidades fonotáticas dos bifones CV, VC, CV, na primeira, segunda, e terceira posição, respectivamente, foram analisadas. Tais distribuições são enviesadas positivamente (Figura 7), com a maioria dos bifones exibindo baixa Pbf.

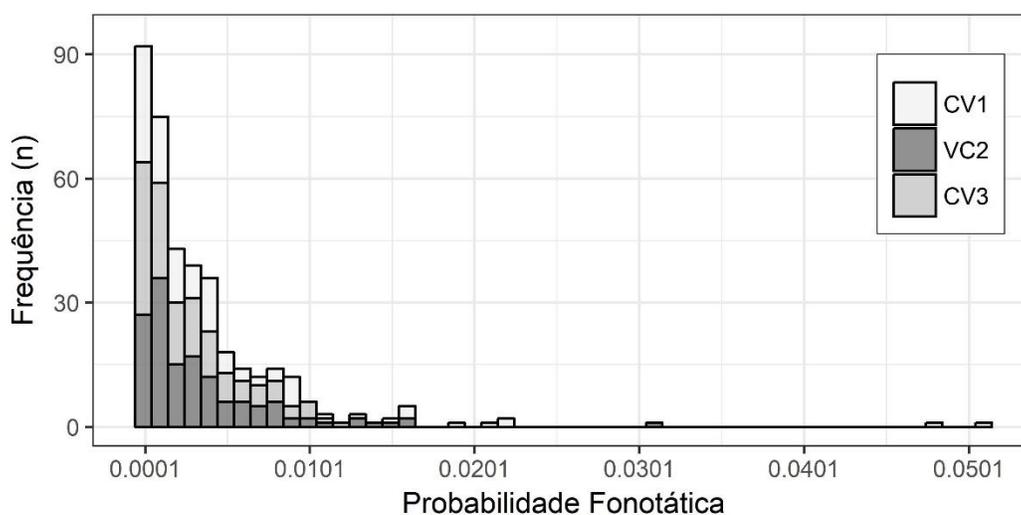


Figura 7. Distribuições das probabilidades fonotáticas dos bifones CV, VC, CV na primeira, segunda e terceira posição das palavras do português brasileiro.

O objetivo principal do controle das probabilidades fonotáticas é controlar quaisquer efeitos indesejados sobre a extração e mapeamento das palavras. Para as línguas que respeitam as características do PB, seleções antagônicas foram realizadas para as

línguas de baixa e alta familiaridade. Para a língua de baixa familiaridade, a Pbf entre os três bifones foi equilibrada para todas as palavras. Quatrocentas e trinta e duas pseudopalavras foram geradas automaticamente pela combinação de bifones com probabilidades iguais a 0,001. Em seguida, 247 pseudopalavras repetidas foram eliminadas. As 185 pseudopalavras restantes foram inseridas no buscador do Léxico do Português Brasileiro (Estivalet & Meunier, 2015¹⁶) e 21 supostas pseudopalavras foram eliminadas por serem palavras existentes no PB. A filtragem das 164 pseudopalavras restantes foi realizada seguindo as três dimensões descritas nos próximos subtítulos (Figura 6, Língua de baixa familiaridade).

Para a língua de alta familiaridade, as palavras foram criadas a partir da combinação das maiores probabilidades fonotáticas possíveis. O desafio foi gerar o número necessário de palavras (6) que não existissem no PB individualmente ou em combinação (parte-palavra), dado que quanto maior a Pbf, maiores as chances de ocorrência na língua natural. A seleção foi iniciada com bifones com probabilidade maior ou igual a 0,01. Trinta palavras foram geradas, porém, 19 foram eliminadas por existirem no PB, e sete por repetirem a mesma sílaba. Em seguida, foram buscados bifones com probabilidade maior ou igual a 0,009. Oitenta e seis palavras foram geradas, porém, 53 foram eliminadas por existirem no PB, e 22 por repetirem a mesma sílaba. O processo se repetiu, diminuindo a proporção de 0,001 por vez, até que 6 palavras com Pbf maior ou igual a 0,006 foram criadas (Figura 6 e Tabela 1).

A primeira língua que seguia as características do inglês foi criada a partir da recombinação dos estímulos utilizados por Mirman et al. (2008). Portanto, as

¹⁶ Buscador disponível em: <http://www.lexicodoportugues.com/>; acessado em 21 de janeiro de 2020.

probabilidades fonotáticas das suas palavras não foram controladas. Para as três outras línguas (Tabela 1), as Pbf das palavras foram controladas seguindo os mesmos passos da seleção das palavras de alta familiaridade do PB. Para tanto, as probabilidades fonotáticas dos bifones CV, VC, CV do inglês americano foram retiradas da calculadora descrita por Vitevitch e Luce (2004)¹⁷ e utilizadas como corpus inicial para a criação de palavras a partir de probabilidades fonotáticas.

A seleção foi iniciada com bifones com probabilidade maior ou igual a 0,006¹⁸. Dez palavras foram geradas. Dessas, três foram eliminadas por existirem no inglês¹⁹, e seis por repetirem a mesma sílaba. Em seguida, foram buscados bifones com probabilidade maior ou igual a 0,005. Trinta e nove palavras foram geradas. Dessas, seis foram eliminadas por existirem no inglês, e 28 por repetirem a mesma sílaba. O processo se repetiu, diminuindo a proporção de 0,001 por vez até que 6 palavras com Pbf maior ou igual a 0,002 foram criadas (Figura 6 e Tabela 1).

Densidade de Vizinhança. A densidade de vizinhança se refere ao número de palavras semelhantes (vizinhos) em uma dada língua. Pelo menos dois cálculos de densidade de vizinhança são amplamente aceitos na literatura linguística, o n de Coltheart e o a Distância Ortográfica de Levensthein (DOL20). O n de Coltheart de uma dada palavra (e.g., gato) é o número de palavras de mesmo tamanho que podem ser produzidas pela mudança de uma unidade da palavra-alvo por vez (e.g., fato, tato, bato, pato, gata, gado)²⁰.

¹⁷ Disponível em: <https://calculator.ku.edu/phonotactic/English/words>, acessado em 20 de janeiro de 2020.

¹⁸ A Pbf inicial de 0,006 dos bifone do inglês é menor do que a Pbf inicial de 0,01 para o PB porque a calculadora de Vitevitch & Luce (2004) utilizada a frequência logarítmica dos bifones do inglês enquanto a calculadora do PB utilizou a frequência linear dos bifones em PB.

¹⁹ O léxico do English Lexicon Project, da Washington University in St. Louis (disponível em: <https://lexicon.wustl.edu/index.html>, acessado em 20 de janeiro de 2020) foi utilizado para buscar as palavras existentes no inglês.

²⁰ Grafemas, ao invés de fonemas, são utilizados para facilitar a exemplificação.

A Distância Ortográfica de Levensthein 20 (DOL20) é a média do número de substituições, inserções, ou exclusões que são necessárias para transformar uma dada palavra em seus 20 vizinhos mais próximos, independentemente do tamanho das palavras de comparação.

Quanto maior o n de Coltheart, mais densa é a vizinhança de uma palavra, e quanto maior a DOL20, menos densa é a vizinhança (Yarkoni, Balota, & Yap, 2008).

Diversas pesquisas demonstram que estímulos com maior densidade de vizinhança são mais facilmente reconhecidos (e.g., Garlock, Walley, & Metsala, 2001; Mainela-Arnold, Evans, & Coady, 2008), aprendidos (e.g., Alt & Plante, 2006; Storkel, Bontempo, Aschenbrenner, Maekawa, & Lee, 2013; Swingley & Aslin, 2007), produzidos (e.g., Munson, Swenson, & Manthei, 2005; Newman & German, 2005), e lembrados (Thomson, Richardson, & Goswami, 2005).

O n de Coltheart e a DOL20 das palavras foram calculadas utilizando o pacote `vwr` para o software R (Keuleers, 2013; R Core Team, 2020). Listas com todas as palavras de quatro letras do PB (Estivalet & Meunier, 2015) e do inglês (Brysbaert & New, 2009) foram utilizadas para o cálculo do n de Coltheart. Os corpora completos, do PB e do inglês, foram utilizados para o cálculo do DOL20.

Para a língua de baixa familiaridade do PB, a densidade de vizinhança das 164 palavras em análise (ver subseção anterior) foi calculada. Buscando controlar possíveis efeitos da densidade de vizinhança no desempenho experimental, 104 palavras com n de Coltheart ≥ 2 (mais densas; e.g., fote, deto, zoro) foram eliminadas, restando 60 palavras.

Por outro lado, para a língua de alta familiaridade do PB e para as línguas 2, 3, e 4 do inglês, a cada Pbf buscada, e após eliminar as palavras existentes no PB ou inglês (ver subtítulo anterior), o n de Coltheart e o DOL20 foram calculados, sempre privilegiando a seleção das palavras com vizinhanças mais densas.

Wordlikeness. *Wordlikeness* se refere à percepção de similaridade entre pseudopalavras e palavras existentes em uma dada língua (e.g., pabo e pato). Pesquisas têm demonstrado que pseudopalavras com maiores graus de *wordlikeness* são mais facilmente processadas, aprendidas e lembradas (e.g., Bailey & Hahn, 2001; Frisch et al., 2000; Gomes, Mendes, Silva, Esteves, & Gomes, 2015).

Diferentemente das medidas descritas até o momento, o grau de *wordlikeness* não é uma estatística exclusivamente linguística: ele resulta do julgamento subjetivo de falantes de uma dada língua. O julgamento de *wordlikeness* foi realizado apenas para os estímulos da língua de baixa familiaridade do PB e ocorreu em duas etapas. Primeiro, o pesquisador avaliou as 60 pseudopalavras restantes e eliminou 12 delas por julgá-las bastante parecidas com palavras do PB (e.g., vuto, koza, nuju). Segundo, 37 adultos, falantes nativos do PB, avaliaram o grau de *wordlikeness* de cada uma das 48 palavras restantes seguindo uma escala *likert* de sete pontos, sendo 7 indicativo de uma probabilidade muito alta de ser uma palavra do PB e 1 impossível de ser uma palavra do PB (Frisch et al., 2000; Apêndice F). Cada palavra foi apresentada na forma escrita por 10 segundos e os avaliadores atribuíam um grau de *wordlikeness* para cada uma delas (Figura 8).

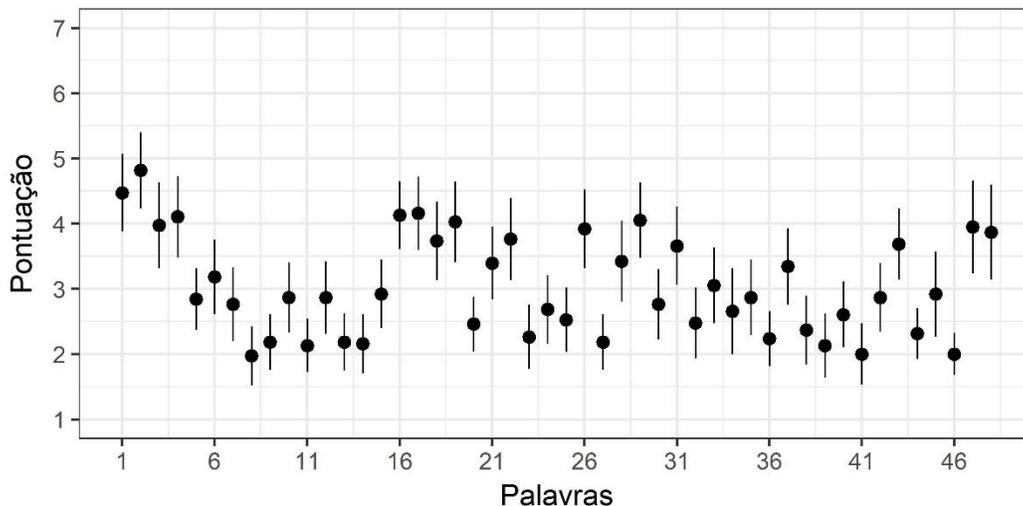


Figura 8. Pontuação média e intervalo de confiança (95%) para cada palavra (48) avaliada na tarefa de *wordlikeness*.

Probabilidade transicional. A probabilidade transicional entre sílabas é a probabilidade de coocorrência adjacente de ao menos duas sílabas (Canistraci et al., 2019; Estes, 2012). A direção da probabilidade pode ser para frente, quando a ocorrência da sílaba antecedente informa a probabilidade de ocorrência da sílaba subsequente (e.g., *ca* informando a probabilidade de *co*, em *caco*), ou para trás, quando a ocorrência da sílaba subsequente informa a probabilidade de ocorrência da sílaba antecedente (e.g., *co* informando a probabilidade de *ta*, em *taco*). A probabilidade transicional (para frente e para trás) entre sílabas é a principal regularidade manipulada em pesquisas de extração de palavras de falas contínuas (e.g., Pelucchi et al., 2009a, 2009b; Perruchet & Desaulty, 2008).

Para a língua de baixa familiaridade do PB, as seis pseudopalavras que não compartilhavam sílabas entre si e que possuíam o menor grau de *wordlikeness* possível foram selecionadas dentre as 48. Em seguida, suas sílabas foram recombinaadas em todas as

posições possíveis para verificar se as parte-palavras resultantes (36) formavam palavras existentes no PB. Três pseudopalavras foram substituídas.

Para as línguas de alta familiaridade do PB e para as línguas 8 e 10 (versões A e B) do inglês, a cada Pbf buscada, as palavras que não eram eliminadas por existirem no PB ou no inglês eram combinadas entre si e as recombinações que existiam nas línguas naturais eram eliminadas. Por fim, as palavras que compartilhavam sílabas eram eliminadas (privilegiando a maior Pbf).

Parte-palavras e não-palavras. Parte-palavras foram geradas a partir da recombinação da segunda sílaba de uma palavra com a primeira sílaba de outra palavra. Com exceção da primeira língua do inglês (cujas parte-palavras foram selecionadas a partir dos estímulos utilizados por Mirman et al., 2008), 30 parte-palavras foram geradas e seis foram selecionadas para cada língua (Tabela 1).

Não-palavras foram geradas para a Língua 10 (versões A e B) do inglês a partir da recombinação da segunda sílaba de uma palavra com a segunda sílaba de outra palavra. Dessa forma, as duas sílabas ocorreram na fala contínua, mas nunca na sequência apresentada (PT = 0).

Línguas. A combinação de palavras gerou as 10 línguas (e cinco versões contrabalanceadas) apresentadas na Tabela 1. A combinação de palavras com diferentes dimensões em diferentes línguas permitiu o teste e, conseqüentemente, o controle de efeitos indesejados de dimensões linguísticas (e.g., Pbf) sobre a extração de palavras de falas contínuas e mapeamento ao longo de tentativas ambíguas. As manipulações e motivos experimentais para cada uma das línguas são descritos no Capítulo 3 e 5.

Gravação dos estímulos auditivos. Durante a gravação dos estímulos, buscou-se evitar a presença de qualquer dica prosódica sistemática que pudesse influenciar na extração das palavras das falas contínuas.

Pesquisas em extração de palavras de falas contínuas utilizam tanto falas sintetizadas (e.g., Cunillera, Laine, et al., 2010) quanto falas naturais (e.g., Estes et al., 2007). O uso de sintetizadores permite a manipulação precisa de cada uma das dimensões relevantes da fala [i.e., coarticulação, F0, intensidade, duração], com a desvantagem de produzir uma fala de aspecto artificial. Na fala natural, por outro lado, o controle das dimensões mencionadas é dificultado. Considerando as vantagens e desvantagens das duas modalidades, falas naturais e sintetizadas foram criadas para a língua de baixa familiaridade do PB e para as línguas do inglês, e falas sintetizadas foram criadas para a língua de alta familiaridade do PB.

Fala natural. Uma falante nativa do PB (28 anos de idade) gravou todas os estímulos que compuseram a fala natural da língua de baixa familiaridade do PB e as Línguas 8 e 9 do inglês (Tabela 1). Todas as gravações e edições foram realizadas no software Praat e exportadas no formato .wav em 24bits²¹ (Boersma & Weenink, 2017); Dal Ben, 2019a). A seguir, são descritas as manipulações da (1) coarticulação, (2) intensidade, e (3) duração da fala. A frequência fundamental (F0) não foi manipulada, tendo sido uniforme entre as sílabas. Os passos da gravação da fala natural são demonstrados na Figura 9.

²¹ A manipulação dos arquivos no Praat ocorre em 64bits. A exportação foi feita em formato .wav por ser o padrão do sistema operacional Windows e em formato 24bits por ser a qualidade mais alta reproduzida pelo software MATLAB, utilizado nos primeiros experimentos do Estudo 1.

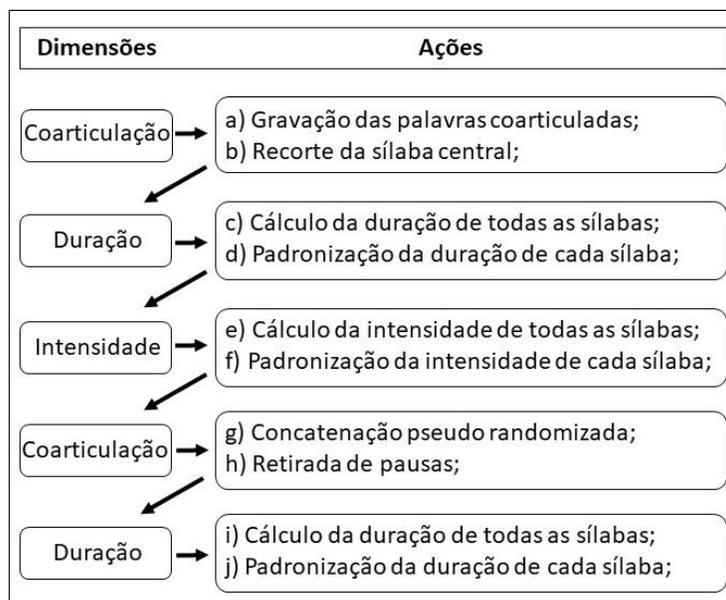


Figura 9. Processo de gravação das falas naturais.

Coarticulação. Uma das dimensões prosódicas de falas naturais é o correlato acústico entre as sílabas que se combinam para formar uma fala contínua (Abaurre & Sandalo, 2017). A coarticulação da fala natural foi realizada em duas etapas.

Primeiro, dois tipos de palavras trissílabas foram criados (1) pelo acréscimo da segunda sílaba de uma dada palavra (e.g., [kire]) antes das demais palavras (e.g., [revura], [rezote]) e (2) pelo acréscimo da primeira sílaba de uma dada palavra depois das demais palavras (e.g., [vuraki], [zoteki]). A combinação exaustiva das primeiras e segundas sílabas de cada palavra com todas as demais palavras gerou 60 palavras trissílabas para cada uma das línguas.

Cada uma das palavras foi lida em voz alta. A falante foi instruída a ler todas as palavras com a mesma F0 ($M = 220\text{Hz}$) e na mesma cadência. As sílabas centrais, coarticuladas com as primeiras e terceiras sílabas (e.g., *vu* em [revura]), foram, então, recortadas e salvas. O recorte foi feito a partir do início e fim das ondas sonoras para a

língua de baixa familiaridade do PB e do final do *pitch contour* da primeira sílaba até o final do *pitch contour* da sílaba central para as línguas do inglês. O corte a partir do *pitch contour* gerou uma melhor co-articulação das sílabas.

A segunda etapa, tanto para a língua de baixa familiaridade em PB e para as línguas em inglês, ocorreu após o controle da intensidade, duração, e concatenação das sílabas de cada língua. Após a concatenação, pequenas pausas entre as sílabas foram retiradas.

Intensidade. A intensidade se refere à amplitude de uma onda sonora. A intensidade média de cada uma das sílabas (já coarticuladas) foi calculada para a língua de baixa familiaridade do PB, em seguida, a intensidade média geral de todas as sílabas foi calculada ($M = 67\text{dB}$) e a intensidade de todas as sílabas foi padronizada para tal valor médio. Para as línguas do inglês, a intensidade de todas as sílabas foi padronizada para 70dB.

As intensidades utilizadas estão alinhadas com as intensidades relatadas na literatura (e.g., 57dB, 60dB, 75dB; Johnson & Jusczyk, 2001b; Pelucchi et al., 2009a; Vlach & DeBrock, 2017). Vale notar que os participantes podiam ajustar o volume do computador durante o experimento, variando, portanto, a intensidade, mas mantendo sua uniformidade ao longo da fala.

Duração. Diferenças nas durações das palavras da fala natural formam o seu ritmo. Buscando controlar possíveis efeitos de diferentes durações das palavras nas tarefas experimentais (e.g., Frost, Monaghan, & Tatsumi, 2017; Shatzman & McQueen, 2006), a duração das sílabas (coarticuladas e com a mesma intensidade) foi controlada em duas etapas.

Para a língua de baixa familiaridade do PB, a duração foi controlada em dois passos. Primeiro, a duração de cada uma das 60 sílabas (coarticuladas e com mesma intensidade) foi calculada e suas durações foram alteradas de modo a atingir a duração média total ($M =$

470ms). Segundo, após a concatenação das sílabas na sequência pseudo-randomizada (descrita na página 54) e após a segunda etapa de coarticulação, a duração de cada uma das 176 sílabas foi calculada novamente e suas durações foram alteradas de modo a atingir a duração média total ($M = 382\text{ms}$). Para as línguas do inglês, a duração de cada sílaba foi de 348ms. Esses valores foram escolhidos levando em conta a duração utilizada nas línguas de alta familiaridade do PB (696ms por palavra) e a literatura relevante (Cunillera, Laine, et al., 2010).

Assim como no controle da intensidade, a duração das sílabas foi padronizada para diminuir a chances de que o ritmo funcionasse como dica para a extração das palavras da fala contínua. Além disso, ela está de acordo com as durações relatadas na literatura (e.g., 696ms, 625ms; 233 e 333ms; 300ms; Cunillera, Laine, et al., 2010; Estes et al., 2007; Frost et al., 2017; Mirman et al., 2008, respectivamente).

Fala sintetizada. Falas sintetizadas foram produzidas para as línguas de baixa e alta familiaridade do PB e para as línguas 7 e 10 do inglês. O sintetizador MBROLA, com a voz feminina br4 para as línguas do PB e com a voz feminina us4 para as línguas em inglês (Dutoit, Pagel, Pierret, Bataille, & van der Vrecken, 1996), foi utilizado. O sintetizador recebe como *input* fonemas (em formato SAMPA; Wells, 1995), durações em ms, e curvas prosódicas (formadas por F0, em Hz, e pela posição, em %, na qual o valor de F0 será atingido).

As dimensões das línguas do PB de baixa familiaridade sintetizadas foram diferentes das dimensões da fala natural. Elas seguiram os valores descritos por Garcia et al. (2018). Cada sílaba durou 277ms, com consoantes durando 100ms e vogais 177ms, com F0 de 218Hz (atingida na metade do fonema), e intensidade média de 77dB (definidas após a sintetização, utilizando o Praat). As dimensões da língua de alta familiaridade do PB se

basearam nos valores descritos por Cunillera, Laine, et al. (2010). Cada sílaba durou 348ms, gerando palavras com duração total de 696ms. Embora superior às velocidades relatadas na literatura (e.g., 300ms, Mirman et al., 2008), tal duração foi preferida considerando o tempo que provavelmente seria necessário para o mapeamento e extração simultâneos no terceiro estudo. A F0 média de 180Hz (atingida na metade do fonema) foi definida após testes com entonações variando entre 220Hz e 160Hz. Por fim, a intensidade média foi definida em 77dB.

As palavras das falas sintetizadas do inglês tiveram a mesma duração das palavras das línguas de alta familiaridade do PB (i.e., duração = 174ms/fone). Porém, o valor de F0 foi alterado para 220Hz, após testes com frequências entre 220Hz e 180Hz. Novamente, o valor máximo de F0 era atingido no meio do fonema (posição = 50%). A intensidade média das línguas do inglês foi de 70dB.

Concatenação, sequência e frequência. Após controlar as dimensões descritas anteriormente (i.e., coarticulação, intensidade, duração, F0) as palavras das falas naturais e das falas sintetizadas, tanto do PB como do inglês, foram concatenadas seguindo uma sequência pseudo-randomizada que respeitou dois critérios: 1) a mesma palavra nunca foi repetida em sequência (e.g., *barabara...*) e 2) a frequência de metade das palavras foi o dobro da frequência da outra metade. Esse balanceamento buscou igualar a frequência das palavras e parte-palavras a serem utilizadas nos testes de extração, uma vez que a ocorrência de uma parte-palavra depende da ocorrência de duas palavras (e.g., Aslin et al., 1998; Endress & Langus, 2017; Apêndice G).

Na literatura, o número médio de repetições de cada palavra nas falas contínuas varia de acordo com a população estudada (sendo menor para crianças em comparação com os adultos; e.g., 45, 300; Saffran et al., 1996a; Saffran et al., 1997). Considerando a

população-alvo da presente pesquisa, adultos, e as frequências relatadas pela literatura (e.g., 78, 200, 300 vezes; Cunillera, Laine, Càmara, & Rodríguez-Fornells, 2010; Mirman et al., 2008; Saffran et al., 1997), as palavras mais frequentes foram repetidas 300 vezes e as menos frequentes foram repetidas 150 vezes. No teste de extração, as palavras menos frequentes foram contrastadas com as parte-palavras formadas a partir da combinação das palavras mais frequentes (i.e., 150 vezes), o que resultava em uma $PT = 0,5$.

No entanto, é importante ressaltar que a frequência das palavras e parte-palavras testadas não foi igualada para as falas da língua de baixa familiaridade do PB. Ao invés disso, a frequência de uma das parte-palavras mais frequentes foi de 90 vezes e das outras duas, 105 vezes (vs. 150 vezes das palavras menos frequentes). As implicações desse desbalanceamento são discutidas no Capítulo 3 (no subtítulo Discussão Geral).

Por razões práticas, um *segmento* com 20 repetições das palavras mais frequentes e 10 repetições das palavras menos frequentes foi criado para cada língua e esse segmento foi repetido 15 vezes, sem pausas, para formar as falas contínuas utilizadas nos experimentos. A duração das línguas naturais de baixa familiaridade do PB foi de aproximadamente 18min e 20s, já a duração da língua sintetizada de baixa familiaridade do PB foi de aproximadamente 12min e 28s. Por outro lado, a duração das falas sintetizadas de alta familiaridade do PB e das falas do inglês (tanto as naturais quanto as sintetizadas) foi de aproximadamente 15min e 39s.

Estímulos visuais. Os estímulos visuais foram resgatados na base de imagens NOUN (*The Novel Object and Unusual Name* – NOUN; Horst & Hout, 2016), que reúne 64 estímulos visuais 3D inventados, naturalísticos e complexos.

Dois conjuntos foram construídos. O primeiro conjunto, utilizado com falantes do PB, buscou garantir um maior nível de discriminabilidade e novidade entre os estímulos

visuais. Foram selecionados os 12 estímulos com maior grau de novidade ($M = 77\%$) dentre os 16 com o maior grau de discriminabilidade ($M = 90\%$) (Tabela 2).

Durante o terceiro estudo com falantes do inglês, além das medidas comportamentais, a direção do olhar e o tamanho da pupila também foram medidas utilizando um rastreador de olhar (EyeLink 1000+). Tais medidas podem dar dicas sobre a força das relações palavra-referente e do esforço cognitivo envolvido no estabelecimento dessas relações. Todavia, os dados do rastreamento de olhar não são apresentados no presente texto. Eles estão sendo analisados (filtragem e análise) e serão apresentados em publicações futuras. Não obstante, para garantir que a luminância dos estímulos visuais estivesse balanceada entre as imagens do conjunto (Willenbockel et al., 2010; Winn Wendt, Koelewijn, & Kuchinsky, 2018), os estímulos foram processados utilizando o algoritmo *lumMatch* da *toolbox* SHINE_color para MATLAB (Dal Ben, 2019b).

Inicialmente, os 12 estímulos já selecionados e testados com os falantes de PB (Conjunto 1, Tabela 2) foram processados utilizando a *toolbox*. Porém, devido à grande variação de luminância entre os estímulos, a maioria deles ficou distorcida. Por exemplo, alguns estímulos ficaram escuros enquanto outros ficaram claros.

Um segundo conjunto de estímulos foi construído. Foram selecionados seis estímulos com maior nível de novidade ($M = 81\%$) e menor nível agregado de saliência (textura e cor, $M = 27\%$). O menor nível de saliência foi utilizado por diminuir as chances de distorções da imagem após o processamento luminoso. Os seis primeiros estímulos (números: 2030, 2033, 2035, 2014, 2048, 2041) foram processados utilizando a *toolbox*. A inspeção visual pelos experimentadores revelou que o estímulo 2033 possuía baixa nitidez após o processamento. Então, ele foi substituído pelo sétimo estímulo mais novo e menos

saliente: 2025. Em seguida, os estímulos foram manipulados de modo a possuírem o mesmo tamanho e o mesmo fundo neutro (cinza).

O grau de saliência (textura e cor) não foi um critério de seleção para os estímulos do primeiro conjunto, utilizados com falantes do PB, mas ele foi bastante próximo da saliência do segundo conjunto ($M = 34\%$ versus $M = 27\%$, respectivamente). De modo semelhante, o grau de discriminabilidade não foi um critério de seleção para os estímulos do segundo conjunto, mas ele foi bastante próximo da discriminabilidade do primeiro conjunto ($M = 83\%$ versus $M = 89\%$, respectivamente). Os estímulos e estatísticas estão apresentados na Tabela 2.

Tabela 2

Número de Identificação na Base de Imagens NOUN (Id), Grau de Discriminabilidade (Disc, %), Novidade (Nov, %), e Saliência Agregada (Sal, %) para cada Estímulo Visual de cada Conjunto

Conjunto 1					Conjunto 2				
Estímulo	Id	Disc	Nov	Sal	Estímulo	Id	Disc	Nov	Sal
	2054	88	97	43		2030	75	91	27
	2005	91	95	47		2035	82	88	30
	2028	89	94	58		2014	83	63	20
	2023	89	88	52		2048	92	88	31
	2002	91	78	34		2041	85	66	22
	2045	90	75	29		2025	85	94	36
	2063	89	75	16					
	2040	88	69	40					
	2037	88	66	20					
	2046	93	66	23					
	2057	93	66	19					
	2062	88	56	36					

Análise de dados

Análises confirmatórias e exploratórias foram realizadas em cada um dos estudos. Com exceção dos três primeiros experimentos, que utilizaram modelos estatísticos clássicos (e.g., teste-t, ANOVA) em dados transformados (de categóricos para proporção contínua), modelos mistos lineares generalizados com efeitos cruzados entre itens e participantes (glmer, pacote lmer4 para R; Bayeen et al., 2008) foram utilizados na análise de medidas categóricas e repetidas (Jaeger, 2008; Young, 2018). Sempre que possível, foi utilizada a estrutura máxima dos efeitos randômicos (Barr et al., 2013). Quando modelos mistos frequentistas não convergiram com a estrutura máxima dos efeitos randômicos, modelos mistos bayesianos foram utilizados (brm, pacote brms para R; Bürkner, 2018). Além disso, os escores médios e seus desvio padrões são descritos sempre que pertinentes (Cumming, 2014).

Tamanho das amostras

O tamanho das amostras utilizadas nos estudos foi definido a priori pelo cálculo do poder estatístico dos designs utilizados nos experimentos e pela análise do tamanho amostral de pesquisas prévias. O software G*Power (Faul, Erdfelder, Lang & Buchner, 2007) foi utilizado para o cálculo do poder. Os seguintes parâmetros foram constantes: família teste-t, teste de Wilcoxon para uma amostra, bidirecional, distribuição logística, alfa = 0,05, e poder = 80%. O tamanho do efeito variou para cada estudo.

Para o Estudo 1 (extração apenas), os tamanhos dos efeitos de Saffran et al. (1996b) e de Saffran et al. (1997) foram estimados²² por serem estudos de extração realizados com adultos em tarefas explícitas e implícitas, respectivamente. O tamanho do efeito da Saffran

²² Cohen's d para uma amostra = (média – nível do acaso) / desvio padrão.

et al. (1996b) foi estimado em $d = 2,17$; com isso, seriam necessários quatro participantes para atingir um poder de 80% em uma tarefa de extração explícita. O tamanho do efeito de Saffran et al. (1997) foi estimado em $d = 0,75$; com isso, seriam necessários 15 participantes para atingir um poder de 80% em tarefas implícitas. Diante disso, e com base na média do tamanho amostral de pesquisas anteriores, nós definimos que o experimento com tarefa explícita (Estudo 1, Experimento 2), deveria ter no mínimo 10 participantes. Já os experimentos com tarefas implícitas (Estudo 1, Experimentos 1, 3, 4, 5, 6, 7, 8, 9) deveriam ter no mínimo 20 participantes por língua. Nesse sentido, para línguas com versões contrabalanceadas (Línguas 3, 7, 8, 9, 10), pelo menos 10 participantes foram testados em cada versão da língua.

Para o Estudo 2 (mapeamento CSWL apenas), o tamanho do efeito da condição 4 x 4 do Experimento 1 de Yu e Smith (2007) com adultos foi utilizado, $d = 1,45$. Com esse tamanho de efeito, seriam necessários no mínimo seis participantes para atingir um poder de 80%. O design da tarefa utilizado no Estudo 2 da presente investigação (2 x 2) foi mais simples do que o design 4 x 4 de Yu e Smith. Logo, ao utilizar o tamanho de efeito relatado pelos autores para a condição 4 x 4, assume-se uma estimativa de poder mais conservadora. Diante disso, e com base na média do tamanho amostral de pesquisas anteriores, nós definimos que os experimentos do Estudo 2 teriam no mínimo 15 participantes cada.

Para o Estudo 3 (extração e mapeamento CSWL em paralelo), o tamanho do efeito da condição *meaningful*, tentativas *target-target*, do Experimento 1 de Cunillera et al. (2010) foi estimado. Esse experimento e condição foram selecionados por terem o design que mais se aproxima do design dos experimentos do Estudo 3 da presente investigação. O tamanho do efeito foi estimado em $d = 1,11$. Logo, seriam necessários no mínimo oito participantes para cada experimento. Diante disso, e com base na média do tamanho

amostral de pesquisas anteriores, nós definimos que os experimentos do Estudo 2 deveriam ter no mínimo 20 participantes cada.

Capítulo 3

Estudo 1 – Extração de Palavras

Contexto

O presente estudo buscou verificar se as palavras selecionadas para formar as línguas de baixa e alta familiaridade, do PB e do inglês (Tabela 1), podem ser extraídas de falas contínuas por meio do rastreamento de TP (Saffran et al., 1997). Dois motivos tornam tal verificação altamente relevante. Primeiro, esse é o primeiro estudo a investigar o processo de extração de palavras com falantes do PB. Todo o processo de seleção e gravação dos estímulos auditivos é inédito e carece da avaliação empírica. Segundo, os desempenhos encontrados no presente estudo servem como linha de base para a investigação dos processos de extração e mapeamento em paralelo, que responde a pergunta de pesquisa da presente investigação (Estudo 3).

O estudo é composto por nove experimentos. Três experimentos foram realizados com as línguas de baixa familiaridade, que respeitaram as características linguísticas do PB. O primeiro testou as Línguas 1 e 2 (Tabela 1), gravadas por um falante natural, ao longo de uma tarefa implícita de extração. Para a presente pesquisa, uma tarefa implícita de extração foi definida tão somente pela ausência de instruções que descreviam a tarefa experimental (descobrir palavras da fala contínua) e pela presença de jogos distratores durante a familiarização com a fala contínua (cf. Saffran et al., 1997). O segundo experimento testou as mesmas línguas, porém em uma tarefa explícita; com instruções descrevendo a tarefa experimental e sem jogos durante a familiarização (Batterink, Reber, Neville, & Paller, 2015). O terceiro experimento testou as mesmas línguas, porém produzidas por um sintetizador, em uma tarefa implícita.

Dois experimentos foram realizados com as línguas de alta familiaridade do PB, utilizando tarefas implícitas de extração de palavras. O primeiro experimento testou se participantes eram capazes de extrair as palavras da Língua 3, versão A e B (contrabalanceamento de palavras e parte-palavras), e da Língua 4 (contrabalanceamento de frequência das palavras em relação à Língua 3, A) (Tabela 1). O segundo experimento testou as Línguas 5 e 6, que tinham palavras e parte-palavras com Pbf equilibrada. As palavras e parte-palavras testadas na Língua 5 possuem a Pbf mais alta possível (antes de se tornarem palavras do PB). Já as palavras e parte-palavras testadas na Língua 6 possuem Pbf mais baixa do que as da Língua 5 (Tabela 1).

Por fim, quatro experimentos foram realizados com as línguas que seguem as características do inglês. Novamente, os experimentos utilizaram tarefas implícitas de extração de palavras (sem instruções e com jogos distratores disponíveis) de falas contínuas com base nas PTs das palavras. Cada experimento testou uma das línguas que seguiram as características do inglês (7 a 10, respectivamente).

Primeiro serão descritos os experimentos realizados com as línguas de baixa familiaridade do PB, seguidos dos experimentos realizados com as línguas de alta familiaridade do PB, seguidos dos experimentos realizados com as línguas do inglês. Ao final, uma discussão geral dos nove experimentos será realizada.

Línguas de baixa familiaridade

Experimento 1

O presente experimento testou, em uma tarefa implícita, se as palavras das línguas de baixa familiaridade podem ser extraídas de uma fala contínua com base nas probabilidades transicionais de suas sílabas (Saffran et al., 1997). Considerando que a PT

será superior nas transições silábicas das palavras ($PT = 1$) do que das parte-palavras ($PT = 0,5$), e que as demais dimensões fonéticas serão controladas, espera-se que as palavras sejam preferidas, acima do esperado pelo acaso, em comparação com as parte-palavras.

Método

Participantes. Trinta adultos ($M = 22$ anos e 7 meses, $SD = 4$ anos e 4 meses; 16 mulheres) participaram do estudo. Todos eram falantes nativos do PB e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no desempenho experimental. Todos os participantes eram estudantes universitários, 23 eram estudantes da graduação (de 14 cursos diferentes) e 7 eram estudantes da pós-graduação (de 6 programas). Eles foram divididos aleatoriamente em dois grupos (15 participantes cada). Cada grupo foi exposto a uma das línguas.

Estímulos experimentais. As Línguas 1 e 2 (Tabela 1) foram utilizadas. Ambas foram gravadas por uma falante natural e seguiram as características descritas no Capítulo 2.

Situação experimental e equipamentos. O experimento foi conduzido em uma sala de coleta de dados do Laboratório de Estudos do Comportamento Humano da UFSCar. Todas as instruções e tarefas experimentais foram apresentadas em um computador com sistema operacional Windows por meio de um *script* criado pelo pesquisador no software MATLAB com o pacote *Psychtoolbox* (Kleiner et al., 2007). Buscando garantir a qualidade dos estímulos auditivos, eles foram apresentados por um fone de ouvido neutro (AKG 740, com decodificador e amplificador Fiio E10). Por fim, de forma a eliminar possíveis distrações na emissão das respostas, um teclado numérico foi adaptado para o experimento.

Além dos equipamentos descritos anteriormente, sete jogos distratores, que exigem atenção para serem resolvidos, foram utilizados (Apêndice H); quais sejam: um quebra

cabeça do tipo *slide puzzle* com 15 peças; uma caixa mágica, que só é aberta com uma posição específica dos dedos; três puzzles de equilíbrio com esferas metálicas (*dexterity puzzle*); um conjunto com 6 peças de madeira que quando encaixadas formam uma estrela; e um conjunto de peças com formas e cores que em determinadas combinações preenchiam um tabuleiro (*iq puzzle*).

Desenho experimental e desenho de tentativas. O experimento foi conduzido em duas fases: familiarização e teste (Figura 10). A familiarização começou com uma instrução escrita para os participantes jogarem os jogos distratores enquanto uma das línguas (1 ou 2, a depender do grupo) era tocada. Durante o teste, um procedimento de escolha forçada entre duas alternativas foi utilizado. Dois estímulos auditivos, separados por 500 ms, eram apresentados junto com a instrução para os participantes indicarem se o primeiro ou o segundo estímulo era mais parecido com as palavras da língua recém apresentada; não havia tempo limite para a escolha. As três palavras menos frequentes de cada Língua foram contrastadas com as três parte-palavras mais frequentes. Cada palavra foi combinada com as três parte-palavras, gerando nove combinações. Cada combinação foi apresentada duas vezes, gerando 18 tentativas. Cinco randomizações das ordens de apresentação das tentativas do teste foram utilizadas.

Entre a fase de familiarização e a fase de teste, um aquecimento com quatro tentativas de escolha forçada entre uma palavra existente no PB e uma pseudopalavra (i.e., bola vs. sibu, tela vs. bafi, cabo vs. guvi, pato vs. tibo) foi realizado.

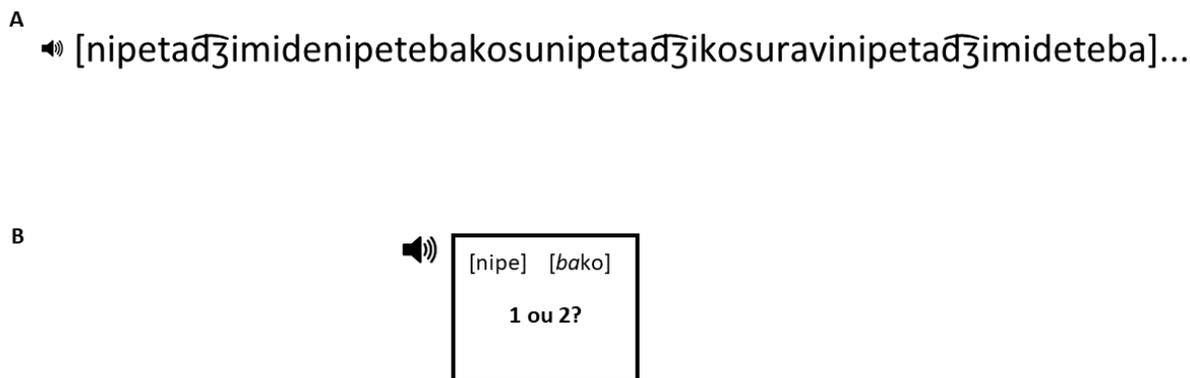


Figura 10. (A) Fase de familiarização com a fala contínua e (B) teste de escolha forçada entre palavra e parte-palavra, do Estudo 1.

Procedimento. No início do experimento, uma música (com a mesma intensidade dos estímulos experimentais, i.e., 67 dB) era tocada junto com a instrução para ajustar o volume do áudio. Após o ajuste do volume, a instrução de que uma língua nova seria tocada por 15 minutos era apresentada junto com a instrução de que, enquanto a língua estivesse sendo tocada, os jogos disponíveis em cima da mesa podiam ser jogados. O participante não era informado de que sua tarefa era descobrir as palavras da língua. Após a familiarização, dava-se início às tentativas de aquecimento com a instrução para pressionar os números 1 ou 2 no teclado para indicar se o primeiro ou segundo estímulo auditivo apresentado correspondia a uma palavra existente no PB. Em seguida, no início do teste, era apresentada a instrução para a seleção dos estímulos mais parecidos com as palavras da língua recém tocadas. Após o teste, o participante foi instruído a estimar o seu desempenho experimental pressionando 1 se estimasse que havia acertado entre 0 e 25% das tentativas, 2 para 25 e 50%, 3 para 50 e 75%, e 4 para 75 e 100%. Após as tarefas experimentais, os participantes foram instruídos a responder um questionário pós-experimental (Capítulo 2, Questionário Pós-experimental; Apêndice C).

Análise de dados. A principal variável dependente foi a taxa de acertos no teste de extração de palavras. Adicionalmente, foram medidas a latência média de acertos e erros, a estimativa de desempenho (auto-relato) e o conhecimento linguístico de cada participante. A taxa geral de acertos foi obtida por meio da divisão do número de acertos pelo número de tentativas, enquanto a taxa de acertos para cada item foi obtida por meio da divisão do número de acertos em cada item pelo número de tentativas (6) nas quais o item foi testado.

A análise por meio de testes estatísticos clássicos (e.g., teste-t, ANOVA) buscou medir se a taxa geral de acertos: 1) diferiu entre as versões da língua; 2) variou de acordo com o gênero do participante; 3) foi maior do que o esperado pelo acaso. Além disso, foram medidas as correlações entre a taxa geral de acerto e: 1) a idade dos participantes; 2) a auto-avaliação; 3) o número de línguas conhecidas; 4) e o maior nível de fluência em língua estrangeira (independentemente da língua); 5) o tempo de moradia no exterior. Por fim, as taxas de acerto de cada item foram testadas contra o acaso e a suas correlações com estatísticas linguísticas (Pbf, densidade de vizinhança, *wordlikeness*) também foram testadas.

Resultados e Discussão

Uma ANOVA de duas vias não encontrou efeitos de versão da língua [$F(1, 26) = 0,566$; $p = 0,459$] ou gênero [$F(1, 26) = 0,302$; $p = 0,587$] na taxa geral de acerto. Com base nesses resultados, os dados obtidos nas duas versões foram analisados em conjunto e a variável gênero não foi incluída em análises posteriores.

Um teste-t bicaudal contra o acaso (0,5) revelou que a taxa geral de acerto ($M = 0,554$; $SD = 0,2$) não foi diferente do esperado pelo acaso, $t(29) = 1,472$; $p = 0,152$; $d = 0,269$; IC 95% (-0,098; 0,631) (Figura 11, Experimento 1). Logo, os participantes não

foram capazes de extrair as palavras da fala contínua a ponto de as discriminarem das parte-palavras na fase de teste.

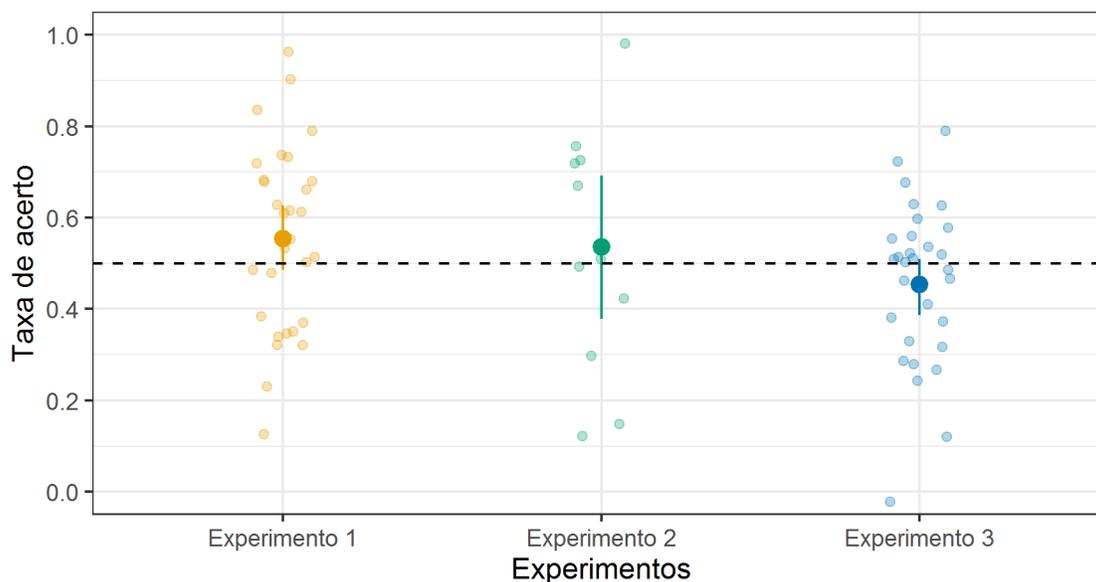


Figura 11. Taxa geral de acerto para cada participante e média geral de acertos para os Experimentos 1, 2, e 3 do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,5).

Não foram encontradas correlações significativas entre as variáveis pessoais (i.e., idade e conhecimento linguístico) e a taxa geral de acerto. Também não foi encontrada correlação entre a taxa geral de acerto e a auto-avaliação, indicando que os participantes não conseguiram estimar o quanto conheciam da língua.

A análise de cada item indicou que apenas uma palavra de cada versão foi aprendida acima do esperado pelo acaso: [reru] para a versão 1, $t(14) = 2,168$; $p = 0,048$; [rozu] para a versão 2, $t(14) = 2,303$; $p = 0,037$. Não foram encontradas correlações entre a taxa de acerto por item e suas estatísticas linguísticas (Pbf: $r = -0,225$; $p = 0,485$; Coltheart's n: $r = -0,09$; $p = 0,817$; DOL20: $r = 0,318$; $p = 0,314$; *wordlikeness*: $r = 0,032$; $p = 0,953$).

O conjunto dos resultados contraria o descrito na literatura (Saffran et al., 1997; Mirman et al., 2008). Os participantes não preferiram palavras (PT = 1) em detrimento das parte-palavras (PT = 0,5). Tal desempenho não pode ser explicado por variáveis pessoais (e.g., conhecimento linguístico) e nem pelas estatísticas linguísticas individuais de cada estímulo (e.g., densidade de vizinhança).

Diante desse cenário, duas explicações para o desempenho são levantadas. Por um lado, é possível que a tarefa distratora (jogos) tenha competido pela atenção dos participantes durante a familiarização com a fala contínua, de modo a interferir com a extração das palavras. Por outro lado, é possível que a co-articulação das sílabas na língua natural não tenha sido discriminável o suficiente para permitir a extração e subsequente preferência acima do esperado pelo acaso na fase de teste. O teste empírico do efeito distrator das atividades foi realizado em um segundo experimento, descrito a seguir.

Experimento 2

Diante da ausência de preferência por palavras (PT = 1) e das possíveis explicações para tal desempenho, o presente experimento foi realizado com o objetivo de investigar se a atenção explícita durante a familiarização com a fala contínua pode influenciar o desempenho na tarefa de extração. Embora haja evidências a favor do processamento implícito das PTs de falas contínuas (Saffran et al., 2007), há evidências de uma maior ativação da atividade neurofisiológica durante tarefas explícitas de extração (Batterink et al., 2015).

No presente experimento, foi realizada uma tarefa explícita de extração, sem jogos distratores durante a familiarização e com instruções sobre o objetivo do experimento. Se a dificuldade em extrair as palavras encontradas no Experimento 1 for produto do nível de

atenção na fala contínua, espera-se que os participantes do presente experimento prefiram palavras (PT = 1), acima do esperado pelo acaso, em comparação à parte-palavras (PT = 0,5), replicando os dados relatados na literatura (Saffran et al., 1996b).

Método

Os estímulos experimentais, a situação experimental, os equipamentos, o desenho experimental, o desenho de tentativas, e a análise de dados foram os mesmos do Experimento 1. As diferenças são descritas a seguir.

Participantes. Onze adultos ($M = 25$ anos e 6 meses, $SD = 4$ anos e 7 meses; 7 mulheres) participaram do estudo. Todos eram falantes nativos do PB e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no seu desempenho na tarefa experimental. Nove participantes eram estudantes universitários, 5 eram estudantes da graduação (de 5 cursos diferentes), 4 eram estudantes da pós-graduação (de 3 programas), e os outros dois eram profissionais graduados. Eles foram divididos aleatoriamente em dois grupos (6 participantes no primeiro grupo e 5 participantes no segundo), cada grupo foi exposto a uma das línguas.

Procedimento. O procedimento foi similar ao procedimento do Experimento 1, com a diferença de que, após ajustar o volume, os participantes eram informados de que ouviriam uma nova língua e instruídos a descobrir as palavras dessa nova língua e de que o seu conhecimento sobre as palavras seria testado ao final da língua. Além disso, os jogos distratores não estavam disponíveis.

Resultados e Discussão

Uma ANOVA de duas vias não encontrou efeitos das versões da língua [$F(1, 8) = 1,689$; $p = 0,230$] ou gênero [$F(1, 8) = 0,298$; $p = 0,6$] na taxa geral de acertos. Com base

nesses resultados, os dados obtidos nas duas versões foram analisados em conjunto e a variável gênero não foi incluída em análises subsequentes.

Um teste-t bicaudal contra o acaso (0,5) revelou que a taxa geral de acerto ($M = 0,535$; $SD = 0,275$) não foi diferente do esperado pelo acaso, $t(10) = 0,427$; $p = 0,679$; $d = 0,129$; IC 95% (-0,468; 0,719) (Figura 11, Experimento 2). Logo, é provável que os participantes não tenham extraído as palavras da fala contínua a ponto de as discriminarem em comparação com parte-palavras na fase de teste. Além disso, não houve correlação entre qualquer variável pessoal (e.g., conhecimento linguístico) e o desempenho geral. A análise de desempenho por item revelou que apenas duas palavras, ambas da primeira versão da língua, foram extraídas acima do esperado pelo acaso (i.e., [reru], $t(5) = 3,841$; $p = 0,012$; e [vura], $t(5) = 3,371$; $p = 0,02$).

Mesmo com instruções explícitas e ausência de jogos distratores, os participantes não preferiram as palavras em comparação às parte-palavras. Novamente, não foram encontradas correlações significativas entre as variáveis pessoais (e.g., conhecimento linguístico) ou estatísticas linguísticas (e.g., densidade de vizinhança) e o desempenho. Diante dos resultados, um terceiro experimento foi realizado para testar se o grau de discriminabilidade produzido pela co-articulação da fala natural pode ter influenciado os desempenhos encontrados.

Experimento 3

O presente experimento teve como objetivo verificar se a discriminabilidade das palavras, que pode ter sido afetada principalmente pela co-articulação da fala natural, pode ter influenciado os desempenhos encontrados nos experimentos anteriores. Conforme descrito no Capítulo 2, falas naturais são menos artificiais (principalmente devido à

entonação variável), porém, a coarticulação de suas sílabas é uma dimensão de difícil controle. Assim, uma fala sintetizada, que possui melhor coarticulação (com o contraponto de possuir um aspecto mais artificial), foi utilizada no presente experimento.

Método

A situação experimental, os equipamentos, o desenho experimental, o desenho de tentativas, o procedimento, e a análise de dados foram os mesmos do Experimento 1. As diferenças são descritas a seguir.

Participantes. Trinta adultos ($M = 22$ anos e 5 meses, $SD = 3$ anos e 3 meses; 12 mulheres) participaram do estudo. Todos eram falantes nativos do PB e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no desempenho experimental. Todos os participantes eram estudantes universitários, 20 eram estudantes da graduação (de 8 cursos diferentes) e 10 eram estudantes da pós-graduação (de 7 programas). Eles foram divididos aleatoriamente em dois grupos (15 participantes para cada grupo), cada grupo foi exposto a uma das línguas.

Estímulos experimentais. Diferentemente dos dois experimentos anteriores, que utilizaram uma fala natural, no presente experimento, a fala contínua e as palavras e parte-palavras utilizadas na fase de teste foram produzidas por um sintetizador. Todo o processo de produção da fala sintetizada e suas dimensões são descritas no Capítulo 2.

Resultados e Discussão

Uma ANOVA de duas vias não encontrou efeitos das versões da língua [$F(1, 26) = 0,409$; $p = 0,528$] ou gênero [$F(1, 26) = 0,689$; $p = 0,414$] na taxa geral de acerto. Em razão desses resultados, os dados das duas versões foram analisados em conjunto e a variável gênero não foi incluída em análises posteriores.

Um teste-t bicaudal contra o acaso (0,5) revelou que a taxa geral de acerto ($M = 0,454$; $SD = 0,173$) não foi diferente do esperado pelo acaso, $t(29) = -1,468$; $p = 0,153$; $d = -0,268$; IC 95% $(-0,630; 0,098)$ (Figura 11, Experimento 3). Logo, provavelmente os participantes não foram capazes de extrair as palavras da fala contínua a ponto de discriminarem entre palavras e parte-palavras na fase de teste. Além disso, não houve correlação entre qualquer variável pessoal (e.g., conhecimento linguístico) e o desempenho geral. A análise de desempenho por item revelou que nenhuma das palavras ou parte-palavras foi extraída acima do esperado pelo acaso.

A ausência de preferência por palavras em comparação com parte-palavras indica que as diferenças de coarticulação e F0 entre a fala natural e a sintetizada provavelmente não afetaram o desempenho encontrado nos experimentos anteriores. Na próxima seção, as manipulações e as possíveis variáveis críticas dos três experimentos são discutidas.

Discussão Geral

Os três experimentos descritos até o momento tiveram como objetivo replicar os resultados descritos por Saffran et al. (1997), de modo a garantir que as palavras previamente selecionadas pudessem ser extraídas com base na PT entre suas sílabas. O primeiro experimento buscou replicar os experimentos de Saffran et al. da maneira mais direta possível, utilizando uma tarefa implícita (sem instruções descrevendo a tarefa experimental e com jogos distratores disponíveis). O segundo experimento utilizou uma tarefa explícita na qual os participantes foram instruídos a descobrir as palavras da nova língua, de modo a medir os efeitos dos jogos distratores na extração das palavras. Por fim, o terceiro experimento utilizou uma fala sintetizada (que ofereceu controle preciso das dimensões fonéticas da fala contínua) ao invés da fala natural, em uma tarefa implícita. Os

resultados dos três experimentos foram nulos e falharam em replicar os resultados relatados na literatura (e.g., Garcia et al., 2018; Saffran et al., 1997).

Análises post hoc dos estímulos experimentais indicam duas possíveis explicações para tais resultados. A primeira explicação, improvável, se fundamenta no desbalanceamento da frequência das parte-palavras da Língua 1. As parte-palavras de alta frequência deveriam ocorrer 150 vezes cada, porém, uma delas ocorreu 90 vezes e as outras duas 105 vezes cada (Capítulo 2). Tal desbalanceamento deveria facilitar a extração das *palavras* testadas em comparação às parte-palavras, uma vez que a frequência superior das palavras (150 repetições) poderia funcionar como uma dica adicional à PT para a extração. No entanto, mesmo com essa dica adicional, os participantes não extraíram as palavras, indicando que, se existente, o efeito dessa dica indesejada não foi forte o suficiente para influenciar o desempenho.

A segunda explicação se fundamenta na diferença entre as probabilidades fonotáticas (Pbfs) e as densidades de vizinhança das palavras e parte-palavras e é composto por dois argumentos. Primeiro, palavras com baixa Pbf (0,003) e densidades de vizinhança podem ter produzido estímulos com estatísticas linguísticas diferentes das encontradas no PB, o que pode ter dificultado o processamento da fala e extração das palavras. Segundo, a recombinação das palavras gerou parte-palavras com Pbfs e densidades de vizinhança (n de Coltheart) em média seis e oito vezes superior às palavras (0,006 vs. 0,003; e 8 vs. 0,5; Tabelas 1). Se por um lado as sequências silábicas que formam as palavras são mais robustas considerando a história recente dos participantes (PT = 1), por outro lado, as sequências silábicas que formam as parte-palavras (PT = 0,5) são mais robustas considerando a história pré-experimental com o PB.

Discussões recentes têm ressaltado o papel da Pbf na extração de palavras (Estes, 2009; Estes, Edwards, & Saffran, 2011; Estes, Gluck, & Grimm, 2016; MacKenzie et al., 2012; Mersad & Nazzi, 2011), e indicam que tais estatísticas se combinam com a PT em tarefas de extração e mapeamento de palavras. Além disso, palavras com Pbf altas (em relação à língua nativa dos participantes) são preferidas nas tarefas de extração em comparação com palavras com probabilidades fonotáticas baixas ou ilegais; apesar da PT determinística nas palavras (Finn & Hudson Kam, 2008; Mersad & Nazzi, 2011).

Diante desse cenário, línguas de alta familiaridade com o PB (alta Pbf e densidade de vizinhança) foram criadas. O processo de criação e os critérios utilizados estão descritos no Capítulo 2. Se a alta familiaridade com a língua nativa dos participantes contribui para a extração das palavras da fala contínua, as palavras das línguas de alta familiaridade deveriam ser preferidas em comparação com as parte-palavras. Isso deve ser verdadeiro quando a Pbf for mais alta para as palavras em comparação com as parte-palavras e quando tais probabilidades estiverem balanceadas entre palavras e parte-palavras. Não obstante, a extração de palavras deve ser dificultada quando as parte-palavras tiverem Pbf superiores às palavras (mesmo que a PT continue a ser determinística nas palavras e probabilística nas parte-palavras). A próxima seção descreve dois experimentos que investigaram essas hipóteses.

Línguas de alta familiaridade

Experimento 4

O presente experimento testou os efeitos do desbalanceamento das probabilidades fonotáticas entre palavras e parte-palavras de duas línguas de alta familiaridade (Língua 3, versões A e B; e 4; Tabela 1). Na Língua 3, versão A, as palavras possuíam Pbf mais altas

do que as parte-palavras; embora a PT continuasse a ser determinística para as palavras ($PT = 1$) e probabilística para as parte-palavras ($PT = 0,5$). Na outra versão (3B), o inverso foi verdadeiro: as parte-palavras possuíam maior Pbf, e continuavam com menor PT.

A segunda língua (Língua 4) foi criada pelo contrabalanceamento da frequência das palavras da Língua 3A. Palavras de alta frequência na Língua 3A, se tornaram palavras de baixa frequência na Língua 4 e vice-versa. As diferenças de Pbf entre palavras e parte-palavras na Língua 4 eram similares à Língua 3A. Logo, cada Língua serviu como controle para possíveis efeitos idiossincráticos das palavras testadas nas demais línguas.

Com base na literatura e nos resultados dos experimentos anteriores, espera-se que as palavras sejam preferidas, acima do esperado pelo acaso, em comparação com parte-palavras na Língua 3A e na Língua 4, mas não na Língua 3B. Ademais, considerando que a tarefa explícita (com instruções diretas e sem jogos distratores) utilizada no Experimento 2 não produziu resultados diferentes dos encontrados nas tarefas implícitas (Experimento 1 e 3), e considerando que os participantes relataram uma experiência mais agradável quando os jogos estavam presentes, optou-se por uma tarefa implícita no presente experimento.

Método

Participantes. Sessenta e um adultos ($M = 21$ anos e 9 meses, $SD = 3$ anos e 1 mês; 43 mulheres) participaram do estudo. Todos eram falantes nativos do PB e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no seu desempenho na tarefa experimental. Dos 61 participantes, 49 eram estudantes da graduação (de 21 cursos diferentes), 10 eram estudantes da pós-graduação (de 9 programas), e dois haviam concluído o ensino médio, mas não frequentavam o ensino superior. Eles foram divididos aleatoriamente em três grupos. Um grupo foi exposto à Língua 3A; outro grupo foi exposto à 3B (21 participantes); e outro grupo foi exposto à Língua 4.

Estímulos experimentais. A Língua 3A, 3B, e 4 foram utilizadas (Tabela 1). Elas foram geradas por um sintetizador e seguiram as características descritas no Capítulo 2.

Situação experimental e equipamentos. A mesma sala de coleta de dados e equipamentos descritos anteriormente foram utilizados no presente experimento. A única diferença foi a substituição do software de apresentação da tarefa experimental: do *Psychtoolbox* (MATLAB; Kleiner et al., 2007) para o *Psychopy2* (Peirce et al., 2019).

Desenho experimental e desenho de tentativas. O experimento teve as mesmas fases dos experimentos descritos anteriormente (familiarização e teste, Figura 10).

Procedimento. O procedimento foi o mesmo do Experimento 1 e 3, descritos anteriormente.

Análise de dados. A principal variável dependente foi a preferência por palavras em comparação às parte-palavras na fase teste. A preferência foi atestada contra o nível do acaso (50%). Considerando a natureza binária das respostas (acerto ou erro), modelos mistos generalizados (Bayeen et al., 2008) foram utilizados seguindo a estrutura máxima dos efeitos randômicos²³. Quando os modelos frequentistas falharam em convergir na estrutura máxima dos efeitos randômicos, modelos mistos bayesianos foram utilizados (Bürkner, 2018).

Variáveis dependentes secundárias incluíram a latência de acertos e erros, a estimativa de desempenho (auto-relato) e o conhecimento linguístico de cada participante. Essas variáveis não foram analisadas, uma vez que não foram encontrados efeitos de diferenças individuais na preferência por palavras ou parte-palavras.

²³ Sintaxe (R): `lme4::glmer(respostas ~ offset(logit(nível do acaso)) + (1 | participantes) + (1 | itens))`

Resultados e Discussão

A Figura 12 apresenta a média geral de acertos para cada Língua (i.e., 3A, 3B, 4). A probabilidade dos participantes escolherem as palavras em detrimento das parte-palavras quando expostos à Língua 3A ($M = 68,05\%$, $SD = 27,50\%$) foi consistentemente superior ao esperado pelo acaso ($b = 1,40$, Wald IC 95% [0,30; 2,50], $z = 2,508$, $p = 0,0122$). Isso sugere que a combinação da PT e da Pbf nas palavras possibilitou a extração das mesmas durante a fase de familiarização.

Por outro lado, a probabilidade dos participantes escolherem as palavras em detrimento das parte-palavras quando expostos à Língua 3B ($M = 43,91\%$, $SD = 23,30\%$) não foi diferente do esperado pelo acaso ($b = -0,28$, IC 95% [-0,79; 0,21], $z = -1,124$, $p = 0,261$). Tal desempenho sugere que a extração de palavras é prejudicada quando estatísticas aprendidas recentemente (PT) são contrastadas com estatísticas mais antigas, provenientes da história pré-experimental dos participantes (Pbf). Outra possibilidade é que os participantes tenham extraído tanto palavras, sob controle das PTs, e parte-palavras, sob controle conjunto das PTs e das Pbf. Porém, quando palavras e parte-palavras foram contrastadas nas tentativas de teste, as Pbf talvez tenham exercido maior controle sobre a preferência. Futuros estudos podem investigar se outras configurações de tentativas de teste (e.g., go/no-go, François et al., 2017) ou mesmo a inclusão de não-palavras sendo contrastadas com palavras e parte-palavras, em tentativas de escolha forçada, podem lançar luz sobre essa questão.

Curiosamente, a probabilidade de escolhas das palavras pelos participantes expostos à Língua 4 ($M = 60\%$, $SD = 22,69\%$), que inverte a frequência das palavras de alta e baixa frequência da Língua 3A, mas que mantém a Pbf mais forte nas palavras, não foi diferente do esperado pelo acaso ($b = 0,49$, IC 95% [-0,19; 1,18], $z = 1,405$, $p = 0,16$). Tal resultado

sugere que a diferença na Pbf e PT entre palavras e parte-palavras pode não ser a única variável a explicar os resultados encontrados no teste da Língua 3A e 3B. Em especial, é possível que características idiossincráticas dos estímulos utilizados naquelas línguas tenham exercido controle sobre a extração.

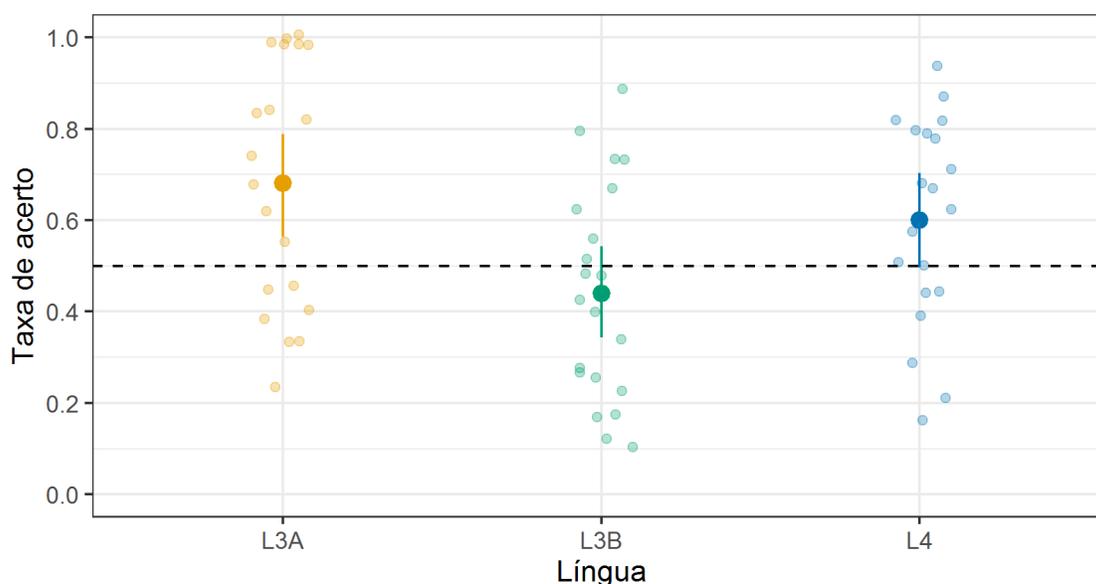


Figura 12. Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 3A, 3B, e 4 do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,5).

A análise²⁴ dos efeitos de cada palavra sobre o desempenho encontrado na Língua 3A revela que um dos estímulos (i.e., [bara]) teve uma probabilidade de escolha superior às demais palavras na Língua 3A ($b = 1,0634$, IC 95% [0,07; 2,04], $z = 2,115$, $p = 0,0344$). O modelo máximo, que controla por preferências individuais, falhou em convergir para a análise da Língua 3B. Logo, um modelo misto linear bayesiano²⁵ foi utilizado. A probabilidade de escolha do mesmo estímulo (i.e., [bara]), porém agora na função de parte-

²⁴ lme4::glmer(acertos ~ palavras + (palavras | participantes))

²⁵ brms::brm(acertos ~ parte-palavras + (parte-palavras | participantes))

palavra, foi superior às probabilidades de escolha das demais parte-palavras ($b = -2,77$, intervalo crível 95% [-6,20; -0,58]). Futuras pesquisas poderão investigar se características idiossincráticas desses estímulos podem explicar tais achados.

O conjunto dos resultados indica que a preferência por palavras na Língua 3A foi controlada pelo conjunto das diferenças de PT e Pbf entre palavras e parte-palavras e pela preferência por um dos itens (i.e., [bara]). O mesmo é válido para explicar a distribuição das preferências entre palavras e parte-palavras dos participantes expostos à Língua 3B.

Tais resultados podem explicar as distribuições de preferência entre palavras e parte-palavras encontrada nos Experimentos 1 a 3, com as línguas de baixa familiaridade. Nelas, palavras tinham Pbf's mais baixas e parte-palavras tinham Pbf's mais altas. Nesse sentido, o conhecimento prévio do PB pode ter guiado a preferência por parte-palavras enquanto que a PT pode ter guiado a preferência por palavras.

Vale ressaltar que um teste paramétrico dos efeitos combinados entre PTs e probabilidades fonotáticas, semelhante ao conduzido por Mersad e Nazzi (2011), pode ser informativo nesse sentido, em especial, considerando que as diferenças de probabilidades fonotáticas entre os itens utilizados na presente pesquisa são menores do que as utilizadas em pesquisas anteriores (e.g., Finn & Hudson-Kam, 2008; Mersad & Nazzi, 2011). Não obstante, o objetivo principal do presente estudo foi garantir que a extração das palavras da língua a ser utilizada na tarefa conjunta de extração e mapeamento ocorresse sob controle principal da PT e não de diferenças na Pbf ou por variáveis idiossincráticas dos estímulos.

Buscando controlar tais efeitos, os estímulos utilizados nas Línguas 3 e 4 foram recombinados de modo a gerar outras duas línguas de alta familiaridade com Pbf's balanceadas entre palavras e parte-palavras (Línguas 5 e 6). Além disso, em uma das línguas, o estímulo [bara] foi utilizado como palavra de baixa frequência. O que permitiu

medir se o seu efeito individual continuaria forte, independentemente das distribuições de Pbf. O próximo experimento testou tais línguas.

Experimento 5

O presente experimento testou os efeitos da PT e da Pbf na extração das palavras de duas línguas de alta familiaridade (5 e 6, Tabela 1). A manipulação crítica envolveu testar palavras e parte-palavras que possuíam Pbf equilibrada (a mais alta possível para a Língua 5 e um pouco mais baixa para a Língua 6), sendo a diferença entre PT a única variável estatística que explicaria preferência por palavras na fase de teste. O uso de uma língua com Pbf maior (Língua 5) e outra com Pbf menor (Língua 6), ambas balanceadas, permitiu testar se o nível da Pbf influenciaria na preferência por palavras ou parte-palavras. Esperava-se que as palavras de ambas as línguas fossem preferidas durante a fase de teste, dada sua maior PT em comparação com parte-palavras.

Além disso, o estímulo [bara] foi uma palavra de baixa frequência na Língua 5, o que permitiu medir se os seus efeitos idiossincráticos, encontrados no Experimento 4, continuariam a influenciar a preferência por palavras quando as Pbf estivessem balanceadas entre os itens testados.

Método

A situação experimental, os equipamentos, o desenho experimental, o desenho de tentativas, o procedimento, e a análise de dados foram os mesmos do Experimento 4. As diferenças são descritas a seguir.

Participantes. Quarenta adultos ($M = 23$ anos e 1 mês, $SD = 6$ anos e 9 meses; 23 mulheres) participaram do estudo. Todos eram falantes nativos do PB e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no seu

desempenho na tarefa experimental. Dos 40 participantes, 33 eram estudantes da graduação (de 19 cursos diferentes) e 7 eram estudantes da pós-graduação (de 7 programas). Eles foram divididos aleatoriamente em dois grupos. Um grupo foi exposto à Língua 5 e o outro grupo foi exposto à Língua 6.

Estímulos experimentais. As Línguas 5 e 6 foram utilizadas (Tabela 1). Ambas foram geradas por um sintetizador e seguiram as características descritas no Capítulo 2. Criticamente, as Pbf's eram balanceadas entre palavras e parte-palavras testadas. Elas eram as maiores possíveis para a Língua 5 e um pouco mais baixas para a Língua 6.

Resultados e Discussão

A Figura 13 apresenta a média geral de acertos para cada língua (5 e 6). A probabilidade dos participantes escolherem as palavras em detrimento das parte-palavras quando expostos à Língua 5 ($M = 69,16\%$, $SD = 26,33\%$) foi consistentemente superior ao esperado pelo acaso ($b = 1,24$, Wald IC 95% [0,44; 2,05], $z = 3,057$, $p = 0,00224$). O mesmo foi encontrado para a probabilidade de preferência por palavras dos participantes expostos à Língua 6 ($M = 68,61\%$, $SD = 20,25\%$, $b = 0,978$, Wald IC 95% [0,12; 1,83], $z = 2,248$, $p = 0,0246$). Tais resultados indicam que quando a Pbf esteve balanceada entre os itens, a PT entre sílabas permitiu a extração das palavras da fala contínua, medida pela preferência por palavras na fase de teste. Além disso, o nível da Pbf (maior possível, Língua 5; um pouco mais baixo, Língua 6) não influenciou no desempenho ($p = 0,79334$)²⁶.

²⁶ lme4::glmer(acertos ~ offset(logit(acaso)) + versão da fala + (1 | participantes) + (1 | itens))

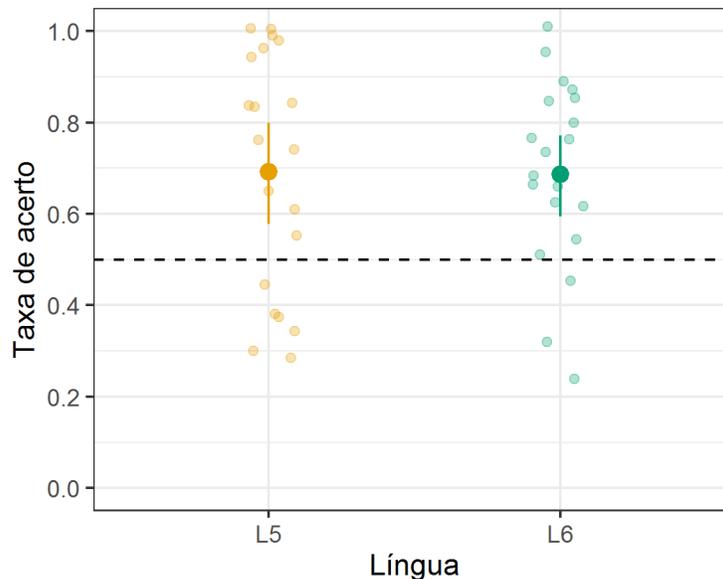


Figura 13. Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 5 e 6 do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,5).

A análise²⁷ dos efeitos de cada palavra sobre o desempenho encontrado na Língua 5 e 6 revelou efeitos de item em ambas as línguas. Para a Língua 5, diferentemente do previsto, a probabilidade de escolha da palavra [bara] não foi maior do que a probabilidade de escolha das demais palavras ($b = 1,27$, Intervalo Crível 95% [-0,74; 4,30]). Por outro lado, a probabilidade de escolha da palavra [sute] foi superior às probabilidades de escolha das demais palavras ($b = 1,70$, Intervalo Crível 95% [0,34; 3,37]). Para a Língua 6, a probabilidade de escolha da palavra [nipe] foi superior às probabilidades de escolha das demais palavras ($b = 1,78$, Intervalo Crível 95% [1,01; 2,72]). Já a probabilidade de rejeição da palavra [tad̥zi] foi maior do que as probabilidades de rejeição das demais palavras ($b = -1,59$, Intervalo Crível 95% [-2,51; -0,72]).

²⁷ brms::brm(acertos ~ palavras + (palavras | participantes))

Diferentemente dos resultados com as Línguas 3A e 3B (Experimento 4), não há diferenças sistemáticas nas Pbf's entre os itens que expliquem tais preferências. Além disso, quando tais palavras (i.e., [sute], [nipe], [tad̥zi]) tinham funções de parte-palavras nas Línguas 3B e 3A, respectivamente, elas não foram preferidas em detrimento dos demais estímulos. Futuras pesquisas poderão investigar se características idiossincráticas desses estímulos podem explicar tais achados.

Não obstante, a preferência geral por palavras em ambas as línguas, independentemente do nível das probabilidades fonotáticas de suas palavras e parte-palavras (mais alto ou mais baixo), indica que a PT entre as sílabas das palavras controlou sua preferência e extração. De modo amplo, os resultados replicam o relatado por Saffran et al. (1997). Apesar da presença de efeitos de itens não seja ideal, a Língua 6 foi utilizada na tarefa de extração e mapeamento simultâneo (Estudo 3, Capítulo 5).

A seguir serão descritos os experimentos (6, 7, 8, 9) realizados com falantes do inglês americano e com as línguas que respeitam as características do inglês (Tabela 1).

Experimento 6

O presente experimento testou os efeitos da PT na extração das palavras de duas línguas que seguem as características do inglês (7A e 7B, Tabela 1). Tais línguas foram construídas a partir da recombinação dos estímulos utilizados por Mirman et al. (2008) (Capítulo 2). Mirman e cols. utilizaram pseudopalavras (PT = 1) dissílabas para investigar a interação entre a extração de palavras e o mapeamento em situação não ambígua com adultos falantes do inglês. Nesse sentido, os estímulos utilizados pelos autores oferecem um ótimo ponto de partida para a criação de uma língua a ser usada tanto na extração (Estudo 1, controle) quanto na extração e mapeamento em situação ambígua (Estudo 3). Esperava-

se que as palavras (PT = 1) fossem preferidas em comparação com as parte-palavras (PT = 0,5) para ambas as línguas.

Método

Os equipamentos, o desenho experimental, o desenho de tentativas, o procedimento, e a análise de dados foram os mesmos do Experimento 5. As diferenças são descritas a seguir.

Participantes. Quarenta e dois adultos ($M = 18$ anos e 1 mês, $SD = 8$ meses; 29 mulheres) participaram do estudo. Todos eram falantes nativos do inglês e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no seu desempenho na tarefa experimental. Todos os 42 participantes eram estudantes da graduação (de 30 cursos diferentes). Eles foram divididos aleatoriamente em dois grupos. Um grupo foi exposto à Língua 7A e o outro grupo foi exposto à Língua 7B.

Situação experimental. O experimento foi conduzido em uma sala experimental do *Infant Language and Perceptual Learning Lab*, na *University of Tennessee*, com isolamento acústico. O arranjo experimental e equipamentos foram os mesmos dos experimentos com falantes do PB.

Estímulos experimentais. As Línguas 7A e 7B foram utilizadas (Tabela 1). Ambas foram geradas por um sintetizador e seguiram as características descritas no Capítulo 2.

Resultados e Discussão

A Figura 14 apresenta a média geral de acertos para cada Língua (7A e 7B). A probabilidade dos participantes escolherem as palavras em detrimento das parte-palavras quando expostos à Língua 7A ($M = 74,74\%$, $SD = 14\%$) foi consistentemente superior ao esperado pelo acaso ($b = 1,4189$, Wald IC 95% [0,14; 2,68], $z = 2,191$, $p = 0,0284$). Porém,

o mesmo não foi verdadeiro para os participantes expostos à Língua 7B ($M = 39,72\%$, $SD = 22,09\%$, $b = -0,4904$, Wald IC 95% $[-1; 0,02]$, $z = -1,861$, $p = 0,0627$).

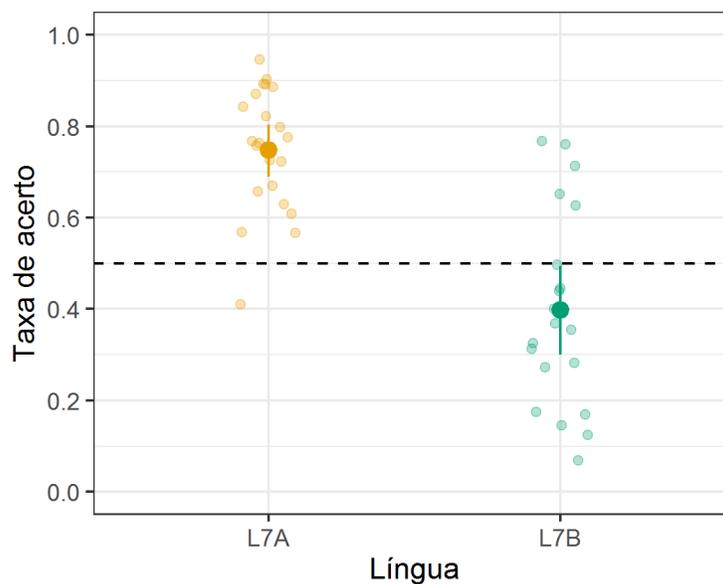


Figura 14. Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 7A e 7B do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,5).

Tais resultados indicam que o desempenho nas fases de familiarização e teste pode ter sido influenciado por outras variáveis que não a PT entre as sílabas das palavras e parte-palavras. A análise das probabilidades fonotáticas médias das palavras (0,0065) e parte-palavras (0,0085) para a Língua 7A (e o contrário para a Língua 7B) revelou um balanceamento razoável entre elas. Se tais diferenças tivessem exercido algum efeito sob a extração e preferência, um padrão contrário ao encontrado seria antecipado, já que as palavras tinham Pbf's médias mais baixas na Língua 7A e mais altas na 7B.

A análise²⁸ das preferências por cada palavra da Língua 7A revela que uma das palavras ([pæbo]) foi preferida consistentemente ($b = 2,70$, Intervalo Crível 95% [2,00; 3,55]), enquanto que as demais (i.e., [pibu], [dotu]) foram rejeitadas consistentemente ($b = -1,10$, Intervalo Crível 95% [-2,01; -0,22]; $b = -2,61$, Intervalo Crível 95% [-3,66; -1,63]; respectivamente). Nenhuma diferença nas preferências foi encontrada pela análise das parte-palavras. Por outro lado, nenhuma diferença na preferência foi encontrada para as palavras da Língua 7B, mas a probabilidade de escolha de uma das parte-palavras ([pæbo]) foi superior às demais ($b = -0,78$, Intervalo Crível 95% [-1,61; -0,06]).

Tais resultados indicam que ao menos um dos estímulos ([pæbo]) teve influência sistemática sob a preferência dos participantes de ambas as línguas. Futuras pesquisas poderão investigar se características idiossincráticas desses estímulos podem explicar tais achados.

Diante de tais resultados, duas novas línguas foram criadas (Língua 8A, 8B). A seleção das palavras dessas línguas seguiu a estratégia de seleção das Línguas 5 e 6 do PB (Capítulo 2). Suas palavras foram compostas por sílabas com alta Pbf no inglês e buscou-se balancear as Pbf entre palavras e parte-palavras testadas de modo que as PTs entre as sílabas informassem os limites das palavras. O próximo experimento testou tais línguas.

Experimento 7

O presente experimento testou os efeitos da PT na extração das palavras de duas línguas de alta familiaridade no inglês (8A e 8B, Tabela 1). Tais línguas foram construídas a partir da seleção e combinação de sílabas com alta Pbf no inglês (Capítulo 2). A gravação

²⁸ brms::brm(acertos ~ itens + (itens | participantes))

das línguas foi natural e todos os passos de controle estão descritos no Capítulo 2. Além disso, a Pbf dos estímulos foi balanceada entre palavras e parte-palavras, deixando a PT como única dica estatística para a extração. Logo, esperava-se que palavras ($PT = 1$) fossem preferidas em comparação com as parte-palavras ($PT = 0,5$) para ambas as línguas.

Método

A situação experimental, os equipamentos, o desenho experimental, o desenho de tentativas, o procedimento, e a análise de dados foram os mesmos do Experimento 6. As diferenças são descritas a seguir.

Participantes. Trinta adultos ($M = 19$ anos e 2 meses, $SD = 1$ ano; 23 mulheres) participaram do estudo. Todos eram falantes nativos do inglês e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no seu desempenho na tarefa experimental. Dos 30, 28 participantes eram estudantes da graduação (de 21 cursos diferentes), 1 era estudante de pós-graduação, e outro preferiu não responder ao questionário pós-experimental. Eles foram divididos aleatoriamente em dois grupos. Um grupo foi exposto à Língua 8A e o outro grupo foi exposto à Língua 8B.

Estímulos experimentais. As Línguas 8A e 8B foram utilizadas (Tabela 1). Ambas foram gravadas por um falante natural e seguiram as características descritas no Capítulo 2.

Resultados e Discussão

A Figura 15 apresenta a média geral de acertos para cada Língua (8A e 8B). A probabilidade de os participantes escolherem as palavras em detrimento das parte-palavras quando expostos à Língua 8A ($M = 60\%$, $SD = 16,29\%$) foi consistentemente superior ao esperado pelo acaso ($b = 0,426$, Wald IC 95% [0,06; 0,78], $z = 2,319$, $p = 0,0204$). Por outro lado, o contrário foi verdadeiro para os participantes expostos à Língua 8B ($M = 20\%$,

$SD = 17,28\%$): eles preferiram parte-palavras acima do esperado pelo acaso ($b = -1,7123$, Wald IC 95% $[-2,59; -0,83]$, $z = -3,806$, $p = 0,000141$).

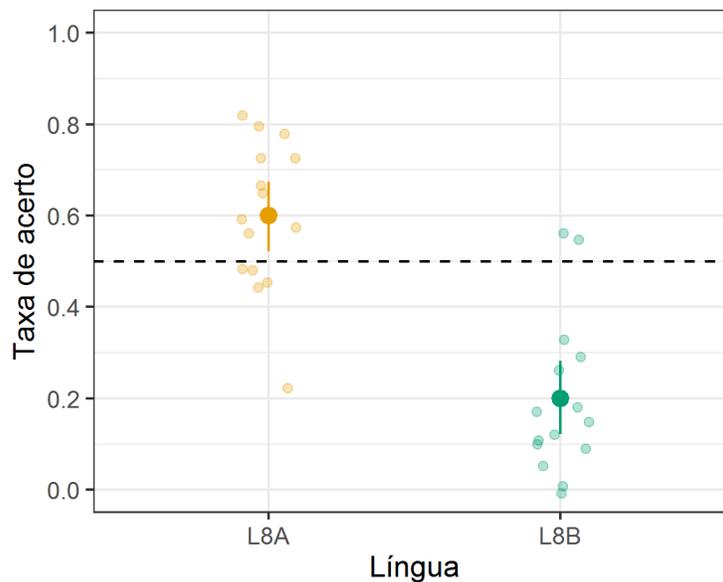


Figura 15. Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 8A e 8B do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,5).

A análise²⁹ das preferências por cada palavra da Língua 8A não revelou nenhuma probabilidade de preferência em particular. Por outro lado, a probabilidade de preferência da parte-palavra [keti] foi maior do que a probabilidade de preferência das demais parte-palavras ($b = -1,04$, Intervalo Crível 95% $[-2,08; -0,11]$). Já para a Língua 8B, uma palavra ([kɛnɔ]) foi rejeitada consistentemente ($b = -2,37$, Intervalo Crível 95% $[-3,47; -1,54]$) e uma parte-palavra ([mæsi]) foi preferida consistentemente ($b = -2,68$, Intervalo Crível 95% $[-4,04; -1,60]$). A ausência de um efeito sistemático dos itens entre as línguas indica que

²⁹ brms::brm(acertos ~ itens + (itens | participantes))

tais preferências podem estar sob controle de variáveis que advém da combinação das sílabas que formam as palavras e parte-palavras na Língua 8B ao invés das propriedades particulares dos estímulos em si. Futuras pesquisas poderão investigar se características idiossincráticas desses estímulos podem explicar tais achados.

Diante de tais resultados, e levando em conta os desafios da gravação natural dos estímulos, uma reanálise das gravações de ambas as línguas foi realizada. Durante a reanálise, foi identificado que um dos fones utilizados (i.e., [ɐ]) é ilegal no inglês. O seu uso gerou uma parte-palavra com fonotática ilegal na Língua 8A (i.e., [kɛnɔ]) e duas palavras ilegais na Língua 8B (i.e., [silɐ] e [kɛnɔ]). Tal status pode explicar a rejeição consistente da palavra [kɛnɔ] na Língua 8B. Por outro lado, esse mesmo item não foi rejeitado quando ele foi apresentado como parte-palavra na Língua 8A. Diferenças nas Pbfs dos itens não oferecem explicação para as demais preferências.

Considerando tais resultados e a presença de um fone ilegal nas línguas, duas novas línguas (9A, 9B) foram criadas pela substituição do fone ilegal por um fone legal. O Experimento 8 foi realizado para testar o efeito de tal substituição.

Experimento 8

O presente experimento teve como objetivo testar os efeitos da PT na extração das palavras de duas línguas de alta familiaridade no inglês (9A e 9B, Tabela 1). Criticamente, tais línguas foram construídas a partir da substituição do fone ilegal [ɐ] pelo fone legal [æ] e do fone [ɔ] pelo fone [o], presentes nas palavras e parte-palavras das Línguas 8A e 8B. Novamente, a gravação de tais fones foi feita por uma falante natural e todos os passos de controle descritos no Capítulo 2 foram seguidos. Tais substituições garantem que todas as palavras e parte-palavras de ambas as línguas possuam Pbfs equilibradas entre si. Nesse

sentido, esperava-se que as palavras ($PT = 1$) fossem preferidas em comparação com as parte-palavras ($PT = 0,5$) para ambas as línguas.

Método

A situação experimental, os equipamentos, o desenho experimental, o desenho de tentativas, o procedimento, e a análise de dados foram os mesmos do Experimento 7. As diferenças são descritas a seguir.

Participantes. Vinte adultos ($M = 19$ anos e 10 meses, $SD = 10$ meses; 14 mulheres) participaram do estudo. Todos eram falantes nativos do inglês e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no seu desempenho na tarefa experimental. Todos os participantes eram estudantes da graduação (de 17 cursos diferentes). Eles foram divididos aleatoriamente em dois grupos. Um grupo foi exposto à Língua 9A e o outro grupo foi exposto à Língua 9B.

Estímulos experimentais. As Línguas 9A e 9B foram utilizadas (Tabela 1). Ambas foram construídas a partir das Línguas 8A e 8B, foram gravadas por um falante natural, e seguiram as características descritas no Capítulo 2.

Resultados e Discussão

A Figura 16 apresenta a média geral de acertos para cada Língua (9A e 9B). A probabilidade de os participantes escolherem as palavras em detrimento das parte-palavras quando expostos à Língua 9A ($M = 45\%$, $SD = 13,21\%$) não foi diferente do esperado pelo acaso ($b = -0,2023$, Wald IC 95% $[-0,51; 0,11]$, $z = -1,256$, $p = 0,209$). Da mesma forma, os participantes expostos à Língua 9B também não escolheram palavras acima do esperado pelo acaso ($M = 30\%$, $SD = 13,65\%$). Pelo contrário, eles preferiram parte-palavras acima do esperado pelo acaso ($b = -0,8735$, Wald IC 95% $[-1,27; -0,47]$, $z = -4,287$, $p < 0,001$).

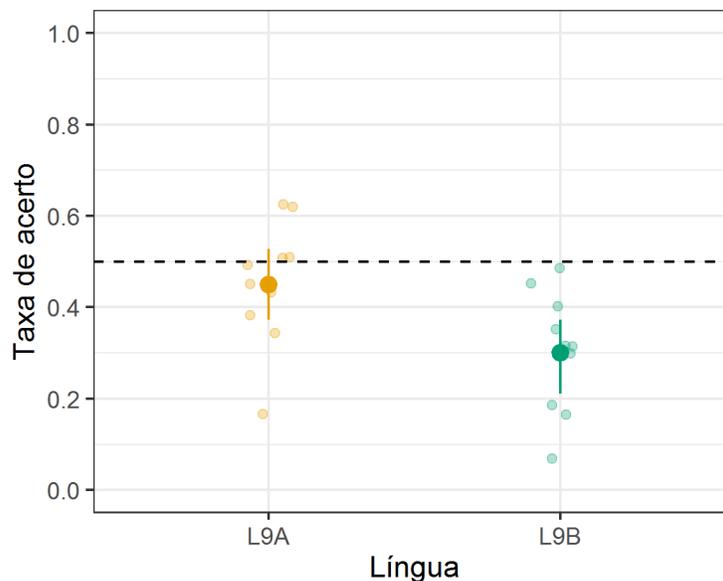


Figura 16. Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 9A e 9B do Estudo 1. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,5).

A análise³⁰ das preferências por cada palavra e parte-palavra da Língua 9A não revelou nenhuma probabilidade de preferência em particular. Por outro lado, os participantes expostos à Língua 8B rejeitaram consistentemente a palavra [kæno] ($b = -1,26$, Intervalo Crível 95% [-2,27; -0,42]). Além disso, eles demonstram preferência consistente pela parte-palavra [mæsi] ($b = -1,44$, Intervalo Crível 95% [-3,14; -0,14]). A ausência de um efeito sistemático dos itens entre as línguas indica que as preferências por cada item podem estar sob controle de um conjunto de variáveis distinto. Futuras pesquisas poderão investigar se características idiossincráticas desses estímulos podem explicar tais achados.

³⁰ `brms::brm(acertos ~ itens + (itens | participantes))`

Diante dos resultados, duas novas línguas foram construídas (10A e 10B). Buscando um maior controle das propriedades fonéticas dos estímulos, um sintetizador foi utilizado na produção dessas línguas. Além disso, uma nova amostra de palavras e parte-palavras, com Pbfs balanceadas e de alta familiaridade, foi selecionada. O Experimento 9 testou tais línguas.

Experimento 9

O presente experimento testou os efeitos da PT na extração das palavras de duas novas línguas de alta familiaridade no inglês (10A e 10B, Tabela 1). Tais línguas tiveram palavras e parte-palavras com Pbfs balanceadas na fase de teste, além de contrastarem palavras com não-palavras ($PT = 0$). A inclusão de não-palavras aumenta a sensibilidade do teste de extração. Se por um lado as altas Pbfs das palavras e parte-palavras podem ser combinadas com a PT para a promover a extração, por outro, o mesmo não será verdadeiro com não-palavras ($PT = 0$). Nesse sentido, palavras devem ser preferidas consistentemente quando comparadas com não-palavras, mesmo que não o sejam quando comparadas a parte-palavras. A gravação das novas línguas foi feita por meio de um sintetizador de fala, de modo a controlar sistematicamente suas características fonéticas. Esperava-se que as palavras ($PT = 1$) fossem preferidas em comparação com as parte-palavras ($PT = 0,5$) e não-palavras ($PT = 0$) para ambas as línguas.

Método

A situação experimental, os equipamentos, o desenho experimental, o desenho de tentativas, o procedimento, e a análise de dados foram os mesmos do Experimento 8. As diferenças são descritas a seguir.

Participantes. Vinte adultos ($M = 18$ anos e 9 meses, $SD = 6$ meses; 9 mulheres) participaram do estudo. Todos eram falantes nativos do inglês e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no seu desempenho na tarefa experimental. Todos os participantes eram estudantes da graduação (de 17 cursos diferentes). Eles foram divididos aleatoriamente em dois grupos. Um grupo foi exposto à Língua 10A e o outro grupo foi exposto à Língua 10B.

Estímulos experimentais. As Línguas 10A e 10B foram utilizadas (Tabela 1). Ambas foram gravadas por um sintetizador e seguiram as características descritas no Capítulo 2.

Resultados e Discussão

A Figura 17 apresenta a média geral de acertos para cada Língua (10A e 10B), separado por tipo de estímulo distrator (não-palavra ou parte-palavra). No geral, a probabilidade dos participantes escolherem as palavras no teste, quando expostos à língua, 10A ($M = 56,66\%$, $SD = 15,22\%$) não foi diferente do esperado pelo acaso ($b = 0,7212$, Wald IC 95% [-0,16; 1,60], $z = 1,598$, $p = 0,11$)³¹. Não obstante, o tipo de estímulo distrator influenciou significativamente a probabilidade de preferência das palavras. Quando palavras foram contrastadas com parte-palavras, a probabilidade de preferência por palavras abaixou significativamente ($b = -0,8293$, Wald IC 95% [-1,47; -0,18], $z = -2,527$, $p = 0,0115$). Quando as parte-palavras foram excluídas da análise, a preferência por palavras ($M = 65\%$, $SD = 14\%$) foi significativamente superior ao esperado pelo acaso ($b = 0,6587$, Wald IC 95% [0,09; 1,22], $z = 2,281$, $p = 0,0225$).

³¹ lme4::glmer(acertos ~ offset(logit(acaso)) + tipo de distrator + (1 | participante) + (1 | itens))

No geral, os participantes expostos à Língua 10B escolheram palavras acima do esperado pelo acaso ($M = 72,22\%$, $SD = 25,25\%$, $b = 1,6933$, Wald IC 95% [0,60; 2,78], $z = 3,052$, $p = 0,00227$). Além disso, uma tendência a um efeito significativo foi encontrada para as parte-palavras na diminuição da preferência por palavras ($b = -0,7504$, Wald IC 95% [-1,51; 0,01], $z = -1,932$, $p = 0,05332$).

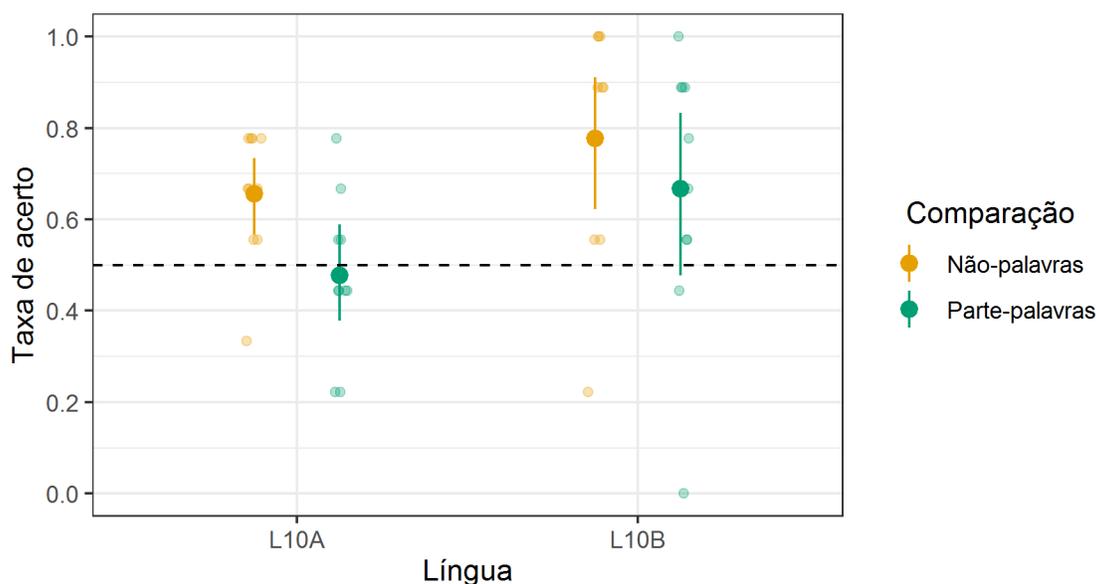


Figura 17. Taxa geral de acerto para cada participante e média geral de acertos para as Línguas 10A e 10B (Estudo 1) por item de comparação (não-palavra ou parte-palavra). As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,5).

A análise³² das preferências por cada palavra da língua 10A revelou uma preferência consistente pela palavra [kæni] ($b = 0,8773$, Wald IC 95% [0,25; 1,49], $z = 2,780$, $p = 0,00543$) e uma rejeição consistente da palavra [potu] ($b = -1,4362$, Wald IC

³² O modelo frequentista: `lme4::glmer(acertos ~ palavras + (palavras | participantes))` convergiu na análise da preferência pelas palavras da Língua 10A. Para as demais análises, o modelo bayesiano: `brms::brm(acertos ~ itens + (itens | participantes))` foi utilizado.

95% [-2,21; -0,65], $z = -3,598$, $p = 0,00032$). Já a análise das preferências dos participantes expostos à língua 10B não revelou nenhuma preferência particular por nenhuma palavra ou parte-palavra. Dada a ausência de um efeito sistemático de um dos itens entre as línguas, é possível que a preferência ou rejeição encontrada para tais itens se deva a um conjunto de variáveis advindo da combinação das sílabas de palavras e parte-palavras.

O conjunto dos resultados indica que os participantes extraíram as palavras da fala contínua, replicando os resultados relatados na literatura (e.g., Saffran et al., 1997). Por outro lado, a diminuição da preferência quando parte-palavras serviram de estímulo comparação nas tentativas teste indicam que: a) o conjunto de PT (0,5) e Pbf foi suficiente para que tais sequências silábicas fossem tratadas como possíveis “palavras” da língua recém-ouvida; e b) a diferença de 0,5 na PT entre palavras e parte-palavras não gerou um representação forte o suficiente para controlar a preferência na fase de teste.

Diante de tais resultados, as Línguas 10A e 10B foram utilizadas na coleta de dados do Estudo 3, que testou a extração e mapeamento simultâneo, com falantes do inglês. Apesar de não ser o cenário ideal, se o modelo de processamento defendido por Räsänen e Rasilo (2015) for robusto, é razoável esperar que, apesar de não serem extraídas em com base em PT apenas, as palavras fossem extraídas em uma tarefa de extração e mapeamento conjunto, com base em PT mais coocorrências com referentes.

Discussão Geral

O presente estudo, composto pelos experimentos 1 a 9, buscou investigar se palavras podiam ser extraídas de falas contínuas com base da PT entre suas sílabas. Os resultados com falantes do PB (Experimentos 1 a 5) demonstraram que palavras podem ser extraídas com base na PT entre suas sílabas, porém, apenas quando as Pbf das palavras e

parte-palavras são balanceadas (Experimento 5). Por outro lado, os resultados com falantes do inglês (Experimentos 6 a 9) demonstram que a combinação de uma alta Pbf com uma alta PT (em comparação com a PT na língua natural) pode ser suficiente para que palavras (PT = 1) e parte-palavras (PT = 0,5) sejam julgadas como bons candidatos a palavras/rótulos da nova língua (Experimento 9).

Embora as Pbf tenham influenciado o desempenho dos participantes brasileiros e norte-americanos, é possível que diferenças nos algoritmos de cálculo das Pbf tenham gerado estímulos com diferentes graus de familiaridade para cada população. O cálculo da Pbf para o PB se baseou na frequência de *types* (Estivalet & Dal Ben, em preparação; Estivalet & Meunier, 2015). Por outro lado, o cálculo da Pbf no inglês foi feito com base na frequência *token* das palavras (Vitevitch & Luce, 2004). Dada a maior sensibilidade da medida *token*, é razoável especular que as palavras que compunham as línguas do inglês talvez fossem mais familiares do que as palavras das línguas do PB. Essa diferença pode oferecer uma possível explicação para a preferência por palavras (PT = 1) e parte-palavras (PT = 0,5) demonstrada pelos norte-americanos, mesmo quando a Pbf estava balanceada entre itens (Experimento 9).

A interação entre PT e Pbf foi um achado inesperado no presente estudo. Futuras investigações que testem a interação entre PT e Pbf parametricamente podem ser especialmente informativas sobre a interação de tais regularidades (cf. Mersad & Nazzi, 2011). Considerando que o objetivo principal do presente estudo é garantir que as palavras utilizadas nas línguas inventadas (que seguem as características do PB e do inglês) podem ser extraídas com base na PT de suas sílabas, as Línguas 6 do PB e 10 (A e B) do inglês foram selecionadas para o teste da extração e mapeamento ao longo de tentativas ambíguas em paralelo.

O teste prévio de tais línguas foi especialmente importante para os falantes do PB, dada a ausência de dados prévios sobre extração com essa população. Nessa linha, antes de prosseguir para o teste conjunto da extração e mapeamento (Estudo 3), é importante garantir que as palavras e parte-palavras que compõem as línguas de alta familiaridade do PB podem ser mapeadas com referentes (estímulos visuais) ao longo de tentativas ambíguas. O próximo estudo faz esse teste.

Capítulo 4

Estudo 2 – Mapeamento de palavras ao longo de tentativas ambíguas

Contexto

O primeiro estudo demonstrou que as palavras da língua de alta familiaridade do PB podem ser extraídas de falas contínuas com base nas PTs de suas sílabas quando suas probabilidades fonotáticas são balanceadas (Estudo 1, Experimento 5, Línguas 5 e 6). O presente estudo teve como objetivo verificar se os estímulos (palavras e parte-palavras) que formam as línguas de alta familiaridade do PB, utilizadas no primeiro estudo, podem ser mapeadas a estímulos visuais abstratos ao longo de tentativas ambíguas (Yu & Smith, 2007). Assim como no Estudo 1, o presente estudo é relevante pelo menos por duas razões. Primeiro, esse é o primeiro estudo a investigar o processo de mapeamento ao longo de tentativas ambíguas com falantes do PB. Segundo, os dados do presente estudo servem como linha de base para a investigação dos processos de extração e mapeamento em uma tarefa simultânea, que responde a pergunta de pesquisa da presente investigação (Estudo 3).

Dois experimentos foram conduzidos. No primeiro experimento, 12 estímulos auditivos foram pareados com 12 estímulos visuais e tais pares foram divididos em dois conjuntos. Somente pares do mesmo conjunto foram apresentados juntos ao longo das tentativas. Não obstante, ambos os conjuntos foram apresentados em paralelo durante a fase de familiarização e foram testados em seguida. Tal divisão de conjuntos mimetiza a diversidade contextual do próximo Estudo (3; extração e mapeamento simultâneo), no qual apenas seis palavras foram pareadas com 6 estímulos visuais.

No segundo experimento, buscando explorar os possíveis efeitos do rastreamento simultâneo de dois conjuntos independentes de pares, as fases de familiarização e teste foram divididas em dois blocos. Cada conjunto foi apresentado separadamente em um dos

blocos, de modo sequencial. Logo, a diversidade contextual em ambos os experimentos foi a mesma, porém, a concorrência entre os conjuntos, presentes no primeiro experimento, não esteve presente no segundo experimento.

Experimento 1

Método

Participantes. Trinta adultos ($M = 22$ anos e 2 meses, $SD = 4$ anos e 6 meses; 25 mulheres) participaram do estudo. Todos eram falantes nativos do PB e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no desempenho na tarefa experimental. Todos os participantes eram estudantes universitários, 22 eram estudantes da graduação (distribuídos entre 9 cursos) e 8 eram estudantes da pós-graduação (distribuídos entre 6 programas). Eles foram distribuídos aleatoriamente em dois grupos (15 participantes cada).

Situação experimental e equipamentos. O experimento foi conduzido na mesma sala utilizada para o primeiro estudo. Assim como no primeiro estudo, as tarefas experimentais foram apresentadas em um computador com sistema operacional Windows por meio do software *Psychopy2* (Peirce et al., 2019). Além dos equipamentos de áudio e de registro de respostas utilizados no primeiro estudo, uma câmera digital Logitech C920 foi posicionada acima do monitor, aproximadamente na altura dos olhos dos participantes, de modo a capturar as direções de seus olhares. As imagens foram capturadas em 720p (alta definição) a 30fps pelo software *Open Broadcaster Software*.

Estímulos experimentais. Os estímulos auditivos (palavras e parte-palavras da Língua 3A e 4, Tabela 1) foram combinados randomicamente com figuras abstratas (Tabela 2), formando 12 pares. Em seguida, eles foram divididos em dois conjuntos. Os estímulos

do primeiro conjunto (i.e., [d̥ʒini], [deta], [pemi], [sute], [viko], [bara]; $M = 0,037$) tinham Pbfs mais altas do que e os estímulos do segundo conjunto (i.e., [nipe], [taḍʒi], [mide], [teba], [kosu], [ravi]; $M = 0,017$). Buscando evitar qualquer relação espúria entre os estímulos auditivos e visuais, novos pares foram construídos pela inversão das combinações entre conjuntos.

Desenho experimental e desenho de tentativas. O experimento teve duas fases: familiarização e teste (Figura 18). Na familiarização, foi utilizada a configuração 2 x 2, com dois pares apresentados a cada tentativa. Cada par foi apresentado seis vezes, resultando em 36 tentativas. Pares de um conjunto sempre foram apresentados com pares do mesmo conjunto, e nunca com pares do outro conjunto. Nesse sentido, a familiarização e o teste eram compostos pela ocorrência de dois conjuntos independentes, cada um com 6 pares. Todos os pares de um conjunto foram combinados entre si, gerando cinco combinações. Quatro combinações coocorreram uma vez (probabilidade de 0,17 cada) e uma combinação (balanceada entre os pares) coocorreu duas vezes (probabilidade de 0,33).

Seguindo a configuração 2 x 2, cada tentativa iniciava com duas figuras apresentadas na tela. Após 950 ms, um estímulo auditivo era tocado (≈ 696 ms), e após uma pausa de 700 ms o segundo estímulo auditivo era tocado (≈ 696 ms), seguido de uma pausa de 950 ms. A duração total de 4 segundos é 2 segundos mais rápida do que a utilizada por Yu e Smith (2007), porém, ela foi preferida por ser mais próxima da duração prevista para cada tentativa do Estudo 3 (1,392 ms). Além disso, ela está em linha com outras pesquisas da área (e.g., Smith & Yu, 2008; Yu & Smith, 2011). A ordem de apresentação dos estímulos auditivos em relação à posição dos estímulos visuais foi randomizada, evitando qualquer correspondência sistemática. Entre as tentativas, um ponto de fixação era apresentado no centro da tela por 1s, de modo a induzir o participante a olhar para o centro

da tela no começo de cada tentativa. Os participantes não foram expressamente instruídos a olhar para o ponto, porém, análises das filmagens indicaram que eles o fizeram.

Na fase teste, um procedimento de escolha forçada entre quatro alternativas foi utilizado. Cada par (12) foi testado duas vezes (total de 24 tentativas). Cada tentativa apresentava quatro figuras, uma figura-alvo (probabilidade de coocorrência com o estímulo auditivo alvo = 1,0) e três distratoras (probabilidade de coocorrência = 0,17). Após 1 s, um estímulo auditivo era tocado. O participante indicava qual dentre as figuras correspondia ao estímulo auditivo. Não havia limite de tempo para a escolha.

Entre a familiarização e o teste, um aquecimento com duas tentativas com estímulos conhecidos dos participantes (i.e., bola, casa, gato, pato) foi realizado para aquecer os participantes com a estrutura e funcionamento do teste.

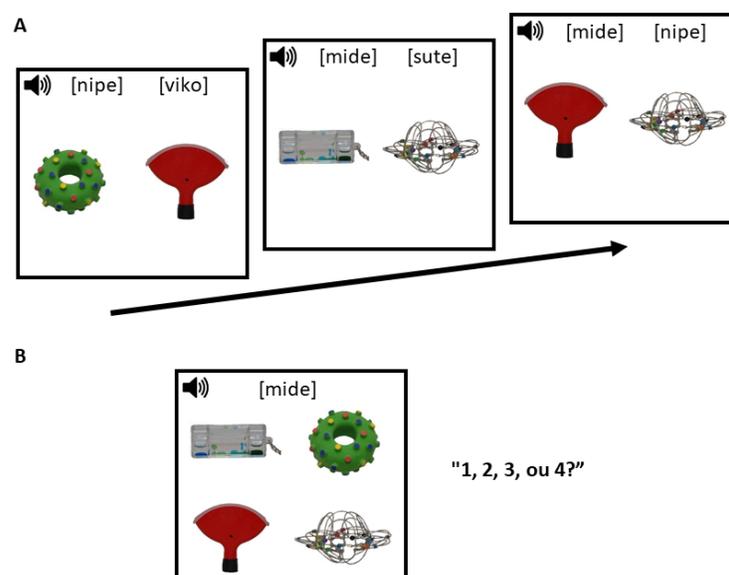


Figura 18. (A) Fase de familiarização (2 x 2) e (B) Teste (escolha forçada entre 4 alternativas) do Estudo 2.

Procedimento. Cada participante foi exposto individualmente às tarefas experimentais. No início do experimento, era apresentada uma música (com a mesma intensidade dos estímulos experimentais, i.e., 77dB) e o participante era instruído³³ a ajustar o volume do áudio. Em seguida, ele era informado de que ouviria novas palavras e veria novas figuras, inventadas, e que sua tarefa era descobrir a relação entre as palavras e as figuras. Ele não era informado de que cada palavra correspondia apenas a uma figura. Após a familiarização, um aquecimento para o teste com dois pares conhecidos (e.g., “pato” + foto de um pato) era realizado. O participante era instruído a selecionar a figura correspondente ao nome falado pressionando os números 1, 2, 3, ou 4. Em seguida, no teste, ele era instruído a fazer o mesmo para os estímulos experimentais. Finalizado o teste, o participante era instruído a estimar o seu desempenho experimental. Ele deveria pressionar 1 se estimasse que havia acertado entre 0 e 25% das tentativas, 2 para 25 e 50%, 3 para 50 e 75%, e 4 para 75 e 100%. Após concluir as tarefas experimentais, o participante respondia um questionário contendo questões sobre o seu conhecimento linguístico e limitações físicas que poderiam ter dificultado a sua participação (Apêndice D).

Análise de dados. A principal variável dependente foi a preferência por figuras correspondentes aos estímulos auditivos apresentados durante as tentativas de teste. A preferência foi atestada contra o nível do acaso (25%) por meio de modelos mistos generalizados que, sempre que possível, tiveram estrutura máxima dos efeitos randômicos (Barr et al., 2013; Bayeen et al., 2008)³⁴. Quando os modelos frequentistas com estrutura

³³ Todas as instruções estão disponíveis no Apêndice I. Vale ressaltar que o número de instruções e o detalhamento das mesmas seguiu o descrito por Yu e Smith (2007).

³⁴ Sintaxe (R): `lme4::glmer(respostas ~ offset(logit(nível do acaso)) + (1 | participantes) + (1 | itens auditivos) + (1 | itens visuais))`

máxima falharam em convergir, modelos mistos bayesianos foram utilizados (Bürkner, 2018).

Variáveis dependentes secundárias incluíram a latência de acertos e erros, a estimativa de desempenho (auto-relato) e o conhecimento linguístico de cada participante. Essas variáveis não foram analisadas, uma vez que não foram encontradas diferenças individuais na preferência por palavras ou parte-palavras.

Resultados e Discussão

A Figura 19 apresenta a média geral de acertos. A probabilidade dos participantes escolherem a figura-alvo em detrimento das figuras distratoras ($M = 65\%$, $SD = 24,57\%$) foi consistentemente superior ao esperado pelo acaso (0,25) ($b = 2,0836$, Wald IC 95% [1,42; 2,73], $z = 6,242$, $p < 0,001$). Tais resultados indicam que os participantes foram capazes de aprender as relações entre estímulos auditivos e visuais, apesar do baixo número de combinações (6 pares em cada conjunto) e da presença de dois conjuntos independentes apresentados paralelamente na mesma tarefa.

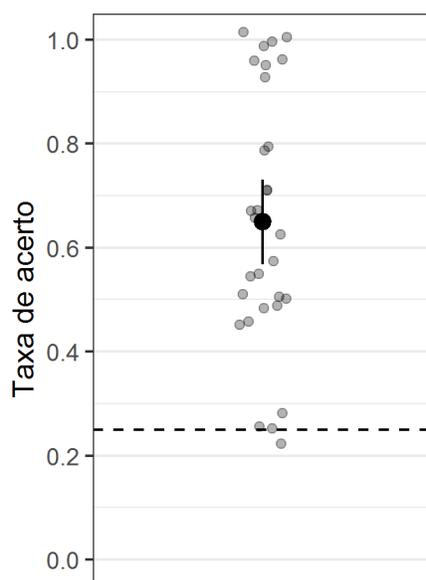


Figura 19. Taxa geral de acerto para cada participante e média geral de acertos para o Experimento 1 do Estudo 2. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,25).

A análise³⁵ das preferências por cada estímulo auditivo revelou uma preferência consistente por [bara] ($b = 2,98$, Intervalo Crível 95% [0,69; 8,77]), [kosu] ($b = 9,70$, Intervalo Crível 95% [1,85; 28,45]), e [ravi] ($b = 3,15$, Intervalo Crível 95% [0,13; 9,14]). Dificilmente, as Pbf's dos estímulos explicam tais preferências, dadas as suas diferenças (0,035; 0,029; 0,013; respectivamente). Futuras pesquisas poderão investigar se características idiossincráticas desses estímulos podem explicar tais achados.

O conjunto dos resultados indica que os participantes aprenderam as relações entre estímulos auditivos e visuais ao longo de tentativas ambíguas (CSWL) de ambos os conjuntos. Tal resultado replica estudos anteriores (e.g., Yu & Smith, 2007) e garante que

³⁵ O modelo bayesiano: `brms::brm(acertos ~ palavras + (palavras | participantes))` foi utilizado.

os estímulos auditivos da língua de alta familiaridade do PB e os estímulos visuais selecionados podem ser mapeados ao longo de tentativas ambíguas.

Apesar dos resultados em linha com o relatado na literatura, nenhum outro estudo, que seja do conhecimento do pesquisador, usou tão poucos pares combinados entre si (6 pares) ou usou dois conjuntos independentes apresentados paralelamente ao longo da familiarização. O uso de poucos pares gera uma baixa diversidade contextual, o que pode dificultar a aprendizagem das relações (Suanda, Mugwanya, & Nami, 2014). Somado a isso, a presença de dois conjuntos independentes de pares na mesma tarefa pode ter gerado uma competição nos processos de atenção e memória (Vlach & DeBrook, 2017). Buscando explorar essas questões, um segundo experimento foi realizado.

Experimento 2

Com base nos resultados do experimento anterior, o presente experimento buscou investigar os possíveis efeitos da presença simultânea de dois conjuntos pequenos (6 pares cada) sendo apresentados paralelamente ao longo de tentativas ambíguas. Para tanto, a tarefa experimental foi dividida em dois blocos. Cada bloco apresentou e testou um conjunto de pares. Se a diversidade contextual baixa (6 pares por Conjunto) influenciou o desempenho encontrado no Experimento 1, então, o desempenho deveria ser similar ao encontrado anteriormente e não deveria haver diferença entre os blocos. Se a presença de dois conjuntos em uma mesma tarefa influenciou o desempenho encontrado no Experimento 1, então, esperava-se que o desempenho no presente experimento fosse superior ao do experimento anterior, porém, sem diferença entre blocos.

Método

Participantes. Doze adultos ($M = 21$ anos e 11 meses, $SD = 3$ anos; 7 mulheres) participaram do estudo presencialmente e outros 32 adultos ($M = 23$ anos e 6 meses, $SD = 4$ anos e 6 meses; 25 mulheres) participaram do estudo à distância. A seleção dos participantes à distância foi similar à seleção dos participantes presenciais. Eles foram recrutados por meio de postagens no grupo de alunos da UFSCar no Facebook; as postagens continham um link para o termo de consentimento livre e esclarecido e para a tarefa experimental.

Todos os participantes eram falantes nativos do PB e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no seu desempenho na tarefa experimental. Eles foram distribuídos aleatoriamente em dois grupos (Grupo 1, $N = 27$; Grupo 2, $N = 17$).

Situação experimental. O mesmo local e equipamentos utilizados no Experimento 1 foram utilizados para a coleta presencial. A coleta à distância foi realizada por meio do software *Psychopy3* (Peirce et al., 2019), hospedado no servidor *Pavlovia*³⁶.

Desenho experimental e desenho de tentativas. Os Conjuntos de pares 1 e 2 foram apresentados e testados sequencialmente. Para o Grupo 1, o primeiro bloco apresentou o Conjunto 1, formado por pares com estímulos auditivos com Pbf mais alta ($M = 0,037$) e o segundo bloco apresentou pares do Conjunto 2 (Pbf mais baixa; $M = 0,017$). O inverso ocorreu para o Grupo 2. A familiarização e o teste seguiram as mesmas configurações (2 x 2, e 4 x 1, respectivamente) e o mesmo número de tentativas por par (6 e 2, respectivamente) do Experimento 1.

³⁶ <https://pavlovia.org/> ; acessado em 03 de Fevereiro de 2020.

Procedimento. O procedimento foi o mesmo do primeiro experimento, com a diferença de que o aquecimento para o teste, com estímulos conhecidos (e.g., pato, bola, gato, casa), foi realizado apenas no primeiro bloco.

Análise de dados. A análise de dados seguiu a mesma estrutura da análise do Experimento 1³⁷.

Resultados e Discussão

A Figura 20 apresenta a média geral de acertos para cada bloco (A), e a média de acertos por local de coleta e ordem de apresentação dos conjuntos (B). No geral, as escolhas dos referentes alvos no primeiro bloco ($M = 64,96\%$, $SD = 23,88\%$) e no segundo bloco ($M = 71,4\%$, $SD = 27,86\%$) foram acima do esperado pelo acaso (0,25) ($b = 1,2966$, Wald IC 95% [0,26; 2,33], $z = 2,455$, $p = 0,0141$). Além disso, a acurácia aumentou entre o primeiro e o segundo bloco ($b = 0,5764$, Wald IC 95% [0,05; 1,10], $z = 2,146$, $p = 0,0319$). Crucialmente, local (laboratório ou à distância) e ordem de apresentação dos Conjuntos (1-2 ou 2-1) não influenciaram o desempenho ($p = ns$).

³⁷ lme4::glmer(acertos ~ offset(logit(nível do acaso)) + blocos + ordem + local + (blocos | participantes))

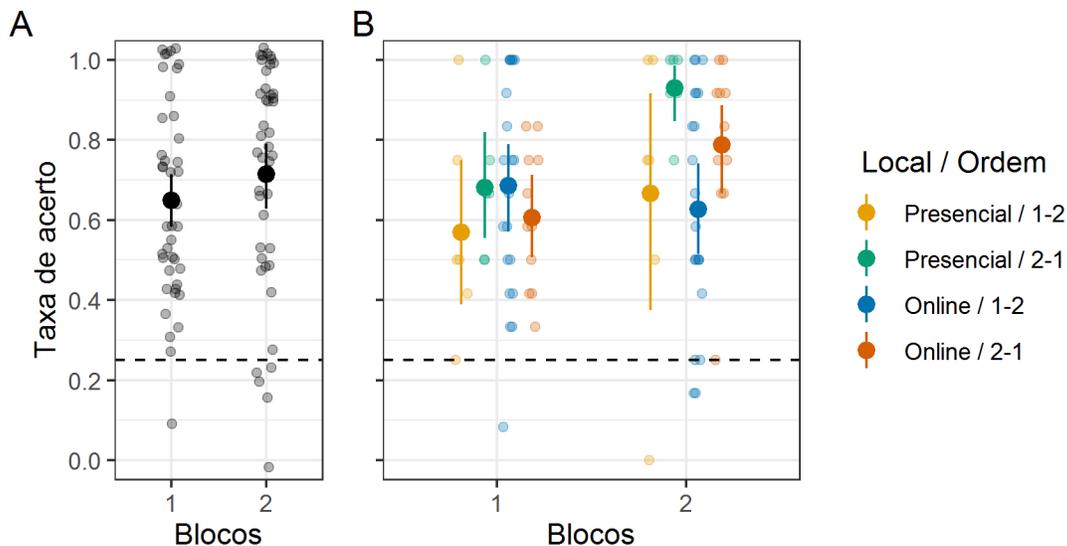


Figura 20. (A) Taxa geral de acerto para cada participante e média geral de acertos para cada bloco do Experimento 2 do Estudo 2; (B) Taxa geral de acerto para cada participante e média geral de acertos para cada bloco por local de coleta e ordem dos conjuntos. A ordem “1-2” indica que o Conjunto 1 (com estímulos com a probabilidade fonotática mais alta) era apresentado no primeiro bloco; o inverso é verdadeiro para a ordem “2-1”. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e a linha pontilhada indica o nível do acaso (0,25).

Os resultados comparáveis entre o primeiro bloco do presente experimento ($M = 64,96\%$, $SD = 23,88\%$) e o encontrado no experimento anterior ($M = 65\%$, $SD = 24,57\%$) indicam que a presença de dois conjuntos de estímulos independentes na mesma tarefa de familiarização não influenciou a probabilidade de escolha dos referentes alvos durante o Experimento 1. Além disso, o aumento na acurácia entre blocos sugere que a diversidade contextual pode não ter influenciado significativamente a habilidade de mapeamento dos pares. Ela sugere, por outro lado, um efeito de *Learning Set* (Harlow, 1949). É provável que a exposição à fase de teste entre os blocos tenha deixado explícita a estrutura da tarefa

(cada palavra correspondendo a um referente), o que pode ter contribuído para o melhor desempenho no segundo bloco.

Embora a diferença não tenha sido significativa, a inspeção visual indica que o desempenho médio dos participantes que foram expostos ao Conjunto 2 (palavras com Pbf's menores do que as do Conjunto 1) no primeiro bloco parece ter sido mais acurado durante o segundo bloco, quando eles foram expostos às palavras com Pbf mais alta. Tal aumento na acurácia pode estar relacionado com a maior facilidade de mapeamento de palavras com alta Pbf (Graf-Estes, 2014) e aos efeitos de *Learning Set* (Harlow, 1949). Isso porque, se a Pbf fosse a única variável explicativa, o desempenho deveria ser maior no primeiro bloco e menor no segundo para os participantes expostos ao Conjunto 1 no primeiro bloco. Não obstante, vale ressaltar que essas são especulações a partir de inspeções visuais e que não houve diferenças significativas entre os desempenhos. Futuras pesquisas com amostras maiores podem trazer mais dados sobre a relação entre Pbf e *Learning Set* na aprendizagem ao longo de tentativas ambíguas.

Discussão Geral

O presente estudo, composto pelos experimentos 1 e 2, buscou investigar se estímulos auditivos de alta familiaridade do PB (extraídos de discursos contínuos no Estudo 1) poderiam ser mapeados a estímulos visuais abstratos ao longo de tentativas ambíguas. O conjunto dos resultados do primeiro e do segundo experimento indica que os estímulos auditivos e visuais podem ser mapeados ao longo de tentativas ambíguas, e isso acontece mesmo quando os dois conjuntos de estímulos são apresentados na mesma tarefa de modo paralelo. Além disso, a separação em dois blocos sucessivos aumentou a taxa de acerto no segundo bloco, provavelmente por um efeito de *Learning Set* (Harlow, 1949).

Tais resultados fornecem a primeira demonstração de que falantes do PB rastreiam coocorrências para mapear palavras e figuras ao longo de tentativas ambíguas. Além de somar para a generalidade de tal processo de aprendizagem estatística, os resultados também servem de linha de base para a investigação conjunta dos processos de extração e mapeamento ao longo de tentativas ambíguas em paralelo, próximo estudo.

O presente estudo não incluiu participantes norte-americanos dada a existência de uma literatura madura em *cross-situational word learning* com falantes do inglês (mais de 50 estudos empíricos; Dal Ben & Souza, em preparação). Tal literatura e a replicação sem maiores surpresas com falantes do PB embasou a decisão de concentrar os esforços de coleta de dados com os falantes do PB.

Capítulo 5

Estudo 3 – Extração e mapeamento ao longo de tentativas ambíguas em paralelo

Contexto

O Estudo 3 investigou o processo de extração de palavras de falas contínuas e mapeamento ao longo de tentativas ambíguas em paralelo e busca responder à pergunta de pesquisa que guia a presente investigação: *Palavras são extraídas de falas contínuas e mapeadas a estímulos visuais simultaneamente ao longo de tentativas ambíguas?*

Vale lembrar que o presente estudo é construído a partir dos dois estudos anteriores. O primeiro estudo demonstrou que as palavras selecionadas para as línguas de alta familiaridade do PB e do inglês podem ser extraídas de falas contínuas com base na PT de suas sílabas (Língua 6, Experimento 5, e Língua 10, Experimento 9; Estudo 1). O segundo estudo demonstrou que as palavras e parte-palavras das línguas de alta familiaridade do PB podem ser mapeadas com estímulos visuais ao longo de tentativas ambíguas, mesmo quando a diversidade contextual é baixa (6 pares). Agora, os dois processos, extração e mapeamento CSWL, são unidos em uma única tarefa. Nenhuma investigação empírica do conhecimento do pesquisador testou tal interação.

Assim como nos estudos anteriores, os experimentos do presente estudo também foram divididos em duas fases. Na fase de familiarização, uma fala contínua foi apresentada junto com uma série de imagens. O arranjo foi feito de tal modo que duas palavras da fala foram sempre apresentadas com duas figuras, seguindo o arranjo 2 x 2 utilizado no Estudo 2 (mapeamento ao longo de tentativas ambíguas apenas). A coocorrência entre as sílabas das palavras (PT = 1) possibilitou a extração das palavras enquanto a coocorrência entre palavras e figuras possibilitou o mapeamento. A fase de teste foi composta por duas tarefas. Primeiro, a extração das palavras foi testada por meio de

tentativas de escolha forçada entre palavras e parte-palavras ou não-palavras. Segundo, o mapeamento foi testado por meio de tentativas de escolha forçada entre quatro figuras.

Em linha com as simulações computacionais de Räsänen e Rasilo (2015), esperava-se que a extração e o mapeamento ocorressem em paralelo; logo, a hipótese era a de que palavras seriam preferidas no teste de extração e figuras-alvo seriam preferidas no teste de mapeamento. Além disso, ainda em linha com as propostas de Räsänen e Rasilo (2015), esperava-se que o desempenho quando a extração e o mapeamento foram feitos em conjunto fosse melhor do que quando a extração e o mapeamento foram realizados separadamente (Estudos 1 e 2). A seguir são descritos dois experimentos que testaram a primeira hipótese, um com falantes do PB e outro com falantes do inglês. Além disso, os resultados de cada experimento são comparados com os resultados encontrados nos Estudos 1, 2; de modo a testar a segunda hipótese.

Experimento 1

Método

Participantes. Vinte adultos ($M = 21$ anos e 1 mês, $SD = 1$ anos e 11 meses; 11 mulheres) participaram do estudo. Todos eram falantes nativos do PB e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no desempenho na tarefa experimental. Todos os participantes eram estudantes universitários, 19 eram estudantes da graduação (distribuídos entre 12 cursos) e um era estudante da pós-graduação.

Situação experimental e equipamentos. O experimento foi conduzido em uma sala similar à utilizada nos Estudos 1 e 2. As tarefas experimentais continuaram a ser apresentadas em um computador com sistema operacional Windows por meio do software

Psychopy2 (Peirce et al., 2019). Foram utilizados os mesmos equipamentos de áudio e vídeo dos experimentos anteriores.

Estímulos experimentais. A Língua 6 (alta familiaridade PB) foi utilizada. Cada palavra da Língua (Tabela 1) foi pareada com uma figura abstrata (Conjunto 1, Tabela 2), formando seis pares de palavras-figuras (Tabela 4). Durante a familiarização, os pares de alta frequência foram apresentados 300 vezes e os pares de baixa frequência foram apresentados 150 vezes. Considerando a exigência da tarefa, a fala contínua de ≈ 15 min foi dividida em cinco segmentos com ≈ 3 minutos cada. Os primeiros e os últimos 500 ms de cada segmento foram esvanecidos para evitar dicas claras sobre o início da primeira palavra e o fim da última.

Tabela 3

Número de Identificação dos Pares (Id), Palavra, Figura, para a Língua 6 do português brasileiro

Id	Palavras L6	Figuras
1	[sute]	
2	[viko]	
3	[bara]	
4	[nipe]	
5	[tad̃zi]	
6	[mide]	

Desenho experimental e desenho de tentativas. O experimento teve duas fases: familiarização e teste (Figura 21). Na familiarização, a fala contínua foi apresentada junto com uma sequência de imagens. A sequência foi construída respeitando a coocorrência determinística entre os pares de palavras e figuras. Para gerar tentativas ambíguas contínuas, quando a primeira palavra da fala foi apresentada, duas figuras foram apresentadas e ficaram expostas até o final da segunda palavra. Quando a terceira palavra foi apresentada, duas novas figuras foram apresentadas, elas ficaram presentes até que a quarta palavra foi apresentada, e assim por diante. As figuras sempre corresponderam às palavras faladas em cada apresentação (pares determinísticos), porém, não havia correspondência sistemática entre a disposição espacial das figuras e a ordem das palavras (2 x 2, conforme o Estudo 2). Para evitar que a apresentação ou retirada das figuras servissem como dicas para o início e o término das palavras, a apresentação ou retirada das figuras foram dessincronizadas em ± 100 , ± 150 , ou ± 200 ms do início ou fim das palavras. Tal arranjo gerou um total de 675 tentativas na configuração 2 x 2. Para cada segmento de ≈ 3 minutos, 135 tentativas foram realizadas (total de 270 pares). Pausas de 5 segundos foram realizadas entre segmentos.

Na fase de teste, duas tarefas foram realizadas. O teste de extração de palavras foi exatamente o mesmo do Experimento 5 do Estudo 1. Foram realizadas 18 tentativas de escolha forçada entre palavras de baixa frequência e parte-palavras de alta frequência. O teste de mapeamento foi exatamente o mesmo do Experimento 2 do Estudo 2. Foram realizadas 12 tentativas de escolha forçada entre 4 figuras, em relação a uma palavra-alvo.

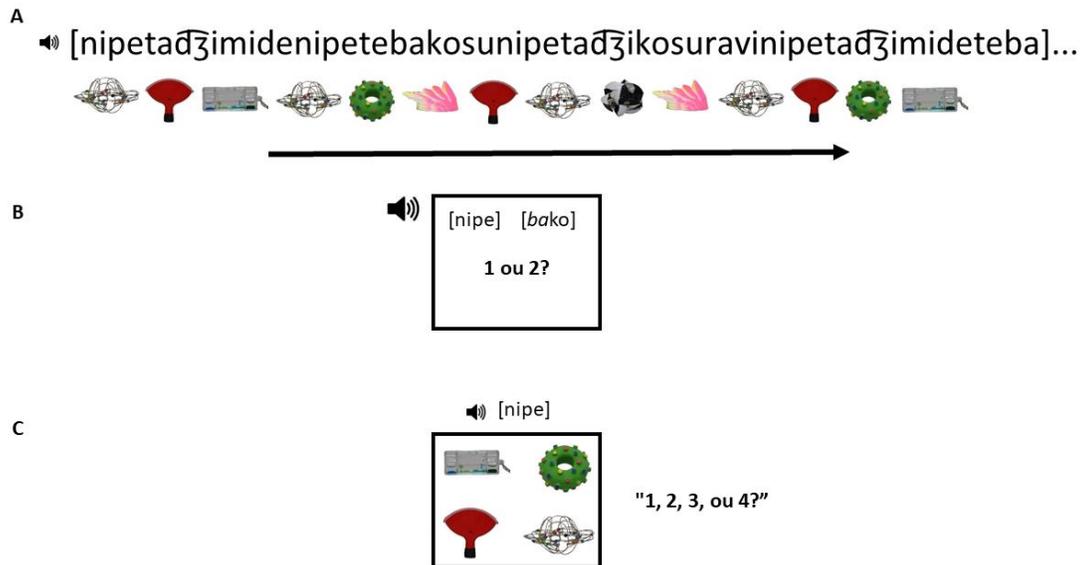


Figura 21. (A) Estrutura da fase de familiarização com a fala contínua e pares; (B) Estrutura do teste de extração de palavras; (C) Estrutura do teste de mapeamento do Estudo 3.

Procedimento. O procedimento foi similar aos Estudos anteriores. Cada participante foi exposto individualmente às tarefas experimentais. No início do experimento, uma música, com a mesma intensidade dos estímulos experimentais (i.e., 77dB), foi apresentada junto com a instrução para ajustar o volume do áudio. Em seguida, os participantes foram instruídos de que uma nova língua seria apresentada junto com figuras novas, e que sua tarefa era descobrir as relações entre as palavras e as figuras. A instrução não informava sobre as correspondências palavras-figuras e nem sobre a estrutura das tentativas³⁸.

Após a familiarização, foi realizado um aquecimento para o teste de extração, com palavras do PB sendo comparadas com palavras inventadas (e.g., “pato” vs. “bafi”). Em

³⁸ As instruções do Estudo 3 estão disponíveis no Apêndice I.

seguida, durante o teste de extração, os participantes foram instruídos a indicar qual dos sons mais se assemelhava com palavras da nova língua recém apresentada (mesmas instruções e estrutura do Estudo 1). Após o teste, os participantes foram instruídos a estimar o seu desempenho (de 0% a 100%). Em seguida, um aquecimento para o teste de mapeamento, com palavras e figuras conhecidas (e.g., “pato” + foto de um pato), foi realizado. Os participantes foram instruídos a selecionar a figura que correspondesse ao nome apresentado, dentre quatro figuras. Por fim, o teste de mapeamento era realizado seguindo a mesma estrutura do aquecimento (mesmas instruções e estrutura do Estudo 2). Após o teste, os participantes foram instruídos a estimar o seu desempenho (de 0% a 25%, de 25% a 50%, e de 75% a 100%). Por fim, o questionário pós-experimental (Capítulo 2) era apresentado.

Análise de dados. As principais variáveis dependentes foram as preferências por palavras no teste de extração e pelos referentes-alvo no teste de mapeamento. Assim como nos estudos anteriores, as preferências foram atestadas contra o nível do acaso (50% e 25%) por meio de modelos mistos generalizados que, sempre que possível, tiveram estrutura máxima dos efeitos randômicos (Barr et al., 2013; Bayeen et al., 2008)³⁹. Quando os modelos frequentistas com estrutura máxima falharam em convergir, modelos mistos bayesianos foram utilizados (Bürkner, 2018).

Variáveis dependentes secundárias incluíram a latência de acertos e erros, a estimativa de desempenho (auto-relato) e o conhecimento linguístico de cada participante. Essas variáveis não foram analisadas, uma vez que não foram encontrados efeitos de diferenças individuais na preferência por palavras ou parte-palavras.

³⁹ Sintaxe (R): `lme4::glmer(respostas ~ offset(logit(nível do acaso)) + (1 | participantes) + (1 | estímulos auditivos ou visuais))`

Resultados e Discussão

A Figura 22 apresenta a média geral de acertos para os testes de extração e mapeamento. A probabilidade de escolha das palavras em detrimento das parte-palavras ($M = 84\%$, $SD = 15\%$) foi consistentemente superior ao esperado pelo acaso ($b = 2,246$, Wald IC 95% [1,37; 3,11], $z = 5,047$, $p < 0,001$). O mesmo foi encontrado para a probabilidade de preferência por referentes-alvo no teste de mapeamento ($M = 79\%$, $SD = 26\%$, $b = 3,255$, Wald IC 95% [2,12; 4,38], $z = 5,639$, $p < 0,001$). Tais resultados confirmam as simulações de Räsänen e Rasilo (2015) ao demonstrar que palavras podem ser extraídas e mapeadas simultaneamente, dando força ao modelo de processamento unificado.

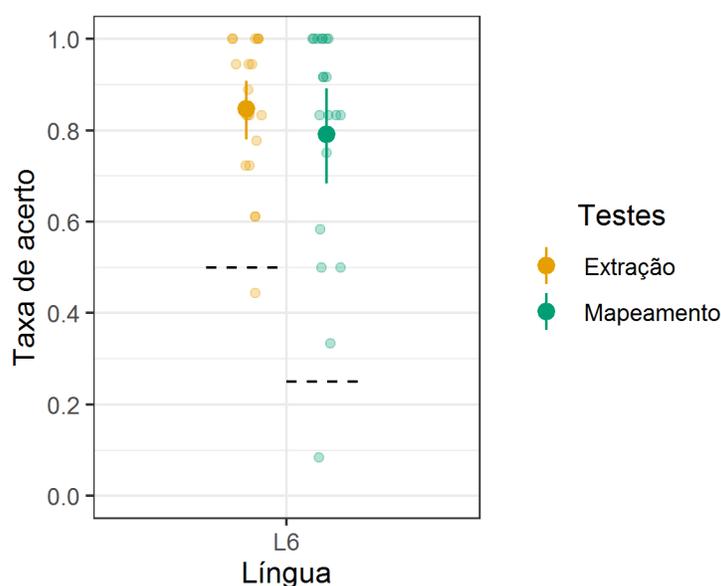


Figura 22. Taxa geral de acerto para cada participante e média geral de acertos para os testes de extração e mapeamento do Experimento 1 do Estudo 3. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e as linhas pontilhadas indicam os níveis do acaso (0,5 e 0,25).

A análise por item⁴⁰ revelou a preferência pela palavra [nipe] no teste de extração ($b = 3,20$, Intervalo Crível 95% [1,72; 5,47]) em detrimento das demais palavras.

Interessantemente, as demais preferências/rejeições encontradas no Experimento 5 (Estudo 1) não ocorreram no presente experimento. Esse resultado indica que a presença de dicas visuais eliminou parte dos efeitos idiossincráticos dos estímulos na extração de palavras. Já para o teste de mapeamento, apenas o primeiro par (Tabela 3) foi preferido consistentemente em comparação com os demais ($b = 5,46$, Intervalo Crível 95% [2,09; 10,53]). É possível que a alta frequência do par na língua (300 repetições) somada à novidade relativamente baixa da figura (66%) possa ter facilitado o seu mapeamento em comparação com os demais itens.

A Figura 23 apresenta a média geral de acertos para os testes de extração e mapeamento para o presente experimento em comparação com os experimentos anteriores (extração ou mapeamento apenas). A comparação⁴¹ da extração encontrada no presente experimento com a extração encontrada Experimento 5 (Estudo 1; Figura 13) revela que a preferência por palavras aumentou consistentemente no presente experimento ($b = 1,2118$, Wald IC 95% [0,40; 2,02], $z = 2,933$, $p < 0,001$). O mesmo foi verdadeiro para a comparação do mapeamento encontrado no presente experimento com o mapeamento do primeiro bloco do Experimento 2 do Estudo 2 (Figura 20) ($b = 1,3763$, Wald IC 95% [0,57; 2,18], $z = 3,355$, $p < 0,001$). O desempenho no Bloco 1 do Estudo 2 foi utilizado para a comparação porque tal bloco mimetiza as condições do presente experimento ao testar 6 pares que foram apresentados combinados entre si durante a familiarização.

⁴⁰ brms::brm(acertos ~ itens + (itens | participantes))

⁴¹ lme4::glmer(respostas ~ experimento + (1 | participantes) + (1 | itens))

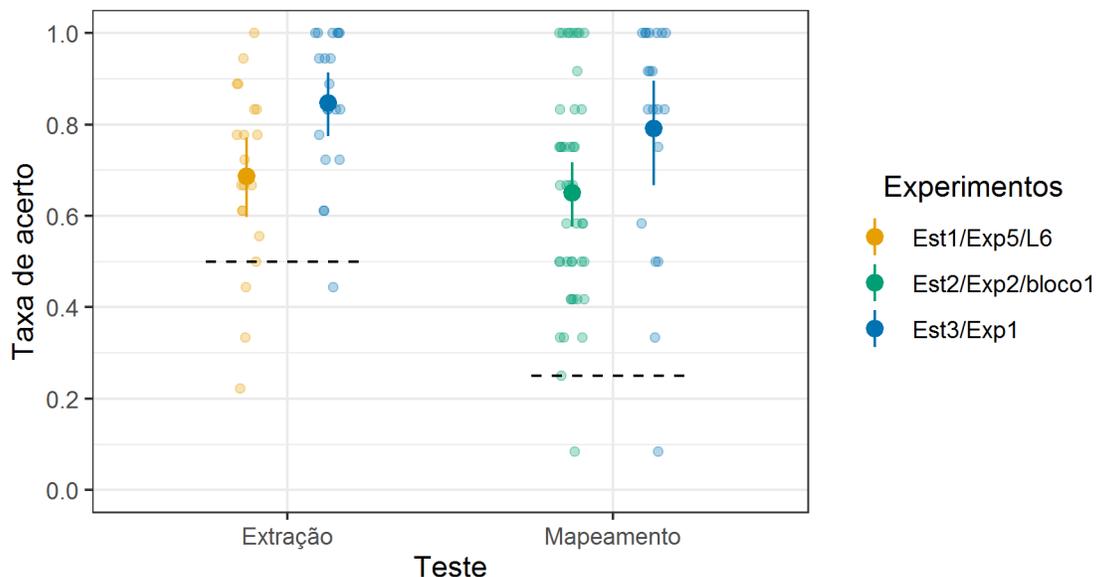


Figura 23. Taxa geral de acerto para cada participante e média geral de acertos para o teste de extração para a Língua 6, Experimento 5 do Estudo 1 e para o Experimento 1 do Estudo 3; e taxa geral de acerto para cada participante e média geral de acertos para o teste de mapeamento para o Bloco 1 do Experimento 2 do Estudo 2 e para o Experimento 1 do Estudo 3. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e as linhas pontilhadas indicam os níveis do acaso (0,5 e 0,25).

Tais resultados confirmam as simulações de Räsänen e Rasilo (2015) ao indicar que a apresentação conjunta de estatísticas que informam sobre os limites das palavras (PT entre sílabas) com estatísticas que informam sobre as relações entre palavras e referentes (coocorrência entre palavras e figuras) ao longo de tentativas ambíguas aumenta a acurácia da extração e do mapeamento em comparação à apresentação de cada dica separadamente. Nesse sentido, os dados encontrados no presente estudo suportam o modelo unificado de extração e mapeamento simultâneo proposto pelos autores.

Experimento 2

O presente experimento estende a investigação dos efeitos da apresentação paralela de dicas estatísticas entre sílabas (PT) e entre palavras e figuras para adultos falantes do inglês. A lógica experimental que motivou o experimento anterior foi a mesma que motivou o presente experimento. Assim como no experimento anterior, esperava-se que palavras fossem preferidas no teste de extração e que referentes-alvo fossem preferidos no teste de mapeamento acima do nível do acaso. Além disso, esperava-se que o desempenho nessas duas tarefas fosse mais acurado do que a extração e o mapeamento em tarefas separadas.

Método

Participantes. Quarenta adultos ($M = 21$ anos e 2 meses, $SD = 5$ anos e 8 meses; 22 mulheres) participaram do estudo. Todos eram falantes do inglês e relataram não possuir limitações físicas (auditivas, visuais e motoras) que interferissem no desempenho na tarefa experimental. Trinta e três participantes eram estudantes da graduação (distribuídos entre 25 cursos) e 7 eram estudantes da pós-graduação (distribuídos entre 7 programas).

Situação experimental e equipamentos. A situação experimental, os equipamentos, o desenho experimental, o desenho de tentativas, e a análise de dados foram os mesmos do experimento anterior. As diferenças nos estímulos experimentais e no procedimento são descritas a seguir.

Estímulos experimentais. As Línguas 10A e 10B (alta familiaridade do inglês) foram utilizadas. Cada uma das palavras das línguas (Tabela 1) foi pareada com uma figura abstrata (Conjunto 2, Tabela 2), formando seis pares de palavras-figuras para cada língua (Tabela 4). Assim como no experimento anterior, durante a familiarização, os pares de alta frequência foram apresentados 300 vezes e os pares de baixa frequência foram apresentados 150 vezes. A fala contínua de ≈ 15 min foi dividida em cinco segmentos com ≈ 3 minutos

cada. Os primeiros e os últimos 500 ms de cada segmento foram esvanecidos para evitar dicas claras sobre o início da primeira palavra e o fim da última.

Tabela 4

Número de Identificação dos Pares (Id), Palavra, Figura, para as Línguas 10A e 10B do inglês

Id	Palavras		Figuras
	L10A	L10B	
1	[siko]	[nipo]	
2	[sumæ]	[tudæ]	
3	[tifæ]	[bikæ]	
4	[kæni]	[mæsi]	
5	[potu]	[koti]	
6	[dæbi]	[fæsu]	

Procedimento. O procedimento foi similar ao do experimento anterior. As pequenas diferenças foram resultantes do uso de um rastreador de olhar durante o experimento (EyeLink 1000+). Antes do início do experimento, o equipamento era ajustado e calibrado para cada participante. Na fase de familiarização, a duração da pausa entre os segmentos da fala contínua era controlada pelos participantes e, uma vez dado o comando, o início do segmento era condicionado ao olhar para um ponto de fixação. Essas alterações buscaram garantir o registro das fixações iniciais (que, conforme teste pilotos, eram facilmente perdidas pela distração ou cansaço dos participantes). Vale ressaltar que os

dados de rastreamento de olhar (fixações e tamanho de pupila) não são apresentados nesse texto. Eles ainda estão sendo analisados (filtragem e análise temporal) e serão apresentados em publicações posteriores. Os dados comportamentais coletados a partir de pressões a teclas (conforme Estudos anteriores) serão apresentados e discutidos a seguir.

Resultados e Discussão

A Figura 24 apresenta a média geral de acertos nos testes de extração e mapeamento. Não houve diferença significativa entre o desempenho dos participantes que receberam créditos acadêmicos e o dos que receberam *gift-cards* como compensação pela participação (Capítulo 2). No geral, a probabilidade de escolha das palavras no teste de extração da Língua 10A ($M = 86\%$, $SD = 23\%$) foi consistentemente superior ao esperado pelo acaso ($b = 4,4008$, Wald IC 95% [2,47; 6,32], $z = 4,490$, $p < 0,001$). Esses resultados se contrapõem ao encontrado com a mesma Língua no Experimento 9 (Estudo 1), que não encontrou preferência por palavras enquanto o tipo de estímulo distrator (parte-palavras ou não-palavras) não foi discriminado na análise de dados. Os resultados do presente experimento sugerem que a presença adicional de dicas visuais em paralelo com probabilidades transicionais (PTs) promove uma extração de palavras robusta. Não obstante, a probabilidade de escolha das palavras diminuiu nas tentativas em que palavras eram comparadas com parte-palavras ($b = -1,6656$, Wald IC 95% [-2,65; -0,67], $z = -3,297$, $p < 0,001$), replicando o padrão encontrado no Experimento 9 (Estudo 2).

O mesmo padrão foi encontrado para a língua 10B. A probabilidade de escolha das palavras ($M = 70\%$, $SD = 27\%$) foi consistentemente superior ao esperado pelo acaso ($b = 1,8963$, Wald IC 95% [0,78; 3,00], $z = 3,348$, $p < 0,001$). No entanto, a probabilidade de escolha das palavras diminuiu nas tentativas em que palavras eram comparadas com parte-

palavras ($b = -0,7787$, Wald IC 95% [-1,33; -0,21], $z = -2,726$, $p = 0,006402$). Tais dados replicam o padrão encontrado no Experimento 9 (Estudo 1).

No teste de mapeamento, a probabilidade de preferência por referentes-alvo ($M = 92\%$, $SD = 17\%$) foi consistentemente superior ao esperado pelo acaso para a língua 10A ($b = 8,352$, Wald IC 95% [2,52; 14,17], $z = 2,811$, $p = 0,00494$) e para a língua 10B ($M = 60\%$, $SD = 34\%$, $b = 2,112$, Wald IC 95% [0,89; 3,32], $z = 3,413$, $p < 0,001$). Tais resultados confirmam as simulações de Räsänen e Rasilo (2015) ao demonstrar que palavras podem ser extraídas e mapeadas simultaneamente.

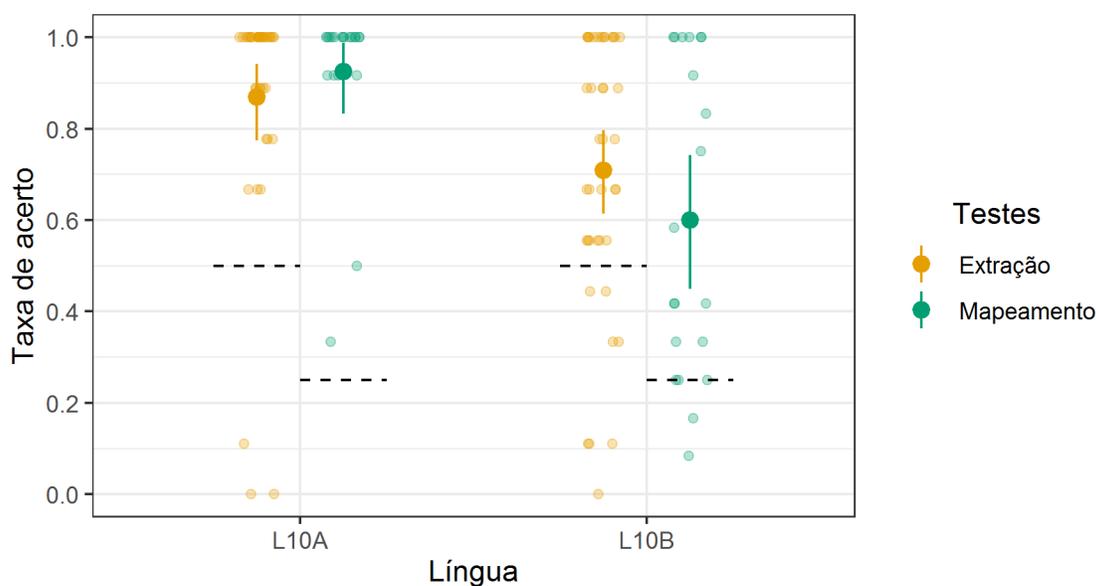


Figura 24. Taxa geral de acerto para cada participante e média geral de acertos para os testes de extração e mapeamento para as Línguas 10A e 10B do Experimento 2 do Estudo 3. As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e as linhas pontilhadas indicam os níveis do acaso (0,5 e 0,25).

Assim como no Experimento 9 (Estudo 1), a análise por item⁴² da língua 10A revelou preferência pela palavra [kæni] ($b = 5,41$, Intervalo Crível 95% [2,77; 9,34]) e rejeição pela palavra [potu] ($b = -2,53$, Intervalo Crível 95% [-5,96; -0,11]) no teste de extração. Nenhuma preferência ou rejeição foi encontrada para a língua 10B.

Já para o teste de mapeamento, três pares foram mais preferidos em relação aos demais na língua 10A: par 1 (Tabela 4; $b = 11,01$, Intervalo Crível 95% [3,30; 25,47]), par 2 (Tabela 4; $b = 96,59$, Intervalo Crível 95% [10,74; 394,60]), e par 6 (Tabela 4; $b = 16,04$, Intervalo Crível 95% [2,07; 49,72]). Provavelmente, a preferência pelos pares 1 e 2 está ligada às suas altas frequências na língua (300 repetições cada); por outro lado, a preferência pelo sexto par pode estar relacionada ao maior índice de discriminabilidade (92%, Tabela 2) desse par em comparação aos demais pares de baixa frequência (82% e 85% para o par 4 e 5; Tabela 4). Para a língua 10B, o par 5 (Tabela 4; $b = 1,77$, Intervalo Crível 95% [0,17; 3,75]) foi preferido. Curiosamente, esse é um par de baixa frequência e a figura não possui nenhuma característica (discriminabilidade, novidade ou saliência) que aparentemente explique essa preferência.

A Figura 25 apresenta a média geral de acertos para os testes de extração do presente experimento em comparação com o Experimento 9 (Estudo 1, extração apenas). A comparação⁴³ da extração encontrada no presente experimento com a extração encontrada no Experimento 9 (Estudo 1; Figura 25) para a língua 10A revela que a preferência por palavras aumentou consistentemente no presente experimento ($b = 2,5119$, Wald IC 95% [1,18; 3,83], $z = 3,719$, $p < 0,001$). A mesma análise para a língua 10B não revelou diferença na extração. A ausência do teste de mapeamento em separado com falantes do

⁴² brms::brm(acertos ~ itens + (itens | participantes))

⁴³ lme4::glmer(respostas ~ experimento + (1 | participantes) + (1 | itens))

inglês (Estudo 2) impossibilita a análise comparativa do mapeamento. Uma coleta futura com tal população poderá informar sobre os efeitos da apresentação simultânea de estatísticas auditivas e visuais sobre o desempenho de mapeamento.

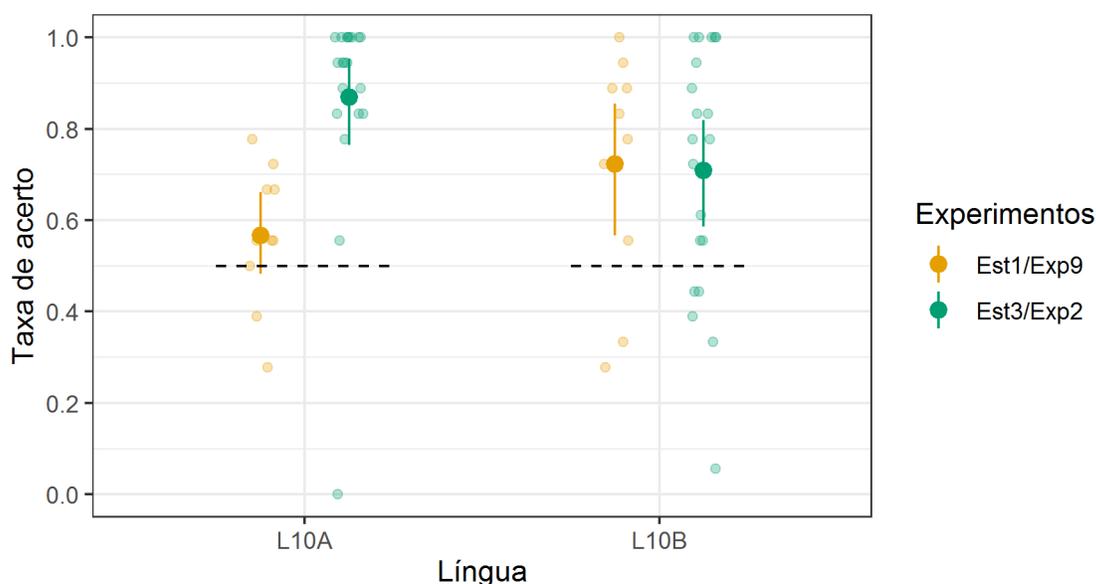


Figura 25. Taxa geral de acerto para cada participante e média geral de acertos para os testes de extração das Línguas 10A e 10B no Experimento 9 (Estudo 1) e no Experimento 2 (Estudo 3). As barras de erro indicam o intervalo de confiança (95%, *bootstrap*) e as linhas pontilhadas indicam os níveis do acaso (0,5).

Tais resultados replicam parcialmente o encontrado com falantes do PB e confirmam, também parcialmente, as propostas de Räsänen e Rasilo (2015). O aumento da extração para a língua 10A no presente experimento indica uma vantagem da apresentação conjunta de dicas estatísticas auditivas (PTs) e visuais (coocorrência de palavras e figuras) para os processos de extração. Por outro lado, os níveis similares de extração para a língua 10B indicam que variáveis idiossincráticas dos estímulos parecem ser combinadas com as demais dicas estatísticas na influência da extração das palavras de falas contínuas. Além

disso, futuras investigações do mapeamento, com os mesmos estímulos, em uma tarefa separada com os falantes do inglês podem oferecer um quadro mais completo sobre os efeitos da apresentação conjunta das dicas estatísticas auditivas e visuais em tal desempenho.

Capítulo 6

Considerações Finais

O presente trabalho buscou investigar os processos de extração de palavras de falas contínuas e de mapeamento ao longo de tentativas ambíguas em paralelo. Os experimentos conduzidos nos Estudos 1 e 2 são os primeiros a investigar os processos de extração e mapeamento ao longo de tentativas ambíguas com falantes do Português Brasileiro. Ao fazê-lo, eles atestam a generalidade dos processos de extração e mapeamento CSWL e servem como controles para o Estudo 3, que investigou a extração e o mapeamento em paralelo.

No Estudo 1, o conjunto de dados sugere que além da influência da PT sobre a extração, dicas estatísticas podem ser combinadas com o conhecimento linguístico pré-experimental dos participantes. Isso foi verdadeiro para os participantes falantes do PB e do inglês. Tais dados chamam atenção para a importância do controle cuidadoso de tais dimensões em pesquisas sobre extração, ao menos com participantes adultos, além de apontar para a natureza eminentemente interativa da aprendizagem estatística (Mersad & Nazzi, 2011; Saffran, 2014; Smith et al., 2014).

No Estudo 2, com falantes do PB, os dados demonstram que palavras (já extraídas) podem ser mapeadas com referentes visuais ao longo de tentativas ambíguas mesmo quando a diversidade contextual é baixa (seis estímulos comparados entre si ao invés de 12 ou mais, como é típico de pesquisas nessa área) e quando dois conjuntos de referentes independentes são apresentados na mesma tarefa. Além disso, quando tais conjuntos foram apresentados sequencialmente, os dados indicaram que aprender a estrutura da tarefa pode ser uma variável que aumenta a acurácia do mapeamento. Tais experimentos foram os

primeiros a investigar o mapeamento CSWL com falantes do PB, o que atesta a generalidade desse processo de aprendizagem estatística.

O Estudo 3 é o primeiro estudo a demonstrar que os processos de extração e mapeamento CSWL podem ocorrer simultaneamente. Os resultados do Estudo 3 estendem investigações anteriores (e.g., Cunillera et al., 2010; Estes et al., 2007; Shukla et al., 2011) e oferecem uma contribuição clara para o desafio de integração de processos de aprendizagem estatística (Johnson, 2016; Saffran, 2014; Smith et al. 2014). Além de demonstrar que tais processos podem ocorrer em paralelo, a comparação do desempenho na tarefa conjunta (Estudo 3) com as tarefas em separado (Estudo 1 e 2) mostrou que a acurácia na extração aumentou durante a tarefa conjunta para falantes do PB e do inglês (mas veja Língua 10B) e a acurácia do mapeamento também aumentou durante a tarefa conjunta para falantes do PB. Tais dados indicam que a presença de informações estatísticas complementares, mesmo que em uma situação mais complexa, pode promover e aumentar a aprendizagem estatística (Räsänen & Rasilo, 2015).

Os dados aqui relatados oferecem dados empíricos para futuros esforços de integração teórico-conceitual nas áreas de aprendizagem estatística e aprendizagem de linguagem (e.g., Armstrong et al., 2017; Bates & Elman, 1996; Johnson 2016; Ornat & Gallo, 2004; Schoneberger, 2010; Yurovsky & Frank, 2015). Embora tal esforço seja altamente desejável e valioso, ele está além do escopo do presente trabalho.

É importante destacar ao menos duas limitações do trabalho. Primeiro, a população investigada (adultos) possui um conhecimento linguístico pré-experimental extenso, o que dificulta o controle de todas as variáveis críticas. Futuras pesquisas com essa população devem ter um cuidado especial com tais variáveis linguísticas, o que exigirá um conhecimento aprofundado da língua nativa dos participantes. Segundo, as tarefas

utilizadas no presente trabalho são simplificações das situações de aprendizagem de linguagem em ambiente natural (Clerkin et al., 2017). O conhecimento produzido sob essas condições pode ou não sobreviver ao teste em situações com maior validade ecológica. Futuras investigações poderão aumentar a validade ecológica ao utilizar, por exemplo, estímulos (Hay et al., 2011) ou situações experimentais mais próximas das naturais (Smith, Yu, & Pereira, 2011). Apesar de tais limitações, as evidências apresentadas pelo presente trabalho podem motivar investigações cuidadosas com adultos e a investigação com populações que estão em processo de aprendizagem da sua primeira língua (bebês) ou com pessoas que estão reaprendendo uma língua, como é o caso de pacientes com afasia (e.g., Peñaloza et al., 2017).

Por fim, o conjunto de experimentos realizados para esta tese aborda, de maneira inédita, a interação entre dois processos complexos de aprendizagem estatística. Ao combinar a extração e o mapeamento ao longo de tentativas ambíguas, a presente tese evidencia o nível de complexidade presente no desenvolvimento da linguagem. O presente trabalho também avança na compreensão desses processos ao investigar populações nunca antes estudadas. Ao mesmo tempo, ele proporciona um novo conjunto de perguntas que podem ser perseguidas por pesquisas futuras.

Referências

- Abaurre, M. B., & Sandalo, F. (2017). Coarticulação e harmonia vocálica. *Cadernos de Estudos Lingüísticos*, 59(3), 487–497. <http://doi.org/10.20396/cel.v59i3.8650984>
- Albano, E. C. (2007). Codificação estatística das categorias fonéticas: vestígio da dinâmica da fala na fonotaxe lexical. *Revista de Estudos Da Linguagem*, 15(2), 9–42.
- Aljasser, F., & Vitevitch, M. S. (2018). A Web-based interface to calculate phonotactic probability for words and nonwords in Modern Standard Arabic. *Behavior Research Methods*, 50(1), 313–322. <http://doi.org/10.3758/s13428-017-0872-z>
- Alt, M., & Plante, E. (2006). Factors that influence lexical and semantic fast mapping of young children with specific language impairment. *Journal of Speech, Language, & Hearing Research*, 49, 941–954.
- Arciuli, J., Torkildsen, J. von K., Stevens, D. J., & Simpson, I. C. (2014). Statistical learning under incidental versus intentional conditions. *Frontiers in Psychology*, 5(JUL), 1–8. <http://doi.org/10.3389/fpsyg.2014.00747>
- Armstrong, B. C., Frost, R., & Christiansen, M. H. (2017). The long road of statistical learning research: past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160047. <http://doi.org/10.1098/rstb.2016.0047>
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of Conditional Probability Statistics by 8-Month-Old Infants. *Psychological Science*, 9(4), 321–324. <http://doi.org/10.1111/1467-9280.00063>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>

- Bailey, T. M., & Hahn, U. (2001). Determinants of Wordlikeness: Phonotactics or Lexical Neighborhoods? *Journal of Memory and Language*, *44*(4), 568–591.
<http://doi.org/10.1006/jmla.2000.2756>
- Barbosa, P. G., & Martins, C. C. (2014). Uma revisão dos estudos sobre a fala dirigida à criança e suas implicações para a aquisição inicial do vocabulário. *Linguagem Em (Dis)Curso*, *14*(1), 195–210.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, E., & Elman, J. (1996). Learning Rediscovered. *Science*, *274*(5294), 1849–1850.
<http://doi.org/10.1126/science.274.5294.1849>
- Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78.
<http://doi.org/10.1016/j.jml.2015.04.004>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Bergelson, E., & Swingle, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258. <http://doi.org/10.1073/pnas.1113380109>
- Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: a meta-analytic approach. *Developmental Science*, *19*(6), 901–917.
<http://doi.org/10.1111/desc.12341>
- Black, A., & Bergmann, C. (2017). Quantifying Infants' Statistical Word Segmentation: A

- Meta-Analysis. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 124–129). Austin, TX: Cognitive Science Society.
- Blanchard, D., Heinz, J., & Golinkoff, R. (2010). Modeling the contribution of phonotactic cues to the problem of word segmentation. *Journal of Child Language*, *37*(03), 487. <https://doi.org/10.1017/S030500090999050X>
- Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
- Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.36. Retrieved from <http://www.praat.org/>
- Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and Me. *Psychological Science*, *16*(4), 298–304. <https://doi.org/10.1111/j.0956-7976.2005.01531.x>
- Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33–B44. [http://doi.org/10.1016/S0010-0277\(01\)00122-6](http://doi.org/10.1016/S0010-0277(01)00122-6)
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. <https://doi.org/10.3758/BRM.41.4.977>
- Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(1), 127–132. <http://doi.org/10.1016/j.cognition.2011.06.010>
- Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395. <https://doi.org/10.32614/RJ-2018-017>
- Cannistraci, R. A., Dal Ben, R., Karaman, F., Esfahani, S. P., & Hay, J. F. (2019).

- Statistical Learning Approaches to Studying Language Development. In J. S. Horst & J. von Koss Torkildsen (Eds.), *International Handbook of Language Acquisition* (p. 586). New York, NY: Routledge, [2019] | Series: Routledge: Routledge.
<https://doi.org/10.4324/9781315110622>
- Cashon, C. H., Ha, O.-R., Estes, K. G., Saffran, J. R., & Mervis, C. B. (2016). Infants with Williams syndrome detect statistical regularities in continuous speech. *Cognition*, *154*, 165–168. <http://doi.org/10.1016/j.cognition.2016.05.009>
- Catania, A. C. (2012). *Learning* (5th ed.). Cornwall-on-Hudson: Sloan Publishing.
- Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*(2), B69–B77.
[http://doi.org/10.1016/s0010-0277\(02\)00233-0](http://doi.org/10.1016/s0010-0277(02)00233-0)
- Choi, D., Black, A. K., & Werker, J. F. (2018). Cascading and Multisensory Influences on Speech Perception Development. *Mind, Brain, and Education*, 1–12.
<http://doi.org/10.1111/mbe.12162>
- Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160055.
<http://doi.org/10.1098/rstb.2016.0055>
- Conway, C. M., & Christiansen, M. H. (2005). Modality-Constrained Statistical Learning of Tactile, Visual, and Auditory Sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 24–39. <http://doi.org/10.1037/0278-7393.31.1.24>
- Corrêa, L. M. S. (2007). O que, afinal, a criança adquire ao adquirir uma língua? *Letras de Hoje*, *42*(1), 7–34.

- Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>
- Cunillera, T., Camara, E., Laine, M., & Rodriguez-Fornells, A. (2010). Speech segmentation is facilitated by visual cues. *Quarterly Journal of Experimental Psychology (2006)*, 63(2), 260–74. <http://doi.org/10.1080/17470210902888809>
- Cunillera, T., Càmara, E., Laine, M., & Rodríguez-Fornells, A. (2010). Words as anchors: Known words facilitate statistical learning. *Experimental Psychology*, 57(2), 134–141. <http://doi.org/10.1027/1618-3169/a000017>
- Cunillera, T., Laine, M., & Antoni, R.-F. (2016). Headstart for speech segmentation: a neural signature for the anchor word effect. *Neuropsychologia*, 82, 189–199. <http://doi.org/10.1016/j.neuropsychologia.2016.01.011>
- Cunillera, T., Laine, M., Càmara, E., & Rodríguez-Fornells, A. (2010). Bridging the gap between speech segmentation and word-to-world mappings: Evidence from an audiovisual statistical learning task. *Journal of Memory and Language*, 63(3), 295–305. <http://doi.org/10.1016/j.jml.2010.05.003>
- Dal Ben, R. (2019a, April 20). Creating an artificial language with natural recordings using Praat: step by step. Retrieved from osf.io/xb3gw
- Dal Ben, R. (2019b). Lum_fun: A set of tools to control luminance of colored images (Version 0.2). [Computer program]. doi: 10.17605/OSF.IO/AUZJY, retrieved from <https://osf.io/auzjy/>
- Dal Ben, R. & Souza, D. de H. (n.d.). Aprendizagem ao longo de tentativas ambíguas: uma revisão da literatura empírica (2007-2018). Manuscrito em preparação.
- de Rose, J., Gil, M. S. C. de A., & de Souza, D. das G. (2014). *Comportamento simbólico: bases conceituais e empíricas* (1st ed.). São Paulo: Cultura Acadêmica.

- DeCasper, A., & Fifer, W. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, 208(4448), 1174–1176. <http://doi.org/10.1126/science.7375928>
- Dixon, L. S. (1977). The nature of control by spoken words over visual stimulus selection. *Journal of the Experimental Analysis of Behavior*, 27(3), 433–442. <https://doi.org/10.1901/jeab.1977.27-433>
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & van der Vrecken, O. (1996). The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes. In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP* (Vol. 3, pp. 1393–1396). IEEE. <http://doi.org/10.1109/ICSLP.1996.607874>
- Endress, A. D., & Langus, A. (2017). Transitional probabilities count more than frequency, but might not be used for memorization. *Cognitive Psychology*, 92, 37–64. <http://doi.org/10.1016/j.cogpsych.2016.11.004>
- Endress, A. D., & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language*, 60(3), 351–367. <http://doi.org/10.1016/j.jml.2008.10.003>
- Estes, K. G. (2009). From tracking statistics to learning words: Statistical learning and lexical acquisition. *Linguistics and Language Compass*, 3(6), 1379–1389. <http://doi.org/10.1111/j.1749-818X.2009.00164.x>
- Estes, K. G. (2012). Infants generalize representations of statistically segmented words. *Frontiers in Psychology*, 3(OCT), 1–13. <http://doi.org/10.3389/fpsyg.2012.00447>
- Estes, K. G. (2014). Learning builds on learning: Infants' use of native language sound patterns to learn words. *Journal of Experimental Child Psychology*, 126, 313–327. <https://doi.org/10.1016/j.jecp.2014.05.006>

- Estes, K. G., & Bowen, S. (2013). Learning about sounds contributes to learning about words: Effects of prosody and phonotactics on infant word learning. *Journal of Experimental Child Psychology, 114*(3), 405–417.
<http://doi.org/10.1016/j.jecp.2012.10.002>
- Estes, K. G., Edwards, J., & Saffran, J. R. (2011). Phonotactic Constraints on Infant Word Learning. *Infancy, 16*(2), 180–197. <http://doi.org/10.1111/j.1532-7078.2010.00046.x>
- Estes, K. G., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can Infants Map Meaning to Newly Segmented Words? *Psychological Science, 18*(3), 254–260.
<http://doi.org/10.1111/j.1467-9280.2007.01885.x>
- Estes, K. G., Gluck, S. C. W., & Grimm, K. J. (2016). Finding patterns and learning words: Infant phonotactic knowledge is associated with vocabulary size. *Journal of Experimental Child Psychology, 146*, 34–49. <http://doi.org/10.1016/j.jecp.2016.01.012>
- Estivalet, G. L. & Dal Ben, R. (n.d.). An online calculator to compute phonotactic and orthotactic probability, and phonological and orthographical neighborhood densities for words and nonwords in Brazilian-Portuguese. Manuscrito em preparação.
- Estivalet, G. L. (2018). Algoritmo de silabação e tonicidade: inventário e frequências das sílabas e estruturas silábicas do português brasileiro. In *XXVII Jornada Internacional do Grupo de Estudos Linguísticos do Nordeste - GELNE*. Recife/PE.
- Estivalet, G. L., & Meunier, F. (2015). The Brazilian Portuguese Lexicon: An Instrument for Psycholinguistic Research. *PLoS ONE, 10*(12), 1–24.
<http://doi.org/10.1371/journal.pone.0144016>
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical Learning in Children With Specific Language Impairment. *Journal of Speech Language and Hearing Research, 52*(2), 321. [http://doi.org/10.1044/1092-4388\(2009/07-0189\)](http://doi.org/10.1044/1092-4388(2009/07-0189))

- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Filippi, P., Gingras, B., & Fitch, W. T. (2014). Pitch enhancement facilitates word learning across visual contexts. *Frontiers in Psychology*, *5*(DEC), 1–8. <http://doi.org/10.3389/fpsyg.2014.01468>
- Filippi, P., Laaha, S., & Fitch, W. T. (2017). Utterance-final position and pitch marking aid word learning in school-age children. *Royal Society Open Science*, *4*(8), 161035. <http://doi.org/10.1098/rsos.161035>
- Finn, A. S., & Hudson Kam, C. L. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, *108*(2), 477–499. <https://doi.org/10.1016/j.cognition.2008.04.002>
- Fitneva, S. A., & Christiansen, M. H. (2017). Developmental Changes in Cross-Situational Word Learning: The Inverse Effect of Initial Accuracy. *Cognitive Science*, *41*, 141–161. <http://doi.org/10.1111/cogs.12322>
- Fontanari, J. F., & Cangelosi, A. (2011). Cross-situational and supervised learning in the emergence of communication. *Interaction Studies*, *12*(1), 119–133. <http://doi.org/10.1075/is.12.1.05fon>
- François, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fornells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, *98*(January 2016), 56–67. <http://doi.org/10.1016/j.neuropsychologia.2016.10.006>
- Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of Wordlikeness: Effects of Segment Probability and Length on the Processing of Nonwords. *Journal of Memory*

and Language, 42(4), 481–496. <http://doi.org/10.1006/jmla.1999.2692>

Frost, R. L. A., Monaghan, P., & Tatsumi, T. (2017). Domain-general mechanisms for speech segmentation: The role of duration information in language learning. *Journal of Experimental Psychology: Human Perception and Performance*, 43(3), 466–476. <http://doi.org/10.1037/xhp0000325>

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences*, 19(3), 117–125. <http://doi.org/10.1016/j.tics.2014.12.010>

Garcia, C., Iozzo, G., Lamirato, K., Ledoux, J., Mu, J., Murdock, K., ... Hartshorne, J. K. (2018, July 6). Replication of Saffran, Newport, & Aslin (1996) Word segmentation: The role of distributional cues, Exp. 1. <https://doi.org/10.31234/osf.io/qsyd2>

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-Acquisition, Word Frequency, and Neighborhood Density Effects on Spoken Word Recognition by Children and Adults. *Journal of Memory and Language*, 45(3), 468–492. <http://doi.org/10.1006/jmla.2000.2784>

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(1), 84–95. <http://doi.org/10.1037/0278-7393.25.1.84>

Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review*, 16(3), 486–490. <http://doi.org/10.3758/PBR.16.3.486>

Goldrick, M., & Larson, M. (2008). Phonotactic probability influences speech production. *Cognition*, 107(3), 1155–1164. <http://doi.org/10.1016/j.cognition.2007.11.009>

Gomes, C. A., Mendes, S. D. C., Silva, M. B., Esteves, C. O., & Gomes, G. C. (2015).

- Efeito de Wordlikeness no Processamento de Não-palavras por Falantes do Português Brasileiro. *Revista de Estudos Da Linguagem*, 23(1). <http://doi.org/10.17851/2237-2083.23.1.195-210>
- Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–135.
[http://doi.org/10.1016/S0010-0277\(99\)00003-7](http://doi.org/10.1016/S0010-0277(99)00003-7)
- Gonzalez-Gomez, N., Poltrock, S., & Nazzi, T. (2013). A “Bat” Is Easier to Learn than a “Tab”: Effects of Relative Phonotactic Frequency on Infant Word Learning. *PLoS ONE*, 8(3). <http://doi.org/10.1371/journal.pone.0059601>
- Haebig, E., Saffran, J. R., & Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry*, 58(11), 1251–1263. <http://doi.org/10.1111/jcpp.12734>
- Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, 56(1), 51–65.
<https://doi.org/10.1037/h0062474>
- Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, 63(2), 93–106.
<http://doi.org/10.1016/j.cogpsych.2011.06.002>
- Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, 48(4), 1393–1409. <http://doi.org/10.3758/s13428-015-0647-3>
- Houston, D., Santelmann, L., & Jusczyk, P. W. (2004). English-learning infants’ segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes*, 19(1), 97–136. <https://doi.org/10.1080/01690960344000143>
- Hu, C. F. (2017). Resolving referential ambiguity across ambiguous situations in young

- foreign language learners. *Applied Psycholinguistics*, 38(3), 633–656.
<http://doi.org/10.1017/S0142716416000357>
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446. <http://doi.org/10.1016/j.jml.2007.11.007>
- Johnson, E. K. (2016). Constructing a Proto-Lexicon: An Integrative View of Infant Language Development. *Annual Review of Linguistics*, 2(1), 391–412.
<https://doi.org/10.1146/annurev-linguistics-011415-040616>
- Johnson, E. K., & Jusczyk, P. W. (2001). Word Segmentation by 8-Month-Olds: When Speech Cues Count More Than Statistics. *Journal of Memory and Language*, 44, 548–567. <https://doi.org/10.1006/jmla.2000.2755>
- Johnson, E. K., & Seidl, A. H. (2009). At 11 months, prosody still outranks statistics. *Developmental Science*, 12(1), 131–141. <http://doi.org/10.1111/j.1467-7687.2008.00740.x>
- Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, 13(2), 339–345. <http://doi.org/10.1111/j.1467-7687.2009.00886.x>
- Johnson, E. K., Seidl, A., & Tyler, M. D. (2014). The Edge Factor in Early Word Segmentation: Utterance-Level Prosody Enables Word Form Extraction by 6-Month-Olds. *PLoS ONE*, 9(1), e83546. <https://doi.org/10.1371/journal.pone.0083546>
- Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences*, 3(9), 323–328. [http://doi.org/10.1016/S1364-6613\(99\)01363-7](http://doi.org/10.1016/S1364-6613(99)01363-7)
- Jusczyk, P. W. (2002). How Infants Adapt Speech-Processing Capacities to Native-Language Structure. *Current Directions in Psychological Science*, 11(1), 15–18.

<https://doi.org/10.1111/1467-8721.00159>

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' Detection of the Sound Patterns of Words in Fluent Speech. *Cognitive Psychology*, 29(1), 1–23.

<https://doi.org/10.1006/cogp.1995.1010>

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, 61(8), 1465–1476.

<https://doi.org/10.3758/BF03213111>

Karaman, F., & Hay, J. F. (2018). The longevity of statistical learning: When infant memory decays, isolated words come to the rescue. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(2), 221–232.

<http://doi.org/10.1037/xlm0000448>

Keuleers, E. (2013). vwr: Useful functions for visual word recognition research. R package version 0.3.0.: <https://CRAN.R-project.org/package=vwr> 41.

Kirkham, N. Z., Slemmer, J. a, & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, 83(2), B35–B42. [http://doi.org/10.1016/S0010-0277\(02\)00004-5](http://doi.org/10.1016/S0010-0277(02)00004-5)

Kisilevsky, B. S., Hains, S. M. J., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., ... Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, 32(1), 59–71.

<https://doi.org/10.1016/j.infbeh.2008.10.002>

Kleiner, M., Brainard, D. H., Pelli, D. G., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, 36, 1–16.

<https://doi.org/10.1068/v070821>

Koehne, J., & Crocker, M. W. (2015). The Interplay of Cross-Situational Word Learning

and Sentence-Level Constraints. *Cognitive Science*, 39(5), 849–889.

<http://doi.org/10.1111/cogs.12178>

Kovelman, I., Wagley, N., Hay, J. S. F., Ugolini, M., Bowyer, S. M., Lajiness-O'Neill, R., & Brennan, J. (2015). Multimodal imaging of temporal processing in typical and atypical language development. *Annals of the New York Academy of Sciences*, 1337(1), 7–15. <http://doi.org/10.1111/nyas.12688>

Krogh, L., Vlach, H. A., & Johnson, S. P. (2013). Statistical Learning Across Development: Flexible Yet Constrained. *Frontiers in Psychology*, 3(JAN), 1–11. <http://doi.org/10.3389/fpsyg.2012.00598>

Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2017). Statistical Learning in Specific Language Impairment: A Meta-Analysis. *Journal of Speech Language and Hearing Research*, 60(12), 3474. http://doi.org/10.1044/2017_JSLHR-L-16-0439

Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science*, 14(6), 1323–1329. <http://doi.org/10.1111/j.1467-7687.2011.01079.x>

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, 38(2), 202–210. <http://doi.org/10.3758/BF03192770>

MacDonald, K., Yurovsky, D., & Frank, M. C. (2017). Social cues modulate the representations underlying cross-situational learning. *Cognitive Psychology*, 94, 67–84. <http://doi.org/10.1016/j.cogpsych.2017.02.003>

MacKenzie, H., Curtin, S., & Graham, S. A. (2012). 12-Month-Olds' Phonotactic Knowledge Guides Their Word-Object Mappings. *Child Development*, 83(4), 1129–1136. <http://doi.org/10.1111/j.1467-8624.2012.01764.x>

- Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2008). Lexical Representations in Children With SLI: Evidence From a Frequency-Manipulated Gating Task. *Journal of Speech Language and Hearing Research, 51*(2), 381. [http://doi.org/10.1044/1092-4388\(2008/028\)](http://doi.org/10.1044/1092-4388(2008/028))
- Marques, L. F. (2008). *Estruturas silábicas do português do Brasil: uma análise tipológica* (Dissertação de Mestrado, Universidade de São Paulo). Disponível em: <http://www.teses.usp.br/teses/disponiveis/8/8139/tde-06082009-163245/pt-br.php>
- Matsuoka, A. (2006). Os sons da fala na aquisição de linguagem - a hipótese do bootstrapping fonológico. *Revista Gatilho, 3*, 1–10.
- Mattys, S. L., & Jusczyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance, 27*(3), 644–655. <https://doi.org/10.1037/0096-1523.27.3.644>
- Mattys, S. L., Jusczyk, P. W., Luce, P. a, & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38*(4), 465–494. <http://doi.org/10.1006/cogp.1999.0721>
- Mayo, J., & Eigsti, I.-M. (2012). Brief Report: A Comparison of Statistical Learning in School-Aged Children with High Functioning Autism and Typically Developing Peers. *Journal of Autism and Developmental Disorders, 42*(11), 2476–2485. <http://doi.org/10.1007/s10803-012-1493-0>
- McGregor, K. K., Rost, G., Arenas, R., Farris-Trimble, A., & Stiles, D. (2013). Children with ASD can use gaze in support of word recognition and learning. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 54*(7), 745–753. <http://doi.org/10.1111/jcpp.12073>
- Mersad, K., & Nazzi, T. (2011). Transitional probabilities and positional frequency

- phonotactics in a hierarchical model of speech segmentation. *Memory and Cognition*, 39(6), 1085–1093. <http://doi.org/10.3758/s13421-011-0074-3>
- Mersad, K., & Nazzi, T. (2012). When Mommy Comes to the Rescue of Statistics: Infants Combine Top-Down and Bottom-Up Cues to Segment Speech. *Language Learning and Development*, 8(3), 303–315. <http://doi.org/10.1080/15475441.2011.609106>
- Mersad, K., & Nazzi, T. (2012). When Mommy Comes to the Rescue of Statistics: Infants Combine Top-Down and Bottom-Up Cues to Segment Speech. *Language Learning and Development*, 8(3), 303–315. <https://doi.org/10.1080/15475441.2011.609106>
- Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. a. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, 108(1), 271–280. <http://doi.org/10.1016/j.cognition.2008.02.003>
- Molina, D. de S. L. (2014). *A aquisição verbal e o processamento morfológico por crianças adquirindo o PB* (Dissertação de Mestrado, Universidade Federal de Juiz de Fora). Disponível em: <https://repositorio.ufjf.br/jspui/handle/ufjf/666>
- Moon, C. M., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: A two-country study. *Acta Paediatrica, International Journal of Paediatrics*, 102(2), 156–160. <https://doi.org/10.1111/apa.12098>
- Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16(4), 495–500. [http://doi.org/10.1016/0163-6383\(93\)80007-U](http://doi.org/10.1016/0163-6383(93)80007-U)
- Munson, B., Swenson, C. L., & Manthei, S. C. (2005). Lexical and phonological organization in children: Evidence from repetition tasks. *Journal of Speech, Language, & Hearing Research*, 48, 108–124.
- Name, C. (2011). Descobrimos novas palavras no fluxo da fala: o impacto da prosódia na

- aquisição lexical. *Gragoatá*, 30, 77–88.
- Name, C. (2012). O que nos dizem os resultados experimentais sobre a percepção da fala pelo bebê. *Veredas Online, especial*(2012), 284–297.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756–766. <http://doi.org/10.1037/0096-1523.24.3.756>
- Nazzi, T., Iakimova, G., Bertoncini, J., Fredones, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, 54(3), 283–299. <https://doi.org/10.1016/j.jml.2005.10.004>
- Newman, R. S., & German, D. J. (2005). Life Span Effects of Lexical Factors on Oral Naming. *Language and Speech*, 48(2), 123–156. <http://doi.org/10.1177/00238309050480020101>
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48(2), 127–162. [http://doi.org/10.1016/S0010-0285\(03\)00128-2](http://doi.org/10.1016/S0010-0285(03)00128-2)
- Ornat, S. L., & Gallo, P. (2004). Acquisition, Learning, or Development of Language? Skinner’s “Verbal Behavior” Revisited. *The Spanish Journal of Psychology*, 7(2), 161–170. <https://doi.org/10.1017/S1138741600004868>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 3. <https://doi.org/10.3758/s13428-018-01193-y>
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009a). Learning in reverse: Eight-month-old

infants track backward transitional probabilities. *Cognition*, *113*(2), 244–247.

<http://doi.org/10.1016/j.cognition.2009.07.011>

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009b). Statistical Learning in a Natural Language by 8-Month-Old Infants. *Child Development*, *80*(3), 674–685.

<http://doi.org/10.1111/j.1467-8624.2009.01290.x>

Peñaloza, C., Mirman, D., Cardona, P., Juncadella, M., Martin, N., Laine, M., & Rodríguez-Fornells, A. (2017). Cross-situational word learning in aphasia. *Cortex*, *93*, 12–27. <http://doi.org/10.1016/j.cortex.2017.04.020>

Perruchet, P., & Desautly, S. (2008). A role for backward transitional probabilities in word segmentation? *Memory & Cognition*, *36*(7), 1299–1305.

<http://doi.org/10.3758/MC.36.7.1299>

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238.

<http://doi.org/10.1016/j.tics.2006.03.006>

Poulin-Charronnat, B., Perruchet, P., Tillmann, B., & Peereman, R. (2017). Familiar units prevail over statistical cues in word segmentation. *Psychological Research*, *81*(5), 990–1003. <http://doi.org/10.1007/s00426-016-0793-y>

Quine, W. A. O. (1960). *Word and Object*. MIT press.

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria.: R Foundation for Statistical Computing.

Ramscar, M., Dye, M., & Klein, J. (2013). Children Value Informativity Over Logic in Word Learning. *Psychological Science*, *24*(6), 1017–1023.

<http://doi.org/10.1177/0956797612460691>

Räsänen, O., & Rasilo, H. (2015). A joint model of word segmentation and meaning

acquisition through cross-situational learning. *Psychological Review*, 122(4), 792–829.

<http://doi.org/10.1037/a0039702>

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition.

Wiley Interdisciplinary Reviews: Cognitive Science, 1, 906–914.

<http://doi.org/10.1002/wcs.78>

Saffran, J. (2014). Sounds and Meanings Working Together: Word Learning as a

Collaborative Effort. *Language Learning*, 64(s2), 106–120.

<http://doi.org/10.1111/lang.12057>

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners.

Developmental Psychology, 39(3), 484–494. <https://doi.org/10.1037/0012->

1649.39.3.484

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old

infants. *Science (New York, N.Y.)*, 274(5294), 1926–1928.

<http://doi.org/10.1126/science.274.5294.1926>

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word Segmentation: The Role of

Distributional Cues. *Journal of Memory and Language*, 35(4), 606–621.

<http://doi.org/10.1006/jmla.1996.0032>

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. a., & Barrueco, S. (1997). Incidental

Language Learning: Listening (and Learning) Out of the Corner of Your Ear.

Psychological Science, 8(2), 101–105. <http://doi.org/10.1111/j.1467->

9280.1997.tb00690.x

Saksida, A., Langus, A., & Nespors, M. (2017). Co-occurrence statistics as a language-

dependent cue for speech segmentation. *Developmental Science*, 20(3).

<http://doi.org/10.1111/desc.12390>

- Santolin, C., & Saffran, J. R. (2018). Constraints on Statistical Learning Across Species. *Trends in Cognitive Sciences*, 22(1), 52–63. <http://doi.org/10.1016/j.tics.2017.10.003>
- Shatzman, K. B., & McQueen, J. M. (2006). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception & Psychophysics*, 68(1), 1–16. <http://doi.org/10.3758/BF03193651>
- Schoneberger, T. (2010). Three Myths from the Language Acquisition Literature. *The Analysis of Verbal Behavior*, 26, 107–131.
- Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences of the United States of America*, 108(15), 6038–6043. <http://doi.org/10.1073/pnas.1017617108>
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, 81, 105–120. <http://doi.org/10.1016/j.jml.2015.02.001>
- Smith, L. B., & Yu, C. (2013). Visual Attention Is Not Enough: Individual Differences in Statistical Word-Referent Learning in Infants. *Language Learning and Development*, 9(1), 25–49. <http://doi.org/10.1080/15475441.2012.707104>
- Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The Developing Infant Creates a Curriculum for Statistical Learning. *Trends in Cognitive Sciences*, 22(4), 325–336. <http://doi.org/10.1016/j.tics.2018.02.004>
- Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends in Cognitive Sciences*, 18(5), 251–258. <http://doi.org/10.1016/j.tics.2014.02.007>
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-

situational statistics. *Cognition*, 106(3), 1558–1568.

<http://doi.org/10.1016/j.cognition.2007.06.010>

Smith, L. B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: The dynamics of toddler visual experience. *Developmental Science*, 14(1), 9–17.

<https://doi.org/10.1111/j.1467-7687.2009.00947.x>

Souza, D. D. H. (2008). As crianças e o mundo das palavras: considerações sobre a pesquisa em desenvolvimento lexical. *Psicologia: Reflexão e Crítica*, 21(2), 195–202.

<http://doi.org/10.1590/S0102-79722008000200005>

Storkel, H. L. (2013). A corpus of consonant–vowel–consonant real words and nonwords: Comparison of phonotactic probability, neighborhood density, and consonant age of acquisition. *Behavior Research Methods*, 45(4), 1159–1167.

<http://doi.org/10.3758/s13428-012-0309-7>

Storkel, H. L., & Hoover, J. R. (2010). An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavior Research Methods*, 42(2), 497–506.

<http://doi.org/10.3758/BRM.42.2.497>

Storkel, H. L., Bontempo, D. E., Aschenbrenner, A. J., Maekawa, J., & Lee, S.-Y. (2013). The Effect of Incremental Changes in Phonotactic Probability and Neighborhood Density on Word Learning by Preschool Children. *Journal of Speech, Language, and Hearing Research*, 56(5), 1689–1700. [http://doi.org/10.1044/1092-4388\(2013\)12-0245](http://doi.org/10.1044/1092-4388(2013)12-0245)

Suanda, S. H., & Namy, L. L. (2012). Detailed Behavioral Analysis as a Window Into Cross-Situational Word Learning. *Cognitive Science*, 36(3), 545–559.

<http://doi.org/10.1111/j.1551-6709.2011.01218.x>

- Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology, 126*, 395–411. <http://doi.org/10.1016/j.jecp.2014.06.003>
- Swingley, D. (1999). Conditional probability and word discovery: A corpus analysis of speech to infants. In M. Han & S. C. Stoness (Eds.). *Proceedings of the 21st Annual Meeting of the Cognitive Science Society* (pp. 724–729). Mahwah, NJ.
- Swingley, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology, 50*(1), 86–132. <https://doi.org/10.1016/j.cogpsych.2004.06.001>
- Swingley, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology, 54*(2), 99–132. <http://doi.org/10.1016/j.cogpsych.2006.05.001>
- Thiessen, E. D. (2010). Effects of simultaneously presented visual information on adults' and infants' auditory statistical learning. *Cognitive Science, 34*, 1093–1106.
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*(4), 706–716. <http://doi.org/10.1037/0012-1649.39.4.706>
- Thiessen, E. D., & Saffran, J. R. (2007). Learning to Learn: Infants' Acquisition of Stress-Based Strategies for Word Segmentation. *Language Learning and Development, 3*(1), 73–100. <http://doi.org/10.1080/15475440709337001>
- Thiessen, E. D., Hill, E. a, & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy, 7*(1), 53–71. http://doi.org/10.1207/s15327078in0701_5
- Thomson, J. M., Richardson, U., & Goswami, U. (2005). Phonological similarity neighborhoods and children's short-term memory: Typical development and dyslexia. *Memory & Cognition, 33*(7), 1210–1219. <http://doi.org/10.3758/BF03193223>

- Tincoff, R., & Jusczyk, P. W. (2012). Six-Month-Olds Comprehend Words That Refer to Parts of the Body. *Infancy*, *17*(4), 432–444. <http://doi.org/10.1111/j.1532-7078.2011.00084.x>
- Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*(1), 126–156. <http://doi.org/10.1016/j.cogpsych.2012.10.001>
- Vitevitch, M. S., & Donoso, A. J. (2012). Phonotactic probability of brand names: I'd buy that! *Psychological Research*, *76*(6), 693–698. <http://doi.org/10.1007/s00426-011-0374-z>
- Vitevitch, M. S., & Luce, P. A. (2004). A Web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 481–487. <http://doi.org/10.3758/BF03195594>
- Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *52*(2), 193–204. <http://doi.org/10.1016/j.jml.2004.10.003>
- Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*(1–2), 306–311. <http://doi.org/10.1006/brln.1999.2116>
- Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217–230. <http://doi.org/10.1016/j.jml.2016.10.001>
- Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development*, *81*(2), 517–527. <http://doi.org/10.1111/j.1467-8624.2009.01412.x>

- Wells, J. C. (1995). Computer-coding the IPA: a proposed extension of SAMPA, 1–18.
- Werker, J. F., & Gervain, J. (2013). *Speech Perception in Infancy*. In P. D. Zelazo (Ed.). *Human Auditory Development*. Oxford University Press.
<http://doi.org/10.1093/oxfordhb/9780199958450.013.0031>
- Werker, J. F., & Hensch, T. K. (2015). Critical Periods in Speech Perception: New Directions. *Annual Review of Psychology*, 66(1), 173–196.
<http://doi.org/10.1146/annurev-psych-010814-015104>
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech-perception - evidence for perceptual reorganization during the 1st year of life. *Infant Behavior and Development*, 7(1), 49–63. [http://doi.org/10.1016/S0163-6383\(84\)80022-3](http://doi.org/10.1016/S0163-6383(84)80022-3)
- Wetzel, L. (2014). Types and Tokens. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring, 20, pp. 1–18). Metaphysics Research Lab, Stanford University.
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, 42(3), 671–684. <https://doi.org/10.3758/BRM.42.3.671>
- Willits, J. A., Seidenberg, M. S., & Saffran, J. R. (2009). Proceedings of the Annual Meeting of the Cognitive Science Verbs are Looking Good in Language Acquisition What Do Infants Count? In *Proceedings of the 31st annual conference of the cognitive science society* (pp. 2570–2575). Austin, TX.
- Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best Practices and Advice for Using Pupillometry to Measure Listening Effort: An Introduction for Those Who Want to Get Started. *Trends in Hearing*, 22, 233121651880086.
<https://doi.org/10.1177/2331216518800869>
- Yang, C. D. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Sciences*,

8(10), 451–456. <http://doi.org/10.1016/j.tics.2004.08.006>

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979.

<http://doi.org/10.3758/PBR.15.5.971>

Young, M. E. (2018). A place for statistics in behavior analysis. *Behavior Analysis: Research and Practice*, *18*(2), 193–202. <https://doi.org/10.1037/bar0000099>

Yu, C., & Smith, L. B. (2007). Rapid Word Learning Under Uncertainty via Cross-Situational Statistics. *Psychological Science*, *18*(5), 414–420.

<http://doi.org/10.1111/j.1467-9280.2007.01915.x>

Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental Science*, *14*(2), 165–180.

<http://doi.org/10.1111/j.1467-7687.2010.00958.x>

Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, *145*, 53–62.

<http://doi.org/10.1016/j.cognition.2015.07.013>

Yurovsky, D., Smith, L. B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science*, *16*(6), 959–966.

<http://doi.org/10.1111/desc.12036>

Yurovsky, D., Yu, C., & Smith, L. B. (2012). Statistical speech segmentation and word learning in parallel: Scaffolding from child-directed speech. *Frontiers in Psychology*,

3(OCT), 1–9. <http://doi.org/10.3389/fpsyg.2012.00374>

Zamuner, T. S., Gerken, L., & Hammond, M. (2004). Phonotactic probabilities in young children's speech production. *Journal of Child Language*, *31*(3), 515–536.

<http://doi.org/10.1017/S0305000904006233>

Apêndice A

Pareceres do Comitê de Ética em Pesquisa em Seres Humanos da Universidade Federal de São Carlos (parecer 1.484.847 e 3.085.914) e da *Institutional Review Board da University of Tennessee, Knoxville* (UTK IRB-14-01842 XP).

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: EXTRAÇÃO DE PALAVRAS E MAPEAMENTO COM ESTÍMULOS VISUAIS AO LONGO DE TENTATIVAS AMBÍGUAS SEQUENCIAIS E SIMULTÂNEAS

Pesquisador: Rodrigo Dal Ben de Souza

Área Temática:

Versão: 2

CAAE: 53493216.0.0000.5504

Instituição Proponente: Programa de Pós-Graduação em Psicologia

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.484.847

Apresentação do Projeto:

Há duas tarefas principais para um indivíduo que aprende uma língua oral: extrair palavras da fala contínua (segmentação de fala) e atribuir significado ou mapeá-las a seus referentes. A presente investigação pretende responder duas perguntas: (1) palavras extraídas de uma fala contínua são preferidas no estabelecimento de relações com estímulos visuais ao longo de tentativas ambíguas?; (2) palavras podem ser extraídas de falas contínuas e relacionadas com estímulos visuais simultaneamente ao longo de tentativas ambíguas? Dois experimentos serão realizados. No primeiro, participarão 50 adultos universitários que ouvirão uma fala contínua inventada, cuja única regularidade será a probabilidade transicional entre as sílabas que compõem suas palavras. Em seguida, eles serão expostos a tentativas ambíguas durante as quais duas palavras ou parte das palavras, presentes na fala contínua, ocorrerão concomitantemente a apresentação de duas figuras abstratas. Por fim, as relações entre as palavras ou parte-palavras e figuras serão testadas. No segundo experimento, participarão outros 50 adultos universitários. Eles ouvirão uma fala contínua inventada ao mesmo tempo em que serão expostos a uma sequência de figuras abstratas. Duas figuras serão apresentadas e ficarão disponíveis com a mesma duração de duas palavras da fala contínua, formando tentativas ambíguas. Em sequência, um teste de extração de palavras será realizado, e, por fim, as relações entre palavras e figuras serão testadas.

Endereço: WASHINGTON LUIZ KM 235

Bairro: JARDIM GUANABARA

CEP: 13.565-905

UF: SP

Município: SAO CARLOS

Telefone: (16)3351-9683

E-mail: cephumanos@ufscar.br

Continuação do Parecer: 1.484.847

A presente investigação, ao combinar a probabilidade transicional e a coocorrência de estímulos ao longo de tentativas ambíguas, pretende contribuir para o avanço do conhecimento atual sobre os papéis de tais regularidades no aprendizado de uma língua. Ao mesmo tempo, a proposta inaugura essa linha de investigação em âmbito nacional, podendo complementar investigações voltadas para outros aspectos do desenvolvimento da linguagem.

Objetivo da Pesquisa:

Objetivo Primário:

Investigar: (1) se palavras extraídas de falas contínuas são preferidas no estabelecimento de relações com estímulos visuais ao longo de tentativas ambíguas; e (2) se palavras podem ser extraídas de falas contínuas e relacionadas com estímulos visuais simultaneamente ao longo de tentativas ambíguas.

Avaliação dos Riscos e Benefícios:

Riscos:

A pesquisa não envolverá exposições públicas identificadas tanto da participação quanto do desempenho dos participantes. O anonimato será sempre preservado. Não obstante, a participação poderá gerar constrangimento e ansiedade para os participantes que julgarem seu desempenho como insatisfatório. Nesses casos, o pesquisador explicará a racional da pesquisa e mostrará dados de pesquisas que fundamentaram a presente investigação de modo a mostrar as variações nos desempenhos como algo esperado das amostras. Se a ansiedade ou constrangimento persistir, o pesquisador entrará em contato com um profissional competente da área para um melhor encaminhamento.

Benefícios:

A pesquisa não produz benefícios diretos aos participantes, os benefícios indiretos envolvem o avanço das linhas de investigação sobre os processos de aprendizagem estatística, provavelmente envolvidos na aprendizagem de palavras e significados, avançando assim o desenvolvimento da ciência brasileira sobre processos linguísticos por meio de uma linha de investigação inédita no Brasil.

Endereço: WASHINGTON LUIZ KM 235
Bairro: JARDIM GUANABARA **CEP:** 13.565-905
UF: SP **Município:** SAO CARLOS
Telefone: (16)3351-9683 **E-mail:** cephumanos@ufscar.br

Continuação do Parecer: 1.484.847

Comentários e Considerações sobre a Pesquisa:

Estudo transversal, quantitativo, não randomizado, sobre aquisição de linguagem.

Considerações sobre os Termos de apresentação obrigatória:

Estão presentes todos os Termos de apresentação obrigatória

Recomendações:

Conclusões ou Pendências e Lista de Inadequações:

Não há nenhuma pendência.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_649126.pdf	18/03/2016 18:19:17		Aceito
Outros	CEPEsclarecimentoPendencia.docx	18/03/2016 18:17:13	Rodrigo Dal Ben de Souza	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	CEPTCLERodrigo.docx	29/01/2016 13:42:37	Rodrigo Dal Ben de Souza	Aceito
Projeto Detalhado / Brochura Investigador	CEPRodrigoProjetoDoutorado.docx	29/01/2016 13:41:29	Rodrigo Dal Ben de Souza	Aceito
Folha de Rosto	CEPfolhaDeRostoAssinada.pdf	29/01/2016 13:40:54	Rodrigo Dal Ben de Souza	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SAO CARLOS, 08 de Abril de 2016

Assinado por:
Ricardo Carneiro Borra
(Coordenador)

Endereço: WASHINGTON LUIZ KM 235
Bairro: JARDIM GUANABARA CEP: 13.565-905
UF: SP Município: SAO CARLOS
Telefone: (16)3351-9683 E-mail: cephumanos@ufscar.br



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: EXTRAÇÃO DE PALAVRAS E MAPEAMENTO COM ESTÍMULOS VISUAIS AO LONGO DE TENTATIVAS AMBÍGUAS SEQUENCIAIS E SIMULTÂNEAS

Pesquisador: Rodrigo Dal Ben de Souza

Área Temática:

Versão: 3

CAAE: 53493216.0.0000.5504

Instituição Proponente: Programa de Pós-Graduação em Psicologia

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.085.914

Apresentação do Projeto:

Resumo:

Há duas tarefas principais para um indivíduo que aprende uma língua oral: extrair palavras da fala contínua (segmentação de fala) e atribuir significado ou mapeá-las a seus referentes. Uma fonte de informação importante para esse aprendiz são as regularidades presentes nessa língua oral, mais especificamente, sua prosódia, organização fonotática, probabilidade transicional. Por exemplo, há evidências de que a maior ou menor probabilidade de uma sílaba ser seguida por outra sílaba (probabilidade transicional) pode informar sobre os limites de uma palavra em uma fala contínua (quando ela começa e quando termina). Os processos de extração de palavras e mapeamento das mesmas aos seus referentes (e.g., estímulos visuais) podem, no entanto, ocorrer sequencialmente ou simultaneamente, por meio de tarefas de emparelhamento entre estímulos. Pesquisas recentes têm demonstrado o papel da coocorrência de estímulos auditivos e visuais no estabelecimento de relações entre tais estímulos, ao longo de tentativas ambíguas. Cada tentativa envolve a apresentação de duas ou mais palavras faladas e duas ou mais figuras. No entanto, há um número ainda muito restrito de estudos investigando esses dois processos conjuntamente: os processos de extração de palavras com base na probabilidade transicional e o estabelecimento de relações entre essas palavras e estímulos visuais com base em suas coocorrências ao longo de tentativas ambíguas. Seguindo essa direção, a presente investigação pretende responder duas perguntas: (1) palavras extraídas de uma fala contínua são preferidas no estabelecimento de

Endereço: WASHINGTON LUIZ KM 235

Bairro: JARDIM GUANABARA

CEP: 13.565-905

UF: SP

Município: SAO CARLOS

Telefone: (16)3351-9683

E-mail: cephumanos@ufscar.br



relações com estímulos visuais ao longo de tentativas ambíguas?; (2) palavras podem ser extraídas de falas contínuas e relacionadas com estímulos visuais simultaneamente ao longo de tentativas ambíguas? Dois experimentos serão realizados. No primeiro, participarão 50 adultos universitários que ouvirão uma fala contínua inventada, cuja única regularidade será a probabilidade transicional entre as sílabas que compõem suas palavras. Em seguida, eles serão expostos a tentativas ambíguas durante as quais duas palavras ou parte de palavras, presentes na fala contínua, coocorrerão com duas figuras abstratas. Por fim, as relações entre as palavras ou parte-palavras e figuras serão testadas. No segundo experimento, participarão outros 50 adultos universitários. Eles ouvirão uma fala contínua inventada ao mesmo tempo em que serão expostos a uma sequência de figuras abstratas. Duas figuras serão apresentadas e ficarão disponíveis com a mesma duração de duas palavras da fala contínua, formando tentativas ambíguas. Em sequência, um teste de extração de palavras será realizado, e, por fim, as relações entre palavras e figuras serão testadas. A presente investigação, ao combinar a probabilidade transicional e a coocorrência de estímulos ao longo de tentativas ambíguas, pretende contribuir para o avanço do conhecimento atual sobre os papéis de tais regularidades no aprendizado de uma língua. Ao mesmo tempo, a proposta inaugura essa linha de investigação em âmbito nacional, podendo complementar investigações voltadas para outros aspectos do desenvolvimento da linguagem.

Objetivo da Pesquisa:

Objetivo Primário:

Investigar: (1) se palavras extraídas de falas contínuas são preferidas no estabelecimento de relações com estímulos visuais ao longo de tentativas ambíguas; e (2) se palavras podem ser extraídas de falas contínuas e relacionadas com estímulos visuais simultaneamente ao longo de tentativas ambíguas.

Avaliação dos Riscos e Benefícios:

Riscos:

A pesquisa não envolverá exposições públicas identificadas tanto da participação quanto do desempenho dos participantes. O anonimato será sempre preservado. Não obstante, a participação poderá gerar constrangimento e ansiedade para os participantes que julgarem seu desempenho como insatisfatório. Nesses casos, o pesquisador explicará a racional da pesquisa e mostrará dados de pesquisas que fundamentaram a presente investigação de modo a mostrar as variações nos desempenhos como algo esperado das amostras. Se a ansiedade ou constrangimento persistir, o pesquisador entrará em contato com um profissional competente da área para um melhor

Endereço: WASHINGTON LUIZ KM 235
Bairro: JARDIM GUANABARA **CEP:** 13.565-905
UF: SP **Município:** SAO CARLOS
Telefone: (16)3351-9683 **E-mail:** cephumanos@ufscar.br



Continuação do Parecer: 3.085.914

encaminhamento.

Benefícios:

A pesquisa não produz benefícios diretos aos participantes, os benefícios indiretos envolvem o avanço das linhas de investigação sobre os processos de aprendizagem estatística, provavelmente envolvidos na aprendizagem de palavras e significados, avançando assim o desenvolvimento da ciência brasileira sobre processos linguísticos por meio de uma linha de investigação inédita no Brasil.

Comentários e Considerações sobre a Pesquisa:

Trata-se de uma solicitação de inclusão de uma terceira fase, virtual em pesquisa que aborda a aprendizagem de idioma com a fala contínua além desta também apresentada simultaneamente a estímulos visuais.

Considerações sobre os Termos de apresentação obrigatória:

Estão presentes todos os Termos de apresentação obrigatória.

Recomendações:

No TCLE a frase: Os pesquisadores não obterão qualquer retorno financeiro ou lucro por sua participação, não explicita que os Participantes da Pesquisa não obterão qualquer retorno financeiro ou lucro por sua participação.

Conclusões ou Pendências e Lista de Inadequações:

Solicita-se a inclusão no TCLE a informação que o PARTICIPANTE da pesquisa não obterá qualquer retorno financeiro ou lucro por sua participação.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_1262093_E1.pdf	27/11/2018 12:14:34		Aceito
Projeto Detalhado / Brochura Investigador	CEPRodrigoProjeto_emenda.pdf	27/11/2018 12:10:07	Rodrigo Dal Ben de Souza	Aceito
Outros	CEP_emenda.pdf	27/11/2018 12:09:22	Rodrigo Dal Ben de Souza	Aceito
TCLE / Termos de Assentimento / Justificativa de	CEPTCLERodrigo.docx	29/01/2016 13:42:37	Rodrigo Dal Ben de Souza	Aceito

Endereço: WASHINGTON LUIZ KM 235

Bairro: JARDIM GUANABARA

CEP: 13.565-905

UF: SP

Município: SAO CARLOS

Telefone: (16)3351-9683

E-mail: cephumanos@ufscar.br



Continuação do Parecer: 3.085.914

Ausência	CEPTCLERodrigo.docx	29/01/2016 13:42:37	Rodrigo Dal Ben de Souza	Aceito
Folha de Rosto	CEPfolhaDeRostoAssinada.pdf	29/01/2016 13:40:54	Rodrigo Dal Ben de Souza	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SAO CARLOS, 15 de Dezembro de 2018

Assinado por:
Priscilla Hortense
(Coordenador(a))

Endereço: WASHINGTON LUIZ KM 235
Bairro: JARDIM GUANABARA **CEP:** 13.565-905
UF: SP **Município:** SAO CARLOS
Telefone: (16)3351-9683 **E-mail:** cephumanos@ufscar.br



THE UNIVERSITY OF
TENNESSEE
KNOXVILLE

September 25, 2018
Jessica Sari Hay,
UTK - College of Arts & Sciences - Psychology

Re: UTK IRB-14-01842 XP

Study Title: Testing the Effects of Experience Both Inside and Outside of the Laboratory on Speech Perception and Word Learning in Adults

Dear Jessica Sari Hay:

The UTK Institutional Review Board (IRB) reviewed your application for **revision** of your previously approved project, referenced above.

The IRB determined that your application is eligible for **expedited** review under 45 CFR 46.110(b)(2). The following revisions were approved as complying with proper consideration of the rights and welfare of human subjects and the regulatory requirements for the protection of human subjects:

- **Approved Changes:**
 - Added Rodrigo Dal Ben de Souza to the application in the role of co-investigator.
 - Removed Rebecca Bauer and Dora Moore from the Research Staff.
 - Changed Daniela's role to Lab/Tech Manager.
 - Added demographic survey.
- **Approved Documents:**
 - Application v1.9
 - post_experimental_quest_en v1.0

Approval does not alter the expiration date of this project, which is 12/25/2018.

In the event that subjects are to be recruited using solicitation materials, such as brochures, posters, web-based advertisements, etc., these materials must receive prior approval of the IRB. Any revisions in the approved application must also be submitted to and approved by the IRB prior to implementation. In addition, you are responsible for reporting any unanticipated serious adverse events or other problems involving risks to subjects or others in the manner required by the local IRB policy.

Finally, **re-approval** of your project is required by the IRB in accord with the conditions specified above. You may not continue the research study beyond the time or other limits specified unless you obtain prior written approval of the IRB.

Institutional Review Board | Office of Research & Engagement
1534 White Avenue Knoxville, TN 37996-1529
865-974-7697 865-974-7400 fax irb@utk.edu

BIG ORANGE. BIG IDEAS.

Flagship Campus of the University of Tennessee System

Sincerely,

Colleen P. Gilrane

Colleen P. Gilrane, Ph.D.
Chair

Institutional Review Board | Office of Research & Engagement
1534 White Avenue Knoxville, TN 37996-1529
865-974-7697 865-974-7400 fax irb.utk.edu

BIG ORANGE. BIG IDEAS.

Flagship Campus of the University of Tennessee System 

Apêndice B

Termo de Consentimento Livre e Esclarecido

Você está sendo convidado (a) a participar da pesquisa Extração de Palavras e Mapeamento com Estímulos Visuais ao Longo de Tentativas Ambíguas Sequenciais e Simultâneas. Esta pesquisa investiga os processos de extrações de palavras de falas contínuas e mapeamento dessas palavras com figuras. Você está na faixa etária de interesse para o presente estudo e por isso está sendo convidado a participar.

Durante a sua participação, você deverá ouvir palavras, olhar para figuras, e selecionar figuras ao ouvir palavras. Você poderá desistir de participar e retirar seu consentimento a qualquer momento; basta informar o experimentador. Sua recusa não trará qualquer prejuízo à sua relação com os pesquisadores ou com a Universidade Federal de São Carlos. Não há benefícios diretos para você, mas a participação é uma oportunidade de contribuir para o avanço da ciência psicológica brasileira. Os possíveis riscos da sua participação são: ficar cansado, entediado, ou constrangido. Se algum desses incômodos ocorrer, o experimentador lhe explicará novamente os objetivos do estudo e, se você desejar, sua participação poderá ser interrompida. Caso ocorra qualquer problema não previsto, o pesquisador o encaminhará para um profissional competente.

Durante toda a pesquisa e publicação, seu anonimato e sigilo serão garantidos. Os dados da pesquisa serão analisados apenas pelos pesquisadores responsáveis e poderão ser submetidos à publicação científica. Os pesquisadores não obterão qualquer retorno financeiro ou lucro por sua participação.

Você receberá uma cópia deste termo no qual consta o telefone e o endereço do pesquisador principal, podendo tirar dúvidas sobre a pesquisa a qualquer momento.

São Carlos, _____ de _____ de 201__

Rodrigo Dal Ben de Souza - Pesquisador responsável

Endereço: Departamento de Psicologia, UFSCar, Rodovia Washington Luiz, Km. 235 –
Caixa Postal 676 – CEP 13.565-905 – São Carlos–SP – Laboratório de Interação Social
(LIS/ UFSCar), Fone: (16) 3351-9591. E-mail: dalbenwork@gmail.com

Declaro que entendi os objetivos, riscos e benefícios da minha participação na pesquisa e concordo em participar. O pesquisador me informou que a pesquisa foi aprovada pelo Comitê de Ética em Pesquisa em Seres Humanos da UFSCar, sediado na Pró-Reitoria de Pós-Graduação e Pesquisa da Universidade Federal de São Carlos, Rodovia Washington Luiz, Km. 235 – Caixa Postal 676 – CEP 13.565-905 – São Carlos – SP – Brasil. Fone (16) 3351-8110. Endereço eletrônico: cephumanos@power.ufscar.br

Nome completo do participante:

Assinatura: _____

São Carlos-SP, _____ de _____ de 201_

Apêndice C

INFORMED CONSENT STATEMENT

Project Title: Testing the Effects of Experience Both Inside and Outside of the Laboratory on Speech Perception and Word Learning in Adults

INTRODUCTION

You are being asked to participate in a research study investigating how people learn from the language they hear in their environment. This research is being conducted in the Infant Language and Perceptual Learning Laboratory under the guidance of Dr. Jessica F. Hay, Assistant Professor in the Department of Psychology.

INFORMATION ABOUT THE STUDY

This study is designed to determine how previous experience with language shapes adults' preferences for and abilities to learn new things about language.

~~Memory Task: If you have signed up for our memory study, you have signed up to come into the lab twice. Your second visit will be in 4 / 24 hours from now. Today you will participate in the familiarization phase of the experiment (see below for details). When you return to the lab for your second visit you will complete either the Word learning task or the 2AFC task (see below for details).~~

Familiarization Phase: In the first part of the experiment you will be presented series of sounds from either an artificial or natural language. While you are listening to the language you will be asked to do an unrelated task. Following this familiarization phase we will assess what you have learned from the language through one of two following procedures:

1. Word (Label) Learning Task: In this phase you will be presented with a series of objects and labels. Your job is to try to learn which label goes with which object. Following training on the label- object pairings you will be asked to identify which visual images match the auditory label by pressing a key on the keyboard and/or by gaze response. Your gaze will be monitored non- obtrusively using an eye tracker.

OR

2. Two-Alternative Forced-Choice Task (2AFC): In this phase you will be presented with two words. Your job is to indicate, using a key press, which word is more similar or most likely to have come from the language that you heard during the familiarization phase of the experiment. After selecting your response, you will be asked to indicate whether you are just guessing, pretty sure of, or absolutely sure of your response.

The experiment is short and takes approximately 30-60 minutes, total. In addition to the experimental procedure, you will be asked to fill out a General Information Questionnaire.

Your total time commitment to this experiment will not exceed 60 minutes across 1-2 testing sessions.

RISKS

There are no known risks to you for your participation in this study. You may choose to stop at any time for any reason.

BENEFITS

This study is designed to contribute to knowledge of how people, in general, learn about language. There are no direct benefits to you for having participated in this study.

CONFIDENTIALITY

Participants are identified on all records solely by a number, insuring confidentiality of the data. Any information that is obtained in connection with this study and that could identify you will remain confidential, and will not be released or disclosed without your further consent, except as specifically required by law. Once signed, this consent form will be stored in a locked filing cabinet in the Infant Language and Perceptual Development Lab for at least 3 years.

COMPENSATION

If you sign-up through the Psychology SONA system, you will receive one hour of credit toward General Psychology course requirements for your participation. Otherwise, you will receive a 10 dollars Amazon gift card for your participation.

EMERGENCY MEDICAL TREATMENT

The University of Tennessee does not “automatically” reimburse participants for medical claims or other compensations. If physical injury is suffered in the course of the research, or for more information, please notify the investigator in charge (Dr. Jessica F. Hay, (865) 974-3329).

If you have any questions about your rights as a participant, contact the Office of Research Compliance Officer at 865-974-7697.

Your signature indicates that you have read this informed consent, have been fully advised of the nature of the procedure and the possible risks and complications involved in it, understand that no compensation will be given in the event of physical injury to you, and give permission for participation of you. You are free to withdraw your consent and discontinue participation in the study at any time without prejudice. A copy of this informed consent form will also be given to you. Any questions, you have concerning the study will be answered by contacting Jessica F. Hay, Ph.D., Department of Psychology, University of Tennessee, Knoxville, TN 37996, (865) 974-0514.

Participant Signature _____ Date _____

Experimenter Signature _____ Date _____

IRB NUMBER: UTK IRB-14-01842 XP

IRB APPROVAL DATE: 05/09/2019

IRB EXPIRATION DATE: 12/24/2019

Apêndice D

Questionário Pós-experimental

Idade: _____ Gênero: () Masculino () Feminino () Outro

Nível educacional: () Superior em andamento, curso: _____
() Superior completo (graduação, pós), curso: _____
() Outro (ensino médio): _____

Você nasceu no Brasil? () Sim () Não, eu nasci no(a) _____

Você já morou em um país estrangeiro? () Não () Sim
Se sim, por quanto tempo e com qual idade? _____

Você é falante nativo (a) do português? () Sim () Não, minha língua nativa é o _____

Além do português, você fala outras línguas? () Não () Sim

Se sim, indique quais são as línguas e o seu nível de proficiência em cada uma delas:

1) Língua: _____

a) Qual era a sua idade quando você começou a aprender essa língua? _____ anos de idade;

b) Qual é o seu nível de proficiência (sendo 1 iniciante e 7 praticamente um falante nativo):

() 1 () 2 () 3 () 4 () 5 () 6 () 7

2) Língua: _____

a) Qual era a sua idade quando você começou a aprender essa língua? _____ anos de idade;

b) Qual é o seu nível de proficiência (sendo 1 iniciante e 7 praticamente um falante nativo):

() 1 () 2 () 3 () 4 () 5 () 6 () 7

3) Língua: _____

a) Qual era a sua idade quando você começou a aprender essa língua? _____ anos de idade;

b) Qual é o seu nível de proficiência (sendo 1 iniciante e 7 praticamente um falante nativo):

() 1 () 2 () 3 () 4 () 5 () 6 () 7

4) Língua: _____

a) Qual era a sua idade quando você começou a aprender essa língua? _____ anos de idade;

b) Qual é o seu nível de proficiência (sendo 1 iniciante e 7 praticamente um falante nativo):

() 1 () 2 () 3 () 4 () 5 () 6 () 7

Você possui alguma condição que afete a sua percepção auditiva (por exemplo, perda auditiva, hipersensibilidade, dificuldade de processamento auditivo)?

() Não () Sim, _____

Você possui alguma condição que afete a sua percepção visual (por exemplo, perda de visão, hipersensibilidade, dificuldade de processamento visual)?

() Não () Sim, _____

Como ficou sabendo da pesquisa: () Facebook () Outro: _____

Apêndice E

Post-experimental questionnaire (n° ____)

Age: _____

Gender: Male Female Other

Educational background: Undergraduate in progress, major: _____

Graduate in progress, major: _____

Undergraduate completed, major: _____

Graduate completed, major: _____

Other (high school completed): _____

Were you born in the United States? Yes No, I was born in _____

Have you ever lived in a foreign country? No Yes

If yes, for how long and how old was you? _____

Are you an English native speaker? Yes No, my native language is _____

Apart from English, do you speak other languages? No Yes

If yes, indicate each language you speak and your level of proficiency for each of them:

1) Language: _____

a) How old were you when you began to learn this language? _____ years old;

b) What is your proficiency level (1 being starter and 7 practically a native)

1 2 3 4 5 6 7

2) Language: _____

a) How old were you when you began to learn this language? _____ years old;

b) What is your proficiency level (1 being starter and 7 practically a native)

1 2 3 4 5 6 7

3) Language: _____

a) How old were you when you began to learn this language? _____ years old;

b) What is your proficiency level (1 being starter and 7 practically a native)

1 2 3 4 5 6 7

4) Language: _____

a) How old were you when you began to learn this language? _____ years old;

b) What is your proficiency level (1 being starter and 7 practically a native)

1 2 3 4 5 6 7

Do you have any hearing impairment (for example, hearing loss)?

No Yes, _____

Do you have any visual impairment (for example, photophobia)?

No Yes, _____

Apêndice F

Ficha de registro de Wordlikeness

Convite

Você está sendo convidado para participar da avaliação dos estímulos da pesquisa intitulada “*Extração de Palavras e Mapeamento com Estímulos Visuais ao Longo de Tentativas Ambíguas Sequenciais e Simultâneas*”. Se você concordar em participar, você irá avaliar o quanto as *pseudopalavras* (não-palavras) apresentadas se assemelham com as palavras que você conhece do português.

As suas respostas serão utilizadas exclusivamente na seleção dos estímulos da pesquisa já mencionada. Suas respostas e identidade serão mantidas em sigilo e a qualquer momento você poderá desistir de participar.

Termo de Consentimento Livre e Esclarecido

Eu, _____, RA: _____
declaro que: a) entendi os objetivos da pesquisa e concordo em participar; b) fui informado(a) de que minhas respostas e identidade serão mantidas em sigilo durante toda a pesquisa; c) posso desistir da participação a qualquer momento.

Informações básicas

Idade: _____

Além do português, você conhece alguma outra língua? () Não () Sim, quais:

Língua: _____ Nível de proficiência: () Iniciante, () Intermediário, () Avançado

Língua: _____ Nível de proficiência: () Iniciante, () Intermediário, () Avançado

Língua: _____ Nível de proficiência: () Iniciante, () Intermediário, () Avançado

Instruções

Cada *pseudopalavra* (não-palavra) será seguida de uma escala com 7 opções, conforme o quadro abaixo:

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baixo-impossível			Médio-Neutro			Alto-possível

Baixo-impossível: essa palavra nunca poderia ser uma palavra em Português;

Médio-neutro: essa palavra é igualmente boa e ruim como uma palavra do Português;

Alto-possível: essa palavra poderia ser uma palavra do Português facilmente.

Para cada palavra, escolha a opção que você julgar mais adequada.

Folha de registro

0 – Exemplo

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baixo-impossível			Médio-Neutro			Alto-possível

1

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baixo-impossível			Médio-Neutro			Alto-possível

2

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baixo-impossível			Médio-Neutro			Alto-possível

Idem até a palavra n. 48.

Apêndice G

Sequência pseudorrandômica das palavras para um dado segmento da fala contínua. Tal sequência foi utilizada na criação de todas as línguas (Tabela 1). Os números 1, 2, 3 representam as palavras com alta frequência (300 repetições), os números 4, 5, 6 representam as palavras com baixa frequência (150 repetições). O quadro deve ser lido por linha.

3	2	1	6	2	1
3	6	4	3	2	1
3	2	5	6	5	4
3	1	3	4	6	1
2	4	1	3	4	5
2	1	2	1	2	6
5	3	2	1	6	5
1	2	4	5	3	2
1	3	6	4	1	3
2	4	3	2	1	3
5	2	3	1	2	3
1	5	3	2	1	3
2	1	4	2	6	1
3	2	3	5	6	4
1	2	5	1	3	6

Apêndice H

Jogos distratores utilizados no Estudo 1 com falantes do PB e do inglês, respectivamente.



Apêndice I

A seguir são apresentadas três tabelas com as instruções dos três Estudos desenvolvidos no presente trabalho. Vale notar que exatamente as mesmas instruções foram traduzidas para o inglês e utilizadas nos experimentos com falantes do inglês.

Tabela 1A.

Instruções apresentadas durante os experimentos do Estudo 1

Número	Instrução
1	“Olá, seja bem vindo! Por favor, coloque o fone de ouvido e pressione a tecla ‘Enter’”
2	“Antes de começar, ajuste o volume. Pressione “-“ para diminuir e “+” para aumentar o volume. Pressione ‘Enter’ para continuar.”
3	“Fique à vontade para ajustar o volume durante todo o experimento. Pressione ‘Enter’ para continuar”
4	“A seguir, você ouvirá uma nova língua! Enquanto ouve a língua, resolva os jogos que estão na caixa ao lado. Você não precisa apertar nenhum botão enquanto ouve a língua. Você terá 15 minutos para resolver os jogos. Pressione "Enter" para começar” <i>OU</i> “A seguir, você ouvirá uma nova língua! Sua tarefa é descobrir as palavras dessa nova língua. No começo, poderá ser difícil descobrir as palavras, mas com o tempo espera-se que a tarefa fique mais fácil. Ao final, o seu conhecimento sobre as palavras será testado! Pressione "Enter" para continuar”*
5	“Muito Bem! Agora coloque os jogos dentro da caixa. A seguir, você vai ouvir duas palavras, uma delas existe no português! Indique qual palavra existe no português. Pressione "Enter" para continuar”**
6	“Lembre-se de que você pode ajustar o volume se necessário. Pressione "Enter" para continuar”
7	“Pressione “1” se a primeira palavra existir no português; Pressione “2” se a segunda palavra existir no português; Pressione “Enter” para continuar”
8	“Ótimo! O treino acabou! Agora você vai ouvir palavras da língua que você ouviu anteriormente. Pressione "Enter" para continuar”
9	“As tentativas terão o mesmo formato do treino. Você ouvirá duas palavras por vez, uma delas é mais parecida com as palavras da língua que você ouviu. Serão apresentadas 18 tentativas (+- 3 minutos). Dessa vez, eu não vou dizer se você acertou ou errou. Pressione "Enter" para continuar”
10	“Pressione “1” se a primeira palavra for a mais parecida; Pressione “2” se a segunda palavra for a mais parecida; Pressione “Enter” para continuar”
11	“Muito bem! O estudo está quase no final! Por favor, indique qual a porcentagem de palavras que você acha que conseguiu identificar. Pressione "Enter" para continuar”
12	“Pressione: 1 = até 25%; 2 = entre 25 e 50%; 3 = entre 50 e 75%; 4 = acima de 75%.”
13	“Agora, indique o quanto você se concentrou nos jogos enquanto ouvia a língua. Pressione "Enter" para continuar”
14	“Pressione: 1 = nada concentrado; 2 = pouco concentrado; 3 = concentrado; 4 = muito concentrado.”
15	“Muito bem! Você terminou o estudo! Você já pode tirar o fone de ouvido. Pressione ‘Enter’ para chamar o pesquisador”

Nota. *a primeira instrução foi apresentada nos experimentos cujas tarefas eram implícitas (experimentos 1, 3, 4, 5, 6, e 7), a segunda instrução foi apresentada no experimento cuja tarefa era explícita (experimento 2).

**Essa instrução não foi apresentada no Experimento 2 (tarefa explícita).

Tabela 2A.

Instruções apresentadas durante os experimentos do Estudo 2

Número	Instrução
1	“Olá, seja bem vindo! Por favor, coloque o fone de ouvido e pressione a tecla ‘Enter’”
2	“Antes de começar, ajuste o volume. Pressione “-“ para diminuir e “+” para aumentar o volume. Pressione ‘Enter’ para continuar.”
3	“Fique à vontade para ajustar o volume durante todo o experimento. Pressione ‘Enter’ para continuar”
4	“A seguir você verá várias figuras e ouvirá várias palavras. A sua tarefa é descobrir as relações entre as palavras e as figuras. Pressione ‘Enter’ para continuar”
5	“As figuras e palavras aparecerão automaticamente. Você não precisa dar nenhum comando. Apenas preste atenção. Lembre-se: a sua tarefa é descobrir as relações entre as palavras e as figuras. Pressione ‘Enter’ para continuar”
6	“Muito bem! Agora, indique qual dentre as quatro figuras corresponde à palavra que você ouvir. Vamos começar treinando com palavras conhecidas. Pressione ‘Enter’ para continuar”
7	“Pressione o número correspondente à pergunta que você vai ouvir. 1, 2, 3, ou 4. Pressione ‘Enter’ para continuar”
8	“Ótimo! O treino acabou! Agora você vai ouvir as palavras e ver as figuras que você viu e ouviu anteriormente. Pressione ‘Enter’ para continuar”
9	“As tentativas terão o mesmo formato do treino. Indique a figura correspondente à palavra que você ouvir. Serão apresentadas 24 tentativas (+- 4 minutos). Pressione ‘Enter’ para continuar”
10	“Muito bem! O estudo está quase no final! Por favor, indique qual a porcentagem de relações que você acha que conseguiu identificar. Pressione ‘Enter’ para continuar”
11	“Pressione: 1 = até 25%; 2 = entre 25 e 50%; 3 = entre 50 e 75%; 4 = acima de 75%.”
12	“Muito bem! Você terminou o estudo! Você já pode tirar o fone de ouvido. Pressione ‘Enter’ para chamar o pesquisador”

Tabela 3A.

Instruções apresentadas durante os experimentos do Estudo 3

Número	Instrução
1	“Olá, seja bem vindo! Por favor, coloque o fone de ouvido e pressione a tecla ‘Enter’”
2	“Antes de começar, ajuste o volume. Pressione “-“ para diminuir e “+” para aumentar o volume. Pressione ‘Enter’ para continuar.”
3	“Fique à vontade para ajustar o volume durante todo o experimento. Pressione ‘Enter’ para continuar”
4	“A seguir você verá várias figuras e ouvirá várias palavras. A sua tarefa é descobrir as relações entre as palavras e as figuras. Pressione ‘Enter’ para continuar”
5	“As figuras e palavras aparecerão automaticamente. Você não precisa dar nenhum comando. Apenas preste atenção. Lembre-se: a sua tarefa é descobrir as relações entre as palavras e as figuras. Pressione ‘Enter’ para continuar”
6	“Muito Bem! A seguir, você vai ouvir duas palavras, uma delas existe no português! Indique qual palavra existe no português. Pressione "Enter" para continuar.”
7	“Lembre-se de que você pode ajustar o volume se necessário. Pressione "Enter" para continuar”
8	“Ótimo! O treino acabou! Agora você vai ouvir palavras da língua que você ouviu anteriormente. Pressione "Enter" para continuar”
9	“As tentativas terão o mesmo formato do treino. Você ouvirá duas palavras por vez, uma delas é mais parecida com as palavras da língua que você ouviu. Serão apresentadas 18 tentativas (+- 3 minutos). Dessa vez, eu não vou dizer se você acertou ou errou. Pressione "Enter" para continuar”
10	“Muito bem! O estudo está quase no final! Por favor, indique qual a porcentagem de palavras que você acha que conseguiu identificar. Pressione "Enter" para continuar”
11	“Muito bem! Agora, indique qual dentre as quatro figuras corresponde à palavra que você ouvir. Vamos começar treinando com palavras conhecidas. Pressione ‘Enter’ para continuar”
12	“Pressione o número correspondente à pergunta que você vai ouvir. 1, 2, 3, ou 4. Pressione ‘Enter’ para continuar”
13	“Ótimo! O treino acabou! Agora você vai ouvir as palavras e ver as figuras que você viu e ouviu anteriormente. Pressione ‘Enter’ para continuar”
14	“As tentativas terão o mesmo formato do treino. Indique a figura correspondente à palavra que você ouvir. Serão apresentadas 24 tentativas (+- 4 minutos). Pressione ‘Enter’ para continuar”
15	“Muito bem! O estudo está quase no final! Por favor, indique qual a porcentagem de relações que você acha que conseguiu identificar. Pressione ‘Enter’ para continuar”
16	“Pressione: 1 = até 25%; 2 = entre 25 e 50%; 3 = entre 50 e 75%; 4 = acima de 75%.”
17	“Muito bem! Você terminou o estudo! Você já pode tirar o fone de ouvido. Pressione ‘Enter’ para chamar o pesquisador”