

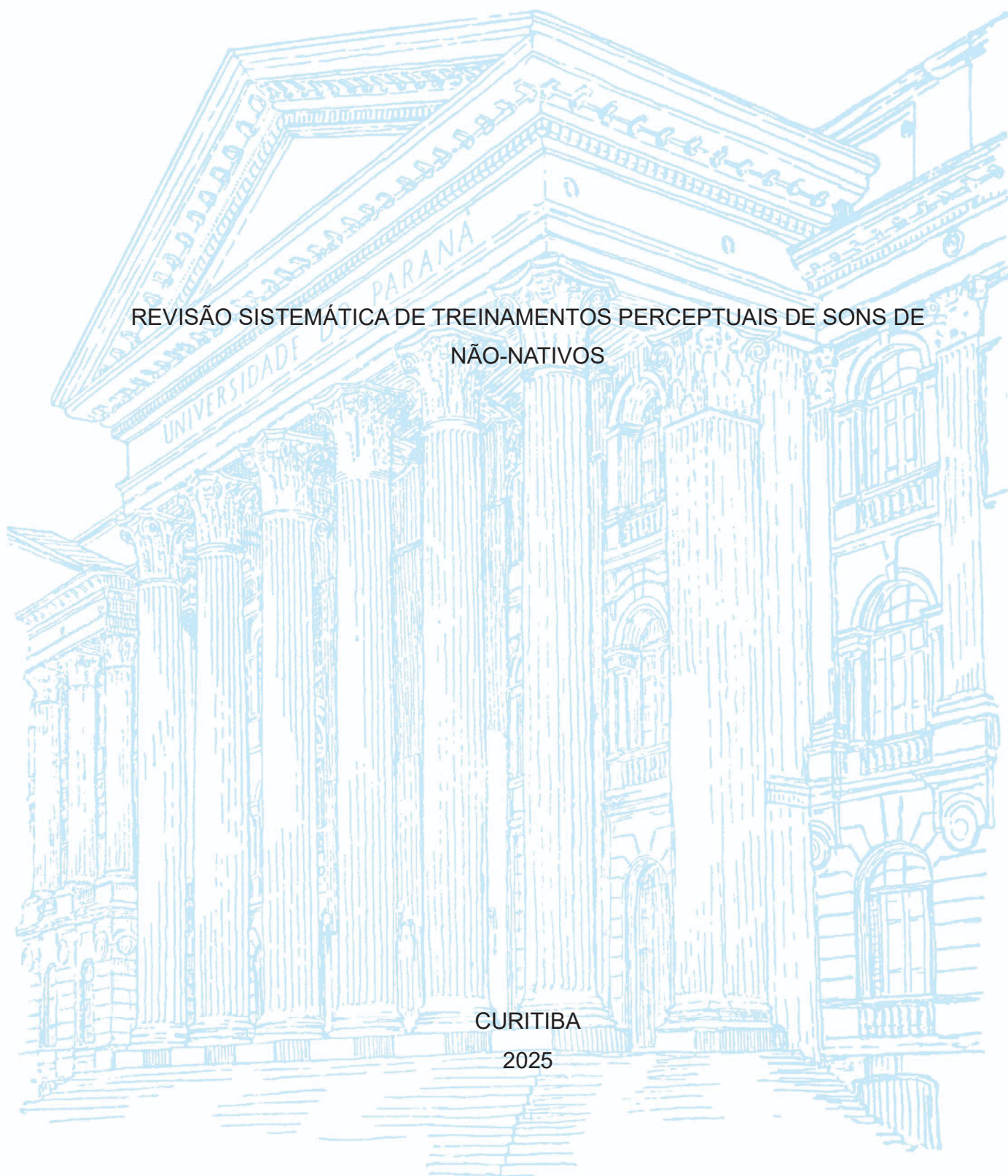
UNIVERSIDADE FEDERAL DO PARANÁ

OTAVIO AUGUSTO RODRIGUES BERNARDO SILVA

REVISÃO SISTEMÁTICA DE TREINAMENTOS PERCEPTUAIS DE SONS DE
NÃO-NATIVOS

CURITIBA

2025



OTAVIO AUGUSTO RODRIGUES BERNARDO SILVA

REVISÃO SISTEMÁTICA DE TREINAMENTOS PERCEPTUAIS DE SONS DE
NÃO-NATIVOS

Tese apresentada ao curso de Pós-Graduação em Letras, Setor de Ciências Humanas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Linguística.

Orientadora: Prof^a. Dr^a. Denise Cristina Kluge

CURITIBA

2025

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIAS HUMANAS

Silva, Otavio Augusto Rodrigues Bernardo
Revisão sistemática de treinamentos perceptuais de sons não nativos.
/ Otavio Augusto Rodrigues Bernardo Silva. – Curitiba, 2025.
1 recurso on-line : PDF.

Tese (Doutorado) – Universidade Federal do Paraná, Setor de
Ciências Humanas, Programa de Pós-Graduação em Letras.
Orientadora: Profª. Drª. Deise Cristina Kluge.

1. Língua inglesa – Estudo e ensino – Falantes estrangeiros.
2. Fonética. 3. Linguística. I. Kluge, Deise Cristina, 1975-. II. Universidade
Federal do Paraná. Programa de Pós-Graduação em Letras. III. Título.

Bibliotecário: Dênis Junio de Almeida CRB-9/2092



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS HUMANAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO LETRAS -
40001016016P7

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação LETRAS da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **OTAVIO AUGUSTO RODRIGUES BERNARDO SILVA**, intitulada: **Revisão Sistemática de Treinamentos Perceptuais de Sons Não Nativos**, sob orientação da Profa. Dra. DENISE CRISTINA KLUGE, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 22 de Maio de 2025.

Assinatura Eletrônica
29/05/2025 11:26:07.0
DENISE CRISTINA KLUGE
Presidente da Banca Examinadora

Assinatura Eletrônica
30/05/2025 13:08:50.0
MARIA LUCIA DE CASTRO GOMES
Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO
PARANÁ)

Assinatura Eletrônica
29/05/2025 15:26:42.0
UBIRATÃ KICKHÖFEL ALVES
Avaliador Externo (UNIVERSIDADE FEDERAL DO RIO GRANDE DO
SUL)

Assinatura Eletrônica
30/05/2025 11:14:15.0
GUSTAVO NISHIDA
Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO
PARANÁ)

Assinatura Eletrônica
03/06/2025 18:23:26.0
ANGELICA CARLET
Avaliador Externo (CHARLES DARWIN UNIVERSITY (CDU))

AGRADECIMENTOS

Devo agradecer muito à minha mãe, Sara Cristina Rodrigues, e meu padrasto, Milton da Silva, por essa introdução ao mundo das letras, através de HQs quando eu tinha cinco anos. Devo agradecer muito aos meus tios e avós por sempre alimentarem o dever e a necessidade de estudar durante toda minha vida.

Em minha graduação, tive contato com inúmeras áreas de estudo: Fonética e Fonologia, Estudos do Texto, Português para Falantes de Outras Línguas, Tradução Técnica e Literária, Teoria Literária e Psicolinguística. Agradeço muito a muitos professores meus do ensino técnico, como Claudio Eugênio Canabrava Barbalho, Juliano Vieira, e todos meus professores da graduação na UTFPR, em especial Malu Gomes, Andressa Albini (que também foi banca do meu mestrado), Jeniffer Albuquerque, Ana Maria Martins, Jaqueline Lindstron, Jaqueline Donada, Ana Valéria, Ana Paula Petriu, Edson Fagundes, Elisa Cordeiro e Miriam Retorta. A disciplina de Introdução à Pesquisa em Linguagem, ministrada então pela Profª Drª Maria L.C. Gomes, certamente foi fundamental para instigar meu interesse em pesquisa. Devo agradecer também aos professores Gustavo Nishida e Denize Nobre pela participação na minha banca de TCC1 e TCC2 na UTFPR. Agradeço muitíssimo às orientações, conselhos e dicas de vida da prof. Denise Cristina Kluge, além da prof. Angélica Foresti Carlet. Aos membros de bancas, professores Ubiratã K. Alves, Gustavo Nishida, Andressa B. Albini, Malu Gomes, Denize Nobre: muito obrigado por tudo!

Após muitas dificuldades pessoais e algumas acadêmicas, defendi o mestrado em 20 de setembro de 2021 e comecei o doutorado na semana seguinte, ao final de 2021, sob a orientação da Profª Drª Angélica F. Carlet e coorientado pela Profª Drª Denise C. Kluge. Também fui aprovado no doutorado da UFPR em 2021/2022, mas optei pela UIC. Cabe notar que em 2021/22 o contexto ainda era pandêmico e era muito difícil conseguir realizar atividades presenciais, inclusive na própria UIC, então este cenário de distanciamento social influenciou nas decisões de pesquisa tomadas no doutorado. Antes de trancar o curso, fui aprovado em uma

banca de qualificação internacional com a presença dos professores Juli Cebrian e Joan C. Mora. Recomecei o doutorado na UFPR no início de 2023, apesar de nunca, neste meio tempo, ter parado de estudar e principalmente pensar na tese.

A partir de 2023, contei com o apoio da pessoa que mais acreditou em mim: Janine Pinheiro Marques. Muito obrigado! De forma direta ou indireta, meus amigos sempre estiveram lá (e foram dezenas), mas principalmente, em ordem cronológica: Lucas Fernandes, Henrique de Mares, Rogger Alves, Rafael Bisol, Adil Coelho, Cleverson Honório, Fernanda Arantes (alô pessoal do BodeBrown, GESD, Tec, AAM, CWB, OSRS, RS3 e Tork, vocês não foram esquecid@s)

“A palavra foi feita para dizer”

Graciliano Ramos

RESUMO

Os Treinamentos Perceptuais de Sons Não-nativos (TP) surgiram para melhorar a percepção das consoantes /r, l/ do inglês por japoneses. Com o tempo, expandiram-se para outros segmentos (como *Voice Onset Time*, vogais, tons linguísticos) e contextos (diferentes línguas nativas/alvo, número de sessões, testes de generalização). Diante disso, realizou-se uma Revisão Sistemática (RS) para responder: "*Qual é o estado da arte dos TP e quais são possíveis novas direções de pesquisa?*". O objetivo principal é reportar e avaliar trabalhos publicados sobre TP, com três metas específicas: analisar sujeitos, experimentos e teorias dos estudos. Com o intuito de responder à pergunta de pesquisa e alcançar o objetivo principal, foram elencados três objetivos específicos: reportar e avaliar acerca dos sujeitos presentes nas pesquisas de TP; reportar e avaliar acerca dos experimentos realizados nas pesquisas de TP; reportar e avaliar acerca de teorias e análises abarcadas nas pesquisas de TP. Os dados das pesquisas foram subdivididos em 16 descritores, como ano de publicação e presença de tecnologia. Os procedimentos metodológicos, descritos em detalhes no capítulo 3, consistiram na elaboração da RS: definir as bases de dados utilizadas para consultas (Google Acadêmico e portal de periódicos da CAPES); analisar somente artigos científicos revisados por pares; refinar as palavras-chave da pesquisa e as buscas avançadas em quatorze buscas totais (Google Acadêmico: Conter: Palavras exatas / Presentes no Título; Portal CAPES: Conter as palavras em: Título e Assunto). Foram filtrados 2744 artigos, que passaram pelos critérios de inclusão: ser trabalhos da área de treinamento, pesquisados em português e inglês; indexados nas bases de dados pesquisadas; e ser de acesso gratuito. Após isso, quatro critérios de exclusão foram aplicados: exclusão de artigos repetidos; exclusão de monografias, dissertações e teses; exclusão de pôsteres e trabalhos apresentados em eventos; exclusão de resenhas/resumos/compilações, ou seja, apenas trabalhos empíricos e inéditos foram incluídos. O número de trabalhos que passaram pelos critérios de inclusão e exclusão foram de 117. Como exemplos de resultados, nota-se um aumento no número de artigos da área de TP (59 trabalhos foram selecionados na década entre 2011 e 2020, e já 34 trabalhos haviam sido publicados entre 2021 e 2024). A maioria

dos trabalhos foca no inglês como a língua-alvo (63,5% de todas as pesquisas contaram com o inglês como língua-alvo, enquanto a língua-alvo que ficou com a segunda colocação apresentou somente 8,5% de incidência). A maior parte dos treinamentos contou com 8 ou 9 sessões, em média, e a maior parte das pesquisas, representadas por 67,5% do total, apresentou algum tipo de feedback (divididos entre imediato, universal, ambos, visual etc.) para o participante do treinamento.

Palavras-Chave: Treinamentos Perceptuais; Treinamentos Perceptuais de Sons Não-nativos; Revisão Sistemática; Fonética e Fonologia; Linguística

ABSTRACT

Perceptual Training of Non-Native Sounds (PTNS, in short PT) was initially developed to improve Japanese learners' perception of English consonants /r, l/. Over subsequent decades and into the 21st century, research expanded to myriad contexts, including diverse segments (e.g., Voice Onset Time in initial consonants, high/low vowels, rounded/unrounded vowels, nasal vowels), suprasegmental factors (e.g., lexical tones in tonal languages, word and sentence stress), varied native/target language pairs, training protocols, delayed retention tests, and generalization tests (e.g., to novel speakers or phonetic contexts). Following an extensive review of theoretical frameworks in humanities, linguistics, phonetics, and phonology, this study conducted a Systematic Review (SR) of PTNS research to address the question: *"What is the state of the art in PTNS, and what are potential new research directions?"* The primary objective was to evaluate, synthesize, and report published academic works on Perceptual Training through this SR, with three specific goals: analysing participant profiles, assessing experimental designs, and reviewing theoretical frameworks. The data was divided into 16 subcategories of analysis, spanning from the date of the publication of the articles, tasks used for training and software used. Methodological procedures involved searches in Google Scholar and the CAPES Periodicals Portal, focusing on peer-reviewed, open-access articles in English/Portuguese. From 2,744 initial articles, 230 met inclusion criteria (e.g., empirical studies on PTNS), and 117 were retained after full-text screening. Findings revealed a surge in PTNS publications (59 studies in 2011–2020; 34 in 2021–2024), with English as the dominant target language (63.5% of studies; second-place language: 8.5%). Most protocols included 8–9 training sessions, and 67.5% provided feedback (e.g., immediate, visual).

Keywords: Perceptual Training; Perception Training of Non-native Sounds; Systematic Review; Phonetics and Phonology; Linguistics

LISTA DE FIGURAS

Figura 1: Curvas desenvolvimentais dos valores médios das durações absolutas produzias pela Participante 1 o grupo Experimental e Gauer (2024)	50
Figura 2: Uma secção da planilha de dados.....	77

LISTA DE QUADROS

Quadro 1: Metodologia da RS de Sella et al. (2013)	59
Quadro 2: Metodologia da RS de Jesus et al. (2020).....	60
Quadro 3: <i>Metodologia da RS de Da Silva e Junior (2017)</i>	61
Quadro 4: Metodologia da RS de Franco et al. (2023).....	62
Quadro 5: Metodologia da RI de Azevedo e Bonfim (2019).....	64
Quadro 6: Metodologia da RI de Galdino e Oliveira Jr. (2023):	65
Quadro 7: Organizações de dados da Revisão Sistemática	81
Quadro 8: Distribuição do ano de publicação dos trabalhos da RS.....	84

LISTA DE GRÁFICOS

Gráfico 1: Distribuição de L1.	87
Gráfico 2: Distribuição das L2:	89
Gráfico 4: Número de participantes nos grupos experimentais	95
Gráfico 5: Número de sessões de treinamento	96
Gráfico 6: Duração (em dias) dos treinamentos perceptuais.	98
Gráfico 7: Distribuição de feedback.....	100
Gráfico 8: Distribuição do tipo de feedback.	101
Gráfico 9: Abordagem de modelos perceptuais por trabalhos de TP	120
Gráfico 10: Os trabalhos de TP apenas citam ou aprofundam modelos teóricos?	121

LISTA DE ABREVIATURAS OU SIGLAS

A / AV	(Treinamento Perceptual auditivo / audiovisual)
BNCC	(Base Nacional Comum Curricular)
CAPES	(Coordenação de Aperfeiçoamento de Pessoal de Nível Superior)
CNPQ	(Conselho Nacional de Desenvolvimento Científico e Tecnológico)
CAPT	(Treinamento de pronúncia assistido por computador)
CVC	(Estrutura silábica Consoante-Vogal-Consoante)
DIS	(Tarefa de discriminação)
ID	(Tarefa de Identificação)
F1/F2/F3	(Primeiro, segundo e terceiro formante)
HVPT	(<i>High Variability Phonetic Training</i>)
L1/LM	(Primeira Língua / Língua Materna)
L2	(Segunda(s) Língua(S))
LH	(Língua de Herança)
LVPT	(<i>Low Variability Phonetic Training</i>)
MQ	(Metanálise Qualitativa)
NLM	(<i>Native Language Magnet Model</i>)
PAM	(<i>Perceptual Assimilation Model</i>)
PB	(Português Brasileiro)
PCNs	(Parâmetros Curriculares Nacionais)
PESAA	(<i>Postech English Speech Assessment and Assistant</i>)

PFOL	(Português Para Falantes de Outras Línguas)
PLE	(Português Língua Estrangeira)
PRISMA	<i>(Preferred Reporting Items for Systematic Reviews and Meta-analyse)</i>
RI	(Revisão Integrativa)
RN	(Revisão Narrativa)
RS	(Revisão Sistemática)
SAC	(Sistema Adaptativo Complexo)
SDC	(Sistema Dinâmico Complexo)
SLM	<i>(Speech Learning Model)</i>
SSBE	<i>(Standard Southern British English)</i>
TF	(Treinamento Fonético)
TP	(Treinamento Perceptual)
TPSN	(Treinamento(s) perceptual (is) de sons não-nativos)
TSDC	(Teoria dos Sistemas Dinâmicos Complexos)
UFPR	(Universidade Federal do Paraná)
UFRGS	(Universidade Federal do Rio Grande do Sul)
UFSC	(Universidade Federal de Santa Catarina)
VOT	<i>(Voice Onset Time)</i>

SUMÁRIO

1. INTRODUÇÃO	18
1.1 Contextualização geral, pergunta de pesquisa e objetivos	18
1.2 Sobre Treinamentos Perceptuais e teorias subjacentes	21
2. REFERENCIAL TEÓRICO	26
2.1 Treinamentos Perceptuais de sons de línguas não-nativas	27
2.2 As teses em TP no Brasil	43
2.3 Teorias da Percepção e Modelos perceptuais	51
2.4 Diferentes metodologias em trabalhos bibliográficos	56
2.4.1 Sella <i>et al.</i> (2013)	59
2.4.2 Jesus <i>et al.</i> (2020)	59
2.4.3 Da Silva e Junior (2017)	60
2.4.4 Franco <i>et al.</i> (2023)	62
2.4.5 Azevedo e Bonfim (2019)	63
2.4.6 Galdino e Oliveira JR (2023)	64
2.4.7 Bicudo (2014)	65
2.4.8 Barbosa <i>et al.</i> (2023)	66
2.4.9 Dias <i>et al.</i> (2015)	67
2.4.10 Saldanha (2016)	67
2.4.11 Chaves (2013)	68
3. METODOLOGIA	70
3.1 Metodologia para o título da pesquisa	70
3.2 Metodologia da Revisão Sistemática	72
3.3 Organização e descrição do conjunto de dados da RS	76
4. ANÁLISE E DISCUSSÃO DOS RESULTADOS	83
4.1 Ano de Publicação	83
4.2 Língua Materna	86
4.3 Língua-Alvo	88
4.4 Foco fonético	91
4.5 Existência de Grupo Controle	92

4.6 Número de Participantes do Grupo Controle	93
4.7 Número de Participantes do Grupo Experimental	94
4.8 Número de sessões.....	95
4.9. Duração do período de treinamento e das sessões	97
4.9.1 Duração do período de treinamento.....	97
4.9.2 Duração das sessões	99
4.10 Existência e tipo de feedback.....	100
4.11 Existência de testes de generalização	104
4.11.1 Generalização para a produção da fala.....	104
4.11.2 Generalização para novos falantes	108
4.11.3 Generalização para novos segmentos, palavras, sílabas ou contextos.....	112
4.12 Existência de testes de retenção	114
4.13 Adoção do paradigma da alta variabilidade.....	117
4.14 Adoção de modelos teóricos perceptuais	120
4.15 Uso da tecnologia.....	126
4.15.1 <i>Softwares</i> de treinamento	127
4.15.2 <i>Softwares</i> auxiliares	128
4.16 Tarefas de treinamento	128
4.17 Considerações metodológicas sobre informações e organizações de artigos de TP	129
5. CONSIDERAÇÕES FINAIS	131
5.1 Considerações sobre os objetivos e Agenda de pesquisa.....	131
5.1.1 Considerações sobre o objetivo específico 1 (sujeitos):	131
5.1.2 Considerações sobre o objetivo específico 2 (experimentos):.....	132
5.1.3 Considerações sobre o objetivo específico 3 (teorias e análises):.....	136
5.1.4 Considerações sobre o objetivo geral e a pergunta de pesquisa.....	136
5.2 Limitações e sugestões	139
REFERÊNCIAS	148
ANEXOS	166

1. INTRODUÇÃO

Esta introdução está dividida em duas etapas: a contextualização geral da pesquisa atual e uma breve apresentação da área de Treinamentos Perceptuais de Sons Não-Nativos (TPSN). Antes, porém, é importante discutir brevemente a nomenclatura dada aos treinamentos, originalmente nomeados apenas de Treinamentos Perceptuais (TP) até a qualificação da tese. Foi um processo exaustivo peneirar os trabalhos de TP L2s utilizando somente os termos elegidos para a pesquisa (o que será descrito no capítulo 3), mesmo ao buscar por filtros, então é importante afirmar que a escolha do título da pesquisa foi feita levando em consideração **como** encontrar uma forma acessível e descritiva para os treinamentos de sons não-nativos. Cabe notar que, como os únicos treinamentos relevantes para a RS são TPSN, opto por continuar usando a sigla TP durante a maior parte da tese, por ser mais sintética e não-ambígua no contexto atual.

1.1 Contextualização geral, pergunta de pesquisa e objetivos

Brevemente resumindo uma história de quase quatro anos, no dia seguinte à minha defesa de mestrado, em setembro de 2021, fui aprovado no doutorado na Universitat Internacional de Catalunya (UIC). A qualificação do doutorado da UIC era no primeiro ano, e fui aprovado na banca, porém enfrentei inúmeras dificuldades pessoais, financeiras, de saúde e voltei ao Brasil, ingressando no doutorado da Universidade Federal do Paraná. Percebi, que, após estudar a bibliografia de TPs por meses, poderia desenvolver um trabalho muito mais interessante para minha curiosidade intelectual e para a área da linguística do que efetivamente realizar um treinamento. Minha ideia inicial era traçar algum tipo de análise que permitisse entender quais teorias e visões de língua subsidiavam os trabalhos de TP. Formulei algumas perguntas de pesquisa e apresentei às orientadoras (na época duas, Angélica Carlet, da UIC, e Denise Kluge, da UFPR), que responderam em aprovação quanto à relevância do tema. Subsequentemente, no final de 2022 e

2023, formulei um objetivo e trabalhei melhor as perguntas de pesquisa e levei ao grupo de pesquisa de Fonética e Fonologia da Associação Nacional de Pós-graduação e Pesquisa em Letras e Linguística (ANPOLL), e ouvi o que diversos especialistas da área me sugeriram para aprimorar o trabalho. Dentre as diretrizes, estava selecionar uma metodologia firme para sustentar a análise, e nominalmente foram citadas Metanálise e Revisão Sistemática (RS). Após estudar dezenas de trabalhos nas áreas de Letras e Linguística que se utilizaram dessas metodologias, ainda estava insatisfeito com o que via: as pesquisas, mesmo extensas metanálises e revisões sistemáticas, careciam muito da discussão diacrônica, para mim. Contactei o Profº Drº Gustavo Nishida para questionar sobre minha sensação de que estas metodologias eram insuficientes para executar o que eu queria no doutorado. Ele me sugeriu estudar Historiografia da Linguística (HL), e assim o fiz. Retornei às orientadoras com relatórios sobre a leitura de diversas metodologias, e, além da Metanálise, RSL e HL já citadas, Revisões Integrativas de Literatura (RIL), Revisões Narrativas / Literárias (RN) e Metanálises Qualitativas (MQ), para ser questionado, assim como fui questionado durante o mestrado, qual era o real objetivo da pesquisa. Ao longo da pesquisa, tornou-se observável que realizar uma Revisão Sistemática bem-feita e comentada poderia, por ora, suprir a necessidade da curiosidade por relatar e descrever a área de TP.

Após essa breve apresentação de percurso, parte-se para aspectos mais práticos que concernem à metodologia da presente tese. Além da contribuição à área de TP, acredito que esta tese pode ser de grande valia para estudiosos de pesquisas bibliográficas e documentais em geral que desejam realizar metanálises, revisões sistemáticas, revisões integrativas ou até mesmo revisões literárias, que são padrão em pesquisas científicas, uma vez que houve um esforço se debruçar sobre as áreas da saúde e das humanas, mais especificamente na linguística e ainda mais especificamente na fonética e fonologia, com o intuito de avaliar como estas pesquisas vêm sendo desenvolvidas. Cabe, a partir destas considerações, passar para aspectos formais da metodologia da tese, endereçando a pergunta de pesquisa e os objetivos.

A pergunta de pesquisa que foi elencada para esta tese é “Qual é o estado da arte da área de TP e quais podem ser novos caminhos de pesquisa?¹”. Deve-se adicionar a esta pergunta de pesquisa um objetivo principal e três objetivos específicos. O objetivo principal desta pesquisa é, através de uma Revisão Sistemática, reportar e avaliar os trabalhos acadêmicos publicados sobre o tema de Treinamentos Perceptuais. Durante a pesquisa, de forma autoral, dividi os dados da pesquisa em 16 descritivos de análise, que posteriormente colaboraram para dividir os objetivos específicos nos três finais (sujeitos, experimentos e teorias), que serão mais bem descritos a seguir:

- Objetivo específico 1: Reportar e avaliar acerca dos sujeitos presentes nas pesquisas de TP com apoio dos descritivos: (2) Língua materna, (5) Existência de Grupo Controle; (6) Número de participantes no Grupo Controle; (7) Número de participantes no Grupo Experimental;
- Objetivo específico 2: Reportar e avaliar acerca dos experimentos realizados nas pesquisas de TP com apoio dos descritivos: (1) Ano de publicação (divididos em cinco categorias: Antes de 1990, 1990-2000, 2000-2010, 2010-2020, 2020- presente); (3) Língua-alvo; (4) Foco Fonético (segmentos ou suprasegmentos do treinamento); (8) Número de Sessões; (9) Duração do período das sessões; (10) Existência e tipo de feedback; (11) Existência de testes de generalização; (12) Existência de testes de retenção. (15) Uso da tecnologia (16) Tarefas de treinamento;
- Objetivo específico 3: Reportar e avaliar acerca de teorias e análises abarcadas nas pesquisas de TP com apoio dos descritivos: (13) Adoção do paradigma da alta variabilidade; (14) Adoção de modelos teóricos perceptuais.

Tendo apresentado brevemente o histórico da pesquisa, os objetivos geral, específicos e a pergunta da pesquisa, apresento brevemente os TPs na subseção 1.2.

¹ Uma tentativa, dentre muitas, de traduzir *new research venues* da língua inglesa.

1.2 Sobre Treinamentos Perceptuais e teorias subjacentes

Antes do início da descrição de trabalhos e da contextualização, é importante familiarizar o leitor com alguns termos. Existem muitas variáveis em uma pesquisa de treinamento (como o foco da pesquisa, duração do período de treinamento e número de sessões, *feedback*, testes de retenção e generalização, apenas para citar algumas). Entretanto, alguns processos fundamentais para a compreensão básica de um TP: O(s) pré-testes, a fase de treinamento e os pós-testes. Os pré-testes e pós-testes podem ser tarefas iguais ou não aos treinamentos, e servem para comparar os informantes antes e depois das sessões de treinamento. Dados relacionados a porcentagens de acertos em testes de identificação (ID, em que o participante seleciona o som que ouviu entre vários) ou discriminação (DIS, em que o participante precisa discernir o som que ouviu entre outros), ou coletas de produção, por exemplo. É possível que certos informantes sejam descartados de uma pesquisa de TP por possuírem um *score* muito alto nos pré-testes, fazendo com que a possível margem de melhora no treinamento seja muito baixa. Para mais, em muitos trabalhos, o número de informantes pode diminuir com o tempo, uma vez que os estudos em TP são longitudinais e as sessões de treinamento acontecem ao longo de semanas ou meses, e os informantes podem desistir ou se desistirem da pesquisa. Para trabalhos com testes de retenção postergada depois do final do período de treinamento de pós-teste, esta questão é ainda mais relevante.

As duas tarefas mais comuns que são aplicadas no treinamento, conforme exemplificadas anteriormente, são as de identificação e discriminação, que são detalhadas em Milan e Kluge (2021). Para a grande maioria das pesquisas descritas aqui, os participantes de língua materna² não viveram ou moraram muitos anos em países que falavam as L2 que estavam sendo treinados. Um exemplo disto é em

² Língua Materna e L1 utilizados de forma intercambiável na presente pesquisa. Língua-alvo e L2 são utilizadas como sinônimos também.

Wang *et al.* (2003), em que os autores exemplificam que os 12 participantes do grupo controle, em uma pesquisa de treinamento de estadunidenses para mandarim, “Nenhum deles jamais viveu em um ambiente falante de mandarim, e a maioria deles não possuía experiências com línguas tonais previamente” (WANG *et al.* (2003, p.114)³. Feito este preâmbulo, a apresentação da área pode ser realizada.

Os TP e *High Variability Phonetic Training* (HVPT), ou treinamento fonético de alta variabilidade), começaram a ser pensados como alternativas às dificuldades de percepção de certos sons de segunda língua. Nos meados das décadas de 1980 e 1990, uma série de trabalhos experimentais de diversas universidades pelo mundo começaram a ser desenvolvidos. Em poucos anos os trabalhos começaram a extrapolar o par /r, l/ do inglês para aprendizes japoneses, que foi o foco das pesquisas iniciais, e se expandiram para diversas línguas maternas e línguas-alvo. Barriuso e Hayes-Harb (2018) propõem uma revisão sobre o HVPT como um método na aprendizagem de segunda língua (L2). A primeira questão abordada pelos autores como um problema básico de percepção foi o chamado "problema de constância", ou como um falante nativo pode identificar múltiplas produções de um som de fala específico mesmo que nenhuma produção seja acusticamente idêntica. No capítulo 4 de resultados, subseção 4.14, que trata dos modelos teóricos perceptuais, expande-se este aspecto da constância através da apresentação das diversas correntes teóricas sobre a percepção da fala.

Na década de 1980, pesquisadores de fonologia e fonética começaram a investigar o porquê de os aprendizes de L2 não apresentavam o mesmo desempenho que os falantes nativos na percepção de diferenças sutis (como na distinção entre pares mínimos, por exemplo, /p/ e /b/) e quais seriam os tipos de variabilidade acústica relevantes para diferenciar contrastes. Já na década de 1990, estudos como o de Logan *et al.* (1991), que explorarei com mais detalhes posteriormente neste texto, descobriram que aprendizes japoneses de inglês como segunda língua poderiam melhorar sua percepção do /r/-/l/ em graus diferentes

³ Original: “None of them had ever lived in a Mandarin-speaking environment, and most of them had no experience with a tone language prior...”

dependendo do ambiente fonético do som dentro de cada palavra. Após esse estudo, outras variáveis foram implementadas, como tarefas de generalização que estenderam a percepção a falantes novos, percepção de novas palavras e até mesmo se os segmentos treinados poderiam ter uma retenção atrasada. Houve vários trabalhos sobre esse par /r/-/l/ por falantes japoneses (Strange e Dittman, 1984; Goto, 1971; Dissosway-Huff et al., 1982; Mochizuki, 1981; Sheldon e Strange, 1982, para citar apenas alguns), e parte dessas pesquisas serão discutidas mais detalhadamente no capítulo 4. Desde 1991, portanto, o HVPT tem sido aceito como um método válido para aprimorar a pronúncia, e muitos estudos de pesquisa foram realizados nas últimas décadas. Existem diferentes possíveis tarefas para se implementar TP. Cabe notar que Qin *et al.* (2021) é um exemplo de trabalho de questiona o paradigma do HVPT contrastando com treinamentos de baixa variabilidade.

Carlet (2017) afirma que aprender uma segunda língua em um ambiente instrucional pode dificultar a aquisição da língua, uma vez que o *input* autêntico ajuda no desenvolvimento da língua, pressupondo que fora da sala de aula os estudantes não estão expostos aos *inputs* das línguas estudadas. A autora afirma que o uso e o *input* da L2 são fundamentais para promover o aprendizado das categorias fonéticas da L2. Por isso, treinar os aprendizes usando *input* autêntico de falantes nativos pode ajudar na melhoria das tarefas com segmentos difíceis. Carlet (2017) estudou como falantes bilíngues catalães/espanhóis aprimoraram sua percepção e produção de inglês avaliando os dois HVPT discutidos anteriormente, ID e DIS, em cinco vogais do inglês (/i ɪ æ ʌ ɜ:/), além de oclusivas iniciais e finais. Todos os grupos foram expostos a palavras inventadas de estrutura CVC e realizaram testes tanto com palavras reais quanto com palavras inventadas. Carlet (2017) descobriu que tanto os treinamentos de ID quanto DIS foram eficazes para melhorar as capacidades de percepção da L2 dos alunos para consoantes, no entanto, para vogais, o ID seria preferível, uma vez que foi o único a melhorar a produção dos sons de vogais treinados, além de mostrar mais retenção e ser favorecido pela avaliação dos alunos sobre qual era o melhor sistema de treinamento.

Além disso, autores sul-americanos e brasileiros também já exploraram possibilidades de TP e HVPTs. Brawerman-Albini *et al.* (2017), por exemplo, focaram na aquisição do acento tônico da língua inglesa por falantes brasileiros. O estudo visava o estresse pré-antepenúltimo (na quarta até a última sílaba), um padrão incomum de ocorrer no português brasileiro. As autoras citam exemplos de como palavras pronunciadas com estresse pré-antepenúltimo podem ser compreendidas por falantes nativos: "*comfortable*" poderia ser entendido como "*come for a table*"; "*productivity*" poderia ser ouvido como "*productive tea*". Os autores compararam 30 estudantes brasileiros de inglês com um grupo de controle composto por 10 pessoas que não foram treinadas, para perceber a evolução do grupo experimental. O teste consistiu em tarefas de ID. Os participantes tiveram que utilizar um software para selecionar o número correspondente à sílaba estressada em cada palavra. Em relação aos resultados, a média de percepção no pré-teste foi de 85,4% e melhorou para 95,9% no pós-teste. A generalização alcançou uma média de 93,4%. No entanto, o grupo de controle não apresentou melhora nos testes de pré, pós e generalização. O grupo experimental também mostrou melhora na produção: de 28,4% de palavras corretas no pré-teste para 57,5% no pós-teste, com 58% de retenção e 59% de generalização. Esses resultados brevemente descritos mostram o quanto o treinamento auxiliou os aprendizes brasileiros a perceberem e produzirem aspectos suprasegmentais do inglês.

Segundo Flege (1995), um estudo típico da área de TP conta com 800 *trials* de treinamento divididos em várias sessões. O autor afirma que ganhos de cerca de 10% nos acertos de sons da L2 são comuns para os participantes não-nativos. Outrossim, já em 1995, Flege identifica que os ganhos dos treinamentos perceptuais podem ser generalizados para novos estímulos, novos locutores (não presentes no treinamento) e os conhecimentos podem ser retidos por até seis meses após o término das sessões.

Como já brevemente mencionado, a presente tese está dividida da seguinte forma: no capítulo 2 se encontra uma descrição mais detalhada de como os TPs surgiram e se desenvolveram, com o respectivo background linguístico da época

além de embasamento teórico, respectivamente, para as teses em TP brasileiras, teorias da percepção e modelos perceptuais e diferentes metodologias em trabalhos de RS. No capítulo 3 encontra-se a metodologia do trabalho, que justifica minhas escolhas antes de expor a metodologia adotada na presente pesquisa. O capítulo 4, de discussões e resultados, é autoexplicativo e discute os dados encontrados. No capítulo 5, as conclusões são apresentadas, retomando a pergunta de pesquisa, o objetivo principal e os objetivos secundários, bem como sugestões de futuras pesquisas são oferecidas através de uma Agenda de Pesquisa.

2. REFERENCIAL TEÓRICO

Como introdução à fundamentação teórica, é importante descrever o processo metodológico que embasou este capítulo. Mattos (2015) descreve que existem três principais métodos para se desenvolver uma pesquisa bibliográfica: a Revisão Narrativa (RN), a Revisão Sistemática (RS) e a Revisão Integrativa (RI). No capítulo 3, uma descrição mais detalhada de cada uma destas é apresentada, bem como exemplos de RS e RI, além de metanálises (e metanálises qualitativas). Por ora, é importante dizer que o processo específico do capítulo 2, de referencial teórico, seguiu uma RN, como é costume para fundamentações teóricas (MATTOS, 2015). Este tipo de revisão permite a inclusão de outras pesquisas de revisão, além de não se utilizar de critérios explícitos para a seleção de fontes.

A maioria das fontes utilizadas no presente capítulo provém de um número elevado de citações nas pesquisas de TP, geralmente pela importância histórica na área ou pela inovação de metodologia. Como a RS da presente tese apenas contou com artigos, como descrito na metodologia no capítulo 3, torna-se apropriado mencionar dissertações, teses e trabalhos que não foram abarcados pela RS aqui. Muitos dos trabalhos que serviram para esta fundamentação teórica estão presentes em manuais como Milan e Kluge (2021) e foram indicados pela minha orientadora ou citados por teses importantes para qualquer trabalho de TP como Carlet (2017). Os trabalhos presentes neste capítulo 2 não necessariamente estão incluídos na RS, ou por serem dissertações ou teses, ou por não terem sido filtrados pelas buscas da metodologia. As tendências mais observadas nas bibliografias preliminares foram seguidas para a revisão bibliográfica inicial. Reafirmo: nem todos os trabalhos presentes na seção 2.1, sobre os treinamentos fonéticos/perceptuais/perceptivos estão na revisão sistemática por puras razões metodológicas: Alguns trabalhos revisados na subseção 2.1 foram sugeridos pela orientadora ou por membros das bancas de qualificação e não foram encontrados seguindo os critérios de inclusão ou exclusão da Revisão Sistemática, bem como alguns trabalhos que entraram na Revisão Sistemática foram utilizados nas sessões

de resultados para discutir dados observados, em vez de descrever o panorama dos TP aqui. Em 2.2 apresento as cinco teses de doutorado desenvolvidas no Brasil por pesquisadoras brasileiras (Denize Nobre-Oliveira (2007), Melissa Bettoni (2008), Andressa Brawerman-Albini (2012), Pollianna Milan (2019) e Luana Gauer (2024)) algumas com mais ou menos detalhes pois vários artigos provenientes de algumas das teses ou mesmos as descrições das teses já estão presentes em outras seções do texto. A subseção 2.3 foca as teorias de percepção e modelos perceptuais, e a 2.4 contém diferentes metodologias em trabalhos bibliográficos de revisão.

2.1 Treinamentos Perceptuais de sons de línguas não-nativas

Houve vários trabalhos explorando como o treinamento poderia estimular uma melhoria na distinção do par /r/-/l/ do inglês por falantes japoneses. Strange e Dittman (1984) implementaram um treinamento de 14 a 18 sessões para oito aprendizes adultos japoneses de inglês, usando uma tarefa de discriminação que oferecia *feedback* imediato. O nível de proficiência dos alunos e o tempo que os participantes estavam vivendo nos Estados Unidos variaram consideravelmente, mas a dificuldade com o par /r/-/l/ foi reconhecida por todos os participantes. É importante notar que, por se tratar de um estudo de 1984, os sujeitos verificaram as respostas corretas que haviam dado no papel (diferentemente do método contemporâneo de *software*, que é mais tecnológico e não requer respostas escritas ou no papel). Mais detalhes sobre a tecnologia adotada nos trabalhos está presente em 4.15, mas, em suma, 116 dos 117 trabalhos da RS utilizaram *softwares* com O pré-teste consistiu em três fases. A primeira parte continha pares mínimos de fala real; a segunda continha estímulos sintéticos como 'rock'/'lock'; a terceira continha estímulos sintéticos como 'rake'/'lake'. Esses estímulos alterados sintéticos foram feitos para que as frequências F1/F2/F3 fossem alteradas de forma proporcional para diferenciar cada token. Após a realização do pré-teste, o treinamento começou. É importante notar que o treinamento possuía apenas os pares 'rock'/'lock', portanto, os outros estímulos apresentados no pós-teste foram avaliados por generalização. Os autores descobriram que houve uma melhoria significativa ao longo das

sessões, com ganhos consideráveis ocorrendo principalmente nas primeiras sessões. Apenas um dos oito sujeitos não melhorou na série *rock-lock*. Também houve melhorias significativas que se estenderam à identificação do par de estímulos 'rake'/'lake' e à discriminação de estranheza para cinco alunos, mas essa mudança positiva não pôde ser notada em relação ao pós-teste de palavras reais. Em relação à pesquisa de Strange e Dittman (1984), Logan *et al.* (1991a) demonstra que é importante considerar diferentes ambientes fonéticos e é por isso que os resultados não se estenderam a palavras reais. Além disso, Logan *et al.* (1991a) indica que testes de discriminação com um pequeno conjunto de tokens podem ser ineficazes para alterar o sistema de percepção do ouvinte.

Jamieson e Morosan (1986) treinaram aprendizes francófonos canadenses de inglês para distinguir o contraste das fricativas dentais sonoras (em palavras como "mother") e surdas (em palavras como "thanks") (/ð/-/θ). Os autores afirmam que os francófonos podem ter dificuldade em diferenciar /ð/ de /d/ e /θ/ de /t/, respectivamente. Os autores investigaram se um regime controlado de treinamento poderia melhorar a percepção dos sujeitos por meio de sessões de treinamento perceptivo. Os autores selecionaram 24 palavras Consoante-Vogal (CV), faladas por um locutor masculino e sintetizadas, consideradas bons exemplares por três ouvintes anglófonos canadenses. Os alunos passaram por um pré-teste de ID e DIS e foram divididos em um grupo controle e um grupo experimental. Em seguida, nos próximos dias, o grupo experimental teve sessões de treinamento e o mesmo teste aplicado antes e depois da sessão de treinamento. Houve melhoria significativa do pré-teste para o pós-teste pelo grupo experimental em apenas 90 minutos de treinamento ao identificar tokens sintéticos. Talvez ainda mais impressionante e importante: houve transferência para *tokens* naturais que não haviam sido treinados anteriormente. Houve, de fato, menos precisão ao considerar períodos mais curtos de fricção nas consoantes. No entanto, as consoantes com fricção mais longa foram aprendidas com uma precisão próxima do máximo possível.

Logan *et al.* (1991b) é uma continuação da pesquisa pelos mesmos autores do mesmo ano. Como explicado pelos autores, encontrou dados interessantes

sobre o potencial de treinar falantes japoneses de inglês como L2 para diferenciar melhor /r/ de /l/. A variabilidade dos estímulos foi uma das principais questões da primeira pesquisa, então Logan *et al.* (1991b) planejou dois experimentos para entender melhor os impactos de diferentes vozes de *input* e diferentes ambientes fonológicos. Os quatro ambientes presentes no primeiro experimento foram: posição inicial de sílaba /r/ e /l/, consoantes isoladas, grupos consonantais e consoantes intervocálicas. Os testes e o treinamento consistiram em 2 tarefas de discriminação de escolha forçada, uma vez que os autores afirmaram que esse tipo de teste combinado com *feedback* imediato ajuda a desenvolver novas categorias fonéticas fortes. O primeiro experimento foi realizado por seis falantes femininas japonesas de inglês. Todas as palavras com ambientes /r/ e /l/ mencionadas acima foram testadas e treinadas. A precisão média passou de 79,96% para 85,57%. Os testes de ANOVA também indicaram que os *clusters* (encontros consonantais) iniciais tiveram o pior desempenho pelo grupo experimental. É também uma descoberta importante deste experimento que os sujeitos puderam responder mais rapidamente em cada sessão e generalizar a aprendizagem para novos falantes no final das sessões de treinamento. Diferentemente do primeiro experimento, que contou com cinco falantes de *input* diferentes e novos falantes para os testes e generalização, este segundo experimento contou apenas com um dos locutores (número quatro) para realizar as sessões de treinamento (isto pode ser caracterizado como um treinamento de baixa variabilidade, por não possuir muitos falantes, o que o distingue de um regime de HVPT. Entre outras descobertas, o experimento 2 revelou as limitações de contar com apenas um locutor: os sujeitos não puderam generalizar a melhoria significativa adquirida no processo de treinamento para novos tokens produzidos por um novo falante e nem mesmo para o mesmo locutor número quatro. Resumindo, Logan *et al.* (1991) enfatizam a importância de ter alta variabilidade em falantes, tokens e ambientes para fornecer uma aquisição robusta de categorias fonéticas. Centenas de pesquisas após esta recomendação e variabilidade exploraram a área do HVPT.

Hirata (1999) comparou o treinamento perceptual a nível de palavra *versus* a nível de sentença para participantes estadunidenses de nível introdutório de

japonês. O foco da pesquisa foram os contrastes de duração da língua japonesa. Os dois grupos experimentais da pesquisa foram: (1) o grupo que realizou treinamento a nível de palavras (contendo 12 participantes); (2) o grupo que realizou treinamento a nível de sentenças (11 participantes), além de um grupo controle de sete participantes, que realizou apenas o pré-teste e o pós-teste. Ao total, foram 10 sessões de treinamento, divididas em cinco semanas. As palavras utilizadas para o treinamento de palavras foram as mesmas do treinamento de sentenças, porém isoladas. A autora dividiu as palavras presentes no treinamento (que variavam entre uma e seis moras, entre categorias fáceis, que contavam com vogais curtas e consoantes individuais, e difíceis, que contavam com fonemas mais complexos, e a análise também se baseou nesta divisão interna). Cabe notar que a estrutura do pré-teste e do pós-teste foi a mesma, contudo as palavras foram alteradas. Com relação ao *feedback*, os participantes precisaram ouvir cada palavra ou erro cometido obrigatoriamente três vezes a cada resposta errada. Hirata (1999) confirmou uma hipótese ao observar que, independentemente do tipo de treinamento, ambos os grupos experimentais melhoraram seus resultados para palavras fáceis. Contudo, para as palavras difíceis, o grupo que foi treinado com palavras melhorou 45,3%, e o grupo treinado com sentenças melhorou 47,7%. Todavia, como a própria autora chama a atenção, ao se examinar estas porcentagens mais atentamente, nota-se que a maior parte da melhora veio do teste com sentenças, e não com palavras. Notou-se também que o grupo controle foi capaz de melhorar significativamente para as palavras fáceis do pré para o pós-teste apenas com as aulas regulares de japonês, porém foi necessário a intervenção do TP para que os grupos experimentais conseguissem melhorar nas palavras difíceis (algo que o grupo controle não foi capaz de atingir). É também interessante notar que esta pesquisa, mesmo tendo sido realizada na década de 1990, subverte a maioria dos trabalhos que estavam sendo realizados em treinar nativos japoneses para os sons do inglês (em especial /r,l/), e treina estadunidenses para sons do japonês.

Hardison (2005) investigou a eficiência do treinamento de prosódia usando ferramentas de computador. A autora realizou dois experimentos. O primeiro

forneceu a falantes nativos de inglês e aprendizes de francês três semanas de treinamento de prosódia usando um *display* de tom em tempo real. Aprofundando-se nos detalhes, o primeiro experimento investigou aspectos quantitativos e qualitativos da língua francesa por meio de um treinamento que mostrava a cada aprendiz um *display* em tempo real de seus contornos de tom. Houve sessões de pré-teste e pós-teste, com 13 sessões de treinamento de aproximadamente 40 minutos cada entre elas. Houve também um questionário preenchido anonimamente que perguntava sobre dificuldades com a segunda língua antes do treinamento, qual era o valor do próprio treinamento, que características do francês eles começaram a notar após o treinamento, entre outras perguntas importantes relacionadas aos procedimentos e ao que foi aprendido. A autora descobriu que o treinamento, além de melhorar a prosódia, também estendeu os ganhos para precisão segmental e até mesmo para novas frases. O segundo experimento consistiu em uma tarefa de recordação de memória usando tokens de inteligibilidade reduzida, e mesmo com essa desvantagem, os aprendizes conseguiram obter uma pontuação positiva de 80% em frases de treinamento.

Wayland e Li (2005) investigaram o efeito dos procedimentos de treinamento perceptivo de Identificação e Discriminação para melhorar o tom *médio versus baixo* no tailandês, por meio de dois intervalos inter-estímulos (ISI, 500ms e 1500ms) por falantes nativos de inglês e mandarim. Os autores explicam que em idiomas tonais, discrepâncias no tom podem distinguir palavras: /ma/, por exemplo, pode significar "mãe" (tom alto), "cânhamo" (tom alto ascendente), "mangueira" (tom baixo descendente) e "vendido" (tom alto descendente). Em idiomas como inglês, português e espanhol, o tom não interfere no significado/categorização lexical. Nessas línguas, por exemplo, o tom pode indicar se uma frase é interrogativa ou afirmativa. Os autores explicam que o tailandês é uma língua com cinco tons, um tom extra em comparação com o mandarim. Tanto os americanos quanto os chineses não tinham experiência anterior com o tailandês e foram divididos em subgrupos que passaram por processos de treinamento de Identificação ou Discriminação. O grupo de controle foi formado por falantes tailandeses. Todos os indivíduos passaram pelas fases de pré-treinamento, treinamento e pós-

treinamento. Os estímulos contavam com cinco pares mínimos dos tons baixos e médios do tailandês. A escolha desses tons foi devido a dois ouvintes de mandarim tendo problemas com esse contraste durante um estudo piloto. Quanto aos resultados, notou-se que, interessantemente, não houve diferenças significativas relacionadas ao tipo de teste que cada grupo passou.

Aliaga-García e Mora (2007) estudaram a percepção e produção de oclusivas orais por estudantes espanhóis/catalães de inglês. Quanto aos valores de VOT (Abramson e Lisker, 1973) esperados na literatura, esses autores afirmam que o VOT curto de catalães e espanhóis está na faixa entre 0ms e 30ms. Por outro lado, o VOT mais longo produzido por falantes nativos de inglês está entre 40ms e 80ms. Os autores realizaram um esquema de treinamento de seis semanas (seis horas de treinamento no total) para melhorar a percepção e produção de nove aprendizes de inglês. As plosivas em ataque silábico /p, b, t, d/ foram os segmentos treinados nesta pesquisa. Quanto ao *feedback*, ele consistiu em *feedback* imediato (logo após cada identificação ser feita), além de um *feedback* global no final da sessão e um *feedback* semanal extra. As habilidades dos alunos foram avaliadas por meio de tarefas de identificação contendo dois pares mínimos que foram segmentados em 15 produções de diferentes durações de VOT cada: um contínuo /p-b/ e /t-d/ foi de 0ms a 70ms. Quanto aos resultados, os valores mostrados na sequência referem-se ao ponto de *crossover* de 50% onde ocorre a transição /p-b, t-d/: os falantes nativos de inglês (sete pessoas no total) mostraram 25,8ms para /p, b/ e 41,33ms para /t, d/. O grupo de controle (nove pessoas no total) mostrou 15ms no pré-teste e 19ms no pós-teste para /p, b/ e 28,33ms no pré-teste e 17,5ms no pós-teste para /t, d/. Quanto ao grupo experimental, os alunos registraram 9,17ms no pré-teste e 15ms no pós-teste para /p, b/, bem como 22,5ms no pré-teste e 16,67ms no pós-teste para /t, d/. Após testes estatísticos, aparentemente o treinamento não foi suficiente para fazer os grupos experimentais expandirem sua percepção do limite das consoantes. Em geral, os valores de VOT para /p, t/ produzidos pelo grupo experimental estavam distantes dos falantes nativos de catalão/espanhol, mas ainda sim longe dos falantes nativos de inglês. Quanto a /b, d/, não foi observada melhoria significativa para o grupo experimental. Aliaga-García e Mora (2007)

também afirmaram que um esquema de treinamento de seis semanas pode ser muito curto para melhorar a percepção tanto quanto melhorou a produção.

Aliaga-García e Mora (2009) realizaram um programa de treinamento de seis semanas em um grupo experimental de 18 nativos catalães/espanhóis e estudantes de inglês. Os segmentos treinados nesta pesquisa foram as oclusivas /p, t, k, b, d, g/ e o contraste entre as vogais /i: - ɪ/, /æ-ʌ-ɑ:/ e /u:-ʊ/. As seis sessões tiveram duração de duas horas e possuíam um sistema de *feedback* semelhante ao já descrito no parágrafo revisando Aliaga-García e Mora (2007). As sessões de treinamento de consoantes consistiram em fornecer aspiração e VOT como uma pista para diferenciar oclusivas sonoras e surdas, diferenças interlinguísticas e contextos para cada consoante, além das fases de aproximação, retenção e liberação da articulação. As sessões de treinamento de vogais consistiram em focar nos contrastes tensos-relaxados e frente-central-posterior, bem como na descrição do movimento da língua e posição dos lábios. As tarefas realizadas por Aliaga-García e Mora (2009) foram múltiplas e complexas: as tarefas de percepção foram ID, DIS, transcrição fonética e exposição a sons de falantes nativos; as tarefas de produção foram descrição visual dos sons, imitação, leitura em voz alta de frases, diálogos e trava-línguas. Além disso, o *input* de falantes nativos consistiu em vários falantes masculinos e femininos diferentes. O *EyeSpeak* foi o *software* usado para promover o treinamento visual de produção. Cabe notar que uma análise mais aprofundada que investigou os *softwares* presentes nos trabalhos abarcados pela RS se encontra na subseção 4.15, que trata especificamente da tecnologia. Cerca de metade dos trabalhos pesquisados relataram qual o(s) *software(s)* utilizados na pesquisa, mas a outra metade simplesmente não reportou nada a respeito deste assunto. Algumas pesquisas, por exemplo, deixaram claro que algum programa computacional foi utilizado na pesquisa, enquanto outras reportaram apenas utilizar caixas de som ou não especificaram se algum *software* foi utilizado. Voltando para os dados de Aliaga-García (2007), a curva de categorização que indicava como os alunos percebiam /p, b, t, d/ indicou que o treinamento melhorou a categorização pelo grupo experimental. Quanto à tarefa de produção, o grupo experimental mostrou um aumento nos valores de VOT de /p/, o que indica um VOT mais longo.

Reis e Nobre-Oliveira (2007) é um estudo exploratório que treinou 11 brasileiros de nível intermediário de proficiência para os contrastes do VOT das plosivas iniciais do inglês, /p, t, k/. As autoras contaram com quatro falantes de inglês para o *input* do treinamento perceptual de alta variabilidade, e os valores de VOT foram mantidos para o inglês e alterados (reduzidos) artificialmente para simular os VOTs do português. Os participantes participaram de uma sessão de curta duração, de cerca de 10min da tarefa de DIS e 8min da tarefa de ID, com *feedback* imediato. Tarefas de produção da fala também foram realizadas. A tarefa de produção apresentou acurácia de 32,82ms antes do treinamento e 42,84ms pós-treinamento (estatisticamente significativo). Todavia, os valores pioraram para a tarefa de discriminação (de 42,82ms para 32,90). As autoras também apresentaram os seguintes valores para os valores do pré-teste e pós-teste para cada uma das consoantes: /p/: 18,55ms para 31,36ms; /t/: 32ms para 36,91ms; /k/: 46,36ms para 55,18ms. Estatisticamente, apenas /p/ e /t/ demonstraram resultados de mudança significativa.

Iverson e Evans (2007) estudaram o treinamento auditivo do inglês para pessoas cuja língua materna era o alemão e o espanhol. Os autores puderam isolar e comparar os efeitos de duas diferentes L1, uma vez que os participantes passaram pelo mesmo processo de treinamento baseado em computador para a mesma língua com a mesma metodologia. Os autores formularam a hipótese, com base em dois modelos de assimilação, que uma vez que o espanhol tem muito menos vogais do que o alemão, seria mais fácil para os falantes de espanhol estabelecerem novas categorias fonéticas. Um adendo, com relação a isto, é que os autores se embasaram em modelos perceptuais como o *Speech Learning Model (SLM)* que será discutido com mais profundidade. Os estímulos de teste utilizados cobriam todo o espaço vocálico do inglês e foram falados por dois falantes nativos britânicos. As vozes usadas para o treinamento foram de cinco pessoas diferentes (caracterizando uma alta variabilidade), todas diferentes das usadas para o teste. Quanto aos sujeitos contidos nesta pesquisa, os falantes de espanhol moravam em um país de língua inglesa (Inglaterra), enquanto os falantes de alemão adquiriram o idioma em sua própria região na Alemanha. Apesar dessas diferenças, 13 falantes

de cada grupo foram testados e tiveram precisão de identificação de vogais semelhante. O treinamento foi realizado em cinco sessões (225 tentativas cada uma) de 45 minutos cada, em um período de 10 dias. Em média, mesmo com uma precisão semelhante pré-treinamento, os falantes de espanhol melhoraram moderadamente (10%) enquanto os falantes de alemão melhoraram consideravelmente (20%). Ainda assim, ambos os grupos melhoraram significativamente sua percepção de vogais do inglês. Uma hipótese oferecida pelos autores é que os falantes de espanhol já estão vivendo em um país de língua inglesa e já estão atuando em um nível quase máximo.

Iverson e Evans (2009) continuam e aprofundam a pesquisa apresentada em Iverson e Evans (2007), descrita no parágrafo anterior. Os autores ainda investigaram falantes de L1 alemão e espanhol adquirindo inglês. Os grupos e a metodologia inicial são semelhantes aos de Iverson e Evans (2007), assim como os valores das melhorias no treinamento (10% para falantes de espanhol e 20% para falantes de alemão). Nessa nova pesquisa, porém, para o primeiro experimento, os autores investigaram o espaço fonológico e onde exatamente estavam as dificuldades: no pré-teste, os falantes de espanhol tiveram dificuldade com /i/-/ɪ/, /a/-/ʌ/-/ɑ/ e /ɒ/-/ɔ/. Também no pré-teste, os falantes de alemão, por outro lado, tiveram dificuldades com /ɛ/-/a/-/ʌ/ e /ɑ/-/əʊ/-/ɒ/-/ɔ/-/aʊ/. Ambos os grupos conseguiram melhorar a percepção nesses grupos de vogais e ditongos-alvo. Os autores calcularam o quão próximos ambos os grupos estavam das vogais do inglês pré e pós-treinamento usando o melhor exemplar de cada indivíduo e aplicando fórmulas matemáticas. Os resultados mostraram que a precisão do melhor exemplar não mudou muito com o treinamento e que indivíduos com melhores localizações de F1/F2 e movimentos de formantes tenderam a se destacar na identificação de vogais naturais. Iverson e Evans (2009) propõem a hipótese de que o treinamento poderia ser superficial, portanto, não poderia beneficiar processos de categorização a longo prazo, além da hipótese mencionada no último parágrafo relacionada aos falantes de espanhol já estarem em seu limite superior de percepção (vivendo na Inglaterra e usando muito mais inglês do que os sujeitos alemães). No segundo experimento, os autores investigaram de perto essas questões. Os falantes de

espanhol passaram por mais dez sessões de treinamento (média de quatro meses após o experimento 1), enquanto os alemães foram contatados um ano depois para realizar uma tarefa de identificação e verificar se houve retenção. Após essas dez sessões extras, os falantes de espanhol finalmente conseguiram alcançar o mesmo nível de precisão que os falantes alemães. Quanto ao fato de a precisão ter melhorado sem uma correspondência com mudanças no melhor exemplar, os autores afirmam que “o treinamento auditivo melhora a capacidade dos sujeitos de aplicar seu conhecimento de categoria existente à fala natural variável” (IVERSON e EVANS, 2009).

Thomson (2011) propõe uma extensão do Treinamento Fonético de Alta Variabilidade (HVPT) para o Treinamento de Pronúncia Assistido por Computador (CAPT). O autor afirma, primeiramente, tomando como base modelos de assimilação como o SLM, que a língua materna e a quantidade e qualidade do input da L2 são fundamentais para explicar a variação de sotaque presente em aprendizes adultos de L2. O autor investigou como o CAPT pode ter importantes contribuições para a melhoria perceptiva na segunda língua. Como conclusão, Thomson (2011) afirma que as ferramentas de computador disponíveis apenas se assemelhavam a um ambiente de sala de aula e, portanto, precisavam ser aprimoradas e tornadas mais acessíveis e úteis para os aprendizes. O experimento apresentado em Thomson (2011) teve 22 falantes de L1 mandarim discriminando dez vogais do inglês canadense (contidas em sílabas /pV/-/bV/) produzidas por dez falantes de inglês canadense. Os sujeitos tinham que escolher uma entre uma série de dez imagens marcantes que representavam cada categoria de vogal. Após o pré-teste, treinamento e pós-teste, houve melhora tanto em contextos treinados quanto não treinados. O autor também afirma que é importante expandir os ramos de CAPT e HVPT para a sala de aula e colocar a pesquisa em prática fazendo isso. Além disso, é perceptível que até 2011, segundo Thomson (2011), apenas dois estudos haviam mostrado que a percepção poderia interferir na produção, então esse campo tinha um grande potencial para pesquisa.

Diversos trabalhos testam diferentes variáveis em um treinamento perceptual. A pesquisa já descrita Hirata (1999), por exemplo, verificou a diferença em colocar apenas palavras ou sentenças inteiras em um treinamento fonético. Por outro lado, Inceoglu (2014), comparou um treinamento audiovisual (AV) com um treinamento perceptual puramente auditivo (A). Já no resumo da pesquisa, o autor afirma que é conhecida a generalização de TP para a produção da fala, e cita Hardison (2003) e Hazan *et al.* (2005) como exemplos de trabalhos que mostram que um treinamento audiovisual é ainda mais eficaz que treinamentos somente perceptuais auditivos. O treinamento contou com seis sessões de 30min, e dividiu os 60 participantes, nativos estadunidenses e aprendizes de francês, em três grupos: com treinamento AV, com treinamento A e grupo-controle, sem treinamento. As tarefas de produção foram avaliadas de duas formas diferentes: através da avaliação de falantes nativos em uma escala de qualidade, e a avaliação de falantes nativos com uma tarefa de identificação forçada. O escopo fonético da pesquisa foi amplo, contando com diversas vogais e consoantes da língua francesa, divididos em 108 palavras CVC. A pesquisa encontrou diversos resultados significativos para os treinamentos AV/ A na percepção e na generalização para a produção da fala. Como alguns exemplos para a percepção, o grupo AV foi de 58,75% para 77,22% para /ɔ/ nasal, por exemplo e de 74,18% para 90,12% na vogal /ɛ/ nasal. Para o grupo experimental A, exemplos de melhora foram notados também para as nasais /ɔ/ e /ɛ/, respectivamente de 60,77% para 72,01% e 78,88% para 88,79%. Para a produção, notou-se que nem todos os segmentos apontaram resultados estatisticamente significativos de melhora, como é o caso de /ɛ/ para todos os grupos.

Hazan *et al.* (2005) é uma importante pesquisa que comparou treinamentos perceptuais apenas de áudio com treinamentos audiovisuais para produção da fala dos participantes. A pesquisa contou com duas partes. Na primeira, 39 nativos de japonês foram testados para os contrastes labiais/labiodentais /v-b-p/ em distinções de áudio, visão e audiovisual e, em seguida, passaram por 10 sessões (aproximadamente 40min cada ao longo de quatro semanas) de treinamento que foram ou do tipo A (auditivo), ou do tipo AV (audiovisual), antes de serem testados

novamente, os autores previram que, como o gesto labiodental não existe na língua japonesa, possivelmente os aprendizes assimilaram /v/ como /b/. Os pares /b/ e /p/ do inglês, que apresentam a grande distinção de aspiração para a contraparte não-vozeada, também poderiam apresentar dificuldades para japoneses, uma vez que a língua japonesa, assim como as línguas latinas, não apresenta aspiração para /p/. O número de tokens para os pré-testes e pós-testes foi de 81, sempre contendo palavras inventadas. Para os treinamentos, todavia, os autores utilizaram palavras reais da língua inglesa, e uma série de 75 *tokens* foram escolhidas ao todo. Os testes para A, AV e visual (V) foram todos de identificação, e o total de respostas foi de 486 por participante (cada bloco de 81 palavras havia sido repetido uma vez por participante). No segundo estudo, 62 nativos de japonês passaram por uma fase de pré-teste similar, e foram treinados também em 10 sessões nas modalidades A (auditiva), AV (audiovisual natural) ou AV sintético (uma modalidade que um rosto sintetizado era sincronizado com sons da fala naturais). Em resumo, no primeiro estudo, notou-se que o treinamento AV foi mais eficaz que o treinamento A para a percepção do trio /v-b-p/. No segundo estudo, todos os grupos treinados apresentaram mais precisão na percepção da distinção de /l-r/, e notou-se que os grupos AV não demonstraram mais ganhos que estudantes do grupo A. Para a produção da fala, contudo, os participantes do grupo AV natural foram os que mais aumentaram sua precisão, o que segundo os autores, indicam que, nos casos em que o contraste fonêmico é saliente o suficiente, treinamentos audiovisuais são mais eficazes do que treinamentos puramente auditivos. Ao se comparar os dois experimentos, nota-se que certas pistas visuais parecem mais difíceis de serem compreendidas do que outras. Para /l-r/ não há mudança estatisticamente significativas no pós-teste após treinos A ou AV, o que foi contrário ao que a pesquisa anterior de Hardison (2003, *apud* HARDISON, 2005) havia demonstrado. Todavia, este resultado está de acordo com o que os autores de Hazan *et al.* (2005) já haviam previsto com relação ao treinamento AV ser mais ou menos eficaz de acordo com a distância do contraste visual que está sendo treinado (exemplo meu: um /p/ se articula de forma muito diferente, visualmente, a /l/, o que permite um amplo contraste visual).

Outras pesquisas, como a de, González-Bueno e Shaw (2011) investigaram a influência de treinamento perceptual na escrita da língua inglesa por 11 aprendizes nativos de espanhol. As autoras realizaram 20 sessões de treinamentos focadas em discriminação e identificação para 11 estudantes de terceiro ano nativos o espanhol que moravam no centro-oeste dos EUA. As sessões foram bissemanais, com foco no par /d/ e /ð/ (palavras como *wordy* e *worthy*), pelo fato de /ð/ não fazer parte do inventário fonético de línguas como o português e o espanhol, causando confusões em palavras que começam com a grafia de “th” em inglês. As sessões finais de treinamento focaram na escrita dos grafemas associados aos fonemas de /d/ e /ð/.

As pesquisas iniciais ocorridas décadas atrás em TP investigaram o par /r-l/ do inglês por falantes japoneses. Entretanto, pesquisas mais recentes continuaram essa investigação. Shinohara e Iverson (2015) investigaram um treinamento de 10 sessões de ID e DIS para o par /r-l/ para 16 crianças japonesas com conhecimento elementar de inglês (saudações e cores foram citadas pelos autores). Um grupo de 12 crianças foi selecionado como controle, por não receber o treinamento e apenas realizar o pré-teste e o pós-teste como comparação. Além da percepção, os participantes da pesquisa produziram tarefas de produção para verificar a fala (as tarefas de produção foram julgadas por falantes nativos). Os autores constataram que houve mudanças positivas na produção dos segmentos /r, l/ pelas crianças japonesas, mas esta pronúncia não pareceu estar alinhada à percepção da pista acústica principal, o terceiro formante, F3, correlato de outros aspectos da forma da cavidade oral. Continuando as pesquisas no /r, l/ do inglês por falantes japoneses, uma pesquisa ainda mais recente feita pelos mesmos autores, Shinohara e Iverson (2021), expandiu o escopo da pesquisa de 2015 ao comparar crianças, adolescentes e adultos. Desta vez, os autores reportaram que crianças e adolescentes obtiveram mais ganhos do que os adultos, sugerindo que pode haver declínio na capacidade de formação de novas categorias fonêmicas com o avanço da idade. Adolescentes demonstraram mais ganhos na identificação, discriminação e sensibilidade perceptual para a pista acústica de F3 do que as crianças, ao que

os autores indicam que talvez seja pelas crianças ainda possuírem uma cognição e consciência fonológica imaturas.

Mokari e Werner (2017), além de realizarem um HVPT e investigarem as generalizações para percepção e produção, investigaram se o resultado da pesquisa estava relacionado às habilidades musicais⁴ dos participantes. A pesquisa contou com 38 azeris⁵ (28 no grupo experimental e 10 no grupo controle) aprendizes de inglês como L2, e as tarefas principais foram de discriminação. Para medir as habilidades musicais dos participantes, três testes do *Wing Measures of Musical Talents* foram utilizados. Estes testes estavam relacionados à habilidade de identificar quantas notas existiam em um acorde, discriminar entre dois acordes, e memorizar sequências de notas musicais. Nota-se que a língua-alvo desta pesquisa era o inglês, uma língua não-tonal, assim como o Azeri, o que pode isolar esta variável. O número de sessões foi de cinco ao todo, divididas entre duas ou três semanas, e cinco falantes da variante SSBE, Standard Southern British *English accent*, ou sotaque padrão do sul da Grã-Bretanha, foram utilizados para o HVPT. O estímulo foi dividido para formar 11 pares mínimos (basicamente entre todas as vogais) do inglês SSBE. Os testes de produção foram julgados por falantes nativos, que precisavam ouvir as produções e indicar qual vogal haviam escutado. Após testes estatísticos, os autores notaram que o regime de treinamento significativamente melhorou a percepção e produção dos participantes do grupo controle. Os testes estatísticos mostraram que, no geral, as habilidades musicais não possuíam relação com os ganhos de percepção e produção. Todavia, ao analisar cada habilidade musical separada, foi notada significância para ganhos de percepção e memória tonal, além de que a memória tonal foi uma preditora dos ganhos de discriminação.

Alves e Luchini (2017) estudaram como um grupo de 24 estudantes argentinos da língua inglesa melhorou sua percepção e produção de oclusivas

⁴ Segundo autores como Okada (2021), habilidade musical é um conceito com pouca assertividade acerca do seu significado: como definir e como mensurar? Na pesquisa analisada de 2017, foram elencados os testes de discriminação entre dois acordes, identificação de notas em um acorde e sequência de notas tocadas.

⁵ Indivíduos provenientes do Azerbaijão.

surdas iniciais após sessões de treinamento fonético. Todos os 24 sujeitos tinham proficiência em inglês de nível C1/C2, o que significa que eram falantes em níveis avançados. Os grupos nesta pesquisa foram um de controle (os participantes não participaram do treinamento), um grupo que passou por três sessões de treinamento e um grupo que, além de participar das sessões de treinamento, recebeu instruções sobre o segmento-alvo. Os autores descobriram que ambos os grupos experimentais se beneficiaram do treinamento em relação à percepção, e o grupo que recebeu instruções explícitas sobre o segmento também melhorou sua produção. Os participantes ouviram 360 tokens de oclusivas iniciais do inglês produzidos por falantes nativos. Os VOTs foram negativos (consoante sonora), positivos (VOT do inglês nativo) e artificialmente reduzidos a zero e precisaram identificar entre /p, b, t, d, k, g/. O teste de produção consistiu em ler uma palavra apresentada através de *slides*. O áudio foi então analisado acusticamente no software Praat. Em relação aos resultados de percepção, o grupo de controle não teve nenhuma diferença estatisticamente significativa nas etapas de coleta de dados. Tanto o grupo de treinamento quanto o grupo de treinamento + instrução, no entanto, apresentou melhores resultados na identificação de VOT positivo, VOT zero e zero artificial. Em relação aos resultados de produção, o grupo de treinamento melhorou sua produção de /p, k/, e o grupo de treinamento + produção melhorou todas as três oclusivas surdas /p, t, k/. Esse último resultado demonstra que instruir explicitamente pode ser útil. Também é se nota a eficiência do treinamento fonético, uma vez que o primeiro grupo conseguiu melhorar duas consoantes sem nem mesmo ser informado sobre o que era o treinamento.

Shum *et al.* (2021) é uma recente pesquisa que contou com 70 participantes - uma quantidade grande ao se comparar com o número médio encontrado em muitas RS. A pesquisa adota o paradigma da alta variabilidade, testes de generalização e retenção postergada do conhecimento. O foco fonético da pesquisa foram as distinções entre plosivas vozeadas e não-vozeadas do inglês, e as sessões foram 20 ao total (com 72 *tokens* por sessão, que duravam em média 20min cada), divididas em um período médio de cinco semanas. Os participantes da pesquisa eram chineses, que reportaram mandarim ou cantonês como suas

respectivas L1, aprendizes de inglês, divididos em crianças, adolescentes e jovens adultos, que estavam aprendendo inglês desde o jardim de infância.

Hwang (2023) é uma das pesquisas mais recentes incluídas nesta revisão literária da bibliografia. Além disto, esta pesquisa é a que contou com um dos maiores números de participantes na presente Revisão Sistemática: 105 vietnamitas, divididos entre 75 no grupo experimental e 30 no grupo controle. Os participantes do grupo experimental foram classificados em níveis de proficiência básico, intermediário e avançados para a língua coreana, foco do treinamento perceptual. Os participantes participaram de cinco sessões de treinamento de 20-40min em um período de duas semanas, e o foco do treinamento foram as plosivas iniciais do coreano, que possuem a característica da duração do *Voice Onset Time* (VOT) em suas contrapartes não-vozeadas. Alguns dados apontaram ganhos de até 20% para consoantes aspiradas (grupo de coreano básico que apresentou 68% de acertos para a aspirada /^h/ no pré-teste para 88% de acertos no pós-teste), além de 20,8% de ganhos na consoante alvéolo-palatal africana /^ç/, sendo que a diferença do pré e pós teste foi de 36% para 56,8%. No grupo intermediário, a consoante /^ɸ/, que equivale à bilabial não-vozeada /p/, apresentou uma melhora de 16,7%, de 60,8% de acertos para 77,5% no pós-teste. Para os falantes avançados, houve melhora de até 12,9% na plosiva alveolar /^ɬ/, equivalente ao /t/, que foi de uma taxa de acertos de 57,5% para 70,4% no pós-teste. Cabe notar que todos os dados apresentados neste parágrafo são estatisticamente significativos, e a pesquisa relatou dezenas de outros resultados significativos para o treinamento. Um dado interessante, que conversa com a pesquisa de Lee e Hwang (2016b), que também utilizou a língua coreana, embora a partir de falantes japoneses, é que se observou melhora na percepção dos sons do coreano por parte de todos os grupos de proficiência, inclusive dos níveis mais avançados. A pesquisa de Lee e Hwang (2016b) reportou que os grupos com menor proficiência indicaram muito mais ganhos entre o pré e o pós-teste⁶.

⁶ Este parágrafo está com o espaçamento inconsistente devido às quebras de páginas.

Uchihara *et al.* (2024) realizaram uma metanálise de 31 estudos de HVPT para a generalização de produção de L2. Esta metanálise oferece *insights* valiosos para quem deseja realizar pesquisas bibliográficas, uma vez que os autores detalham minuciosamente as etapas de construção da metanálise: a busca nas bases de dados escolhidas de 1649 estudos, além da adição de 81 estudos de outras fontes, seguido da eliminação de estudos duplicados, leitura de cada artigo para investigar e avaliar a pertinência, até que sobraram 31 trabalhos para a metanálise. Dentre os critérios de exclusão utilizados estiveram: serem empíricos, possuir duas ou mais vozes (falantes diferentes) de *input*, reportar a precisão da produção da fala, ter o *design* pré-pós teste e grupo-controle, entre outros. Os trabalhos selecionados possuíram entre 17 e 43 participantes, sendo que os ganhos de pronúncia para os grupos experimentais foram, em média, entre 4,5% e 10,5%. Todavia, a metanálise não encontrou evidências fortes o suficiente para retenção postergada da produção da fala e nem para generalização para itens que não estavam presentes no treinamento. Com relação às possíveis razões que permitiram os ganhos na produção da fala estão: idade e contexto de aprendizado dos participantes; o que esteve presente no treinamento (duração do período de treinamento e duração das sessões, por exemplo, o que está presente nos objetivos específicos da presente tese) e tipos de testes de produção. Os autores apresentam a reflexão que talvez seja um erro promover a produção da fala em estágios em que as categorias perceptivas ainda não estão totalmente consolidadas: as percepções equivocadas podem levar às produções equivocadas. Os autores também ressaltam que os 31 trabalhos presentes em sua metanálise contabilizaram somente HVPT com generalização para a produção da fala, e não treinamentos perceptuais que continham **também** treinamento para a produção da fala.

2.2 As teses em TP no Brasil

Além de apresentar o referencial teórico geral, é importante inspecionar brevemente a história das pesquisas de TP no Brasil. Por serem pesquisas de doutorado, em que a variável de tempo (que geralmente é o nêmesis da maioria dos pesquisadores) não é tão crítica, opta-se por expor o caminho através do tempo de

cada um destes trabalhos (praticamente 20 anos se passaram entre Nobre-Oliveira e Guaer, respectivamente, primeira tese sobre o assunto publicada no Brasil, em 2007, e a mais recente, em 2024). Nota-se que muitos destes trabalhos foram publicados em formato de artigos e estão descritos e discutidos no capítulo de resultados da RS. Nota-se que o primeiro trabalho brasileiro de TP foi o de Nobre-Oliveira (2007), na Universidade Federal de Santa Catarina (UFSC), seguido por o de Melissa Bettoni (2008), também da UFSC. Estes trabalhos abriram caminho para que a área de TP se expandisse para o Brasil, muito embora os dois trabalhos seguintes sejam da Universidade Federal do Paraná (UFPR) e depois da Universidade Federal do Rio Grande do Sul (UFRGS), o que, salvo engano, significa que seria mais adequado dizer que se espalhou pela região sul do Brasil. Até onde nossos conhecimentos nos permitem, estas são as cinco e únicas teses empíricas e TP no Brasil, todas produzidas na região sul.

Nobre-Oliveira (2007) treinou 29 brasileiros (sendo mais sete pertencentes ao GC, além de dois nativos de inglês estadunidense também utilizados como controle) estudantes dos terceiros, quartos e quintos semestres dos cursos de Secretariado Executivo e Letras Inglês de uma Universidade Federal no Sul do Brasil para vogais (/i/, /I/, /ɛ/, /æ/, /ʊ/ e /u/ em sílabas CVC) da língua inglesa. Os estudantes reportaram nível intermediário de proficiência na língua, mas não haviam morado no exterior. Os grupos experimentais foram divididos entre o *SynS*, de 15 pessoas, que receberam TP com estímulos sintetizados, enquanto o grupo *NatS*, de 14 pessoas, receberam TP com estímulos naturais. A pesquisa também contou com um teste de produção após o treinamento, além de um teste de retenção postergada após um mês do fim do período de treinamento. O treinamento em sala de aula contou com uma parte teórica de propriedades articulatórias, de 40min, e uma parte prática, de 50min, no laboratório de línguas. As vogais anteriores foram treinadas na primeira semana, as posteriores na segunda e ambas na terceira semana. Um CD com *software* para quatro tarefas de treinamento foi passado para que os estudantes fizessem em casa, sendo duas tarefas de identificação e duas de discriminação. Como resultados para percepção, o pré-teste do GC apontou 43 acertos, com 42 no pós-teste, enquanto para os GEs os acertos foram de 42 para

51. Os testes de Wilcoxon demonstraram que houve significativamente menos erros na categorização de /i/, /æ/ e principalmente nas duas vogais arredondadas posteriores, /ʊ/ e /u/. Houve também um aumento de acertos significativos na percepção de /ɛ/. Ao comparar os dois grupos experimentais, a autora observou que ambos os grupos performaram significativamente com mais acertos independentemente do estímulo: a porcentagem de acertos de antes/depois para o *SynS* foi de 61,9% para 76% e do *NatS* de 58,5% para 69,4%. A autora também cita pesquisas que indicam que o tempo de treinamento fora do ambiente controlado interfere nos resultados, e esta variável foi uma limitação de seu estudo. O teste de generalização para estímulos naturais pelo grupo com estímulos sintetizados também demonstrou diferenças significativas. Quanto à generalização para produção, os testes estatísticos indicaram diferenças significativas na produção do contraste /i/ - /I/ pelos GEs, mas houve muita diferença entre os dois grupos experimentais, uma vez que o grupo *SynS* apresentou resultados muito mais promissores que o grupo experimental *NatS*. Os testes estatísticos de retenção postergada indicaram que os efeitos na percepção se mantiveram significativos após um mês do treinamento.

Bettoni (2008) estudou encontros consonantais iniciados com /s/ por falantes de nível avançado de inglês nativos de português brasileiro. A autora indicou que a aquisição dos /s/ *clusters* é difícil tanto em L1 como em L2, o que justifica o interesse de pesquisa. A autora afirma que frequentemente esta demografia adiciona uma vogal epentética quando estruturas como essas são produzidas. Dependendo do idioleto do falante, pode haver a palatalização do /s/. A pesquisadora, além de testar a eficácia do treinamento, verificou se houve generalização para a produção da fala e aplicou um pós-teste de retenção postergada oito meses após o fim do treinamento. O design do treinamento consistiu em tarefas de discriminação entre duas opções com *feedback* imediato. O número de sessões foi de seis, com variações de duração e quantidades de blocos de treinamento. A autora realizou um incremento na dificuldade do estímulo (quantidade de falantes, tipo de encontros consonantais e contrastes de sons). Os participantes foram 23, com 200h de instrução formal em inglês, e foram testados, para a produção da fala, em uma

entrevista, uma leitura de parágrafo, uma leitura de uma história e leitura de sentenças-veículo. Os *inputs* da pesquisa foram produzidos por três estadunidenses. Os testes foram de discriminação e identificação, e a análise da produção da fala foi feita através do *software* Praat. Com relação à percepção da fala, apesar de os testes estatísticos não terem apontado significância, os *scores* do contexto de silêncio foram maiores (90,57%) que o contexto de consoante surda (85,74%), que foram maiores que o contexto de vogal (78,8%) que, por fim, foram maiores que o contexto de consoante sonora (76,76%), o que era esperado pelas hipóteses do trabalho. Com relação aos resultados para a produção da fala, a variação individual foi maior no pré-teste do que no pós-teste, e o único participante que não apresentou resultados mais precisos para a percepção no pós-teste já havia apresentado um resultado quase perfeito no pré-teste. O ganho médio de precisão na percepção da fala foi de 8,41%, o que correspondia 59% do máximo de possível ganho de *score*. Os testes estatísticos também demonstraram que o encontro consonantal /s/+ soantes foi significativamente mais difícil de ser produzido que o encontro /s/ + plosivas. O teste de Wilcoxon para retenção postergada da produção da fala indicou que os resultados de produção, que aumentaram 22,92% no pós-teste foram mantidos após oito meses. Para a percepção, todos os sujeitos que ainda poderiam aumentar seu *score* o fizeram, e os que haviam obtido 100% de precisão (P2, P4, P8, P10 e P12) mantiveram este *score*.

Brawerman-Albini (2012) se diferencia de partida dos trabalhos mencionados até agora por treinar questões suprasegmentais de acentuação pré-paroxítona do inglês por estudantes brasileiros. Este padrão de acentuação é encontrado em palavras complexas do inglês quando um sufixo neutro é adicionado a uma sílaba que já é proparoxítona. Brawerman (2001; 2006, *apud* BRAWERMAN-ALBINI, 2012) já havia notado dificuldades por parte dos brasileiros com relação a este fator suprasegmental. Na tese de 2012, a autora contou com 30 participantes estudantes de letras-inglês de uma universidade federal no sul do Brasil, sendo 20 assinalados para o GE e 10 para o GC. Os alunos já possuíam cerca de 400h de exposição à língua inglesa, sendo falantes considerados avançados. Os *inputs* da pesquisa foram produzidos por um homem e uma mulher estadunidenses com

décadas de vivência no Brasil. As tarefas elegidas para o experimento foram de identificação com *feedback* imediato, e as cinco sessões levaram entre 10min e 35min para serem completadas por cada participante, totalizando um tempo total de treinamento entre 50min e 175min dependendo da velocidade do participante. Para resultados de percepção, o GE aumentou seu índice de acerto de pré-proporaxítonas de diferentes números de sílabas para 95,9%, aumento de 4,1% em relação ao pré-teste, o que foi estatisticamente significativo. O GE também apresentou uma taxa de acerto significativamente maior para a generalização da percepção (com três novas vozes femininas): a taxa foi de 93,4%, aumento de 6,6% em comparação ao pré-teste. Entretanto, ao se comparar estatisticamente GE com o GC, não foram encontradas diferenças significativas, o que era esperado, uma vez que os membros do GC possuíam 140h a mais de instrução que o do GE e demonstraram aumentos no índice de acerto de palavras acentuadas. O GE aumentou, de fato, seu índice de acertos, mas não o suficiente para ser indicado pelo teste estatístico. A autora também notou que os valores iniciais de acertos do GC eram de 94,3% e subiram para 96%, o que demonstrou que só repetir os pré e pós-testes não foi suficiente para melhorar estatisticamente a percepção. Os resultados gerais dos pós-testes de produção da fala indicaram que o GE produziu, em média 57,5% das palavras de forma correta. Com o intuito de comparação, os valores no pré-teste foram de apenas 28,4% correto. O teste de retenção indicou que as estatísticas de produção se mantiveram após os dois meses: 58% de produções corretas. O teste de generalização de produção também foi bem-sucedido, uma vez que indicou 59% de produções corretas. Os testes de Wilcoxon com correção de Bonferroni demonstraram estatisticamente essas mudanças entre o pré-teste e os demais. Para o GC apenas o teste de generalização apresentou melhora na taxa de acertos em comparação aos outros, com o pico de 48,4% de produções corretas.

Milan (2019) se diferencia de imediato dos demais trabalhos nesta subseção por lidar com treinamento de heterotônicos da língua espanhola. Este tipo de palavra é grafada de forma muito próxima ou igual às da língua portuguesa, mas possuem acentuação diferente em sílaba tônica. No total, foram selecionadas 56

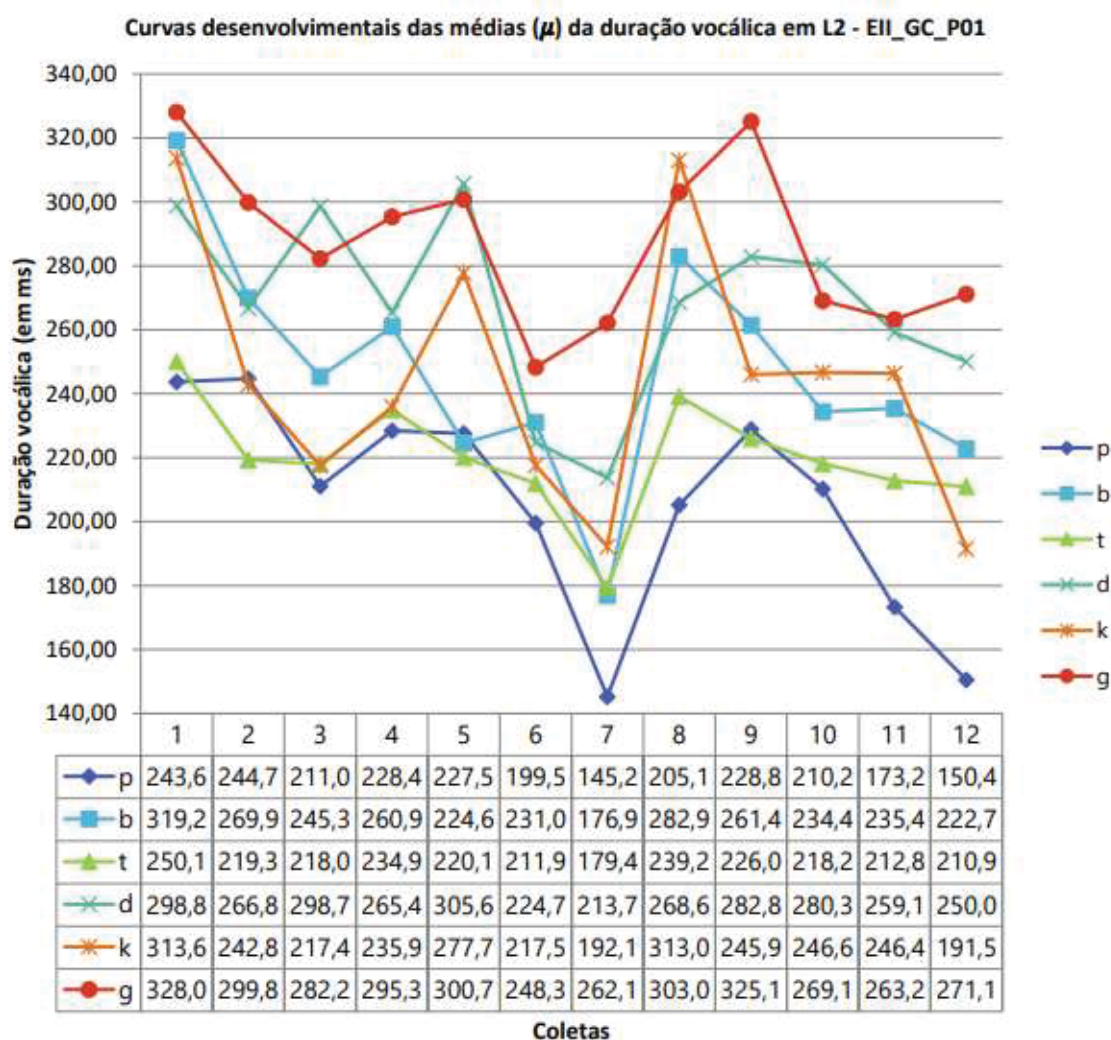
palavras para o treinamento, 40 para os testes (pré, pós e retenção, teste este aplicado entre 42 e 58 dias após o final do período de treinamento) e 19 para generalização para novas palavras. Os locutores da pesquisa foram quatro mexicanos, dois cubanos e dois hondurenhos. Os participantes da pesquisa foram alunos de três turmas de graduação em Letras Espanhol de uma universidade federal no sul do Brasil, possuindo entre 90h e 360h de exposição ao espanhol (estes grupos foram divididos em ordem crescente de proficiência, respectivamente LE1T e LE3T) em sala de aula), sendo que 10 alunos passaram pelo treinamento e três não passaram (LE1 sem treinamento, LE1ST e LE3 sem treinamento, LE3ST). O treinamento em si contou com duas sessões aplicadas em dias distintos. O *software* TP foi utilizado para tarefas de identificação com *feedback* imediato. Como este trabalho é extenso e minucioso (até para um trabalho na área de TP, que já é naturalmente extenso e minucioso), apenas alguns resultados e conclusões serão reportados. Com relação às correlações entre o pré-teste e o pós-teste para os grupos experimentais, os testes de *Kruskal-Wallis* não apontaram diferenças significativas para a percepção, o que não significa que os grupos não acertaram mais: O LE1 de pré-teste passou de 65% para 80% de acerto (15% de diferença), enquanto o grupo LE3 passou de 87% para 88%. Analisando os individualmente os participantes do grupo LE1, nota-se que houve um grande desvio-padrão entre os participantes (chegando a DP=20). Com relação a relação de percepção/produção, o teste de correlação de *Spearman* indicou que o grupo LE3 não apresentou correlação entre o que produziu e percebeu no pré-teste. Já para o grupo LE1 sem treinamento, a correlação foi positiva, forte e significativa: a produção correta dos heterotônicos em frases com contexto deste grupo não esteve associada à percepção correta deles, o que foi justamente o contrário do que aconteceu com o grupo LE1 com treinamento. Com relação aos resultados do pós-teste de produção de heterotônicos em frases com contexto, o grupo LE1 com treinamento acertou 74%, enquanto o LE3 com treinamento acertou 89%. Os testes de *Kruskal-Wallis* demonstraram diferenças significativas entre os grupos, embora não foi possível de detectar de onde essas diferenças emergiram. As médias dos grupos foram: LE1ST: 38%; LE1T: 74%; LE3ST: 86%; LE3T: 89%. Para as frases veículo, as médias de

produção foram: LE1ST: 55%; LE1T: 81%; LE3ST: 90%; LE3T: 90%. Novamente, o resultado dos testes estatísticos foram os mesmos do que para o teste com frases em contexto. Para o pós-teste de produção, os valores de acerto foram os seguintes: LE1ST: 60%; LE1T: 85%; LE3ST: 93%; LE3T: 91% (o grupo sem treinamento acertou 2% a mais do que o grupo com treinamento, o que também chamou atenção da autora). O único grupo que apresentou correlação positiva, forte e significativa, entre o que produziu e o que percebeu em frases-veículo foi o LE1 com treinamento, ou seja, ter produzido corretamente esteve associado a perceber corretamente estes heterotônicos no teste de percepção. Com relação aos testes de generalização da produção, foi constatado que os grupos com maior experiência acadêmica foram capazes de generalizar os resultados de produção em frases com contexto. Para a generalização da percepção, notou-se novamente que todos os participantes dos grupos de LE3, com ou sem treinamento, tiveram um bom desempenho e se mantiveram próximos às médias do grupo, enquanto para os grupos de LE1 houve gigantescas diferenças entre participantes e um grande DP total (27% para o LE1ST). Para os testes de retenção de produção em frases com contextos (aplicados aproximadamente 1,5 mês após o fim do treinamento), os testes de *Kruskal-Wallis* e *Post Hoc de Mann Whitney* demonstraram que os únicos grupos que apresentaram diferenças significativas nos acertos foram o LE1 com (84% de média) e sem treinamento (49% de média).

Gauer (2024) é a pesquisa brasileira mais recente na área de TP e, por ser uma tese de doutorado, e não um artigo, não pode ser incluída na RS por não obedecer aos critérios de inclusão. Esse trabalho, assim como o de Milan (2019), é extremamente extenso e detalhado e este parágrafo é apenas uma brevíssima apresentação do total. Todavia, diversas questões presentes na tese de Gauer (2024) são importantes para a discussão da área. A pesquisadora treinou estudantes da educação básica (ensino médio) brasileiros para os contrastes de vozeadas/não-vozeadas das plosivas no inglês através de um regime de treinamento de cinco sessões através de cinco dias, além de testes de generalização (inclusive para a fala) e retenção (10 semanas após o fim do treinamento). Os 48 participantes foram divididos em grupos experimentais (GE1

que recebia o treinamento perceptual e o GE2, que recebia, além do treinamento perceptual, a instrução explícita das formas-alvo) e um grupo controle (GC). Em uma subseção de análise de Trajetórias desenvolvimentais dos participantes do estudo, a autora analisa individualmente dois participantes de cada um dos grupos. Ao se analisar a participante 1 do grupo controle, nota-se que as curvas de valores médios de durações absolutas das vogais adjacentes às consoantes-alvo do estudo oscilam muito, como observado na figura 1:

Figura 1: Curvas desenvolvimentais dos valores médios das durações absolutas produzidas pela Participante 1 o grupo Experimental e Gauer (2024)



Fonte: Gauer (2024, p.298)

A questão da grande variabilidade das durações médias produzidas pela participante 1 do grupo experimental não é deixada de lado pela autora. O trabalho, que se utiliza das lentes da complexidade através da teoria dos Sistemas Dinâmicos Complexos (SDC), afirma que o fato da participante ser a que possui menor grau e proficiência em língua inglesa pode significar maior instabilidade nas produções na L2, por razões relacionadas à experiência e autoconfiança. Mais à frente na análise, ao estudar os dados de produção de outra participante do grupo controle, a autora também nota oscilações nos subsistemas linguísticos da participante, o que pode suscitar a questão de quanto os dados dos participantes que estão nos grupos experimentais podem sofrer variações não controladas pelos treinamentos/experimentos. Neste sentido Gauer (2024) citou fatores como a pandemia de Covid-19 que impactou de grande forma a vida dos participantes da pesquisa.

2.3 Teorias da Percepção e Modelos perceptuais

Com a intenção de narrar de forma bastante sintetizada a história do estudo da percepção da fala passando brevemente pelo gerativismo através da Teoria Motora, além da Teoria do Realismo Direto, pega-se emprestado de Nishida (2012) informações importantes acerca destes desenvolvimentos. Após isso, apresento os modelos perceptuais mais citados pelos trabalhos do *corpus* da presente revisão sistemática: o *Speech Learning Model*, o *Perceptual Assimilation Model* e o *Native Language Magnet Model*. Seguindo os pressupostos apresentados por Alves (2024), no livro *Percepção e produção dos sons de línguas não nativas: fundamentos teóricos e questões de investigação no contexto brasileiro*, busca-se conectar os pressupostos dos modelos perceptuais às concepções teóricas dos modelos de percepção.

Primeiramente, discutindo visões de percepção psicoacústicas provenientes de autores como Pisoni (1973, 1974) e Carney (1977), tem-se a pressuposição que os sons são identificados como iguais ou diferentes por uma decisão binária. Estas visões psicoacústicas, como descreve Alves (2024), partem de primitivos acústicos

para a diferenciação dos sons, e as representações categóricas são selecionadas a partir do continuum de produção da fala. Os ouvintes utilizam sua memória de curto prazo para comparar o estímulo auditivo com as unidades propriedades auditivas armazenadas na memória de longo prazo. O modelo do *Speech Learning Model*, que será descrito com mais atenção à frente, está aliado a esta visão psicoacústica.

Desde a década de 1950, pesquisadores colaboradores dos laboratórios *Haskins* se perguntavam quais pistas acústicas eram responsáveis pela distinção dos sons da fala. Os pesquisadores descobriram que a questão era muito mais complexa do que parecia, uma vez que a relação entre o sinal acústico e a articulação da fala não era exatamente proporcional: uma única característica acústica não dá conta de descrever uma articulação do som. Destas observações, duas teorias de épocas diferentes buscam explicar a percepção da fala de modo articulatório: a Teoria Motora da Percepção da Fala (TM), (LIBERMAN, MATTINGLY, 1985, *apud*. NISHIDA, 2012) e a Teoria do Realismo Direto da Percepção da Fala (TDR), por Fowler (1996, *apud* NISHIDA, 2012). Segundo Nishida, a TM tem origem por volta dos anos de 1950, no contexto apresentado no início do parágrafo. A versão mais conhecida da teoria data de 1985, e por isso é a elegida para ser referenciada. A teoria se baseia em uma noção de gestos pretendidos da fala, ou seja, gestos que seguem uma dicotomia entre o fonético (realizado) e o fonológico (o idealizado), portanto o conjunto de traços que forma a fala deve ser enxergado como um padrão de gestos percebidos. Ademais, a TM também pressupõe que a produção e a percepção da fala são indissociáveis, o que deixa transparecer características epigenéticas que remetem ao gerativismo, conectando isto com a visão de língua inaugurada por Chomsky. Como explica Nishida (2012), a TM explica a fala de forma modular, ou seja, como a relação entre o gesto articulatório e o sinal acústico não é completa, é necessário adotar os gestos pretendidos como primitivo de percepção.

A TDR, por sua vez, parte das premissas que os gestos de fala não são idealizados, mas sim os próprios componentes fonológicos de uma sentença. Por

outro lado, em vez de os indivíduos possuírem um módulo específico de percepção (como a TM propõe, indicando esta visão gerativista), eles percebem os sons da fala de maneira real e direta. A TDR nega módulos especializados para a fala, abordando a questão de forma realista e evolucionista. Conforme explicado por Alves (2024), o *Perceptual Assimilation Model*, que será discutido mais à frente, se encontra alinhado com as premissas da TDR, através de uma concepção ecológica e que vincula o acústico ao articulatório.

O *Speech Learning Model* (SLM) e sua versão revisada (SLM-r) serão explicados através de pesquisas como Flege (2002, 2003, 2007) e Flege e Bohn (2021). Nota-se que, como o artigo que apresentou o SLM-r é recente, é esperado que pesquisas utilizando este modelo levem alguns anos para se popularizarem. A versão original do SLM postulava que a capacidade neurológica de um indivíduo permanecia plástica para a aquisição de línguas estrangeiras, e pesquisas como Flege *et al.* (2002) compararam imigrantes e aprendizes tardios de línguas estrangeiras para observar as diferenças da aquisição em idades mais avançadas. Outrossim, o SLM possui uma visão de que a fonética e a fonologia se interseccionam e são separadas de forma tênue e difícil de discernir. Consoantes e vogais que são próximas, mas não iguais, entre L1 e L2, como traz Silva (2021), podem ser estabelecidas na mesma categoria fonológica e podem ser facilmente separadas mais cedo na vida. Esta característica de observar diferenças individuais em uma segunda língua para estudar o período crítico (etário) no que concerne à aquisição de segundas línguas, é muito comum nos modelos antigos do SLM (e os que são citados nas pesquisas presentes nesta Revisão Sistemática). A versão revisada e atualizada, segundo Silva (2021), no entanto, tem por objetivo os aprendizados sequenciais das L2 após a consolidação da L1. Para explicar alguns preceitos do *Perceptual Assimilation Model* (PAM, de Best, 2005 e o PAM-L2 de Best e Tyler, cabe a comparação com alguns postulados do SLM. Segundo Florencio (2021), em um quadro comparativo entre o SLM e o PAM, enquanto o SLM postula que os mecanismos e processos utilizados no aprendizado dos sons da L1 se mantêm intactos através da vida e podem ser utilizados para as aquisições de L2, o PAM postula que os sujeitos extraem invariáveis dos gestos articulatórios

provenientes dos sinais acústicos, em vez de formarem categorias por pistas acústicas. Florencio (2021) também distingue que o SLM afirma que aspectos dos sons da fala são especificados como representações (denominadas categorias fonéticas) na memória de longo prazo, enquanto o PAM afirma que o ouvinte capta os gestos articulatórios diretamente, e através do aprendizado perceptual começa a detectar invariáveis no sinal do som. Uma explicação mais direta é que o SLM explica a questão através de representações de categorias fonéticas, enquanto o PAM propõe uma percepção direta dos sons da fala.

Neste ponto, cabe retomar importantes distinções trazidas por Nishida (2012). O autor, ao investigar as teorias de percepção da fala, afirma que a percepção da fala pode ser encarada de forma auditiva ou de articulatória. Na primeira, os argumentos são de que existem invariantes acústicos que permitem a categorização de unidades fônicas distintivas, o que parece se aproximar bastante dos preceitos trazidos pelo SLM. Da segunda forma, a articulatória, os argumentos apontam no sentido de apresentar os gestos como o primitivo perceptual. Estes gestos podem ser observados como abstratos, como traz a Teoria Motora da percepção da fala (TM), ou de forma real, como traz a Teoria do Realismo Direto (TDR). Esta última, encontra-se mais alinhada aos preceitos do PAM. Além disso, trabalhos como o de Gauer (2024), que foi publicado após o início da escrita da presente tese, mas que conta com uma metodologia robusta e execução de um treinamento perceptual com efeitos significativos, trabalham com teorias de visão de línguas, como, no caso a dos Sistemas Dinâmicos Complexos (SDC), o que mostra que é possível alinhar outras teorias, que vão além dos modelos como SLM, PAM e NLM.

Outros pontos que diferenciam o SLM do PAM são relacionados ao período crítico de aquisição. Conforme apresenta Florencio (2021), enquanto o SLM propõe que as categorias fonéticas na infância e suas propriedades evoluem com o passar do tempo e da aquisição de novas L2, a base da L1 sempre vai existir. O PAM, por outro lado, percebe os sons da L2 por meio dos gestos articulatórios, que são refinados continuamente suas percepções através da vida. Como os fonemas da L2

podem ser diferentes, os ouvintes precisam descobrir novas invariantes para abarcar a nova categoria fonológica descoberta. Ademais, o SLM postula que falantes de mais de uma língua batalham para criar contraste entre as categorias fonéticas da L1 e das L2, porque estas categorias coexistem em um mesmo espaço fonológico. Por outro lado, o PAM não concorda com este postulado e afirma que o aprendizado da língua depende de forma crucial na relação dos espaços da L1 e das L2.

O terceiro modelo mais presente da bibliografia incluída nesta RS é o *Native Language Magnet Theory*, de Kuhl, 1995 (NLM). De acordo com Leong *et al.* (2018), o modelo NLM propõe que as categorias fonéticas existem e são consolidadas na aquisição da L1. Estas categorias, então, agem como ímãs que aumentam a similaridade percebida de sons que estão próximos aos das categorias consolidadas. Este efeito magnético é o que causaria dificuldades para ouvintes diferenciarem sons acusticamente similares, mas linguisticamente diferentes entre línguas. O que o NLM propõe, nestes casos, como já descrito, é que é importante que o falante supere estas interferências entre L1 e L2 e crie categorias. Além disso, Nunes (2021), afirma que o NLM postula que a percepção de línguas não-nativas ocorre dentro das fronteiras das categorias estabelecidas pela L1.

Olson (2019) citou o SLM, PAM e NLM para discutir os achados de sua investigação. O objetivo da pesquisa era a produção da fala dos estudantes, que foi treinada através de um regime de três sessões com feedback visual, e contou com 25 nativos do inglês americano de nível intermediário de proficiência na L2 para as três consoantes plosivas não vozeadas do espanhol (/p, t, k/). O autor categoriza os modelos do SLM, PAM e NLM como “*segment-by-segment basis*”, ou seja, que possuem base em cada segmento per se, o que destoa de abordagens como a de Clements (1985) e Sagey (1986), *apud* OLSON (2019) que propõe que os aprendizes podem adquirir séries de conhecimentos que podem ser aplicados para múltiplos fonemas. Como Olson (2019) observou que o grau de mudança de fonemas que não foram treinados foi fortemente relacionado com o grau de mudanças de fonemas treinados, o estudo revelou que as generalizações

ocorreram de forma a demonstrar que o processo de aquisição seguiu linhas como a de Clements (1985) e Sagey (1986), mais voltadas ao nível do subsegmento.

Além do PAM, SLM e NLM, outros modelos foram citados, como o *Motor Theory of Speech Perception* por Shinohara e Iverson (2015), trabalho descrito na subseção 4.15. Em Shinohara e Iverson (2015), os autores investigaram os efeitos de 10 sessões de HVPT em crianças japonesas para o par de consoantes do inglês /r, l/, e utilizam preceitos da teoria, como a capacidade inata de perceber como gestos articulatórios, para discutir seus dados.

A seguir, descrevo as diferentes metodologias encontradas em trabalhos documentais, com o intuito de fortalecer a bibliografia da presente tese.

2.4 Diferentes metodologias em trabalhos bibliográficos

O objetivo desta seção é discutir trabalhos bibliográficos na grande área da linguística, especificamente na fonética/fonologia, analisando criticamente a metodologia aplicada por cada autor para entender como se dão os processos de execução de trabalhos como de Revisões Sistemáticas (RS), Revisões Integrativas (RI), Metanálises e Metanálises Qualitativas (MQ). É importante destacar que **não** é intenção do presente trabalho aplicar juízo de valor às pesquisas aqui citadas. Julguei insuficientes, geralmente por partir de uma grande quantidade de dados, as descrições e narrativas das pesquisas bibliográficas que estudei. O que esperava encontrar, falando de forma mais explícita, era uma discussão ampla e entranhada que fosse além da mera descrição de compilados de dados levantados. Nota-se também que, uma vez que as pesquisas descritas nesta seção não são a respeito de treinamentos fonéticos/perceptuais, a maioria das pesquisas não será descrita em profundidade, tendo um foco incidido especificamente nas seções de metodologia. Cabe notar, como observado por um membro da banca de qualificação, que todos os trabalhos presentes nesta subseção são em português (sendo um de Portugal e os demais do Brasil). Este fato se dá, talvez inconscientemente, por ter priorizado pesquisas que mostram o estado da arte em

terras nacionais. Realizei as buscas nos periódicos CAPES e google acadêmico sem me atentar que só foram feitas em português, e os trabalhos encontrados foram estes presentes a seguir. Algo que ficou evidente nesta pesquisa é que a área de ciências humanas em geral carece muito deste tipo de pesquisa, tão comum nas ciências médicas. Espero que este trabalho ajude a mudar isso.

Mattos (2015), em um trabalho publicado pela UNESP, descreve os principais tipos de revisão de literatura. A primeira, denominada de narrativa, não se utiliza de critérios explícitos e sistemáticos para a busca e tece análises críticas. Como forma de contextualizar a área de TP na presente tese, realizei uma revisão narrativa para encontrar trabalhos que já levantaram dados sobre a área (como Barriuso e Hayes- Harb, 2018, e Milan e Kluge, 2021). Além disto, a Revisão Narrativa não aplica estratégias sofisticadas e exaustivas, e fica sujeita à subjetividade dos autores. É recomendada, e frequentemente utilizada, para realizar fundamentações teóricas, como, conforme mencionado, foi feito para começar o presente trabalho.

De um ponto de vista baseado em teses com extensas e expansivas discussões teóricas, como Albuquerque (2019), Perozzo (2017) e Nishida (2012), cogitou-se a possibilidade de desenvolver uma tese que iria além de uma revisão sistemática: uma tese que estabelecesse narrativas e discussões aprofundadas a partir de uma pesquisa bibliográfica.

Na presente revisão sistemática existem trabalhos publicados entre 1968 até 2024. Em uma pesquisa geral através do google acadêmico, por “treinamento perceptual L2 revisão sistemática” em português, não foi encontrado nenhum trabalho sobre isto ainda em 289 resultados. Em inglês, fazendo uma busca ampla com 29300 resultados, encontram-se publicações recentes, e nenhuma delas uma tese, tentando dar conta do tema: Rato e Oliveira (2023) com uma RS, Yao *et al.* (2025) com uma metanálise e Zhang *et al.* (2021) com uma RS, além dos artigos de revisão como o já citado Barriuso e e Heyes-Harb (2018) ou um manual como o de Milan e Kluge (2021). Apesar de serem muito comuns na área da saúde, pesquisas de RS, RI, metanálises e MQ continuam escassas nas áreas de humanas, letras,

linguísticas e, mais especificamente ainda, fonética e fonologia. Estes trabalhos que compilam grandes quantidades de informação e oferecem análises e sugestões, até Agendas para cada área, são essenciais, e alguém precisava realizar este trabalho no Brasil.

A RS, por sua vez, é mais, em um sentido de senso comum da palavra, científica. Uma RS testa hipóteses e busca responder as perguntas formuladas, sintetizando resultados de estudos primários. Os métodos são sistemáticos e as perguntas são claramente formuladas. Nas áreas da saúde, é considerada a evidência científica de maior grandeza para tomadas de decisão. Como a presente pesquisa não se trata da área de saúde, estimo que uma RS possa colaborar no sentido de levantar discussões importantes sobre as práticas em TP.

Uma Revisão Integrativa, por sua vez, possui uma junção de várias metodologias combinadas. O rigor metodológico de uma RI mantém os critérios da RS, porém o método permite cominar dados de literatura empírica e teórica que podem ser utilizados para identificar lacunas em áreas de estudos, além de revisar teorias e realizar análises metodológicas. Souza *et al.* (2010), ao discutir o que é uma RI e como realizar uma RI na área da saúde, indicam seis etapas necessárias ao se utilizar a epistemologia Cochrane⁷: elaboração de uma pergunta norteadora, busca ou amostragem na literatura, coleta dos dados, análise crítica dos estudos, discussão dos resultados, apresentação da revisão. Assim como na RS, é necessário definir as bases de dados, termos pesquisados, idiomas e as datas da pesquisa, bem como as ordenações por procedência, título, autores, periódico, considerações e temática. Como um prelúdio das pesquisas em linguística e fonética que analisei para a presente tese, é possível afirmar que a maioria dos trabalhos de RS e RI investigados foram muito semelhantes em metodologia.

Dentre as pesquisas estudadas na fase de seleção de metodologia, encontram-se artigos com discussões acerca das escolhas metodológicas de cada

⁷ Esta epistemologia estuda a natureza, dos métodos e da validade do conhecimento científico produzido pela Cochrane, uma organização internacional que produz e dissemina revisões sistemáticas de alta qualidade em saúde.

autor. Novamente, insisto que juízos de valores não são aplicados aqui. Nas próximas subseções, serão apresentadas pesquisas nas áreas de linguística e fonética que realizaram diferentes trabalhos teóricos para se ter uma base do que foi realizado antes em contextos similares.

2.4.1 Sella *et al.* (2013)

Sella *et al.* (2013), em uma RS de comportamentos pré-linguísticos e os primeiros comportamentos de linguísticos de cegos congênitos, executaram a metodologia da maneira descrita no Quadro 1.

Quadro 1: Metodologia da RS de Sella et al. (2013)

Critérios de Filtragem	a) Periódicos presentes nas bases: Google Acadêmico, PubMed, Psycinfo e Scielo; b) Seleção de palavras-chave; c) Critérios técnicos (e.g: cegueira congênita)
Organização	Organização em nove fases acerca dos dados bibliográficos e dos conteúdos do artigo.
Ações posteriores à filtragem	Descrição em cinco etapas sobre os artigos que passaram pelas etapas anteriores.

Fonte: O autor.

Após todas as etapas da RS, as autoras contaram com 19 artigos para a análise. As conclusões do trabalho são, resumidamente, a necessidade de mais publicações nacionais e de melhores definições operacionais dos comportamentos-alvo trabalhados. As autoras também indicam que o Google Acadêmico (scholar.google.com) foi o indexador com mais resultados e com os resultados mais acessíveis de pesquisa.

2.4.2 Jesus *et al.* (2020)

Jesus *et al.* (2020) realizaram uma RS sobre o uso de aplicativos no ensino da língua inglesa. Os autores realizaram os procedimentos ilustrados no Quadro 2 para a RS, baseados em Galvão e Ricarte (2019), ou seja, uma modalidade que busca entender a lógica de um *corpus* de dados a partir do seguimento de protocolos específicos.

Quadro 2: Metodologia da RS de Jesus et al. (2020)

Critérios de Filtragem	a) Periódicos presentes na base do CAPES; b) entre os anos de 2015-2020; c) seleção de palavras-chave; d) operadores lógicos.
Ações posteriores à filtragem inicial	Filtros adicionais: a) Artigos publicados em português; b) artigos avaliados por pares; c) artigos com visualizações gratuitas; d) exclusão de artigos repetidos.

Fonte: O autor.

Jesus *et al.* (2020) encontraram, após as filtragens, 13 artigos, organizados posteriormente em um quadro de dados. Como conclusão, os autores foram capazes de identificar diversos aplicativos utilizados para ensino de inglês em sala de aula, bem como as formas com que professores e alunos se relacionam com os aplicativos e o ensino de tradução, escrita, vocabulário, entre outros itens. Os autores também encontraram empecilhos no uso dos aplicativos e afirmam a importância da adoção crítica de recursos tecnológicos.

2.4.3 Da Silva e Junior (2017)

Da Silva e Junior (2017) realizam uma RS através de um recorte de dissertações de mestrado que trabalharam com multiletramentos e tecnologias digitais. Os autores, no que se refere à RS, descrevem a importância de definir critérios precisos, rigorosos e sistemáticos para a seleção das fontes, além de definir perguntas de pesquisa pontuais e que exijam protocolos criteriosos para busca e seleção. Há também menção que este tipo de pesquisa é muito comum na área da saúde. Ademais, os autores definiram, primeiramente, o problema e a pergunta de

pesquisa. Depois, os autores definiram questões secundárias derivadas da pergunta principal. A seguir, no Quadro 3, os critérios de inclusão e exclusão da pesquisa foram apresentados. Por fim, as bases de dados e os descritores foram apresentados. De 289 trabalhos pesquisados, 18 chegaram após o último filtro da RS.

Quadro 3: *Metodologia da RS de Da Silva e Junior (2017)*

Critérios de inclusão	a) Estudos de conclusão de Mestrado sobre Multiletramentos com uso de tecnologias digitais; b) Estudos produzidos nos últimos seis anos (2011 a 2016); c) Dissertações escritas em Língua Portuguesa (Brasil); d) Dissertações disponíveis nas bases de dados selecionadas (Capes, Biblioteca digital brasileira de teses e dissertações); e) Estudos realizados na Educação Básica, Ensino Técnico, Tecnológico, Superior ou cursos livres, na modalidade presencial; f) Estudos empíricos enfocando práticas e experiências pedagógicas com multiletramentos e tecnologias digitais no ensino de línguas estrangeiras
Critérios de exclusão	a) Estudos que não contemplem os critérios de inclusão; b) Estudos de natureza teórica; revisão de literatura; estado da arte, pesquisa bibliográfica

Fonte: O autor.

Os autores, após isso, discutem, partindo de gráficos de estatísticas descritivas, os anos de publicações das dissertações, as grandes áreas das quais os trabalhos são provenientes, em quais repositórios as pesquisas foram encontradas, quais metodologias foram utilizadas em cada pesquisa, quais instrumentos foram utilizados para a coleta de dados, o nível de ensino focado por cada pesquisa, e, por fim, qual língua estrangeira foi alvo da investigação. Após a descrição do *corpus*, os autores respondem à pergunta de pesquisa principal, bem

como as secundárias, investigando quais práticas de multiletramentos com tecnologias digitais foram usadas na prática.

2.4.4 Franco *et al.* (2023)

Franco *et al.* (2023) realizaram uma RS do bilinguismo português/Libras no Brasil. Os autores partiram de pressupostos teóricos histórico-culturais vigotskyanos, compreendendo a surdez em uma perspectiva sociocultural e de diferença linguística, em vez de doença, e as teoria enunciativo-discursiva bakhtiana. Ao versar sobre a RS, os autores citam Galvão *et al.* (2018) para definir este tipo de investigação como secundário, uma vez que depende de publicações prévias, e citam algumas características essenciais para um bom desenvolvimento, como modos de identificação e seleção de estudos relevantes, análise crítica da metodologia dos estudos selecionados e a síntese dos resultados encontrados. Os autores desenvolveram a RS seguindo as etapas presentes no Quadro 4:

Quadro 4: Metodologia da RS de Franco et al. (2023)

Critérios de filtragem	a) Delimitação da base de dados: Scielo; b) delimitação das áreas de conhecimento: Ciências Humanas e Linguística, Letras e Artes; c) seleção de fontes de indexações públicas; d) seleção de termos e operadores lógicos; e) Seleção das datas de publicação (2015-2021).
Critérios de inclusão	a) Conter os termos: “bilinguismo, educação de surdos, libras, mediação didática, metodologia, língua brasileira de sinais e educação bilíngue”
Critérios de exclusão	a) Conter os termos “revisão sistemática, língua inglesa, língua de sinais americana-ASL, língua de sinais francesa-LSF.”

Fonte: O autor.

Após os procedimentos detalhados no Quadro 4, os artigos foram importados para o *software* Rayyan, e os 113 trabalhos encontrados foram lidos. Destes 113, 42 foram descartados por “população errada, desenho de estudo errado, contexto equivocado, tipo de publicação errada.” (FRANCO *et al.*, 2023). Cabe, aqui, um comentário: a presente pesquisa de doutorado se alinha com a visão de não aplicar juízos de valores nas pesquisas selecionadas para RS, e, em vez de descartar, preferiu-se comentar criticamente os dados para que pesquisas futuras possam ser mais robustas. Ademais, os 71 artigos restantes de Franco *et al.* (2023) foram classificados, em três tabelas, por ano de publicação, periódicos em que foram publicados, afiliação acadêmica, autoria, categoria analítica, público-alvo, área, recurso didático. Descrições das pesquisas entrelaçadas com discussões teóricas foram realizadas a cada apresentação de tabelas de dados, enriquecendo a apresentação da descrição quantitativa encontrada.

2.4.5 Azevedo e Bonfim (2019)

Azevedo e Bonfim (2019) realizaram uma RI em trabalhos de variação linguística no Brasil. As autoras se embasam metodologicamente em Ganong, Broome, Whittemore e Knalf (que possuem diversos trabalhos metodológicos publicados em anos como 1987, 1998 e 2005), e citam Souza, Silva e Carvalho (2010) para afirmar que a RI é a mais ampla das revisões, por permitir estudos experimentais e não experimentais, além de dados de literatura teórica e empírica. As autoras utilizam as seis etapas prescritas por Melo *et al.* (2011) para uma realização de RI, diferentes das aqui trazidas por Souza *et al.* (2010) e descritas antes da subseção x.1 deste capítulo. Aqui, as seis etapas são:

“A identificação do problema (que determina quais estudos serão incluídos), o estabelecimento de critérios de inclusão e exclusão de artigos (isto é, a seleção de amostra, que deve garantir a representatividade), a definição das informações que devem ser retiradas dos artigos previamente selecionados

(coleta dos dados), a análise das informações obtidas de maneira crítica, a interpretação dos resultados (onde serão comparados os dados evidenciados na análise com os existentes no referencial teórico) e por fim, a apresentação da revisão.” (AZEVEDO E BONFIM 2019, P.60)

Através do Quadro 5 é possível observar como Azevedo e Bonfim (2019) organizaram sua metodologia de pesquisa:

Quadro 5: Metodologia da RI de Azevedo e Bonfim (2019).

Critérios de inclusão	a) Conter os termos: “bilinguismo, educação de surdos, libras, mediação didática, metodologia, língua brasileira de sinais e educação bilíngue”
Critérios de exclusão	a) Conter os termos “revisão sistemática, língua inglesa, língua de sinais americana-ASL, língua de sinais francesa-LSF.”

Fonte: O autor.

Percebe-se que Azevedo e Bonfim (2019), mesmo não deixando claro o processo de filtragem, explicam detalhadamente os critérios de inclusão e exclusão.

2.4.6 Galdino e Oliveira JR (2023)

Galdino e Oliveira Jr. (2023) realizaram uma RI a respeito de trabalhos de prosódia e síntese de fala. Os autores se utilizaram da recomendação PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-analyses*): definição de uma pergunta norteadora, elaboração de critérios de exclusão e inclusão, síntese das informações dos estudos e apresentação da revisão. A pergunta norteadora definida pelos autores foi “Como a prosódia tem sido considerada em trabalhos que visam o aprimoramento da síntese de fala?”. A base de dados selecionada pelos autores foi o Google Acadêmico (*scholar.google*), porque, segundo os autores, é

um dos indexadores mais utilizado para busca de periódicos acadêmicos e uma base de dados com pluralidade de conteúdo. Os trabalhos pesquisados foram somente em português brasileiro, produzidos no Brasil, e com os seguintes operadores: (Prosódia OR entoação OR “frequência fundamental”) AND (“*text-to-speech*” OR TTS OR “síntese de fala” OR “síntese da fala”). Os autores incluíram artigos, monografias, teses, dissertações entre os anos de 2010 e 2021, e excluíram produções como livros e resenhas. Os autores também notaram que a metodologia utilizada foi uma adaptação de Ursi (2005, *apud* GALDINO e OLIVEIRA JR, 2023). A síntese da metodologia de Galdino e Oliveira Jr. (2023) pode ser visualizada no Quadro 6.

Quadro 6: Metodologia da RI de Galdino e Oliveira Jr. (2023):

Recomendação PRISMA	a) Definição de uma pergunta norteadora; b) Elaboração de critérios de exclusão e inclusão; c) Síntese das informações dos estudos; d) Apresentação da revisão.
Critérios de inclusão	a) Estar incluso da base de dados do <i>Google Scholar</i> ; b) ser filtrado pelos operadores lógicos escolhidos; c) ser um trabalho acadêmico como artigo, monografia, tese e dissertação.
Critérios de exclusão	a) Ser um trabalho acadêmico como livro, resenha e resumo; b) ser um trabalho repetido na busca

Fonte: O autor.

2.4.7 Bicudo (2014)

Bicudo (2014) traz importantes reflexões a respeito das contribuições da Metanálise para pesquisas qualitativas. Segundo a autora, as pesquisas qualitativas têm se sobressaído às pesquisas quantitativas em ciências humanas, uma vez que existe uma busca pela contextualização dos fenômenos ou das problemáticas envolvidas. Bicudo (2014) apresenta que as ciências exatas e da natureza se

pautaram em modelos provenientes da física e da matemática para encontrar soluções específicas para problemas, aplicando um rigor para comprovar estas soluções como certas. Esses estudos quantitativos passaram a incorporar ferramentas estatísticas e probabilísticas. Estudos de Metanálise se propuseram, então, a combinar esses resultados. O objetivo dos trabalhos de Metanálise seria, portanto, integrar dois ou mais trabalhos de uma área com o objetivo de síntese. Outrossim, as metanálises qualitativas visam esta síntese, porém com interpretações e análises baseadas em dados primários. Bicudo (2014) considera metanálises de forma geral importantes para pesquisas qualitativas por dois argumentos distintos: seja no sentido de aumentar a visibilidade e o impacto das pesquisas qualitativas em andamento, seja por atingir proposições analíticas mais elevadas e generalizar conhecimentos.

2.4.8 Barbosa *et al.* (2023)

Barbosa *et al.* (2023) realizaram uma Metanálise das produções científicas nas áreas da educação e linguagem. Esta pesquisa, segundo os autores, integra uma investigação maior e é um recorte dos dados. Para embasar sua Metanálise qualitativa, os autores utilizaram Sacardo e Sousa (2008) para levantar os trabalhos do corpus, e uma análise embasada em Megale (2015). Os pressupostos da Metanálise qualitativa foram avaliar as pesquisas, reconhecendo os conhecimentos produzidos, sem julgamentos, mas colocando em voga a teoria e os métodos dominantes no tema de estudo escolhido. Os exatos procedimentos metodológicos no sentido de termos de busca, utilização de operadores lógicos e bases de dados não ficaram explícitos no texto. As informações fornecidas pelos autores foram de que as pesquisas foram produzidas entre os anos de 2010 e 2015, que artigos, teses e dissertações foram selecionados e seus resumos, palavras-chave e títulos foram lidos para uma melhor filtragem, e cerca de 81 trabalhos foram selecionados (02 artigos, 22 teses e 57 dissertações).

2.4.9 Dias *et al.* (2015)

Dias *et al.* (2015) realizaram uma metanálise sobre o uso de memes (no sentido contemporâneo da palavra, utilizados na web 2.0). A pesquisa, de caráter bibliométrico, selecionou o Google Acadêmico como base de dados para procurar trabalhos. Foram selecionados artigos publicados entre os anos de 2005 e 2015, através da seleção da palavra-chave “meme”, e que esta palavra estivesse presente na primeira página do artigo. Houve um critério de inclusão que requeria que as pesquisas fossem publicadas no Brasil por pesquisadores brasileiros. Foram encontrados 23 artigos que haviam satisfeito os critérios da pesquisa, e foram realizadas análises referentes aos anos de publicação, locais de publicação e autores teóricos citados nos artigos.

2.4.10 Saldanha (2016)

Saldanha (2016) realizou uma revisão bibliográfica de metanálises em educação a distância. Torna-se nítida a diferença entre um trabalho de revisão e as metanálises propriamente ditas, no que se refere à ausência da metodologia criteriosa, justificada e exaustiva. O autor cita diversas metanálises realizadas com diferentes metodologias distintas, mas com o mesmo propósito. Dentre as pesquisas citadas estão, a cargo de exemplos, Ritzhaput *et al.* (2010), que compreendeu 517 trabalhos analisados através do método métrico de *co-word analysis* no tema de teleconferências em EaD e Lee *et al.* (2004), que investigaram 383 artigos publicados entre 1997 e 2002, examinando os principais tópicos de pesquisa, métodos e as tendências de citação em EaD. Saldanha (2016), ao final da primeira seção do artigo, nota que nem todos os trabalhos internacionais foram incluídos em sua análise, o que é uma limitação do trabalho. Convém, como citado na introdução da presente pesquisa, lembrar-se que é impossível desenhar um mapa do tamanho do mundo real. A reflexão e a lembrança que trabalhos são recortes é sempre importante. Ademais, Saldanha (2016) traz uma reflexão de Moore e Kearsley (2007) sobre a dificuldade de se realizar pesquisas documentais

na internet, uma vez que o pesquisador precisa tomar cuidado com as palavras-chave utilizadas para realizar as buscas. Este tópico de Saldanha (2016), é fundamental na presente tese, uma vez que, apenas em português, foram necessárias buscas para treinamentos fonéticos, perceptuais e perceptivos para chegar ao *corpus* de Saldanha (2016). Saldanha (2016) também afirma que trabalhar com poucos descritores pode ser perigoso, e uma boa busca para palavras-chave é requerida.

2.4.11 Chaves (2013)

Chaves (2013) realizou uma metanálise das vogais pretônicas no português brasileiro. Seu embasamento teórico citou, primeiramente, a linha de Norris e Ortega (2006), no sentido de que este tipo de trabalho ganhou o *status* de pesquisa na década de 1970 e que, através de uma síntese, busca descobrir perguntas como qual a natureza epistemológica e os procedimentos metodológicos das pesquisas da área investigada, a frequência e os tipos de estudos pesquisados, qual a finalidade da própria síntese da pesquisa. Chaves (2013) cita Cooper e Hedges (2009) para diferenciar trabalhos como “síntese de pesquisas”, “revisão de literatura” e “revisão sistemática”. Segundo os autores citados, o trabalho de revisão de literatura busca apenas pesquisar com base nos materiais anteriormente publicados, enquanto a síntese de pesquisas busca integrar uma pesquisa empírica para criar generalizações, identificar teorias, analisar criticamente a pesquisa e tentar resolver conflitos e identificar questões centrais para os futuros pesquisadores. Chaves (2013), então, se utiliza de Gene Glass (1976) para se referir à análise estatística do conjunto de dados individuais compilados, com a intenção de integrar as informações. A pesquisa de Chaves (2013), portanto, utiliza-se de uma síntese de pesquisa para aplicar o método meta-analítico. Neste trabalho, a característica crítica e sugestiva está presente. As questões orientadoras da pesquisa de Chaves (2013) foram:

- Selecionar trabalhos relacionados com o tema;

- Fazer a triagem dos trabalhos para a Metanálise;
- Identificar as abordagens recorrentes;
- Organizar as informações por categorias.

Os procedimentos aplicados, em seguida, foram:

- Extrair resultados das pesquisas analisadas;
- Realizar a Metanálise dos resultados;
- Identificar a característica das pesquisas.

A tese de Chaves (2013) traz mais uma importante contribuição para a pesquisa atual: a citação de Matt e Cook (2009), que afirmam que metanalistas, e amplo aqui para a presente Revisão Sistemática, não podem contar com a amostragem e a conclusão, uma vez que as amostras são limitadas e as inferências devem ser cuidadosas com base nos dados encontrados. Como um exemplo disto na presente tese, devo adiantar a análise de dados no que concerne à falta de trabalhos antigos na RS: geralmente, trabalhos de antes dos anos 2000 podem não estar digitalizados ou estarem mais suscetíveis e barreiras de *pay wall*, dificultando o acesso ao público.

De forma geral, nota-se, com base nesta subseção de revisão, a grande variedade metodológica presente nas pesquisas em linguística e fonética no que se refere aos trabalhos de revisões integrativas, sistemáticas e Metanálise. Os trabalhos são raros, nessas áreas (sendo muito mais comuns nas áreas da saúde) e sua maioria não segue exatamente os mesmos passos para chegarem às análises. No entanto, em todas essas pesquisas, nota-se a procura por um padrão da busca de o estabelecimento de perguntas de pesquisa e/ou hipóteses, a seleção de quais indexadores utilizar para compor o *corpus*, a fixação de critérios de inclusão e exclusão para selecionar trabalhos, e as diferentes formas de se analisar dados. Foi um processo interessante investigar todas as pesquisas brevemente descritas aqui, e mais as que não foram incluídas como revisão de literatura, pois foi possível observar como cada autor e autora selecionou seus métodos e os aplicaram, gerando trabalhos distintos, mas com premissas parecidas, de sintetizar e analisar criticamente uma quantidade de dados massiva.

Descritas algumas importantes pesquisas da área de TP e HVPT, as teses em TP escritas no Brasil por brasileiras, as teorias da percepção e modelos perceptuais e as diferentes metodologias em trabalhos de revisão no capítulo 2, busca-se apresentar os aspectos metodológicos da presente tese no capítulo 3.

3. METODOLOGIA

Este capítulo descreve qual a metodologia adotada no presente trabalho. Em primeiro lugar, na subseção 3.1 há uma discussão sobre como um novo título para a pesquisa foi eleito após a qualificação do doutorado. Uma discussão acerca da metodologia da revisão sistemática é tecida na subseção 3.2, seguida de uma descrição e apresentação de como o conjunto de dados da pesquisa está organizado na subseção 3.3.

3.1 Metodologia para o título da pesquisa

Como descrito brevemente na introdução e discutido com mais profundidade na próxima seção, 3.2, que trata da metodologia da revisão sistemática, o processo para encontrar os trabalhos que compõe esta RS foi árduo. Seria menos difícil, contudo, se as pesquisas em TP possuísem uma nomenclatura mais individualizante e não ambígua. Por sugestão da banca de qualificação e, posteriormente em consulta com a orientadora do trabalho e o manual de Milan e Kluge (2021), duas possibilidades foram consideradas:

- Treinamentos Perceptuais de sons de L2 (Perceptual training of the L2 sounds)
- Treinamentos Perceptuais de sons de línguas não-nativas. (Perceptual training of non-native sounds)

O resultado das buscas simples no google acadêmico e periódico CAPES já indicaram resultados mais refinados, sendo eles:

- (1) Português no *scholar.google* (treinamento perceptual sons L2): 684 resultados.
- (2) Inglês no *scholar.google* (*perceptual training L2 sounds*): 37000 resultados.
- (3) Português nos periódicos da CAPES (treinamento perceptual L2): 4 resultados.
- (4) Inglês nos periódicos da CAPES (*perceptual training L2*): 140 resultados.
- (5) Português no *scholar.google* (treinamento perceptual não-nativo): 144 resultados.
- (6) Inglês no *scholar.google* (*perceptual training non-native*): 35,6k resultados.
- (7) Português nos periódicos da CAPES (treinamento perceptual não-nativo): 0 resultados
- (8) Inglês nos periódicos da CAPES (*perceptual training non-native*): 163 resultados.

Nota-se, ao comparar os itens do google acadêmico, ou seja, 1, 2, 5 e 6, que os termos em inglês L2/*non-native* quase não apresentam diferenças em questões numéricas para as buscas, enquanto em português, L2 e não-nativos, os resultados para “treinamentos perceptuais sons L2” apresentam 4,75 vezes mais resultados em buscas simples. Os resultados foram similares em proporção ao número de buscas para os periódicos CAPES (0 e 4 e 144 e 163, em português e inglês, respectivamente), o que indica que, ao menos numericamente, as opções são equivalentes.

Apesar de L2 significar Segunda Língua, o termo pode ser usado para aprendizado de quaisquer línguas além da nativa de uma pessoa. O termo pode ser problemático para pessoas que estão aprendendo terceiras, quartas e mais línguas, assim como LE (ampliando o debate), que significa Língua Estrangeira, mas pode

ser uma língua do próprio país que um falante desconhece (como algum brasileiro falar alguma língua da área Tupi ou Macro-jê: são línguas que a maioria dos brasileiros não falam, mas seriam “estrangeiras”, por esta nomenclatura). Por sua vez, o termo “não-nativo” abarca ambas as definições sem cair nas questões discutidas.

O nome “Treinamento Perceptual de Sons Não-Nativos” (TPSN), ou “Perceptual Training of Non-Native Sounds” (PTNS) é consideravelmente mais longo do que simplesmente Treinamento Perceptual, porém é mais descritivo e específico para saber do que se trata o trabalho. Antes da qualificação, o título estava muito genérico “Revisão Sistemática de Treinamentos Perceptuais”. Este título inicial foi uma inspiração estilística baseada em Nishida (2012), com seu título sintético de tese e elogio a títulos curtos. O título atual ainda está, todavia, em um tamanho relativamente curto ao se comparar com outros trabalhos da área de Fonética e Fonologia.

3.2 Metodologia da Revisão Sistemática

Após estudo e leitura extensiva de diversos trabalhos bibliográficos na área da linguística, e, quando possível, especificamente nas áreas da fonética e fonologia (parcialmente descritos na seção 2.3), a leitura e o conhecimento necessário para delimitar a metodologia da revisão sistemática da presente pesquisa foram compilados. A seguir, encontram-se os detalhes do procedimento realizado:

- (1) O primeiro passo foi definir quais bases de dados seriam usadas na pesquisa. Optei por escolher duas bases de dados. A primeira foi o Google acadêmico (*Scholar.Google*), por abarcar uma quantidade praticamente infinita de pesquisas. A parte negativa é que, pelo número de pesquisas ser muito grande a cada busca, a filtragem precisa ser de uma qualidade muito alta. Por exemplo: houve a necessidade de filtrar mais de 500 artigos em somente uma das buscas avançadas desta tese. Como segunda base de dados, o portal de periódicos da CAPES foi eleito, por

ser uma ferramenta oficial de um órgão governamental brasileiro, e possuir uma qualidade altíssima de periódicos indexados (principalmente, mas não só, brasileiros).

- (2) Optou-se por analisar somente Artigos publicados em revistas científicas revisadas por pares. Teses e dissertações, além de não serem publicações aprovadas por *corpus* editorial de revistas, são muito extensas para análises profundas. Além disso, geralmente os conteúdos de teses e dissertações são publicados em formato de artigos antes ou depois de suas defesas.
- (3) A seguir, foi necessário escolher as palavras-chave de pesquisa, e os refinamentos das buscas avançadas. As buscas avançadas foram necessárias para a viabilidade da pesquisa. Por exemplo: ao se pesquisar “*Perceptual training*” no Google Acadêmico, foram encontrados 2,240,000 resultados em 0,13s de pesquisa. Dentre os dois milhões de resultados encontrados, estavam trabalhos de áreas como a fisiologia do exercício, fisioterapia e neurologia, que não se configuram no escopo da atual pesquisa.
 - a. As palavras-chave escolhidas para o trabalho foram sete no total, o que gerou quatorze buscas (sete em cada base de dados): “Treinamento Fonético”, “Treinamento Perceptual”, “Treinamento Fonológico”, “Treinamento Perceptivo”, “*Phonetic Training*”, “*Phonological Training*”, “*Perceptual Training*”. É importante destacar que as expressões “Treinamento fonológico” e “*Phonological Training*” não geraram nenhum trabalho incluído nesta pesquisa, uma vez que os termos são usados exclusivamente para a área de fonoaudiologia. Trabalhos com sujeitos possuidores de distúrbios da fala são importantes, mas são foco da fonoaudiologia.

Uma questão interessante que deve ser discutida com mais profundidade tem relação à nomenclatura dos trabalhos neste ramo. Os termos “perceptivo” (apontado por um membro da banca de qualificação como possivelmente uma

tradução errônea do espanhol, mas que mesmo assim se encontra em pesquisas, como Lorencio, 2021), “fonético” e “perceptual” podem ser encontrados em várias pesquisas com objetivos semelhantes: estabelecer um regime de treinamento para melhorar a percepção da fala de uma língua estrangeira, que pode contar ou não com alta variabilidade de informantes, ter ou não *feedback*, possuir ou não testes para verificar a retenção do conhecimento, possuir ou não segmentos ou suprasegmentos não testados etc. Em inglês, os termos “*perceptual training*” e “*phonetic training*” também podem ser utilizados para designar a mesma coisa. Apesar de não ser um problema em si, isso pode dificultar a localização de trabalhos por parte de pesquisadores da área. Hipoteticamente, caso uma pessoa leiga faça suas buscas utilizando apenas um dos termos, uma grande variedade de trabalhos ficará de fora de seu radar.

b. As pesquisas avançadas realizadas foram:

- Google Acadêmico: Conter: Palavras exatas / Presentes no Título.
- Portal CAPES: Conter as palavras em: Título e Assunto.

Dessa forma, foi possível reduzir o número de trabalhos para 2744 artigos.

Foram lidos títulos e resumos de 2744 trabalhos antes do próximo passo, definir os critérios de inclusão.

c. Os critérios para inclusão foram três:

- Serem trabalhos da área de fonética/fonologia sobre treinamento perceptual/perceptivo/fonético, pesquisados em português e inglês;
- Serem indexados nas bases de dados do Google Acadêmico e Portal CAPES.
- Serem de acesso gratuito/aberto, para que toda comunidade linguística tenha acesso às pesquisas que aqui estão sendo analisadas sem barreiras de *paywall* ou dependência de logins institucionais.

(4) Além disto, quatro critérios de exclusão foram aplicados nos 2744 artigos:

a. Excluir artigos repetidos;

- b. Excluir monografias, dissertações e teses;
- c. Excluir pôsteres e trabalhos apresentados em eventos.;
- d. Excluir resenhas/e ou resumos e/ou compilações (apenas trabalhos empíricos e com dados inéditos foram incluídos)

Destes 2744 trabalhos, após a leitura de título e resumo e aplicação dos critérios de exclusão, restaram 230 artigos.

(5) Após o download e leitura de cada um desses 230 artigos, o número final de trabalhos que passaram por todos os critérios de inclusão e exclusão foram 117. O número foi de 230 para 117 pois muitos artigos baixados estavam somente disponíveis em partes (por exemplo: a metodologia e os resultados estavam disponíveis de forma gratuita, mas o resto do trabalho exigia pagamento), ou por estarem repetidos, mas em outras línguas, ou por simplesmente estarem corrompidos ou serem teses /dissertações disponíveis em sites de periódicos. O número final de artigos selecionados para a RS foi, portanto de 117.

Cabe notar, também, que, embora a terapia fonológica não seja o foco da presente pesquisa, e todos os resultados das buscas por “treinamento fonológico” e “*phonological training*” foram descartadas da RS por pertencerem ao campo da fonoaudiologia, é importante mencionar que o treinamento também possui relevância nesse campo (BOWEN, 1998). Esse autor fornece as bases para um *workshop* de terapia fonológica. Existe um *framework* específico a ser seguido (identificação de um ambiente fonológico regular e um anormal, pré-testes, pós-testes e treinamento) para melhorar o sistema fonológico da criança. Bowen (1998) relata que as sessões podem ser realizadas uma vez por semana por 50 minutos, cada sessão, durante dez semanas com uma pausa de dez semanas. Além disso, os exercícios de casa são fundamentais. Entre alguns procedimentos de intervenção estão a distinção de pares mínimos, confronto de homofonia, inovações lexicais e gramaticais e conscientização da relação fonema-grafema. É interessante observar como diversas pesquisas em ambas as áreas, de fonoaudiologia e

fonética, seguem desenvolvimentos ou metodologias parecidas, ainda que para fins distintos.

Uma vez descrita a revisão literária sobre TP e feita a distinção entre os treinamentos fonéticos/perceptuais/perceptivos dos treinamentos fonológicos, represento a pergunta de pesquisa norteadora da presente tese: “Qual é o estado da arte da área de TP e quais podem ser novos caminhos de pesquisa?”

3.3 Organização e descrição do conjunto de dados da RS

Após a filtragem dos artigos, houve um processo de organização interno. Os artigos foram organizados dentro de uma planilha digital do *software Microsoft Excel* e foram organizados internamente por 16 categorias que nomeei “descritivos”. Estas categorias serviram, a posteriori, para ajudar a alcançar os objetivos específicos, que buscam investigar os sujeitos, os experimentos e as teorias e análises dos TPs, mas suas análises (de cada descritivos) são individuais e autorais (o que significa que outro pesquisador poderia ter elegido outros descritivos).

Os 16 descritivos são:

- (1) Ano de publicação (divididos em cinco categorias: Antes de 1990, 1990-2000, 2000-2010, 2010-2020, 2020- presente)**;
- (2) Língua materna**;
- (3) Língua-alvo**;
- (4) Foco Fonético (Segmentos ou suprasegmentos do treinamento)**;
- (5) Existência de Grupo Controle*;
- (6) Número de participantes no Grupo Controle*;
- (7) Número de participantes no Grupo Experimental*;
- (8) Número de Sessões**;
- (9) Duração do período das sessões**;
- (10) Existência e tipo de feedback**;
- (11) Existência de testes de generalização**;
- (12) Existência de testes de retenção**.
- (13) Adoção do paradigma da alta variabilidade***;

- (14) Adoção de modelos teóricos perceptuais***.
- (15) Uso da tecnologia**;
- (16) Tarefas de treinamento**

* Descritivo alinhado ao objetivo específico 1 (sujeitos)

** Descritivo alinhado ao objetivo específico 2 (experimentos)

*** Descritivo alinhado ao objetivo específico 3 (teorias e análises)

Considerando todas as possibilidades de informação de todas as pesquisas por estes filtros, foram contabilizados 2106 dados a partir de 117 trabalhos, uma vez que os descritivos 9 e 15 foram subdivididos em dois para maior detalhamento. A planilha com todas as descrições de cada um dos 117 trabalhos pode ser encontrada no Apêndice 1 desta pesquisa. Cada tópico listado acima será apresentado nesta seção e mais bem descrito e discutido nos resultados.

É possível observar a organização subsequente da pesquisa no quadro através da figura 2:

Figura 2: Uma secção da planilha de dados.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000
Category	Ano	L1	L2	Segmento	Particip	Control?	Control N	Num	ses	Tempo	Post/Ref	HVPT	(usc	Feedback	Gen	Theory																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							

indica o ano de publicação do artigo, com um código de cores organizado por décadas. As colunas seguintes se referem às línguas iniciais e línguas-alvo dos TP. A seguir, as colunas apresentam o foco fonético (segmentos ou suprasegmentos) o número de participantes nas pesquisas, a existência, ou não de grupo controle, o número de participantes no grupo controle, o número de sessões do treinamento perceptual, o tempo (em semanas) da duração do período de treinamento, sem contar os testes de retenção postergada, a existência de testes de retenção e o tempo em meses que ele ocorreu após o fim do período de treinamento. Por fim, as últimas colunas indicam a adoção, ou não, de HVPT, se existisse e qual é o tipo de feedback, se retoma a teoria e quais teorias o artigo retoma. O guia completo de leitura da planilha está no anexo da tese, após o capítulo 5 de considerações finais.

Ao se discutir o descritivo 1, referente aos anos de publicação dos trabalhos, tem-se a seguinte distribuição, em um total de 117 trabalhos:

- Antes de 1990 (1 trabalho).
- 1990-2000 (3 trabalhos).
- 2000-2010 (20 trabalhos).
- 2010-2020 (59 trabalhos).
- Depois de 2020 (39 trabalhos).

Para o descritivo 2, referente à L1, é importante destacar que, dentre estes 117 trabalhos, nem todos estavam em língua portuguesa ou inglesa, os idiomas utilizados para a pesquisa. Um estava em espanhol, um em francês e dois em coreano. Utilizei-me de inteligência artificial para traduzir os artigos em coreano, e eu mesmo traduzi os artigos em espanhol e francês. Além disso, três trabalhos partiram de mais de duas línguas materna, sete de diversas L1, enquanto a maioria dos outros (n=107) partiu apenas de uma. As descrições mais detalhadas e discussões concernentes às línguas maternas podem ser encontradas no capítulo de resultados.

Para o descritivo 3, referente às línguas-alvo dos treinamentos, nota-se que 65.48% (n=74) do total dos trabalhos focava o inglês, e apenas 34,51%(n=43) no

restante das línguas. A apresentação detalhada dos dados pode ser encontrada no capítulo de resultados.

Com relação ao descritivo 4, que abarcou o foco fonético do treinamento, 80,2% dos trabalhos tinham o foco na fonética segmental, ou seja, sons isolados como /p/ ou /t/, enquanto os outros 19,8% possuíam como alvo a fonética suprasegmental, ou seja, mais de um fonema, ou aspectos como prosódia, ritmo ou tom de fala (muito comum em trabalhos com línguas tonais), por exemplo.

- A descrição do foco do treinamento:
 - Consoantes [n=51] 43,59%
 - Vogais [n=34] 29%
 - Suprasegmentais [n=24] 19,8%, incluindo *lexical stress*, *tonality*, *blending*.
 - Vogais e consoantes [n=9] 7.7%

Para o descritivo 5, tem-se que um total de 47 trabalhos (total n=117), 40,17%, não possui grupo controle. A média de participantes nos grupos-controles foi de 16,27 pessoas, enquanto nos grupos experimentais foi de 30,29 pessoas.

As descrições mais detalhadas dos descritivos 6 e 7, que tratam, respectivamente, de números de participantes no grupo controle e no grupo experimental, podem ser encontradas no capítulo de resultados.

O descritivo 8 tratou do número de sessões presentes nos TPs. Ao se analisar o total de pesquisas (n=117), vê-se que a média de sessões foi de nove sessões (8,99), e a mediana de seis sessões. Os dados mais detalhados deste descritivo são apresentados e discutidos no capítulo de resultados

Quanto ao descritivo 9, que tratava da média da duração total do período de treinamento, tem-se a média de tempo total de duração, que foi de 22,5 dias, e a mediana de 14 (baseados em n=75 trabalhos que forneceram a informação). Estes resultados são descritos mais detalhadamente e discutidos no capítulo de resultados. A média da duração das sessões foi de 33min, sendo n=72.

Com relação ao descritivo 10, relacionado ao *feedback*, presente nas pesquisas, tem-se que a maioria (67,5%) das pesquisas forneceu algum tipo de *feedback* aos participantes do regime de treinamento. Estes dados estão mais detalhados e discutidos no capítulo de resultados.

O descritivo 11, referente aos trabalhos que possuíam testes de generalização para itens novos (estímulo, locutores, generalização para fala) pelos participantes, conta com os seguintes resultados: 36 trabalhos (30,77%, $n=117$) realizaram testes de generalização. Os tipos de testes de generalização mais comuns foram: novos falantes, novos segmentos e para a produção de fala. No geral, os testes para novos falantes buscavam verificar se a alta variabilidade dos *inputs* permitiria aos informantes generalizarem seus conhecimentos para outras vozes diferentes onde houve generalização para novos segmentos/suprasegmentos, foi verificado também se o treinamento melhorava estes aspectos linguísticos não treinados. Já os testes de generalização para a fala, buscavam verificar se um treinamento perceptual poderia provocar mudanças na produção da fala. Cada um destes tipos de pesquisa com generalização será descrito e discutido com mais profundidade no capítulo de resultados.

Sobre o descritivo 12, que se refere à existência de testes de retenção, 24 estudos ($n=117$, 20,51%) possuíam testes de retenção postergada do conhecimento, ou seja, um ou mais pós-testes aplicados semanas ou meses após o final dos experimentos de treinamento. A média total de tempo para os testes após o fim do treinamento perceptual foi de 5,83 meses. No entanto, o trabalho de Bettoni (2022) investigou os participantes de um TP 12 anos após o treinamento perceptual, o que mais que dobrou a média de todos os trabalhos. Se os trabalhos forem contabilizados sem este dado de tempo, a média em meses, após o TP que os experimentos postergados foram realizados foi de 2,78 meses. A mediana, para ambos os casos, foi de 2 meses.

Com relação ao descritivo 13, 54 trabalhos ($n=117$), ou 46,15%, adotam o treinamento fonético de alta variabilidade (HVPT). As implicações e premissas de trabalhos com alta e baixa variabilidade serão discutidas no capítulo de resultados.

O descritivo 14, que tratou da adoção de modelos teóricos, foi, durante a fase de leitura e estudo teórico sobre TP da presente pesquisa, o foco da análise. Os dados sobre as adoções e discussões teóricas de modelos como o SLM, PAM e LNM são apresentados com mais detalhes no capítulo de resultados.

O descritivo 15 buscou verificar quais os *softwares* utilizados pelos autores das pesquisas da RS. Apenas um trabalho, o mais antigo de 1968, utilizou-se de fitas cassetes e papel e caneta explicitamente. Porém, como está mais detalhado na subseção, praticamente metade dos autores não deixou claro quais programas computacionais foram utilizados para realizar o treinamento.

O descritivo 16, último da pesquisa, investigou quais as principais tarefas presentes nos TPs: quase 85% de todos os trabalhos adotaram tarefas de identificação (ID), sejam elas acompanhadas de tarefas de discriminação (DIS) ou não.

Um resumo das diferentes organizações de dados da Revisão Sistemática pode ser observado no Quadro 7:

Quadro 7: Organizações de dados da Revisão Sistemática

Descritivo	Organização
1 Ano de publicação	Dividir em 5 períodos (Antes de 1990, 1990-2000, 2000-2010, 2010-2020, 2020- presente)
2 Língua materna	Identificar L1 do trabalho
3 Língua-alvo	Identificar L2(s) do trabalho
4 Foco Fonético	Identificar os alvos fonéticos (segmentais ou suprasegmentais)
5 Existência de Grupo Controle	Identificar existência de grupo controle (o número 0 indica que não houve grupo controle, o número 1 indica que existiu grupo controle)
6 Número de participantes no Grupo Controle	Identificar, em números, quantos participantes estavam neste grupo.

7 Número de participantes no Grupo Experimental;	Identificar, em números, quantos participantes estavam neste grupo.
8 Número de Sessões	Identificar, em números, quantas sessões de treinamento foram realizadas.
9 Duração do período das sessões	Identificar, em número de dias, quanto durou no total o treinamento (excluindo testes de retenção). A subseção 9.2, que expande a análise, investigou a duração em minutos das sessões em si do treinamento.
10 Existência e tipo de <i>feedback</i>	Identificar se houve <i>feedback</i> e, caso sim, qual foi ele.
11 Existência de testes de generalização	Identificar se houve testes de generalização, e, caso sim, para o quê.
12 Existência de testes de retenção	Identificar se houve testes de retenção, e após quanto tempo (em meses) do período de treinamento.
13 Adoção do paradigma da alta variabilidade	Identificar se houve adoção de HVPT =1, o u não, =0.
14 Adoção de modelos teóricos perceptuais	Identificar se houve adoção de modelos teóricos perceptuais, e quais modelos foram implementados.
15 Uso da tecnologia	Identificar quais <i>softwares</i> são utilizados nos TPs (uma vez que 116/117 utilizaram <i>hardwares</i> de computadores, tablets ou afim, o foco foi incidido no <i>software</i> ,
16 Tarefas de treinamento	Identificar quais são as tarefas (ID, DIS ou outras) presentes nos TPs

A seguir, encontram-se os resultados e discussões do trabalho.

4. ANÁLISE E DISCUSSÃO DOS RESULTADOS

Conforme descrito na Metodologia do trabalho, ao descrever o processo de organização dos dados da RS, os trabalhos foram subdivididos em 16 descritivos, que serão analisados e discutidos com mais profundidade neste capítulo.

Os descritivos foram:

- (1) Ano de publicação (divididos em cinco categorias: Antes de 1990, 1990-2000, 2000-2010, 2010-2020, 2020- presente);
- (2) Língua materna;
- (3) Língua-alvo;
- (4) Foco Fonético (Segmentos ou suprasegmentos do treinamento);
- (5) Existência de Grupo Controle;
- (6) Número de participantes no Grupo Controle;
- (7) Número de participantes no Grupo Experimental;
- (8) Número de Sessões;
- (9) Duração do período das sessões;
- (10) Existência e tipo de *feedback*;
- (11) Existência de testes de generalização;
- (12) Existência de testes de retenção.
- (13) Adoção do paradigma da alta variabilidade;
- (14) Adoção de modelos teóricos perceptuais.
- (15) Uso da tecnologia
- (16) Tarefas de treinamento

A análise e os discussão de cada descritivo se encontram nas próximas seções.

4.1 Ano de Publicação

Esta subseção investiga os resultados observados do descritivo 1, sobre anos de publicação. Dos trabalhos que passaram pelos critérios de inclusão e

exclusão na presente revisão sistemática, o Quadro 8 retoma os dados da quantidade de trabalhos incluídos na RS publicados por cada década.

Quadro 8: Distribuição do ano de publicação dos trabalhos da RS

Décadas de Publicação	Número de trabalhos na RS
Antes de 1990	1
1991-2000	3
2001-2010	20
2011-2020	59
2021-2024	34

Fonte: O autor.

A seguir, apresentam-se algumas hipóteses e análises referentes à distribuição irregular dos anos de publicação da pesquisa.

Primeiramente, nota-se a ausência, em grande medida, de trabalhos publicados antes dos anos 2000, e principalmente, antes da década de 1990. O trabalho mais antigo desta RS é o de Mueller e Niedzielski (1968), e existem muitos outros trabalhos das décadas de 1970 e 1980 (que não passaram nos critérios de inclusão e exclusão da presente RS) que são basilares para a área, já citados no embasamento teórico do presente trabalho, como Jamieson e Morosan (1986). Atribuo dois possíveis fatores à baixa representatividade de trabalhos mais antigos nesta RS:

I) Os trabalhos iniciais da área, no século XX, não são brasileiros e nenhum foi escrito em português e inglês, e, talvez por isso, não foram digitalizados e/ou indexados com tanta facilidade nos mecanismos de buscas que foram utilizados na presente pesquisa (portais de periódicos do CNPQ e Google Acadêmico). Duas pesquisas em coreano foram incluídas na RS, mas são exceções. Conforme mencionado anteriormente, o primeiro trabalho de TP no Brasil foi Nobre (2007), desenvolvido na UFSC. A área de pesquisa ainda era pequena e foi lentamente crescendo nos programas de pós-graduação no Brasil, principalmente na região sul.

II) Aliado ao fator número I, estes artigos mais antigos podem ter sido mais difíceis de serem encontrados de forma gratuita, em revistas que podem ser facilmente acessadas por qualquer pessoa ou mesmo por logins institucionais. Revistas importantes nas áreas de linguística e fonética/fonologia cobram valores exorbitantes para a leitura (ou acessos fechados que os tornam inacessíveis).

O embasamento teórico da pesquisa apresentou pesquisas históricas para o desenvolvimento da área, e é algo positivo ter uma maioria de trabalhos de 2010 até a data da realização da RS (início de 2024).

Outro ponto de vista interessante que concerne às datas de publicação destes trabalhos é observar que a área está crescendo e em efervescência. 59 trabalhos incluídos foram da década passada, e, tendo realizado o trabalho de filtragem para a RS em janeiro e fevereiro de 2024, mais de 34 trabalhos de 2021 até o início de 2024 foram incluídos no corpus. Esta tendência de publicação mostra que a década atual pode ultrapassar a década passada ainda mais no número de trabalhos publicados. Um trabalho de 2024 muito robusto, defendido na Universidade Federal do Rio Grande do Sul, precisou ficar de fora da RS, justamente por ter sido defendido pouco tempo depois da realização da pesquisa e por não estar em formato de artigo ainda. Essa pesquisa (Guaer, 2024), encontra-se descrita na seção 2.2.

Através do estudo das datas de publicação, é possível observar algumas tendências de tipos de estudos realizados. Por exemplo: os artigos incluídos na RS que estudaram o par /r-l/ da língua inglesa por aprendizes de brasileiros foram encontrados para os anos em ordem cronológica: 1997, 2000, (3x) 2005, 2006, 2007, 2016, 2019 e 2021. Observa-se que há demanda e estudos que investigam este par na década de 2020, mas eles se concentraram em décadas passadas. Os 13 trabalhos que estudaram tons de línguas como cantonês e mandarim, por sua vez, até possuem representantes em décadas passadas (2001, 2005 e 2003), mas os outros 10 são de anos entre 2010 e 2024, demonstrando uma emergência em trabalhos com este enfoque: 2010, (2x) 2013, 2014, (2x) 2016, 2018, (2x) 2019, 2021). Um caso interessante é observar os três trabalhos que estudaram *stress*,

lexical stress ou *connected speech*: todos os trabalhos são relativamente recentes ocorrendo nos anos de 2017, 2018 e 2023. Os 10 trabalhos com foco específico no VOT também estiveram representados em várias décadas e com mais publicações mais recentes nos anos de 2007, 2015, 2017, (2x) 2019, (2x) 2020, (2x) 2022, 2023.

Assim como apresentado no item 4.15, que trata da presença de tecnologias em pesquisas de TP, é importante ressaltar que, após os anos 2000 e 2010, houve mais acesso às tecnologias digitais acessíveis e/ou gratuitas por grande parte da comunidade científica da linguística através, respectivamente, de programas como o Praat (Boersma, 2001) e TP (Rauber *et al.*, 2012).

Reitera-se, para estas análises, que as ferramentas de pesquisa e os critérios de inclusão e exclusão podem ter grande peso aqui. O número de trabalhos totais da pesquisa poderia ser maior caso fossem escolhidos mais indexadores, ou fossem incluídas teses e dissertações (preferencialmente não publicadas, para que trabalhos repetidos não fossem poluir a RS).

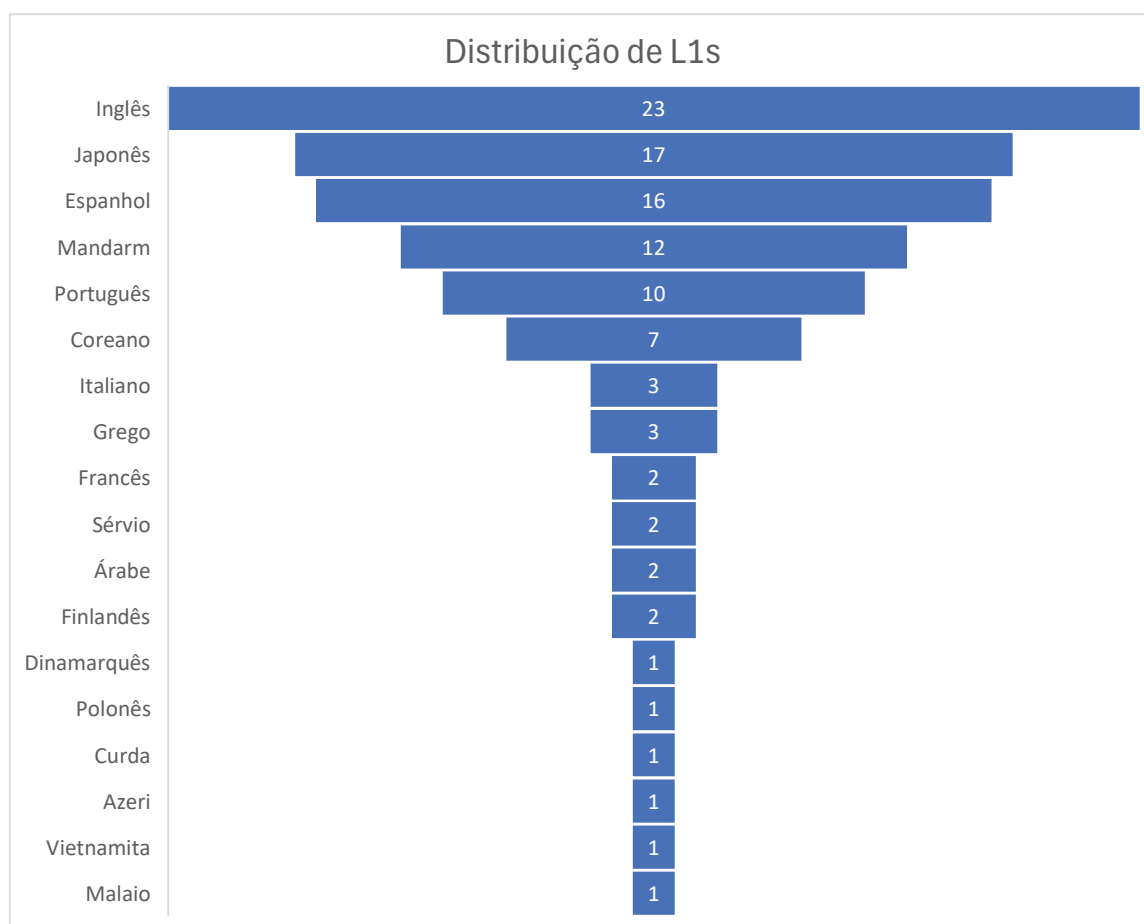
4.2 Língua Materna

Conforme descrito na subseção 3.3. ao discutir o descritivo 2, sobre L1, as línguas maternas participantes dos 117 trabalhos da pesquisa foram observadas. A seguir, descrevo as L1 dos trabalhos do corpus.

- Dos 107 trabalhos com uma L1: Inglês [n=23]; Japonês [n=17]; Mandarim [n=12]; Espanhol (n=16, sendo n=10 bilíngues espanhol/catalão) ; Português europeu/brasileiro [n=10]; Coreano [n=7] Italiano [n=3]; Grego [n=3]; Francês [n=2]; Sérvio [n=2]; Árabe [n=2]; Finlandês [n=2]; Dinamarquês [n=1]; Polonês [n=1]; Língua Curda [n=1]; Língua Azeri [n=1]; Vietnamita [n=1]; Malaio [n=1].

Estes dados podem ser observados através do gráfico 1:

Gráfico 1: Distribuição de L1.



Fonte: O autor.

Contabilizando-se as línguas inglesa, japonesa, espanhola, coreana, portuguesa e o mandarim apenas, tem-se 73,65% de todos os trabalhos apenas partindo destas línguas, com todas as outras cobrindo apenas 27,35% dos trabalhos. Investigar como outras línguas interagem entre si com treinamentos e metodologias diferentes, como trazem Alves e Kampff (2020), é importante para se ter noções de como falantes de diferentes linguagens desenvolvem suas percepções de fala para outras línguas além das mais recorrentes.

Nas décadas iniciais dos estudos na área de TP, conforme trazem Bradlow *et al.* (2008), os trabalhos mais comuns eram relacionados a falantes estadunidenses, nativos de inglês, sendo treinados para diferenças perceptuais entre /r, l/. Aos

poucos, linguistas começaram a aplicar novos treinamentos para outras linguagens, segmentos, ou fatores como tons em línguas tonais. Os 10 trabalhos contabilizados que partiam da língua portuguesa somaram português europeu e brasileiro, e não chegam nem a 9% do total de trabalhos incluídos na RS. É importante que pesquisadores brasileiros utilizem nativos do português para mais treinamentos, e, conforme será apresentado na subseção a seguir, não somente para a língua inglesa. É preciso ir além, testar novas possibilidades, partir de novas línguas

Trabalhos como Iverson e Evans (2009), por exemplo, partiram de duas diferentes línguas maternas para o aprendizado de inglês. Os autores treinaram ambos espanhóis e alemães, levando em conta os inventários fonéticos de suas línguas maternas. Este tipo de pesquisa pode ser interessante por comparar dois universos diferentes com a mesma metodologia de treinamento, e observar como os subsistemas linguísticos dos participantes de cada língua podem ser alterados. Nas pesquisas de Iverson e Evans, os falantes alemães apresentaram, com o mesmo número de sessões, mais facilidade que os falantes espanhóis, mas essa diferença foi mitigada com mais sessões para os espanhóis (este trabalho foi descrito com mais detalhe no capítulo 2 de revisão bibliográfica). Nota-se que, no pré-projeto de doutorado para a Universidade Internacional da Catalunha, era nossa ideia também comparar falantes nativos de línguas diferentes (brasileiros nativos de PB e catalães nativos de catalão e espanhol).

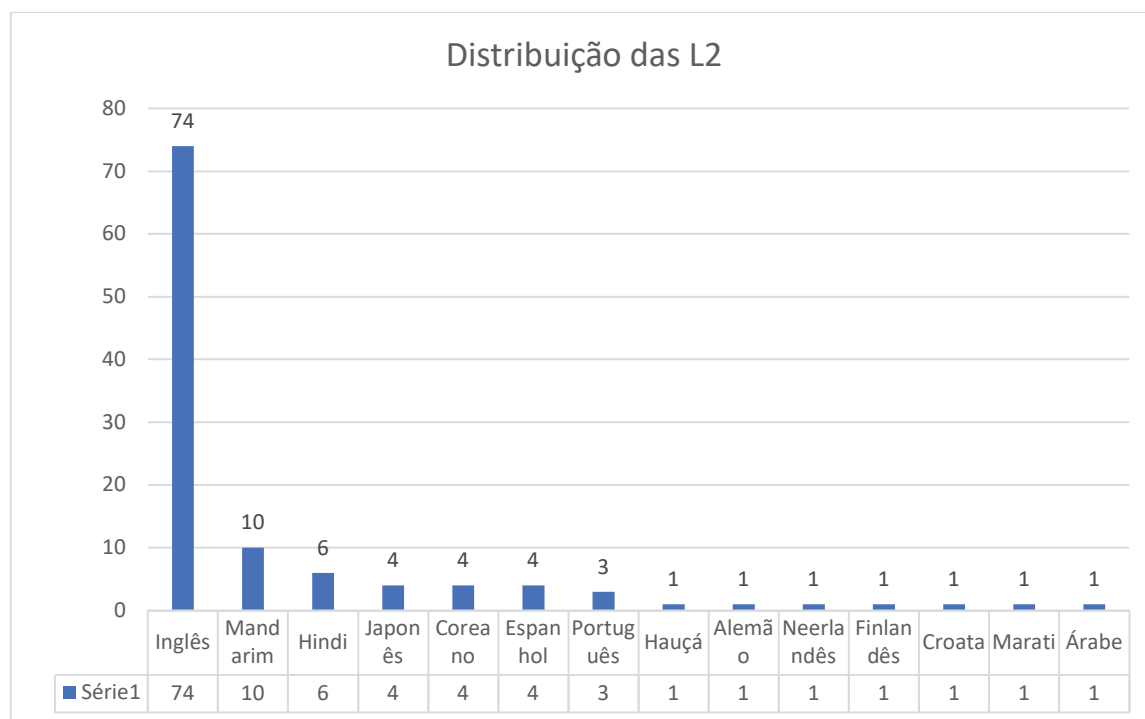
4.3 Língua-Alvo

O descritivo 3 se refere às L2 dos TPs. Assim como na subseção anterior, as línguas-alvo dos TP também foram analisadas com profundidade.

- Seguem os descritivos das principais L2 foco de treinamento (n=105): Inglês [n=74]; Mandarim [n=10]; Hindi [n=6]; Japonês [n=4]. Coreano [n=4]; Espanhol [n=4]; Português [n=3]. Hauçá [n=1]; Alemão [n=1]; Neerlandês [n=1]; Finlandês [n=1]; Croata [n=1]; Marati [n=1]; Árabe [n=1]

É possível observar a distribuição das L2 através do gráfico 2.

Gráfico 2: Distribuição das L2:



Fonte: o autor.

Caso a informação de que mais de 70% das pesquisas partiram das mesmas línguas maternas pode ter parecido exagerada, a distribuição de pesquisas para línguas estrangeiras é ainda mais. A segunda língua com maior representatividade, mandarim, representou 10 pesquisas, um 8,5% do total. A primeira, a língua inglesa, ficou com 63,25% de todo o foco. Sabe-se que o inglês é uma língua global, utilizada como língua franca e falada por bilhões de falantes, nativos e não nativos (Bernardo-Silva, 2021), porém é importante ter em mente que outras línguas são importantes para questões comerciais locais e por questões culturais. Por exemplo, na própria América do Sul: todos os países, com exceções de Suriname, Guiana e Guiana francesa, tem o espanhol como língua oficial e/ou principal. O Brasil, uma das maiores economias do mundo e situado na América do Sul, possui centenas de milhões de falantes de português. Documentos oficiais como os Parâmetros Curriculares Nacionais, PCNs, e a Base Nacional Comum Curricular, BNCC, podem

variar conforme as orientações de ensino de língua estrangeira, mas não deve ser esquecido que o inglês não é a única língua do mundo, mesmo que seja necessário traduzir este parágrafo o inglês no futuro para que mais pessoas possam refletir sobre isso

A língua espanhola possui uma proximidade lexical com o português, porém o aprendizado desta língua não deve ser tomado como certo. Existem diferenças fonológicas, falsos-cognatos, entre outras questões que podem ser estudadas. Milan (2019), por exemplo, estudou a percepção e produção de heterotônicos por brasileiros aprendizes de espanhol. Palavras como *democracia* possuem sílabas tônicas diferentes em português e espanhol, e casos como estes podem ser interessantes para se pesquisar. Existem instituições públicas, como a Universidade Tecnológica Federal do Paraná, que contam com programas de ensino de língua portuguesa para falantes de outras línguas (Português como Língua Estrangeira, Português para Falantes de Outras Línguas, Português para migrantes). Estes ambientes podem ser ótimas fontes de participantes que possam estar interessados em passarem por regimes de treinamento.

Outro fator interessante de ser discutido é que diversos países no mundo todo presenciam imigração para suas terras: Segundo o site [GOV.BR](https://gov.br), acessado em 15/02/2025, por exemplo, o Brasil registrou a entrada de 1,7 milhões de habitantes novos entre 2010 e 2024, principalmente de países próximos como Venezuela, Haiti e Bolívia. Ora, com este grande fluxo de imigração e grande quantidade de cursos de português para migrantes/imigrantes/português para outras línguas, estabelecer treinamentos fonéticos como auxílio na aquisição da língua portuguesa pode ser uma alternativa para aumentar o número de L2 no “cardápio” de possibilidades de TP. Ademais, em muitos lugares, a própria população local faz cursos de línguas como o espanhol e o crioulo haitiano para melhor comunicação com a população de imigrantes.

Conforme apontado por um membro da banca de qualificação, o interesse entre o inglês como L2 parece ligado ao comércio internacional e a comoditização

do ensino de línguas: este discurso está muito presente na mídia em se tratando de aprender inglês, por exemplo.

4.4 Foco fonético

Como já descrito brevemente ao se discutir o descritivo 4, referente ao ano de publicação, é possível observar o que se está ou se estava pesquisando dependendo do ano de publicação.

A imensa maioria das pesquisas (80,2%) foca os segmentos, e não fatores suprasegmentais, como *sentence stress* e tons. Muitas pesquisas que não enfocam seguimentos estudam o treinamento de tons. Entretanto, o número de línguas tonais que existem é enorme, e as pesquisas encontradas na RS são somente para línguas como o chinês (mandarim e cantonês). O prestígio e a quantidade de falantes destas línguas podem explicar a abundância de trabalhos com dialetos do chinês, mas não explicam a falta de pesquisa em outras línguas tonais como línguas amazônicas ou vietnamita. Na RS, encontra-se somente uma pesquisa que partem do vietnamita (Hwang *et al.* 2023), sendo os falantes treinados para coreano, mesmo com o país possuindo quase 100 milhões de habitantes e um PIB de mais de 400 bilhões de dólares. Simon *et al.* (2021) afirmam que o estudo sistemático de um amplo escopo de contrastes em línguas além do inglês são necessárias para entender o verdadeiro impacto de métodos de treinamento como o HVPT. Questões suprasegmentais como investigadas na pesquisa de Brawerman-Albini *et al.* (2017) são fundamentais. Como as autoras apresentam, fonética é frequentemente esquecida nas aulas de inglês, e as questões de acentuação de palavra e sentença ainda mais. No espanhol estas questões podem aparecer também, como visto em Milan (2019).

Uma vantagem de existirem dezenas de pesquisas focadas em certos segmentos, como vogais, distinção das líquidas /r-l/ e VOT, é que existe ampla evidência bibliográfica da pertinência dos TP para uma percepção da fala mais acurada. Todavia, uma vez que estas pesquisas estão sendo realizadas há

décadas, novos, e porque não antigos pesquisadores, estão convidados a investigar novas possibilidades de pesquisa. Como exercício criativo, sugere-se prosódia do português brasileiro e monotongação de ditongos do inglês estadunidense.

4.5 Existência de Grupo Controle

Para o descritivo 5, tem-se que um total de 47 trabalhos (total n=117), ou 40,17%, não possui grupo controle.

Ademais, Huensch (2016), em uma pesquisa que será mais detalhada na subseção 4.11.3, contou com grupo controle, porém este grupo realizou um treinamento focado em alvos distratores, que não eram o foco da pesquisa (palatais em coda).

Como um forte argumento relacionado à importância de um grupo-controle, Lee e Hwang (2016b), em sua pesquisa de HVPT, notaram que o grupo controle havia melhorado cerca de 2,4% do pré-teste para o pós-teste, enquanto os grupos experimentais haviam melhorado 7%. Os autores comentaram que:

“os participantes do grupo controle receberam educação formal na escola e em institutos privados durante o período de treino (de quatro semanas), assim como os participantes do grupo experimental. Isto foi, portanto, importante, para confirmar se a exposição diária à língua inglesa teria efeitos significativos nas performances dos participantes”. (LEE e HWANG, 2016, p. 1862)⁹

Questionar o que realmente pode ser controlado em uma pesquisa é um debate válido, uma vez que os participantes em pesquisas de TP são frequentemente universitários que possuem diversas aulas em inglês, além das tecnologias atuais possibilitarem a comunicação imediata, leitura, digitação, acesso

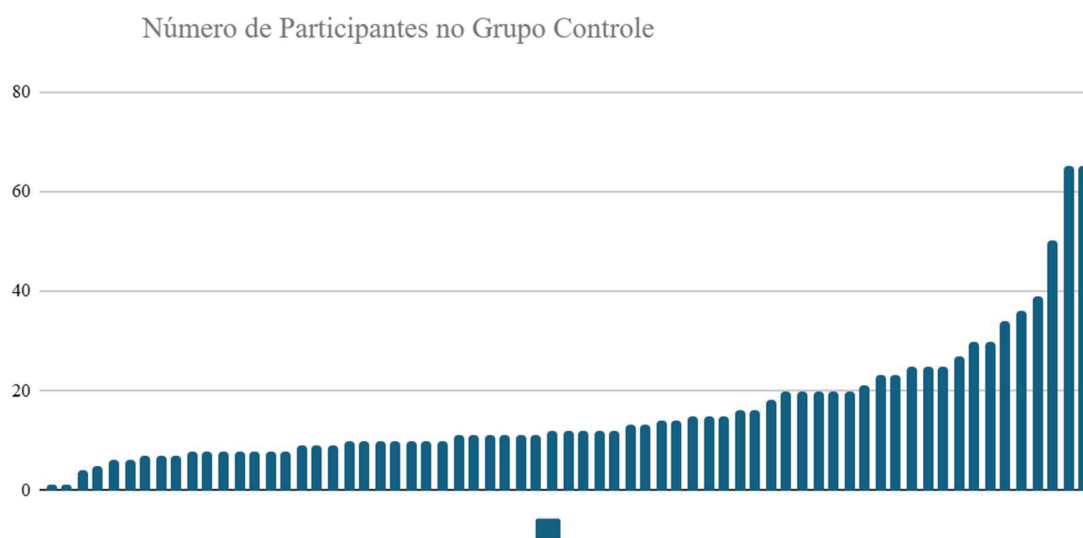
⁹ Original: “The subjects in the control group received formal and private English education both at school and at the private institute during the training period, just like the subjects in the experimental groups. It was therefore important to confirm whether or not daily exposure to English had significant effects on the subjects’ performance.”

a mídias como filmes e videogames, entre outros recursos, que podem aumentar o tempo de *input* de um aprendiz de língua estrangeira. No entanto, é interessante (talvez fundamental, talvez necessário?) comparar os resultados do grupo experimental com o de alguma amostragem de pessoas que não passaram pelo regime de treinamento. Mesmo que as porcentagens de acertos de identificação de algum segmento hipotético possam ter aumentado, pode ser difícil afirmar o quanto do aumento dos acertos foi em decorrência do treinamento *per se*, sem a comparação com um grupo-controle. Apresenta-se, a seguir, o número de participantes dos grupos-controle.

4.6 Número de Participantes do Grupo Controle

Com relação descritivo 6, que tratou do número de participantes no grupo controle, tem-se que a média foi de 16,3 participantes, e a mediana foi de 12 participantes, conforme ilustra Gráfico 3.

p. 94)



Fonte: O autor.

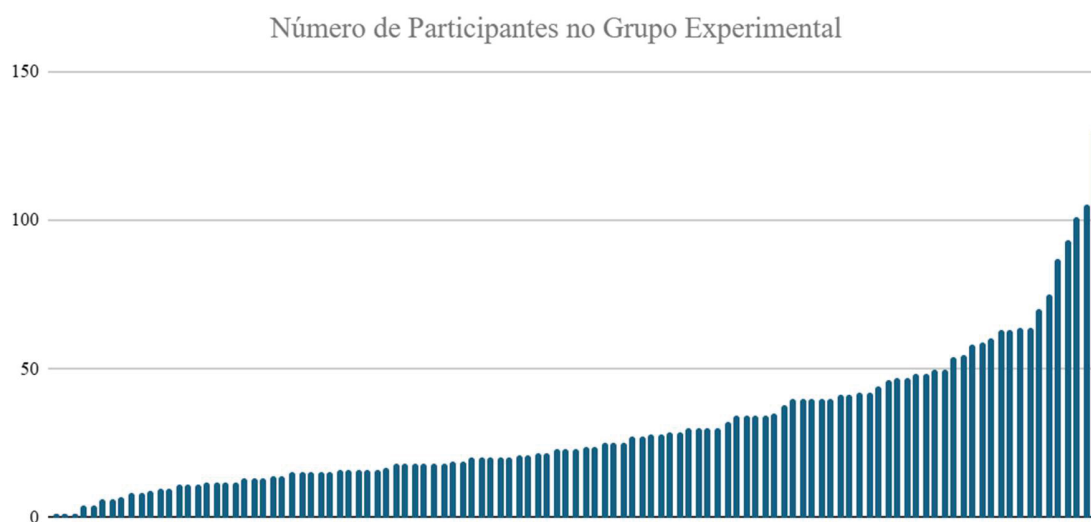
O gráfico de participantes no grupo controle, assim como gráfico de participantes no grupo experimental, é menos abrupto e com uma normalidade maior em suas distribuições, caso comparados aos gráficos, como, por exemplo, do número de participantes nos grupos experimentais e controles ou número de sessões, que apresentavam muitos valores nas extremidades. A média e a mediana de participantes estando próximas, 16 e 12, respectivamente, também indica um certo padrão para o tamanho de grupos-controle.

Recomenda-se, aos pesquisadores, a existência de um grupo controle com um número maior de participantes (a maioria das pesquisas contou com o número de participantes do grupo controle de pelo menos 1/3 ou metade do grupo experimental). Não é possível saber o que os estudantes estão investigando e aprendendo em suas rotinas de estudo, mas é possível garantir, ao menos, que eles não estão expostos ao treinamento fonético que os grupos experimentais estão. Esta comparação com relação a passar ou não pelo treinamento é interessante, e permite a existência de pelo menos uma comparação. Alguns trabalhos possuem mais de um grupo experimental, mas não dispensam o uso de um controle. Iverson e Evans (2009), por exemplo, possuíam grupos experimentais de duas nacionalidades / L1 diferentes, enquanto Alves e Luchini (2017) se utilizaram de um grupo experimental que, além do treinamento perceptual, passaram por sessões de instrução explícita dos conhecimentos que estavam sendo treinados.

4.7 Número de Participantes do Grupo Experimental

Já para o número de participantes nos grupos experimentais, abarcado pelo descritivo 7, a média foi de 30,3 pessoas, e a mediana de 23 pessoas. Nota-se uma tendência de que o número de participantes nos grupos controles é aproximadamente metade do número de pessoas nos grupos experimentais. Os números de participantes nas pesquisas podem ser observados no Gráfico 4.

Gráfico 3: Número de participantes nos grupos experimentais



Fonte: O autor.

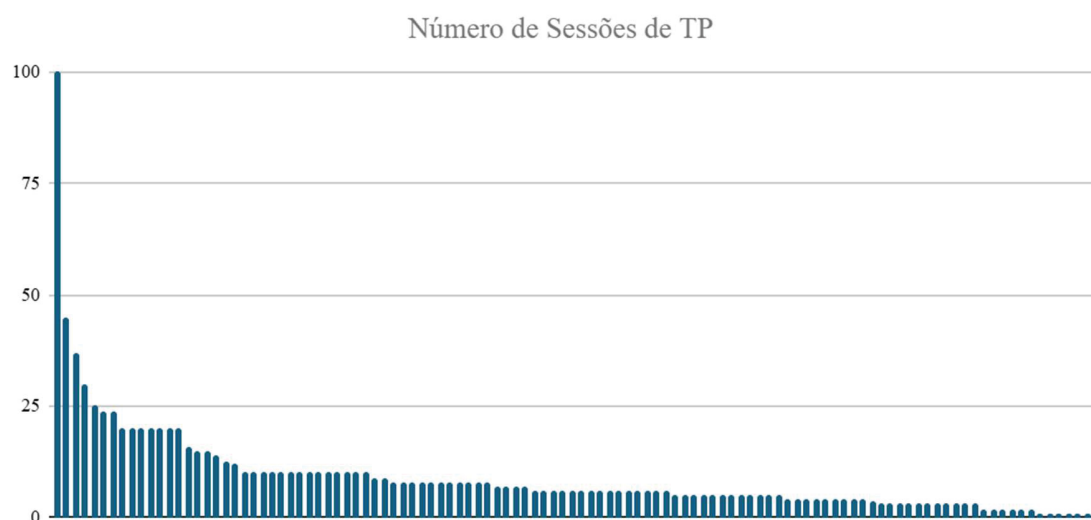
Como descrito na subseção anterior, os gráficos de participantes nos grupos experimentais e controle não são tão abruptos como os gráficos de números de sessões e de participantes no grupo controle. A média e a mediana deste conjunto de dados também está próxima, o que nos permite afirmar que, geralmente o número de participantes em grupos experimentais flutua entre 20-30 pessoas, enquanto este número está entre 10-20 para os grupos-controles.

4.8 Número de sessões

Conforme descrito ao se apresentar, na sessão de metodologia, o descritivo 8, referente o número de sessões, tem-se que a média de sessões nos TP foi de aproximadamente nove sessões, enquanto a mediana foi de seis. Porém, este número não fornece por si muitas informações, uma vez que uma sessão de treinamento varia de pesquisa para pesquisa, podendo durar, por exemplo, 20min, ou até mais de 2h. Além disso, o número de tarefas realizadas em cada pesquisa varia muito, podendo conter, por exemplo, 20 tarefas de identificação por sessão, ou até 100. Nota-se também a grande discrepância com relação aos gráficos de

participantes de grupos experimentais e controle: aqui, existem valores muito mais *outliers* às médias totais. O número pode ser observado através do gráfico 5:

Gráfico 4: Número de sessões de treinamento



Fonte: O autor.

É importante observar que, caso os dois trabalhos da RS com o maior número de sessões fossem desconsiderados para a análise (Burnham (2014) e Bradlow *et al.* (1997), com 100 e 45 sessões, respectivamente, a média de sessões cai de 8,99 para 7,89. A maioria dos trabalhos se encontra entre quatro e oito sessões, sendo seis a mediana. Muitos trabalhos apresentaram treinamentos “relâmpago” com uma, duas ou três sessões, e, como descrevem autores como Alves e Luchini (2017), é considerada uma limitação destes trabalhos a pouca quantidade de sessões. Nota-se, porém, que não são muitos trabalhos que “exageram” na quantidade de sessões, e, dada a quantidade de trabalhos que chegou a resultados significativos para a percepção de língua estrangeira, não parece que seja adequado submeter os participantes a muitas dezenas ou até centenas de sessões de treinamento.

Após a análise do número de sessões das pesquisas contidas na presente RS, não fica evidente a necessidade da realização de muitas (sete ou oito, por exemplo) sessões de treinamento, preferencialmente ao longo de um espaço de

tempo que exercite a memória de longo prazo dos participantes. Como traz Carlet (2019), certos aspectos fonéticos são mais difíceis de treinar que outros, portanto, exigindo mais tempo. Como exemplificado por diversas pesquisas, até treinamentos de curta duração podem trazer benefícios aos estudantes, porém um trabalho diacrônico de semanas ou meses é muito mais interessante para observar mudanças em médio/longo prazo e discutir modelos como os SDC¹⁰.

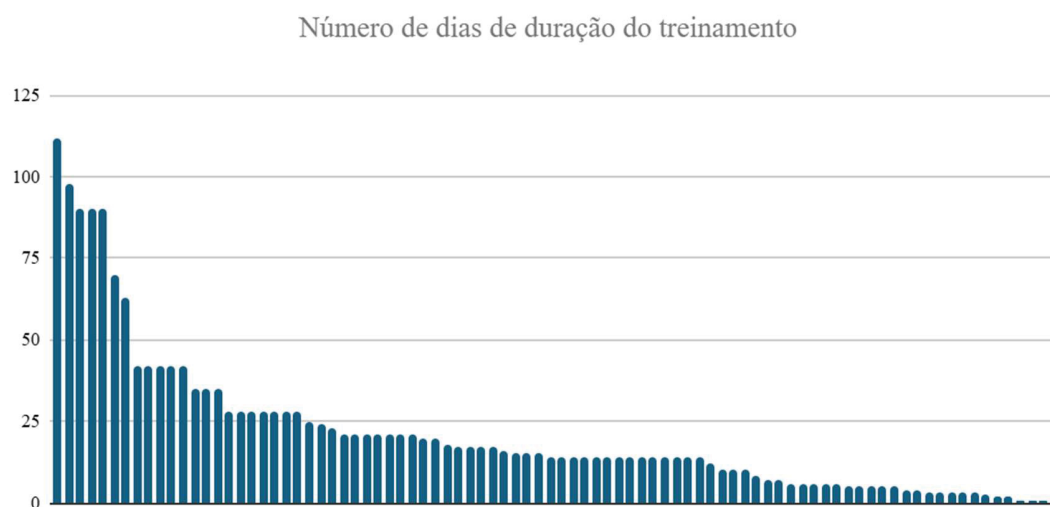
4.9. Duração do período de treinamento e das sessões

Esta subseção é dividida em duas partes, ambas derivadas do descritivo 9. A primeira trata da duração total do período do treinamento, independentemente da frequência de sessões de TP e da duração de cada sessão. Já a segunda, trata da duração média, em minutos, das sessões nos trabalhos de TP.

4.9.1 Duração do período de treinamento

A duração, em dias, do período de treinamento, foi de 22,5 (Gráfico 6). Foi necessário realizar operações aritméticas e aproximações para chegar neste valor, uma vez que, em muitas pesquisas, esse valor não contabilizava os testes postergados dos trabalhos que ocorreram após meses do final do regime de treinamento. Esse número, assim como o número de sessões de cada trabalho, não pode ser analisado sem contextos detalhados, uma vez que alguns trabalhos executavam os treinamentos todos os dias, outros uma vez por semana, outros dia sim, dia não, por exemplo. É interessante avaliar cada um dos 75 casos que ofereceram a duração do período de treinamento detalhadamente para contabilizar o número de horas totais de cada regime de treinamento.

Gráfico 5: Duração (em dias) dos treinamentos perceptuais.



Fonte: O autor.

Diferentemente do caso do descritivo 8, de número de sessões, desta vez vários trabalhos se estenderam por uma quantidade de dias maior que a média, até de mais de 60 dias, como é o caso de trabalhos como Mueller e Niedzielski (1968) e Donnellan *et al.* (2019). Essas pesquisas possuíam também um considerável número de sessões, 25 e 14, respectivamente. Cabe notar que, em alguns casos, mesmo pesquisas que possuíam apenas uma sessão de treinamento, houve intervenções pedagógicas ou outros tipos de atividades que aumentaram o número de dias contabilizados na pesquisa. É o caso, por exemplo, de trabalhos como Qin e Zhang (2019), com uma sessão e dois dias, e Qin *et al.* (2021), com uma sessão e três dias. Um padrão muito comum observado no estudo do número sessões e duração do período de sessões dos TP foi a de uma sessão de treinamento por semana. Com recorrência, os participantes de pesquisas de TP são estudantes de programas de letras/linguística/literatura e afins e podem ser voluntários de certas disciplinas e/ou cursos de fonética e fonologia.

Exemplos de pesquisas que seguiram este exemplo de aproximadamente uma sessão por semana foram Fouz-González e Monpean (2021), com quatro sessões em um mês, Bettoni-Techio e Koerich (2007), com seis sessões ao longo

de um mês e Abdullahi (2023), com três sessões em três semanas. Outro formato muito comum são os treinamentos diários durante uma ou duas semanas, conforme observado, por exemplo, em Rocha *et al.* (2020), com cinco sessões em seis dias, Gong *et al.* (2021), com quatro sessões em seis dias, Li *et al.* (2019), com quatro sessões em quatro dias e Zhang *et al.* (2013), com seis sessões em seis dias.

Uma possível sugestão para futuros pesquisadores que tenham a possibilidade de desenvolver TP com diferentes durações de tempo é entrevistar os participantes para descobrir suas opiniões com relação à duração das pesquisas, uma vez que conforto e motivação são fatores importantes no aprendizado. O perfil subjetivo de cada um também pode ser um fator importante na duração da sessão, uma vez que nem todos os pesquisadores controlam o tempo gasto em cada token, tarefa ou bloco de treinamento. É importante notar também que a possibilidade de participantes abandonarem ou perderem sessões é maior quanto mais longo for o processo de treinamento. Os dados presentes nas pesquisas com relação à taxa de evasão são insuficientes para análises mais aprofundadas a respeito de uma relação entre a duração do treinamento e a taxa de evasão.

4.9.2 Duração das sessões

Apenas analisar a duração do período de treinamento sem levar em conta quantas sessões acontecem por dia/semana e qual a duração destas sessões não é informativo o bastante. Por este motivo, apresento a seguir as médias de durações das sessões de treinamento.

- 45 trabalhos (38,46%) não reportaram quão longas foram cada sessão, enquanto 72 trabalhos (61,54%) informaram a duração de cada sessão.

A duração média das sessões de treinamento para os trabalhos da RS foi, em minutos, de 244min, aproximadamente 4h. Outrossim, foi realizado um cálculo

para aferir qual foi a média de minutos por sessões no total. O número obtido foi de 33min por sessão, sendo a moda de 30min.

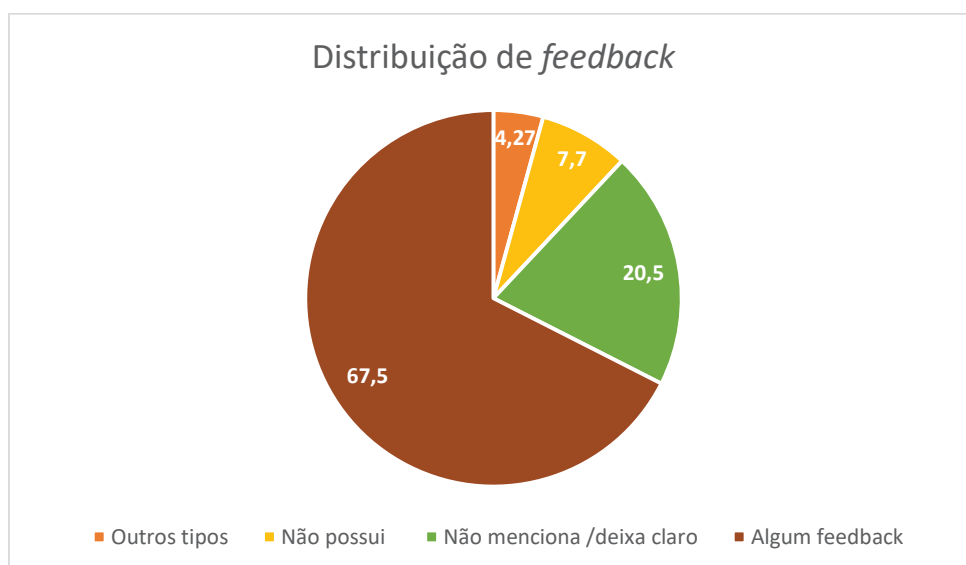
4.10 Existência e tipo de feedback

O descritivo 10 deu conta de descrever o feedback adotado, ou não, pelas pesquisas de TP. Conforme descrito na Metodologia, 67,5% dos trabalhos forneceram *feedback* aos participantes da pesquisa. Seguem-se dados mais detalhados:

- n=24 (20,5%) não mencionam ou não deixam claro o tipo de feedback (talvez tenha existido, porém não foi reportado);
- n=9 (7,7%) não possuem *feedback* (ou seja, afirmaram não existir *feedback*);
- n=5 (4,5%) ofereceu *feedbacks* diferentes dos mais recorrentes (imediato ou universal), como visual, ou ultrassom, individual, somente dos erros cometidos, etc.
- n=79 (67,5%) oferecem algum tipo de *feedback*, dos quais:
 - n=48 (41%) ofereceram *feedback* imediato
 - n=16 (13,67%) ofereceram *feedback* imediato e *feedback* universal ao final;
 - n=8 (6,58%) mais de um tipo de *feedback* foi oferecido, mas eles não foram detalhados nos pormenores;
 - n=7(5,98%) ofereceu algum *feedback*, mas não foi detalhado qual especificamente;

As informações descritas acima sobre a distribuição de *feedback* podem ser no gráfico 7:

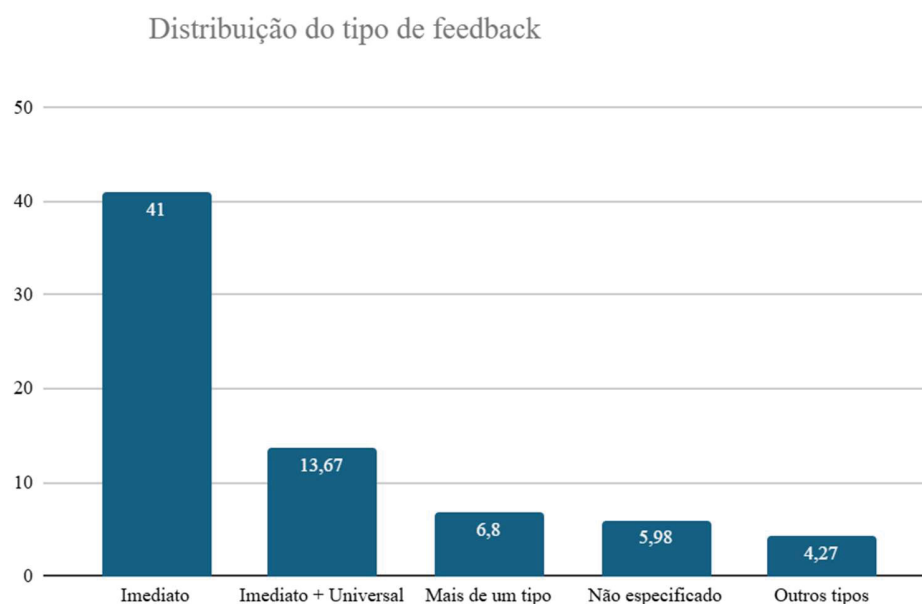
Gráfico 6: Distribuição de *feedback*



Fonte: O autor.

Da porcentagem de pesquisas que reportam uso de *feedback*, a distribuição pode ser encontrada no gráfico 8.

Gráfico 7: Distribuição do tipo de feedback.



Fonte: o autor.

O tipo de *feedback* nas pesquisas não é um tópico que é discutido a fundo nas pesquisas presentes na RS, de forma geral. Tajjima *et al.* (2007) apenas declaram que o *feedback* (imediatos) presente na pesquisa era fornecido após cada *trial* do treinamento (e que este *feedback* não estava presente nos pré e pós-testes). Golestan e Zatorre (2009), apesar de não aprofundarem as discussões relacionadas ao *feedback*, citam Cohen (2007) para argumentar que a existência de *feedback* é bem-vinda durante o treinamento. Quin e Zhang (2019), ao fornecerem *feedback* após tarefas de ID do treinamento, descreveram com mais detalhes que o modelo de retorno às respostas dos estudantes continha uma resposta positiva, no caso de acerto, que indicava textualmente e com a cor verde a escolha correta, enquanto as respostas erradas vinham em texto, acompanhadas da cor vermelha e com a indicação de qual era a resposta correta. Chandrasekaran *et al.* (2010), ao defenderem o uso de *feedback* na pesquisa, citam McClelland *et al.* (2002) para justificar os efeitos benéficos do *feedback*, que, como padrão da maioria das pesquisas, era fornecido no período de treinamento, mas não nos (pós)testes. Nesta pesquisa, o *feedback* fornecido foi apenas descrito como sendo “certo/errado”. Thomson (2012) também citou McClelland *et al.* (2002) como fonte para afirmar os efeitos positivos de um *feedback* imediato. O autor não descreveu exatamente a natureza do *feedback* imediato que foi provido para os participantes da pesquisa, mas demonstrou acreditar no potencial de promover aprendizado.

Mora *et al.* (2022) aprofundaram um pouco mais a discussão teórica no que concerne a presença de *feedback* em pesquisas que combinaram percepção e produção. Os autores citaram Herd *et al.* (2013) como fonte que corrobora o uso do *feedback*, e Lu *et al.* (2015) e Thorin *et al.* (2018) como fontes que não corroboram este aspecto (os autores afirmam que o *feedback* foi indiferente ao resultado, e não que atrapalhou, é importante explicitar). Referente ao treinamento e à percepção, Mora *et al.* também citaram McClelland *et al.* (2002) para corroborar a eficácia do *feedback* na melhora dos acertos de percepção. Em todo caso, a pesquisa de Mora *et al.* (2022) se propôs a aplicar um procedimento de HVPT de discriminação ao processamento pré-lexical, identificação e repetição imediata, afirmando acreditarem ser os primeiros a explorarem esta possibilidade de metodologia de

treinamento. Flege (1995) utilizou *feedback* apenas para os grupos experimentais de falantes de mandarim, enquanto os nativos de inglês realizaram tarefas sem a existência de *feedback*.

É importante notar que certos trabalhos oferecem *feedback* somente nos testes. Outros, somente nos treinamentos. Outros ainda, nos testes e nos treinamentos. Não se observou uma regra no que tange a existência e o tipo de *feedback*, apenas que a maioria dos trabalhos se utilizou de algum tipo desta ferramenta, e que a maioria contou com *feedback* imediato e/ou universal. Como uma nota pessoal que julgo pertinente acrescentar, como alguém que participou de um regime de treinamento, achei muito mais agradável realizar o TP com *feedback*, uma vez que, no meio de centenas de *trials* fica desmotivador não saber como você está se saindo ou se ou o que você está errando exatamente. Recomenda-se, aos pesquisadores, fornecerem *feedback* imediato e global para os participantes das sessões de treinamento. É muito mais fácil para o participante saber o que está errando, e, portanto, chegar a como melhorar com um *feedback* explícito.

O *feedback* global, ao final da sessão ou do treinamento, pode ser muito útil para o estudante avaliar seus erros no geral e se preparar melhor para a próxima sessão ou teste. Muitas pesquisas citaram McClelland *et al.* (2002) como fonte principal na argumentação de que o *feedback* promove aumento de acertos. Porém, como muitos paradigmas na ciência, é necessário replicabilidade de pesquisas, para que paradigmas como este não sejam simplesmente aceitos e nunca mais testados. Pode ser possível, por exemplo, comparar dois grupos experimentais, um com *feedback* e outro sem, com o intuito de avaliar se houve alguma diferença entre os grupos. Como já exemplificado pela minha própria experiência pessoal, é possível coletar dados no que tange à opinião dos participantes para verificar o que eles acham mais agradável na hora de realizar o treinamento.

Ueda e Hashimoto (2019), uma pesquisa que será descrita com mais detalhes em 4.11.2, 4.11.3 e 4.14.2, ao não encontrar resultados estatisticamente significativos para certos aspectos da pesquisa, comentam sobre a forma com que

o *feedback* foi passado. Segundo os autores, a maioria dos trabalhos de TP são realizados em um laboratório e com *softwares*, o que facilita a inserção de um *feedback* automático e corretivo, o que, por sua vez, estimula os participantes a continuarem melhorando a partir do reconhecimento dos erros. A pesquisa de Ueda e Hashimoto (2019), todavia, foi realizada em um contexto escolar, e o *feedback* era apresentado em uma tela para todos os participantes verem, e cabia a cada um reconhecer (ou não) se haviam errado e o que haviam errado, o que tornava o processo de *feedback* menos eficiente.

4.11 Existência de testes de generalização

Conforme descrito na metodologia, o descritivo 11 resume que 36 trabalhos (30,77% do total da RS) apresentaram testes de generalização em suas pesquisas. A maioria desses testes foram para testar novos falantes, novos segmentos, palavras, sílabas ou contextos e verificar se os testes de percepção possuem implicações na produção da fala.

Nas próximas subseções, apresentam-se dados mais aprofundados com relação a distribuição dos tipos de generalização: 6 trabalhos investigaram a generalização para a produção da fala; 16 trabalhos investigaram a generalização para novos falantes; 26 pesquisas investigaram a generalização para novos segmentos, palavras, sílabas ou contextos. Alguns trabalhos, como Wong (2012), verificaram mais de um tipo de testes de generalização na mesma pesquisa.

4.11.1 Generalização para a produção da fala

Conforme descrevem Barriuso e Hayes-Harb (2018), Pisoni *et al.* (1993) é uma continuação da pesquisa de Logan *et al.* (1991), que treinou participantes japoneses aprendizes de inglês para os pares de consoantes /r, l/. A pesquisa de 1993 adicionou testes de generalização para novas palavras produzidas por vozes conhecidas e novas, e os participantes da pesquisa conseguiram generalizar os

estímulos. Porém, em um experimento que o regime de treinamento contou apenas com um falante na fase de treinamento, o mesmo efeito de generalização para novos falantes não foi observado. Esta observação, diga-se de passagem, instaurou a preferência observada por centenas de pesquisas na área em adotar a alta variabilidade como paradigma de input, preferindo a existência de vários locutores para as sessões de treinamento.

Somente uma minoria dos trabalhos estudados verificou a generalização de TP para a produção da fala. As formas mais comuns de se analisar um teste de produção da fala são por análise acústica e medição de características acústicas do som (como durações), como em Alves e Luchini (2017), ou pela avaliação de falantes da língua-alvo, como em Bradlow e Pisoni (1997). Ao todo, foram seis estudos, os quais são descritos e posteriormente debatidos nos parágrafos seguintes. Nas Considerações Finais do trabalho, mais discussões concernentes à estas formas de avaliar as produções podem ser encontradas.

Bradlow e Pisoni (1997) é o trabalho mais antigo desta subsecção. Os autores verificaram a generalização para a produção da distinção do /r, l/ do inglês por falantes japoneses. Os grupos da pesquisa contaram com 11 participantes no grupo experimental e 12 participantes no grupo controle. A metodologia seguiu a estabelecida por Logan *et al.* (1991) e Yamada (1993), contando com pré-teste, treinamento e pós-teste. O número de sessões foi de 45, espalhadas entre três e quatro semanas. Com relação à percepção, o índice de identificações corretas foi de 65% no pré-teste para 81% no pós-teste. Para a produção, os valores foram de 67,5% para 73,3%. A mudança na percepção é visível por teste estatístico com resultados significativos, porém, para a produção, não. Os autores concluíram que, por este estudo, não foi possível generalizar a melhora significativa na percepção dos sons para a produção da fala. Os autores também notaram que somente um participante apresentou melhora significativa na percepção, porém piora na produção. É importante notar que, mesmo não sendo estatisticamente significativo, houve mudança de 5,3% na média da produção. Cabe retomar a metanálise recente

de Uchihara *et al.* (2024) que apontou que os ganhos de produção da fala em pesquisas de HVPT estavam entre 4,5% e 10,5.

Alves e Luchini (2017) realizaram um treinamento perceptual para brasileiros estudantes de inglês para as plosivas iniciais não-vozeadas /p, t, k/. Os 30 participantes da pesquisa foram divididos em três grupos: um grupo controle, um grupo que passou pelo treinamento perceptual de três sessões, cada uma durando cerca de 30min, e outro grupo que, além do treinamento perceptual, passou pela instrução explícita do que estava sendo trabalhado. A tarefa escolhida para a pesquisa foi de Identificação com *feedback* imediato. Os padrões de plosivas para a percepção de produção foram de VOT negativo (o que, no português, indica uma consoante vozeada), VOT positivo (aspirada), VOT zero e VOT com zero artificialmente cortado. Os autores encontraram que, para a generalização do treinamento perceptual para a produção da fala, apenas no Grupo 2, que recebeu as sessões de treinamentos e instrução explícita, houve diferenças significativas para a produção de /p/. Contudo, como na pesquisa de Alves e Luchini (2020), que será descrita abaixo, as médias de produção de /k/ no pré-teste já apresentavam valores semelhantes a um padrão nativo, deixando pouca margem para alterações significativas. Para /t/, apesar de haver grandes diferenças nas médias de valores de duração do VOT, os testes estatísticos não indicaram diferenças significativas. Os autores sugerem, para a área, pesquisas com metodologias distintas, como contendo coleta de dados de forma espontânea (além da forma comum controlada, ou seja, leitura de textos ou sentenças-veículo), com as instruções implícitas ou explícitas, para verificar quais os efeitos de cada tipo de abordagem pedagógica.

Alves e Luchini (2020) realizaram um treinamento perceptual para nativos de português brasileiro aprendizes de inglês, com foco na presença ou ausência de VOT nas plosivas iniciais. O total de participantes foi de 30, e os grupos foram três, divididos da mesma forma descrita em Alves e Luchini (2017), ou seja, um grupo que recebeu treinamento, outro que recebeu treinamento e instrução explícita (uma oficina de 3 partes de 100min cada parte) e um terceiro que serviu de controle, não recebendo nem treinamento e nem instrução explícita. O número de sessões

também foi de três. Com relação à generalização para fala das consoantes /p, t, k/, o grupo de treinamento não apresentou diferenças significativas, embora o grupo que tenha passado pelo treinamento e pela instrução explícita apresentou diferenças significativas para /p/, e, surpreendentemente, o grupo controle, apresentou alterações significativas para /t/. Os autores notaram, assim como em Alves e Luchini (2017), que muitos valores, como as produções de /k/ pelos grupos que passaram pelo treinamento, aumentaram a duração do VOT (ou seja, produzido mais próximo ao padrão da língua inglesa), embora os testes estatísticos não tenham indicado que o aumento tenha sido significativo. Os autores comentam que o pequeno número de participantes foi uma limitação do trabalho e uma possível razão para isso.

Brawerman-Albini *et al* (2017) realizaram cinco sessões de treinamento perceptual para aprendizes brasileiros da língua inglesa, com foco no acento silábico de palavras pré-proparoxítonas, padrão incomum na língua materna dos participantes da pesquisa. Os grupos da pesquisa foram um experimental, com 30 participantes, e um controle, com 10 participantes. O *design* da metodologia contou com um pré-teste de produção, uma aula introdutória, um pré-teste de percepção, o treinamento, dois pós-testes (contendo generalização com novas palavras em um e com novas palavras e falantes em outra), além, de um teste de retenção após dois meses. Com relação à generalização do treinamento perceptual para a fala, tem-se que o total de palavras pronunciadas corretamente foi de 28,4% no pré-teste para 59% no teste de generalização, que, segundo os testes estatísticos, ainda se manteve no teste de retenção tardio de dois meses. Curiosamente, o grupo controle também apresentou diferenças significativas no teste de generalização, indo de 28,4% para 59%. De toda forma, os autores notaram que, apesar dos grupos demonstrarem valores parecidos no pré-teste, os testes estatísticos demonstraram melhoras significativas e demonstram a eficácia do treinamento.

Alves e Kampff (2020) realizaram um treinamento perceptual para brasileiros aprendizes de inglês, com foco no VOT das plosivas iniciais. A pesquisa será descrita com mais detalhes na subseção 4.12, sobre testes de retenção postergada.

Com relação à generalização para a produção da fala, os autores não encontraram evidências suficientes para comprovar que isso ocorreu para /k/ e /t/. Uma hipótese elencada pelos autores é de que os valores de produção de /k/ e de /t/ no pré-teste já se encontravam dentro do esperado por falantes nativo, dada uma “semi-aspiração” das consoantes plosivas. Para /p/, porém, foram encontradas diferenças significativas pelos testes de Friedman, que, no segundo teste postergado, acabaram por diminuir. Os autores levantaram a hipótese de que o motivo desta não generalização fossem, talvez, pelas mudanças na própria percepção não terem sido tão significativas. O pequeno número de sessões (três) também pode ser considerado um fator limitante.

De forma geral, nota-se que a generalização de TP para a produção da fala apresentam indícios modestos. A maioria das pesquisas aponta pequenas mudanças, que nem sempre são estatisticamente significativas. É importante verificar este aspecto com mais pesquisas, e, se possível, com treinamentos com mais participantes no grupo experimental, mais sessões de treinamento e com testes de retenção meses ou anos depois das sessões de treinamento.

4.11.2 Generalização para novos falantes

Diferentemente das análises apresentadas na subseção anterior sobre generalização para a produção da fala, como o número de pesquisas que estudaram generalização para novos falantes¹¹ foi alto (n=16), então foi necessário fazer um recorte das pesquisas mais relevantes e ainda não descritas no trabalho até agora.

Wang *et al.* (2003) conduziram um trabalho que treinou estadunidenses na percepção dos tons de mandarim e a implicação do treinamento na produção da fala. Os participantes foram, ao todo, 16 estadunidenses nativos de inglês, divididos

¹¹ Esta tese utiliza “falantes” e “locutores” como sinônimos: as vozes de *input* para testes ou treinamento perceptual.

pela metade nos grupos experimental e controle. O grupo experimental passou pelo período de duas semanas de treinamento perceptual. Para a tarefa de julgamento da produção da fala, foram recrutados 82 nativos de mandarim, da região de Pequim, acostumados com o sistema *pinyin*¹², utilizado na pesquisa. O número de sessões da pesquisa foi de oito, distribuídas ao longo de duas semanas. Os estímulos foram divididos em 80 palavras existentes no mandarim divididas em grupos de 20 para quatro tons diferentes, sendo que metade desses estímulos apareceram no treinamento, assim como nos testes (estímulos “velhos”) e a outra metade, os “novos”, foram utilizados para testes de generalização para novos segmentos. Para a produção da fala, os participantes demonstraram um aumento no *score* de 57% no pré-teste para 75% no pós-teste para os estímulos “velhos”, enquanto, para os novos, o aumento foi de 55% para 68%. Os testes ANOVA indicaram efeitos significativos para o teste e para o grupo em que os participantes estavam, o que permite afirmar que a performance no pré-teste e no pós-teste foi diferente para os grupos experimentais. Testes t pareados também indicaram a diferença no pré e pós-teste para o GE. Para os dados de percepção, notou-se, primeiramente, uma discrepância imensa entre os níveis iniciais de precisão no pré-teste (entre 24% e 77%) e na melhora no pós-teste (entre 3% e 31%). Cabe notar, neste ponto, que muitas pesquisas descartam participantes no início do treinamento por eles já possuírem muita precisão no começo da pesquisa. Como já descrito, a pesquisa apresentou diversos testes de generalização (inclusive para a produção da fala, com julgamento feito via testes de identificação realizados por falantes nativos de mandarim), porém, nesta subseção, o foco se dá para generalização para novos falantes. Os autores reportaram uma melhora de 21% na percepção, que se delongou para o teste de retenção realizado seis meses após as sessões de treinamento. O número de participantes no grupo experimental, que passou por duas semanas de treinamento, foi de oito, enquanto o grupo controle contou com oito pessoas. As médias de generalização, como já mencionado, foram de 21% no pós-teste com os mesmos estímulos, 18% para novos estímulos, e 25% para novos

¹² Sistema de romanização da escrita mandarim.

falantes. O teste de retenção após seis meses ainda detectou uma melhora de 21% ao se comparar com o pré-teste.

Bettoni-Techio e Koerich (2007) treinaram falantes nativos de português brasileiro para encontros consonantais com /sC/ da língua inglesa. A pesquisa contou com seis sessões de treinamento (cada uma variando entre um e 24 blocos de estímulos, cada um variando entre 16 e 64 *trials*), testes de generalização e retenção. A dificuldade dos estímulos aumentou com o tempo, seguindo a lógica: /s/+ plosivas, /s/ + soantes, com contextos anterior de silêncio, consoantes, e, finalmente, vogais. O ganho médio de acertos, após todas as sessões, foi de 12,3%, já considerando a generalização para um falante novo, o que correspondeu a 48,1% da possibilidade de melhora total (que era de 25,7%, com base no pré-teste). Nota-se que, para algumas palavras, a generalização para o locutor novo foi até mais alta do que para o locutor treinado (por exemplo: em *small*, a porcentagens de acerto para o locutor treinado foi de 77,8%, enquanto para o locutor novo a porcentagem de acerto foi de 88,9%).

Rato (2014) investigou os efeitos de um HVPT na percepção de três pares de vogais (anteriores alta e média e posteriores arredondadas) do inglês por falantes nativos de português europeu. Os participantes eram adultos, calouros de um curso de Letras, e recebiam quatro horas de instrução por semana. A maioria dos estudantes se encontrava no nível de proficiência intermediário B1. A pesquisa contou com cinco sessões de duração de uma hora cada, contendo as etapas de instrução explícita articulatória/visual dos segmentos-alvo, e dois tipos de tarefas: DIS e de ID. Os participantes que seguiram até o final da pesquisa foram 30 (19 no grupo experimental e 11 no grupo controle). O *feedback* desta pesquisa foi imediato e visual, através do *software* TP. As melhoras observadas para o grupo experimental foram, no pós-teste, respectivamente, para os três pares de vogais, 17,37%, 31,84% e 19,38%. Os testes t, na análise de generalização, revelaram que o GE teve os maiores ganhos de acertos na identificação nos dois pares de vogais frontais, enquanto o par de vogais posteriores apresentou ganhos mais modestos.

Wong (2015) investigou os efeitos de um HVPT na percepção e produção dos pares de vogais altas anteriores e posteriores da língua inglesa, /ɪ/-/i:/-/ʊ/-/u:/, por participantes cantoneses. O grupo experimental contou com 35 pessoas de Hong Kong, com inglês como L2, que passaram por 20 sessões de ID de HVPT distribuídas durante 10 dias. Os testes estatísticos apontaram melhoras significativas na identificação das vogais, com uma melhora do par /ɪ/-/i:/ sendo significativamente maior do que a do par de vogais posteriores (9,79% maior, mais precisamente). Os testes estatísticos também apontaram para uma melhora estatisticamente significativa nos testes de generalização, seja para novos locutores ou novas palavras, com níveis similares aos obtidos quando apresentados com palavras já treinadas e locutores já conhecidos. Todavia, notou-se que, novamente, o par de vogais frontais apresentou melhores resultados do que o par de vogais posteriores.

Ueda e Hashimoto (2019) treinaram 24 participantes japoneses, aprendizes de inglês como L2, para percepção e produção de três pares de consoantes /b/-/v/, /l/-/r/, e /s/-/θ/, durante seis semanas em treinamentos de 10 minutos por sessão. De antemão, os autores apontam que a baixa quantidade de tempo dispendida no treinamento pode ter sido um dos fatores que colaboraram para não gerar exatamente os resultados esperados. Os autores apontaram que houve melhoras em percepção e produção para os pares /b/-/v/ e /s/-/θ/ (aproximadamente 6%, incluindo nos testes de generalização de palavras conhecidas com um locutor novo, porém não houve melhoras para o par /l/-/r/, além de nenhum teste de generalização para palavras novas.

Li *et al.* (2019) utilizaram um HVPT híbrido para treinar nativos de japonês com altos níveis de proficiência para os contrastes de nasais finais do mandarim. O grupo experimental (n=8) obteve melhora de 13% na identificação de nasais finais do mandarim, enquanto o grupo-controle (n=8) permaneceu com os mesmos valores. Os valores iniciais de identificação das nasais foi 70% no pré-teste para 83% no pós-teste realizado após quatro dias do treinamento (cada um com uma sessão de uma hora e contendo 76 *trials*), enquanto os valores no pós-testes para

as generalizações foram de 86% para novas palavras, 81% para novos locutores e 82% para novas frases/ sentenças. Os autores concluem que novas representações fonológicas foram estabelecidas para que os participantes tenham conseguido generalizar os conhecimentos do treinamento perceptual para novos estímulos, falantes e sentenças.

Notou-se, durante as Subseções 4.11.1 e 4.11.2, que foi um pouco difícil evitar mencionar as generalizações para novos estímulos. Isso aconteceu porque estas foram os tipos de testes de generalização mais comuns encontrados nos trabalhos da RS. A subseção seguinte apenas compila um pouco destas informações.

4.11.3 Generalização para novos segmentos, palavras, sílabas ou contextos

Assim como na subseção anterior, o número de pesquisas que investigou a generalização para novos segmentos, palavras, sílabas ou contexto foi de $n=26$ trabalhos. Para que a descrição não se estenda em demasiado, uma seleção de pesquisas foi feita para apresentar como se deram estes testes de generalização. Como a generalização para novos segmentos, palavras, sílabas e contextos é a mais comum encontrada na RS, algumas pesquisas descritas nas subseções anteriores de 4.11 já trouxeram testes de generalização para novos estímulos.

Ueda e Hashimoto (2015) afirmaram que trabalhos que apresentaram generalização bem-sucedida para novos estímulos forneceram entre 15 e 22,5 horas de treinamento para seus 24 participantes. Os autores afirmam que seu treinamento de seis semanas, que contou apenas com 10min por semana para três pares de consoantes, totalizando apenas uma hora de treinamento no total, provavelmente foi curto demais para que houvesse generalização para novos estímulos. Os autores também citam a pesquisa de Chukharev-Hudilainen e Levis (2018), que não obteve generalização com 10-100 minutos de treinamento e 12 contrastes de sons. Li *et al.* (2019), em seu HVPT focado em mandarim e com participantes de alto grau de proficiência nativos da língua japonesa, partiram de

resultados de pré-teste de 70% de identificação para nasais finais, para 83% no pós-teste e uma generalização de 86% para novas palavras.

Huensch (2016) realizou um treinamento de oito sessões de 20min cada, com a idiossincrasia de não se utilizar de segmentos, e sim de um discurso contínuo, e observou a generalização para novas estruturas silábicas e domínios gramaticais. Os participantes eram nativos de coreano e aprendizes de inglês. O grupo experimental (que continha 12 participantes) foi submetido à treinamento de palatais em coda, enquanto o grupo controle (também com 12 participantes) fez o treinamento, mas com um alvo não controlado. Além dos pré-testes e pós-testes de percepção, tarefas de produção com leituras de slides foram executadas. O parágrafo da tarefa de produção incluída sons como /n/, /l/ ou /ɹ/ (em palavras como *pinch*, *mulch*, *perch*,) antes da palatal, enquanto o treinamento só possuía codas monossilábicas (como na palavra *fish*). A generalização foi observada, para o grupo experimental, para novas estruturas silábicas (codas simples vs complexas), mas não para outros domínios gramaticais (como o final *ed* de morfemas do inglês). Nas conclusões do trabalho, a autora ainda retoma a pesquisa de Wong (2014), que buscou observar as diferenças em treinamento de apenas uma palavra ou em sentenças-veículo, e resultados mais favoráveis ao treinamento em sentenças-veículo foi observada.

Como conclusão desta seção, recomenda-se, aos pesquisadores, realizarem testes de generalização para novos segmentos/suprasegmentos e/ou vozes, e, se possível, testes de generalização para a fala. Assim, como a argumento de que os testes de retenção são importantes porque os conhecimentos adquiridos pelo treinamento precisam ser úteis a longo prazo, não é prático treinar todos os segmentos ou suprasegmentos para um aprendiz de língua adicional. Os testes de generalização para novos segmentos ou vozes podem ser pensados na elaboração da pesquisa e na coleta de vozes para o treinamento, e não acrescentam muito trabalho à pesquisa, trazendo grandes benefícios e insights no que concerne à eficácia de um regime de treinamento. Os testes de generalização da fala, conforme descrito nos resultados, são raros de serem encontrados, apesar de serem os mais

interessantes e trabalhosos. Nota-se também que, na verificação de generalização para a fala, cinco das seis pesquisas que estudaram isso foram desenvolvidas pelos mesmos grupos de pesquisadores, brasileiros (ALVES e LUCHINI, 2017; ALVES E KAMPFF, 2020; BRAWERMAN-ALBINI et al., 2017, BETTONI-TECHIO e KOERICH, 2007). Isso prova que existem pesquisadores competentes e estudos interessantes em terras nacionais. Seguindo por esta tangente, o software TP também foi desenvolvido em terras nacionais e utilizado por pesquisadores de todo o globo, sendo a ferramenta para se desenvolver testes de percepção por muitos trabalhos da presente RS (fonte).

4.12 Existência de testes de retenção

O descritivo 12 tratou dos testes de retenção postergada. Tem-se que 24 trabalhos (20,51% do total) contaram com testes de retenção. A média de tempo para testes de retenção foi de dois meses. Curiosamente, alguns destes trabalhos foram os mesmos que contaram com testes de generalização, mas não todos. Uma parcela destes trabalhos será doravante descrita para que a discussão dos testes de retenção postergada ocorra.

Alves e Kampff (2020) realizaram um treinamento perceptual com a intenção de averiguar os efeitos na percepção e produção de brasileiros aprendizes de inglês, com foco fonético incidente no VOT das plosivas iniciais. Os testes de retenção postergada foram dois: um realizado um mês após o fim do treinamento (retenção postergada 1), e outro após três anos do final do treinamento (retenção postergada 2). Os grupos controle (que não participou da etapa de treinamento) e experimental contaram com 10 participantes em cada um para os testes iniciais, treinamento e pós-testes imediatos, e o número total de sessões foi de três, distribuídas ao longo de uma semana após a realização do pré-teste. Os participantes eram alunos das três turmas do primeiro semestre de graduação de Letras, e apenas 15 alunos, sendo sete do grupo experimental, ficaram até o final para os testes de retenção postergada. Os estímulos de áudio foram provenientes

de seis falantes nativos de inglês norte-americano residentes no Brasil (três homens e três mulheres), e metade dos *tokens* produzidos por estes falantes tiveram uma manipulação na edição para que a aspiração fosse cortada e o padrão de VOT zero artificial fosse obtido. A tarefa escolhida foi a de identificação com *feedback* imediato, e a tarefa de produção consistia na leitura, em uma sala silenciosa, de um arquivo .ppt com palavras como “*pit, team, tick, pit*”, sendo cada uma falada duas vezes, e os dados foram subsequentemente analisados acusticamente com o *software* Praat. Foi possível observar, para a tarefa de identificação, diferenças significativas apresentadas em três parâmetros do pré-teste para o pós-teste para o grupo experimental: VOT negativo, Zero natural e Zero artificial. O VOT positivo foi apenas marginalmente significativo, e, para o GC, não houve diferenças significativas. Nos testes *post-hoc* de Wilcoxon, com correção de Bonferroni, realizado apenas para o GE (uma vez que o GC não apresentou nenhuma diferença significativa entre os pré-testes e pós-testes), o padrão de VOT Zero Natural apresentou diferenças marginalmente significativas entre pré-teste e pós-teste, pré-teste e pós-teste postergado 1 e pré-teste e pós-teste postergado 2. Com relação às diferenças intergrupos, testes *Mann-Whitney* foram rodados e indicaram diferenças na identificação no pós-teste, postergado 1 postergado 2. Na prática, isso significa que os grupos eram estatisticamente semelhantes apenas antes do treinamento acontecer. Para as tarefas de produção, os valores de duração do VOT foram alterados significativamente nos seguintes casos: /p/ no GE marginalmente para /k/ no GC. A diferença significativa encontrada para /p/ no GE foi entre o pré-teste e o teste postergado 1. Apresentada a pesquisa, observa-se que o teste de retenção postergada 1 apresentou os picos de acuidade de identificação para o padrão de VOT Zero Artificial quando os valores de VOT de /p/ atingem o pico. Todavia, os autores afirmam que muitos índices acabaram por diminuir do teste postergado 1 (um mês depois) para o postergado dois (três anos depois do treinamento).

Wang *et al.* (2003) treinaram estadunidenses para a percepção de tons do mandarim, contando com testes de generalização, implicações na produção da fala e retenção postergada. A pesquisa já foi descrita com mais detalhes na subseção

4.11.2, referente à generalização para novos falantes. Conforme já descrito, dentre os resultados estatisticamente significativos encontrados pela pesquisa, estão a retenção após seis meses do treinamento, na qual os participantes apresentaram a mesma porcentagem de *score* que no pós-teste imediato (21% de melhora na percepção dos tons de mandarim pelo grupo experimental).

A pesquisa de Rato (2014), já descrita na subseção 4.11.2, investigou os efeitos de um HVPT na percepção de três pares de vogais da língua inglesa por falantes nativos de português europeu. A pesquisa contou com um treinamento composto de 210 tokens, sendo que 84 novas palavras foram utilizadas para o teste de generalização do conhecimento. Com a intenção de averiguar a retenção postergada do experimento, um novo teste foi realizado dois meses após o fim das sessões de treinamento. Os resultados estatísticos dos testes ANOVA indicaram que o treinamento de alta variabilidade aumentou a porcentagem de três contrastes de vogais do experimento para os GE, mas não para o GC. Ademais, os resultados significativos continuaram sendo válidos para o teste de retenção postergada: /æ/-/ɛ / pré-teste 66.40, pós-teste 81.32, retenção: 83.77; /i/-/ɪ/ 55,79 no pré-teste, 85 no pós-teste, 87,63 para retenção; /u/-/ʊ/ 64,74 no pré-teste, 78,6 no pós-teste, 84,12 no teste de retenção, todos com $p < 0.001$ no teste estatístico.

Recomenda-se aos pesquisadores, sempre que possível, aplicação de testes de retenção postergada do conhecimento. Em pesquisas de mestrado, em que o tempo é curto, é quase impossível coletar dados meses depois do treinamento. Porém, em pesquisas de doutorado ou em programas de pós-graduação com produções contínuas, é muito importante verificar se os conhecimentos adquiridos nos TP continuam fixados por meses e/ou anos após as experiências. As habilidades de percepção e produção são importantes a longo prazo, e não apenas pelos dias e/ou semanas que o treinamento acontece. Nota-se também que muitas pesquisas da bibliografia realizaram estes testes apenas poucos meses depois da finalização do treinamento (Abdullahy *et al.* 2023 e Hong, 2017, dois meses, Leong *et al.*, 2013 e Hu *et al.*, 2016, três meses, Alves e Luchini, 2017 e Shum *et al.* 2021, um mês, a cargo de exemplo), e que é difícil manter e/ou recuperar o contato com

os participantes do treinamento após anos, porém, sempre que possível, agrega-se à área verificar se os ganhos proporcionados pelos treinamentos perceptuais podem ser notados após anos.

4.13 Adoção do paradigma da alta variabilidade

O descritivo 13 tratou da adoção, ou não, da alta variabilidade. Observou-se que 54 trabalhos (n=117), ou 46,15% do total das pesquisas presentes na RS adotaram o HVPT como metodologia da pesquisa. Conforme já narrado na discussão teórica da área de TP, a partir de uma série de pesquisas do começo da década de 1990, em especial de Logan, Pisoni e colegas (1991 e 1993), chegou-se à conclusão de que a presença de uma maior variedade de vozes nos treinamentos promovia mais ganhos na percepção e poderia promover a generalização do conhecimento para novos falantes, estímulos, produção da fala e até na retenção postergada dos ganhos. Como o número de trabalhos que trouxeram o HVPT para o debate é muito grande, dá-se preferência para descrever uma amostragem destas pesquisas aqui, priorizando trabalhos ainda não mencionados até aqui no texto.

Lee e Hwang (2016b) treinaram, através de seu HVPT, 40 nativos coreanos por quatro semanas para um gradiente de vogais e consoantes da língua inglesa. Além dos 40 estudantes pertencentes a dois grupos, a pesquisa contou com um grupo controle de 20 pessoas. Os estudantes tinham entre 11 e 12 anos de idade, e a diferença entre os grupos experimentais era que o segundo grupo, além de receber o HVPT por quatro semanas, recebeu também instrução explícita via um curso de treinamento (*Postech English Speech Assessment and Assistant - PESAA*). A média de proficiência dos participantes foi avaliada entre B1 e B2 (falantes intermediários de inglês). Os estímulos do pré e pós teste foram diferentes dos do treinamento, e basicamente todos os sons da língua inglesa (11 vogais e 24 consoantes) foram incluídos na pesquisa. Para o treinamento, foram separados oito sets de pares mínimos, quatro para vogais e quatro para consoantes. A variabilidade de falantes preconizada pelo HVPT se deu pela introdução de um locutor novo a cada sessão de treinamento. Como resultados gerais, observou-se que o grau de

discriminação de sons pelo grupo controle melhorou 2,4%, enquanto a dos grupos experimentais foi de 7%. Ao se comparar estatisticamente ambos os grupos experimentais, notou-se que não houve efeito do tipo de treinamento (apenas o HVPT ou aliado à instrução explícita do PESAA. Ademais, os pesquisadores dividiram os 40 participantes dos grupos experimentais em dois grupos de proficiência, e notou-se que o grupo de menor proficiência obteve maiores ganhos. O grupo de menor proficiência apresentava porcentagens de acerto iniciais de 57% (entre 47% e 68%), e melhoraram 9,1% após o HVPT; o grupo de maior proficiência apresentava porcentagens de acerto iniciais de 75,1% (entre 70% e 83%), e melhoraram em média 4,8% após o HVPT.

Isasa (2022) realizou um estudo exploratório do contraste /e/ e /ei/ do espanhol para falantes de espanhol como Língua de Herança (LH) em comparação com os estudantes de L2 no contexto de sala de aula. A pesquisa também comparou a efetividade do HVPT, contendo múltiplas fontes dos *inputs* de fala, com treinamento regular de fonética com baixa variabilidade (LVPT). Primeiramente, a autora define os falantes de uma LH como pessoas com características sociolinguísticas (culturais, familiares, de imigração etc.) que impõe outra língua sobre estes indivíduos. Exemplos como este podem ser observados em imigrantes hispânicos nos EUA, que convivem com famílias que falam espanhol em casa, porém entram em contato com o inglês nas escolas e subsequentemente no trabalho e em toda a sociedade. O Brasil possui muitas pessoas nestas condições: imigrantes haitianos, libaneses, sírios, venezuelanos, que continuam mantendo suas línguas em seus lares e comunidades, mas precisam do português para se integrar à sociedade brasileira. Além dos “Falantes de Herança”, também existem os “Aprendizes de Herança”, que estudam ou reaprendem as línguas herdadas pelo contexto sociolinguístico às quais estão expostas. Isasa cita autores como Godson (2004) e Rao e Ronquest (2015) para explicar que, de um ponto de vista de contexto fonológico, os falantes de LH apresentam certas vantagens em comparação ao falante típico de L2 que nunca teve contato com a língua estudada. Mesmo o contato com a língua em questão sendo diminuído ou cortado, não é suficiente para fazer o indivíduo simplesmente perder a aquisição desta LH. Para este estudo, Isasa contou

com quatro grupos de participantes de uma universidade do *Midwest* dos EUA, divididos em: HVPT-Her (participantes LH treinados com alta variabilidade), LVPT-Her (participantes LH treinados com baixa variabilidade), HVPT-L2 (nativos estadunidenses treinados em alta variabilidade) e LVPT-L2 (nativos estadunidenses treinados em baixa variabilidade). Os participantes dos grupos Her eram imigrantes (primeira geração) ou filhos de imigrantes de diversos países latinos como Cuba, México e Porto Rico. Cabe notar, no que concerne a diferença entre /e/ e /ei/, que muitos sotaques norteamericanos tendem a monotongar o ditongo /ei/ para algo próximo a /e/, o que poderia dificultar a percepção deste contraste do espanhol. Isasa (2022) discute, após apresentar os testes estatísticos, que ambos os participantes LH e L2 foram capazes de discriminar perceptualmente /e/ de /ei/ do espanhol, embora com precisões bem diferentes entre si (91,39% para os LH e 80,14% para os L2). Ao comparar os falantes LH aos nativos de espanhol utilizados como GC na pesquisa (97,14% de precisão na distinção), não houve diferenças significativas, enquanto o grupo de L2 apresentou diferenças estatisticamente significativas ao GC.

Recomenda-se, aos pesquisadores, a adoção do paradigma da alta variabilidade, ou testar e discutir com dados práticos e teoricamente os benefícios de treinamentos com baixa variabilidade. Geralmente, quando um modelo ou metodologia é inserida em uma área, após a fase de incertezas e rejeições, a maior parte dos pesquisadores adota aquele paradigma sem refletir melhor ou aplicarem testes mais aprofundados. Apesar de diversos autores, como Aliaga-Garcia e Mora (2009), Rato (2014) e Carlet e Cebrian (2019), encontrarem ótimos prognósticos para os trabalhos de HVPT, é sempre preciso testar e questionar, replicar metodologias para confirmar suas eficácias e inovar com propostas novas. Qin *et al.* (2021) questionam o paradigma da alta variabilidade, adotado por um número massivo de pesquisas após a década de 1990, no sentido de testar se estes inputs com falantes de diversas idades, sexos e sotaques diferentes realmente auxiliam em alterações significativas nos resultados dos participantes de treinamentos.

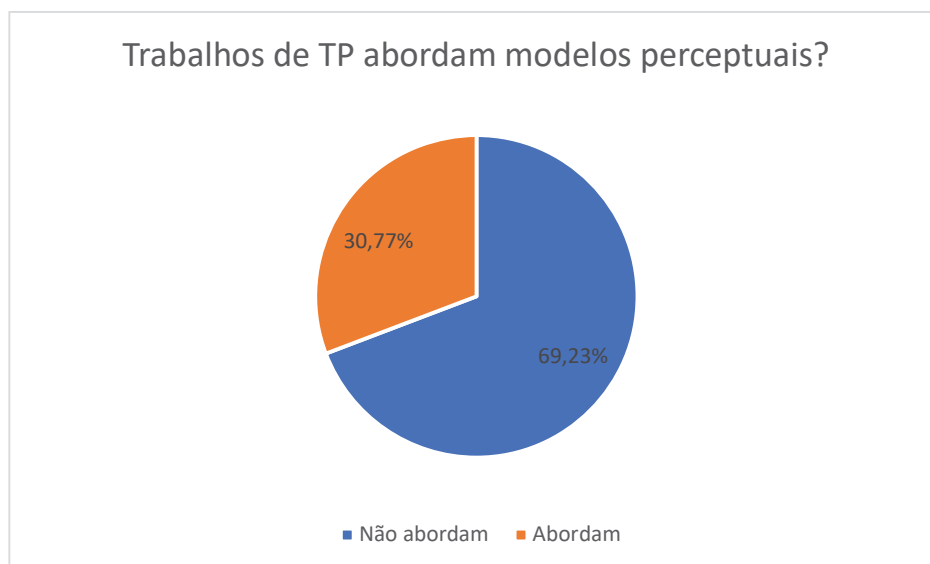
4.14 Adoção de modelos teóricos perceptuais

O descritivo 14 tratou sobre a presença ou não de modelos teóricos e perceptuais nas pesquisas de TP. Os três principais modelos teóricos observados nesta pesquisa foram o SLM, o PAM e o NLM¹³. 36 trabalhos (30,77%) apresentaram *qualquer* menção a teorias fonológicas/ modelos perceptuais;

- 34 destes 36 citaram o SLM;
- 29 destes 36 citaram o PAM;
- 11 destes 36 citaram o NLM;
 - Destes 36 estudos com menções teóricas:
 - 17 estudos (47,22%) apenas citam modelos;
 - 19 estudos (52,78%) retomam, aplicam ou testam modelos teóricos.

A seguir, represento, no Gráfico 9, os dados do descritivo 14.

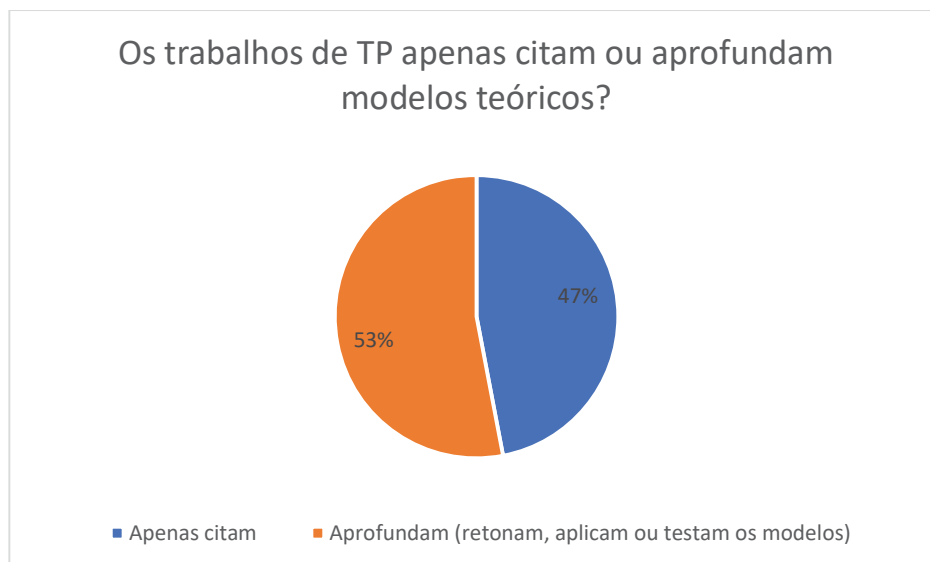
Gráfico 8: Abordagem de modelos perceptuais por trabalhos de TP



¹³ Não discriminei aqui quando os autores citavam o PAM/PAM-L2 e SLM/SLM-r, uma vez que, apesar destas versões revisadas serem mais atuais e demonstrem a evolução do pensamento dos autores, ainda mantêm suas bases.

Fonte: O autor.

Gráfico 9: Os trabalhos de TP apenas citam ou aprofundam modelos teóricos?



Fonte: O autor.

Primeiramente, apesar dos TP serem estudados e praticados desde as décadas de 1960, 70 e 80, apenas cerca de 30% dos trabalhos pelo menos mencionam (e praticamente metade *apenas* menciona) modelos teóricos de percepção da fala ou de aquisição de segunda língua.

Praticamente metade das pesquisas incluídas na RS apenas citaram modelos de percepção ou de aquisição fonético/fonológica como os descritos nesta subseção. Todavia, é importante observar como os 19 estudos aprofundam a discussão dos dados de um TP com a teoria. Como na seção 4.11, relacionada aos testes de generalização, seria por demasiado extenso descrever cada um dos 19 trabalhos, por isso foi feito um recorte de seis pesquisas, além da adição de pesquisas já descritas em outras subseções.

Lacabex *et al.* (2008) citaram os modelos do SLM e do PAM para discutirem seus dados. Os autores estudaram a capacidade de falantes de L1 espanhol identificarem a schwa em inglês enquanto passavam por dois tipos de treinamento: articulatório e auditivo. Os autores ofereceram um treinamento de doze semanas que ocorreu ao longo de três meses, sendo as três primeiras sessões de

conscientização, as quatro seguintes contendo o segmento prático e as últimas, revisões. Os alunos tinham aulas de inglês três horas por semana e estavam, em média, 80% motivados para melhorar. O treinamento perceptivo consistiu em exercícios de discriminação, enquanto o treinamento de produção proporcionou ao grupo pistas visuais e articulatórias, além de feedback fornecido pelo responsável pelo treinamento. Quanto aos resultados, os autores descobriram que o treinamento articulatório contribuiu para a percepção, e as habilidades perceptivas se generalizaram para contextos em que não foram instruídos. No entanto, este trabalho ecoa com Iverson e Evans (2007, 2009) de certa forma, no sentido de que a melhoria para novas categorias fonéticas (sons não nativos) não foi tão grande quanto o esperado e teorizado por modelos de assimilação como o Speech Learning Model (SLM), de Flege (1995) e Perceptual Assimilation Model (PAM), de Best and Tyler (2007).

Florencio (2021) é uma das pesquisas que apresenta e discute de forma mais aprofundada os dados de um TP alinhado aos modelos do SLM e do PAM. A pesquisadora investigou os efeitos de uma sessão de treinamento fonético com várias tarefas visando vogais /ɪ/ e /i:/; /ʊ/ e /u:/; /æ/ e /ɛ/; /ʌ/ e /ɑ/ para 15 participantes brasileiros aprendizes de língua inglesa. A pesquisa contou com testes de identificação, discriminação, e identificação ortográfica de palavras, provenientes de três *corpora*, o que foi realizado por 15 brasileiros natos (sem grupo controle), de diferentes níveis de proficiência. Florencio (2021) apontou que os resultados obtidos em sua pesquisa eram condizentes com o SLM e o PAM para a terceira parte do treinamento (identificação ortográfica), na qual as assimilações das vogais foram percebidas como pertencentes às mesmas categorias fonológicas. A autora concluiu que os pares /ɪ/ e /i:/ foram assimilados ao inventário da L1 (PB) como /i:/, o par /ʊ/ e /u:/ como o som do português, /u:/, e par /æ/ e /ɛ/ como o som /ɛ/, da L1. A autora também afirmou que os participantes mais jovens (em idade adolescente) da pesquisa teve uma porcentagem de acertos maior que o grupo de adultos. Novamente, Florencio (2021) levantou o SLM para discutir que, muito embora a idade não seja determinante para aquisição de uma L2, Flege constatou em

pesquisas de 1981, 1987, 1988, 1995 2009 que os indivíduos mais jovens geralmente apresentavam menos sotaque estrangeiro.

Carlet e Cebrian (2019) trazem os modelos do PAM, SLM e NLM em sua pesquisa. Os autores investigaram se um regime de TP de cinco sessões de 30 minutos espalhadas em um período de 10 semanas influenciaria a percepção e produção de vogais do inglês por nativos de espanhol e catalão. Os participantes foram 52 ao todo e divididos em três grupos (os que foram treinados com identificação, discriminação e um grupo controle). Esta pesquisa adotou o paradigma da alta variabilidade para as vozes presentes, adotando quatro locutores diferentes para os estímulos, além de testar generalização para novos falantes, palavras reais (uma vez que os participantes foram treinados com palavras inventadas) e novas palavras inventadas, além de testar a retenção dois meses depois do final das sessões de treinamento. Os modelos do NLM e do SLM são trazidos à discussão ao se analisar dois fatores referentes aos resultados. Apesar de, de forma geral, os autores haverem reportado resultados significativos de melhora na percepção e produção do inglês por parte dos participantes (ambos os grupos experimentais performaram significativamente melhor que o grupo controle), notou-se que o teste de ID foi mais eficaz que o de DIS para a produção. Além disso, apenas o teste de ID promoveu generalização para palavras reais. O SLM e o NLM postulam que os ganhos de percepção necessitam vir antes dos ganhos de produção, e os participantes de sua pesquisa apresentaram melhoras na percepção mais acentuadas do que na produção, o que deixa uma pista que percepção e produção parecem se desenvolver de forma diferente.

Aliaga-Garcia (2010) citou os modelos do PAM e do SLM para discutir os dados de sua pesquisa. A autora investigou os efeitos de treinamentos de ID e DIS em um sistema de alta variabilidade para 64 nativos de catalão/espanhol na língua inglesa. Os sons vocálicos treinados foram 11 vogais (/i ɪ e æ ʌ ɑ ɜ ɒ ɔ ʊ u/, e os resultados apontaram que o tipo de treinamento possuía interação significativa com o tipo da vogal, a condição da vogal (seja ela natural ou sintetizada), e a dispersão dos erros para cada tipo de vogal examinada. A autora citou o postulado do PAM

que prevê que as vogais da L2 que acusticamente são parecidas com às da L1 tendem a serem erroneamente assimiladas em uma categoria já existente da L1. Além disso, a autora também utiliza o SLM para citar os espaços fonológicos das categorias da L1 e que o aprendizado de sons muito diferentes aos de categorias já formadas é mais fácil. A autora encontrou que os acertos para vogais naturais do grupo de ID aumentaram 10,61% após 10 sessões de treinamento, dentre outros resultados significativos.

Leong *et al.* (2018) citam o SLM, PAM e NLM em sua investigação. Os autores realizaram um treinamento de alta variabilidade de três sessões de 20 minutos com alvo na língua inglesa para 28 malaio multilíngues. Os autores reportaram que os resultados da pesquisa foram significativos e se mantiveram em um teste de retenção realizado seis meses após as sessões de treinamento. Os autores se utilizaram do PAM e do NLM para explicar que os falantes nativos de mandarim/malaio assimilaram as vogais /ɛ/-/æ/ na mesma categoria pela falta de diferenciação destes sons no inglês malaio. Segundo os autores, utilizando-se do SLM para explicar isto, uma distância entre os contrastes facilita a dissimilação de categorias. Ademais, este pressuposto do SLM corrobora que as três sessões de discriminação foram refletidas na acurácia das respostas dos participantes e na fixação destas novas categorias.

Savo e Peltola (2019) citaram o SLM e o PAM como modelos de percepção para sua pesquisa. Os autores treinaram, em um regime de curta duração, 13 falantes de árabe para as vogais finlandesas /y/ e /ø/ presentes em palavras inventadas. O regime de treinamento durou cerca de meia hora, dividido em blocos de estímulos. A pesquisa contou com alta variabilidade, por possuir inputs de quatro falantes diferentes. Os autores discutiram o SLM no sentido de avaliar que, em uma primeira hipótese, /y/ é um som similar a um som que existe no inventário fonético do árabe, e acaba sendo colocado dentro da categoria fonética de /u:/. Em uma segunda hipótese, os autores apenas mencionaram que /y/ é um som novo e por isso deve ser aprendido como não-assimilável, mas notam que a primeira hipótese é a mais forte. Para /ø/, os autores concluem que tanto para o SLM como para o

PAM, o som é classificado como não-assimilável, e está localizado entre os sons de /i/ e /a/ que estão presentes no inventário fonético nativo do árabe, porém longe demais para ser assimilado para qualquer um destes inventários. Os autores mencionam, inclusive, que alguns falantes buscaram categorizar /ø/ como /o/, muito embora o próprio som de /o/ não esteja presente no inventário fonético do árabe, o que ressalta que vogais com arredondamento de lábios podem afetar a percepção e produção de uma maneira diferente das vogais sem arredondamento de lábios.

A pesquisa de Wong (2015), descrita na Subseção 4.11.2, ao descrever como ocorre a generalização para novos falantes, notou que, em seu regime de treinamento realizado com 35 participantes nativos de cantonês e aprendizes de inglês, o par de vogais anteriores mostrou consistentemente melhores resultados na percepção e na produção do que o par de vogais posteriores. A autora, na discussão dos resultados, apontou que isso iria contra o preceito do SLM que previa que fonemas da L2 mais próximos da L1 seriam mais difíceis de serem assimilados. Contudo, a autora apontou que isso talvez se deva ao fato de que o par de vogais posteriores é muito mais raro e, pela sua ocorrência pouco frequente na fala, pode ser mais difícil para que os participantes consigam distinguir na prática. A pesquisa de Ueda e Hashimoto (2019), também descrita em 4.11.2, que havia treinado falantes japoneses para três pares de consoantes da língua inglesa, também se utilizou de modelos perceptuais para a discussão dos resultados da pesquisa. Os autores analisam que o PAM assumia que o par /r-/ do inglês seria assimilado para a mesma categoria por falantes japoneses, segundo o preceito do *single-category assimilation*, enquanto os outros dois pares estudados, que apresentaram melhoras estatisticamente significativas, /b-/v/ e /s-/θ/, caem na categoria de *category-goodness difference*, sendo mais fáceis de serem adquiridas pelos participantes japoneses. Ademais, os autores citam a pesquisa de Sakai e Moorman (2018), uma Metanálise que afirmou que os TP seriam mais eficientes para consoantes plosivas do que para consoantes líquidas.

Como foi possível observar a partir da amostragem descrita nos parágrafos anteriores, não existem muitas discussões aprofundadas relacionando teorias e

modelos e a prática, pelo menos não em artigos acadêmicos. Cabe notar que muitos dos artigos etiquetados como aqueles que discutem os modelos, apenas cita, seus postulados uma ou duas vezes em toda a sessão de resultados. A questão do número de páginas que as revistas acadêmicas exigem pode ser uma variável importante nesta discussão, uma vez que discutir teorias requer espaço. Apesar disso, recomenda-se, aos pesquisadores, mais discussões teóricas, sempre que possível. Existem diversos modelos de aquisição e desenvolvimento de segunda língua, e, muitas vezes, tem-se a impressão de que estes modelos são mencionados apenas para serem mencionados, e não são testados ou discutidos. É interessante pegar o conjunto de dados e estatísticas da pesquisa e tentar interpretar com as lentes dos modelos e teorias adotadas, e não apenas citar o ponto de partida. É compreensível, no entanto, que muitas vezes as revistas e livros exigem um máximo de páginas ou de caracteres para uma publicação, porém muitos trabalhos parecem vazios por simplesmente executarem um TP sem mais nenhuma base exceto a própria metodologia em si. Como apontado por um membro da banca de qualificação, quiçá seria interessante averiguar a relação da presença de discussões teóricas de concepções de línguas e modelos com os anos de publicação. Porém, não foi possível estabelecer correlação: mesmo notando que a maioria dos trabalhos que apresentavam discussões teóricas eram, de facto, das décadas de 2010 e 2020, a maioria dos trabalhos presentes na RS também são destas décadas.

4.15 Uso da tecnologia

O descritivo 15 tratou da tecnologia nos trabalhos de TP. A verificação de como as tecnologias estavam presentes nos trabalhos tiveram como base teórica, principalmente, pesquisas como a de Paiva (2008, 2019): é importante entender que as tecnologias estão sempre em mudança, desde as primeiras formas de escrita até os *smartphones* contemporâneos. De acordo com essa autora, é comum que as tecnologias sejam recebidas com desconfiança e temor e apenas depois

sejam incorporadas às práticas pedagógicas. Deste modo, estando ciente de que TP em papel e caneta possuem tecnologia também, diferenciam-se aqui os TP que contaram com *softwares* computacionais.

- 65 (55,55%) das pesquisas reportaram quais *softwares* estavam utilizando.
- 52 (44,45%) das pesquisas omitiram / sonegaram os *softwares* que estavam utilizando.

Como será discutido com mais profundidade na subseção 4.17, muitos trabalhos indicam que estão utilizando algum *software*, mas simplesmente não especificam qual é. Alguns trabalhos reportaram que utilizaram o computador apenas como caixa de som (Mildner e Tomić 2007, por exemplo, deliberadamente falam que usaram caixas de som) para reproduzir os estímulos do treinamento, o que teoricamente implica ao menos na utilização do *software* de música do sistema operacional, mas este trabalho é sobre TP, e não informática. Muitos dos trabalhos, é importante reafirmar, realmente não reportam quais *softwares* foram utilizados em suas pesquisas.

4.15.1 *Softwares* de treinamento

Segue-se uma apresentação dos *softwares* que mais foram reportados (n=42) nas pesquisas que efetivamente indicaram o *software* utilizado

- 12 vezes TP (18,47%)
- 6 vezes Praat (9,23%)
- 5 vezes Psychology Software Tools Inc (7,7%)

Ademais, 17 programas foram citados apenas uma vez: Paradigm; EAC; WAVES1.ESPS; MS Access; Matlab; R Core Team; PESAA; Bliss; PsychoPy; DmDx; Exbuilder; Ultrasonix SonixTablet; Goldwave; Adobe Audition 1.5; PERCY; Sanako Study Student 7.20.

Seis pesquisas relataram que desenvolveram *software* próprio, mas não revelaram detalhes.

Conforme discutido em 4.1, era previsível que a maioria dos trabalhos adotasse programas computacionais em seus treinamentos, uma vez que uma alta porcentagem de trabalhos após os anos de 2010 foram incluídos na bibliografia. Nota-se também que o programa TP (Rauber, Rato, Kluge, Santos e Figueiredo, 2012), desenvolvido no Brasil por pesquisadores brasileiros, está presente em muitos trabalhos de outras línguas e utilizado por pesquisadores de outras nacionalidades.

4.15.2 *Softwares* auxiliares

Além dos softwares reportados aqui como principais, ou seja, utilizados para o TP, foram mapeados também 40 softwares¹⁴ auxiliares (ou secundários) utilizados principalmente para tarefas de análises acústicas, estatísticas ou tratamento dos inputs utilizados nas pesquisas:

- 24 (60%) utilizaram o Praat como *software* auxiliar.
- 3 (7,5%) utilizam o SPSS como *software* auxiliar.

Outros programas computacionais utilizados foram: EFET; Bliss Baldi (part of CSLU); Cool edit 2000 (reportado duas vezes); WAVES1/ESPS; Adobe Audition 1.5 (reportado duas vezes); Core Psych / R; BESA; HLSyn; Excel; Goldwave.

4.16 Tarefas de treinamento

Por fim, o descritivo 16 tratou das tarefas de treinamento presentes nos TP. Conforme discutido no referencial teórico do trabalho, as principais tarefas presentes nas fases de testes e treinamentos são de ID (identificação, na qual o participante escuta um estímulo e precisa escolher a opção correta, dentre várias

¹⁴ Apesar deste número parecer pequeno (apenas 40 em 117), afirmo que pesquisei profundamente as pesquisas desde os resumos até nas referências para achar indícios dos softwares que foram utilizados, e mesmo assim apenas 65 das 117 reportaram programas computacionais principais de treinamento e 40 das 117 reportaram estes programas auxiliares.

opções apresentadas a ele) e DIS (discriminação, na qual o participante escuta estímulos e precisa escolher se são iguais ou diferentes. Geralmente, autores chamam esta tarefa de AX *discrimination*). A distribuição das tarefas de treinamento nas pesquisas da RS pode ser observada a seguir, em um universo de 117 pesquisas:

- No total, 93 pesquisas (84,54%) utilizam tarefas de ID, sendo que 57 (51,82%) delas utilizam somente tarefas de ID, 2 (1,82%) ID mais uma tarefa extra e 34 (30,91%) ID e DIS na mesma pesquisa;
- 49 (44,54%) pesquisas utilizam tarefas de DIS, sendo que apenas 15 (13,63%) utilizam apenas testes DIS em seus trabalhos.
- 9 (8,18%) não deixam claro quais tarefas estão presentes no trabalho, embora as tarefas em si sejam mencionadas, porém faltam descrições detalhadas.

Nota-se a preferência pelas tarefas de identificação e que quase 1/3 das pesquisas adotam mais de um tipo de tarefa (geralmente ID e DIS), mas que apenas uma minoria dos trabalhos opta apenas por tarefas de discriminação de sons. Carlet (2017), por exemplo, inquiriu seus participantes acerca de qual tarefa (identificação ou discriminação) cada um julgava mais estimulante. Os participantes preferiram os testes de identificação e consideraram os de discriminação mais difíceis.

4.17 Considerações metodológicas sobre informações e organizações de artigos de TP

Durante todo processo de compilação e a análise de dados da presente tese, certas dificuldades encontradas estão conectadas com a dificuldade de encontrar informações importantes nos artigos estudados. Divido estas dificuldades em três:

- (a) Falta de informações importantes (como, por exemplo, não informar quantos participantes a pesquisa possuía);
- (b) Informações imprecisas ou incompletas (como, por exemplo, mencionar que utiliza *software*, porém não especificar qual foi eleito);
- (c) Informações presentes, porém espalhadas de forma desordenada pelo artigo (como, por exemplo, reportar o número de participantes apenas no resumo, informar o número de sessões na metodologia e o número de integrantes do grupo controle nos resultados).

Pesquisas como a Wang *et al.* (2001) apresentam o modelo do computador, do microfone, mas não o nome do *software* utilizado na pesquisa, por exemplo. E houve dezenas neste mesmo padrão. Duas vezes, ao pesquisar por “min, minutes, hour, hora, minutos”, encontrou-se trabalhos que reportaram que as sessões duraram *a few minutes* (alguns minutos). Pergunta-se: quantos minutos são alguns minutos? É possível estimar, como muitos pesquisadores fazem, de 5-10min, para fins de cálculo, contabilizei 7.5min por sessão em casos como estes, utilizando aritmética simples. Porém em pesquisas que reportaram que as sessões duraram *a few minutes* ou *a few hours*, os dados foram simplesmente descartados, porque esta descrição vagante é insuficiente para ser aqui quantificada. Jeong e Cho (2015) é uma pesquisa escrita em língua coreana e presente na RS: traduzindo o *paper* com o *Google Translate* achei, facilmente, informações importantes, como traduzindo semana (주) e *software* (소프트웨어) e com a presença de palavras-chave importantes no alfabeto latino. Nem a barreira de uma língua oriental com outro sistema de alfabeto impossibilitou a pesquisa das informações necessárias para os descritivos 4.9.2 e 4.15. Todavia, para achar informações simples e essenciais para estudos em TP (como número de sessões, durações das sessões e duração do período de treinamento) em alguns artigos escritos em inglês, foi necessário até mesmo aplicar regras de aritmética em tabelas que continham as informações sobre o treinamento. Para exemplificar, um trabalho não mencionava

quantos minutos as sessões duravam, apenas havia uma tabela com o dia da semana e a quantidade de minutos de treinamento, necessitando que eu mesmo realizasse a soma de vários dias com minutagens diferentes para obter a informação. Sugiro, portanto, que os futuros pesquisadores da área deixem as informações de seus respectivos treinamentos acessíveis para que os pares possam consultar os dados mais facilmente. Pesquisas como Zhang *et al.* (2013) foram especialmente difíceis de decifrar, uma vez que os dados estavam espalhados por todas as partes do artigo.

Outras discussões e retomadas de objetivos e da pergunta da pesquisa se encontram a seguir, no Capítulo 5.

5. CONSIDERAÇÕES FINAIS

Apresentados os resultados e as discussões, é possível partir para a retomada da pergunta de pesquisa e objetivos específicos. Além dessas considerações, proponho uma Agenda de pesquisa para a área de TP. Por fim, descrevo as limitações do presente trabalho e sugestões para as próximas pesquisas

5.1 Considerações sobre os objetivos e Agenda de pesquisa

No resumo e introdução do trabalho, a pergunta de pesquisa, o objetivo principal e os objetivos específicos foram apresentados nesta ordem. Proponho, para a conclusão do trabalho, revistar cada um em ordem inversa, começando pelos objetivos específicos:

5.1.1 Considerações sobre o objetivo específico 1 (sujeitos):

Retoma-se, primeiramente, o primeiro objetivo específico: “reportar e avaliar acerca dos sujeitos presentes nas pesquisas de TP”

Após a revisão sistemática, parece essencial possuir GC de nativos/fluentes na língua estrangeira ou de participantes não treinados para avaliar se o treinamento possui alguma diferença entre grupos (que pode ser testada posteriormente através de testes estatísticos como um teste t).

Recomenda-se, como parte de uma Agenda de pesquisa, uma proporcionalidade entre número de participantes e descrições quanti/qualitativas. Ou seja, não há problema em possuir um baixo número de participantes no grupo experimental, desde que a descrição qualitativa seja minuciosa e cuidadosa, preferencialmente discutindo os dados estatísticos de mudanças na percepção/produção do estudante a cada sessão/semana do treinamento. Por outro lado, é extremamente importante pesquisas com uma grande quantidade de alunos, uma vez que isso permite uma maior robustez estatística para validar o treinamento como efetivo. É sabido que as pesquisas de TP acabam por ter evasão de alunos por serem pesquisas longitudinais, geralmente contendo alunos de universidades que podem trancar/abandonar os cursos, ou até mesmo começarem novos estágios e/ou trabalhos. Neste sentido, ter muitas dezenas ou até mesmo centenas de alunos ajudam a ter um grande n amostral caso a evasão se concretize. Nota-se também que não é prática aceita no Brasil pagar os participantes, mas quiçá devesse ser uma luta para que isso acontecesse, uma vez que os participantes podem ficar mais motivados ao exercer este trabalho (que além de aumentar suas taxas de percepção de sons de línguas não-nativas, poderia os remunerar pela ciência).

O objetivo específico 1 foi bem-sucedido, apesar da maioria das pesquisas não terem muitos dados analisando individualmente os participantes (como Milan, 2019) e o foco do objetivo foi ter apenas descrito o número e perfil dos participantes.

5.1.2 Considerações sobre o objetivo específico 2 (experimentos):

Retoma-se, primeiramente, o segundo objetivo específico: “reportar e avaliar acerca dos experimentos realizados nas pesquisas de TP”

Com relação aos quatro primeiros descritivos, referentes, respectivamente, ao ano de publicação, língua materna, língua-alvo e foco fonético, as escolhas de cada pesquisador são subjetivas e devem ser respeitadas. Apesar de haver uma dominância de pesquisas com o inglês como alvo de segunda língua, mais línguas podem ser pesquisadas, e os treinamentos podem ser feitos a partir de outras línguas (como Português para francês ou vice-versa, por exemplo). O foco fonético também pode ser expandido. Existe uma maioria de pesquisas em segmentos, mas questões de prosódia e de supra segmentos também são importantes, quiçá mais importantes para que um falante compreenda de forma efetiva uma língua estrangeira.

Vale lembrar também, como já mencionado na Metodologia, que trabalhos de revisão e Metanálise são um recorte e que não podem ter a arrogância de, de forma ditatorial, dizer como os trabalhos futuros devem ser. Estas são sugestões encontradas após o estudo de centenas de trabalhos da área, e que buscam fortalecer a área dos treinamentos como um todo.

Sugere-se, aos pesquisadores, uma normalização de um protocolo de pesquisa. Apesar de dezenas de artigos citarem a necessidade de novas pesquisas experimentais e a necessidade de testar novas metodologias, alguns requisitos mínimos poderiam ser sugeridos para que pesquisas com o mínimo de rigor e critério pudessem ser efetuadas.

A seguir, deixo uma sugestão do que uma pesquisa básica em TP poderia ter, com base no estudo aprofundado de centenas de trabalhos da área e de 16 descritivos de como as pesquisas são realizadas:

- O treinamento deve levar em conta o foco fonético/segmentos em questão para saber a necessidade da quantidade de sessões envolvidas. Um número arbitrário pensado como modelo, nas que não precisa ser seguido à risca, seria de oito sessões, realizadas em oito dias diferentes ou em oito semanas diferentes, para que haja tempo de consolidar o conhecimento adquirido (diversos trabalhos comentaram sobre a limitação ser a o número de sessões ou a duração do período de sessões);

Além disso:

- O treinamento deverá fornecer *feedback* imediato, além de universal;
- Caso possível, o treinamento deverá conter testes de generalização para novos segmentos/ e/ou falantes e/ou para a produção da fala devem ser realizados;
- Caso possível, o treinamento deverá realizar coletas postergadas para verificar a retenção do conhecimento.

A literatura aponta para que haja *feedback*, preferencialmente imediato e universal ao final da sessão de treinamento. Com relação aos testes de generalização e de retenção postergada, sabe-se que geralmente o tempo é uma variável em falta para a maioria dos pesquisadores. Por isso, aproveitar períodos como doutorados de quatro anos é ideal para que haja tempo de testes de retenção postergada.

Recomenda-se, como parte de uma Agenda de pesquisa, uma melhor apresentação dos objetivos e da descrição da pesquisa não só no título e no resumo, como também no corpo dos textos. Não foi incomum precisar inferir ou deixar de fora algum dado de pesquisa pelos autores não terem escrito de forma explícita dados como duração do período das sessões, número de participantes do grupo controle ou os tipos de testes ou treinamentos realizados. Trabalhos como Fouz-González e Monpean (2021) e Hardison (2005) são exemplos de artigos com informações fáceis de serem acessadas e interpretadas pelo leitor. Como diz a epígrafe da presente pesquisa, pegada emprestada de Graciliano Ramos: “A palavra foi feita para dizer”. Isso fica evidente pelo descritivo (15), sobre uso da tecnologia: praticamente metade dos trabalhos não deixou claro qual o programa computacional foi utilizado. Pelo menos cinco indicaram que estavam utilizando um *software* produzido com a finalidade de atender aquela pesquisa, mas não forneceram nenhum detalhe.

Sugere-se como Agenda de pesquisa uma padronização para o nome da área ou quiçá uma junção de nomes. Como discutido em 3.1, foi realizada uma metodologia específica para se nomear a presente tese: Revisão Sistemática de Treinamentos Perceptuais de Sons Não-nativos (TPSN). Existe também a possibilidade de adotar Treinamento Perceptual de Sons de L2 (TPSL2), ou até mesmo Treinamento Fonético Perceptual (TFP), o que já diferenciaria estas pesquisas dos treinamentos somente de pronúncia, uma vez que estaria explícito no título a característica perceptual. Uma das partes mais trabalhosas desta pesquisa, além de resumos e títulos que careciam de informações importantes com relação às pesquisas que estavam sendo realizadas, foi precisar procurar pelos três (seis, com a língua inglesa) termos. Na área da fonoaudiologia, por exemplo, todos os trabalhos com os termos “*Phonological training*” e “Treinamento fonológico” indicavam nitidamente à qual área o trabalho pertencia, e a leitura dos resumos dos trabalhos confirmavam que a área de estudo era realmente a fonoaudiologia. A busca pelos termos da fonoaudiologia não gerou dados para a revisão sistemática, porém foi útil para perceber o quão importante é a adoção de um termo fixo. Trabalhos estudados para o embasamento teórico e para tecer a metodologia da presente pesquisa, como Dias *et al.* (2015) e Saldanha (2016) descrevem a importância da precisão das palavras-chave e da pesquisa, o que fica mais difícil com um ramo com diversas nomenclaturas.

Pesquisas como Huensch (2016) e Wong (2014) sugerem que talvez seja mais eficaz realizar treinamentos perceptuais com sentenças-veículos ou estruturas frasais maiores, além de palavras individuais. Pode ser uma boa ideia mais pesquisas para comparar as diferenças entre treinamentos de somente palavras soltas e de estruturas maiores. Huensch (2016) traz teorias como a do PAM para embasar esta sugestão, no sentido de que estruturas maiores podem fortalecer criações de categorias mais robustas.

Ademais, conforme descrito nas diversas seções e subseções do capítulo de resultados, recomenda-se a realização de um número robusto de sessões, como o número sugerido de oito, além da existência de grupos-controle e com pelo menos

metade do número de participantes dos grupos experimentais, além da existência de *feedback*, aplicação de testes de retenção e generalização.

O objetivo específico 2 também foi bem-sucedido e de forma mais robusta que o primeiro e o terceiro, uma vez que grande parte dos descritivos de análise foi concentrada neste objetivo. Acredito que foi possível propor novos caminhos através das recomendações de Agendas de pesquisa para a área.

5.1.3 Considerações sobre o objetivo específico 3 (teorias e análises):

Retoma-se, aqui, o terceiro objetivo específico: “reportar e avaliar acerca de teorias e análises abarcadas nas pesquisas de TP.”

Apesar de dezenas de pesquisas adotarem o paradigma da variabilidade, pesquisas como a de Quin *et al.* (2021) desafiam esta vertente teórica, e debates mais profundos podem ser traçados em futuras pesquisas. Os pesquisadores devem discutir mais modelos teóricos de percepção de fala e aquisição de língua, uma vez que tais modelos são frequentemente citados de forma vaga e sem profundidade. Acredito que o objetivo específico 3 tenha sido bem-sucedido, apesar de que de forma amarga, por constatar que um número muito pequeno das pesquisas realmente se preocupa em aplicar e/ou testar teorias e modelos teóricos na prática, tendo a imensa maioria dos pesquisadores apenas mencionado conceitos para justificar a aparição de dados. Novamente, nenhuma pesquisa é obrigada a testar modelos teóricos caso tenham citados eles, mas ficam, para novos caminhos de pesquisa, a possibilidade de discussões mais aprofundadas neste sentido.

5.1.4 Considerações sobre o objetivo geral e a pergunta de pesquisa

Após verificar os objetivos específicos, que foram alcançados através da minha organização interna de dados proposta pela presença dos 16 descritivos, retoma-se o objetivo principal e a pergunta de pesquisa:

- Objetivo principal: Reportar e avaliar os trabalhos acadêmicos publicados sobre o tema de Treinamentos Perceptuais.
- Pergunta de pesquisa: Qual é o estado da arte da área de TP e quais podem ser novos caminhos de pesquisa?

É possível afirmar que o objetivo principal da pesquisa foi alcançado: milhares de dados provenientes de centenas de pesquisas foram reportadas e avaliadas através dos objetivos específicos. Para tal, foi necessário se debruçar sobre diferentes metodologias, como RS, RI, metanálises e MQs, o que, inclusive, pode colaborar para futuros pesquisadores de pesquisas bibliográficas de diversas áreas. A pergunta de pesquisa, por sua vez, também foi respondida: a área de TP está entre suas cinco décadas e primeiro século, mas com uma visível pluralidade de trabalhos. A área de TP vem sendo desenvolvida desde a segunda metade do século XX e busca, através de técnicas de exposição dos participantes a estímulos (áudio)visuais, melhorar a percepção da fala de L2s, através de aspectos segmentais ou suprasegmentais. Estes estímulos podem ser de alta ou baixa variabilidade, apresentados por meio de palavras-chave ou sentenças, contarem ou não com apoio visual, instrução explícita a respeito de quais são os focos da pesquisa e contar com diferentes maneiras de fornecer *feedback*. Estas pesquisas muitas vezes contam com grupos-controle de estudantes que realizam os pré-testes e pós-testes, mas não são submetidos ao regime de treinamento, e podem realizar testes de generalização para novos estímulos (palavras, frases, segmentos que não foram treinados), para a produção da fala (que geralmente pode ser avaliada de forma acústica, através de *softwares*, ou da avaliação de falantes proficientes ou nativos da L2), ou para locutores novos que não estavam presentes no treinamento. As pesquisas também podem mencionar ou discutir de forma mais aprofundada certos aspectos teóricos, tais quais a Teoria dos Sistemas Dinâmicos Complexos (TSDC), e modelos perceptuais como o SLM, o PAM e o NLM. Possíveis novos

caminhos foram apresentados, e só podemos esperar os próximos desdobramentos dos TPs nas próximas décadas.

5.2 Limitações e sugestões

Com relação às limitações dos estudos resenhados pelo trabalho atual, observa-se, primeiramente, que muitas pesquisas descritas aqui passaram por diversas limitações: número pequeno de participantes, falta de financiamento, falta de tempo pela pressão da academia em "produzir para não parecer". Esta pesquisa também possui um escopo e limites. Sugiro, para os pesquisadores de trabalhos bibliográficos que virão, ampliarem o escopo dos critérios de inclusão que foram utilizados na presente RS, incluindo mais bases de dados, ou, quem sabe escolhendo teses e dissertações, em vez de artigos, para verificar se os resultados encontrados são parecidos. O acesso a periódicos geralmente não é fácil, mesmo filiados a instituições de ensino, o que pode dificultar a exploração do conhecimento e o estudo de forma geral. Todavia, é sempre importante lembrar que critérios assertivos precisam ser estabelecidos para uma pesquisa como esta, pois, de outra forma, a tarefa da RS se torna impossível.

Com relação às limitações do trabalho atual em si, tem-se que, após uma revisão das referências bibliográficas sugeridas pela banca (qualificação e defesa), também percebi que algumas versões de dois trabalhos inseridas na análise da RS eram *pre-prints*, ou seja, versões anteriores à publicação efetiva do trabalho. Estes trabalhos foram citados com os anos do *pre-print* e contabilizados desta forma nas estatísticas. Reconheço a falta de atenção no momento da seleção, e, sinceramente, não estava tão atento com relação às discrepâncias dos anos de publicação e estes trabalhos serem versões prévias no momento da análise. Acredito que devam existir mais possibilidades de descritivos e discussões que estão além das que fui capaz de implementar aqui, e mais pesquisas estão convidadas a expandir o escopo. Além disso, é importante reconhecer que somente incluir artigos é uma limitação desta pesquisa, uma vez que, apesar de consideravelmente em menor número, as teses possuem muito mais aprofundamento que artigos, uma vez que artigos geralmente possuem limitação de número de páginas ou caracteres. Muitos artigos são provenientes ou são

constituintes das teses, embora o trabalho completo e fechado como um todo sempre está presente nas teses em si.

Futuros pesquisadores também são convidados, como colaboração à Agenda de pesquisa, a analisarem a área com outras lentes metodológicas: através de Revisões Integrativas, que contam também com trabalhos não-experimentais, e poderia integrar trabalhos teóricos ou de revisão bibliográfica ao *corpus*, como Barriuso e Hayes-Harb (2018) para encorpar o conjunto de dados e partir de uma visão crítica já realizada por outros pesquisadores; metanálises, para compilar exaustivamente cada variável numérica dos trabalhos de TP; metanálises qualitativas, para, além desta compilação exaustiva, discutir à fundo cada ponto. No anexo da presente tese encontram-se as planilhas de dados que utilizei para fazer a RS, bem como um link com acesso para todos os artigos utilizados, em PDF e de fácil acesso. Mais análises e continuações das análises presentes aqui podem ser realizadas. Quem sabe a área de TP possa virar uma área.

Conforme o trabalho de Inceoglu (2014) sugere, futuros pesquisadores também são convidados a realizarem treinamentos audiovisuais, e não apenas perceptuais, uma vez que pesquisas como a própria Inceoglu (2014), Hardison (2005) e Hazan *et al.* (2005) demonstram que o treinamento AV pode ser ainda mais eficaz que um treinamento somente perceptual. Como descrito na pesquisa de 2005, estes autores apresentam a hipótese que contrastes visualmente mais distantes (como /p/ e /l/) podem apresentar maiores diferenças no aprendizado do que pares muito próximos (como /b/ e /v/ e /l/ e /r/). Pesquisas que comparem treinamentos audiovisuais com pares de consoantes mais distantes ou mais próximas visualmente são bem-vindas para elucidar as questões de quão mais efetivos são os treinamentos AV em relação aos treinamentos perceptuais tradicionais. Ademais, os treinamentos audiovisuais também contam com a habilidade individual dos participantes em conseguir realizar a leitura labial do que está sendo dito e realmente conectar essa informação visual com o que está sendo ouvido. Talvez, para alguns aprendizes, sempre seja mais interessante um treinamento apenas perceptual, é o que concluem Hazan e equipe. Hardison (2005),

por sua vez, ao treinar estadunidenses para a prosódia do francês em 13 sessões de 40min, com a parte visual representada pelos contornos de *pitch* como *feedback*, contribuem para a prática pedagógica, aliados às práticas de *input* frequentes, vocabulário contextualizado, entre outras intervenções educacionais. Inceoglu (2014), ao treinar 60 nativos estadunidenses aprendizes de nível intermediário de francês para três vogais nasais da língua latina, notou que, após seis sessões de meia hora de treinamento perceptual, os efeitos na produção da fala dos participantes de ambos os grupos experimentais, A e AV, foram significativamente mais precisos em relação ao GC. A autora afirma que isso confirma a literatura no sentido de validar uma simples tarefa de treinamento perceptual para melhorar a produção da fala. Ademais, os efeitos encontrados para o treinamento audiovisual foram ainda melhores do que os efeitos para treinamentos somente perceptuais, com produções chegando a 90% de precisão em uma vogal nasal do francês. Em suma, os pesquisadores da área estão convidados a produzir mais trabalhos que comparem os treinamentos perceptivos e audiovisuais, uma vez que são poucas pesquisas encontradas na RS com este tema, porém que aparentam muito potencial para a área.

Através do estudo das datas de publicação dos trabalhos, observações interessantes puderam ser tecidas, como observar que não é mais tão comum o estudo do par /r-l/ da língua inglesa por aprendizes japoneses, que, segundo a ampla bibliografia de estudo da área, como Milan e Kluge (2021), foram os trabalhos inaugurais da área de treinamento perceptual. É interessante observar como trabalhos com enfoque suprasegmental só começaram a ficar populares na última década (como os três trabalhos de *stress*, *lexical stress* e *connected speech*, que começaram a partir de 2017).

Nota-se que praticamente todas, ou talvez absolutamente todas, as pesquisas de TP contam com *input* fornecido por falantes nativos, embora saiba-se que desde a década de 1980 (KACHURU, 1985) o número de falantes de inglês (L2 de cerca de 70% das pesquisas da RS) é muito maior fora dos países em que o inglês é a língua nativa. Fica a reflexão e provocação para as próximas pesquisas:

já não estaria na hora de realizar treinamentos para falas não-nativas, uma vez que são muito mais comuns? Empresas multinacionais possuem empregados de todas as partes do mundo, e pessoas que trabalham no mundo corporativo precisam se comunicar em inglês, espanhol, mandarim etc., com pessoas nativas e, no caso do inglês, com uma grande quantidade de falantes não-nativos. Como evidência anedótica da minha observação de amigos próximos que trabalham em departamentos globais, pelo menos metade das pessoas que trabalham em empresas internacionais não são nativas de inglês e passam a maior parte do tempo se comunicando em inglês com outros não-nativos. O questionamento a respeito da “supremacia” da fala nativa não foi feito durante a própria revisão sistemática, mas fica claro que a maioria, para não falar todos, dos autores desta área parecem não pensar na possibilidade de incluir treinamentos perceptuais que extrapolem este paradigma de colocar os falantes nativos em um pedestal e com a necessidade de todos os falantes não-nativos imitarem suas produções (principalmente produções, mas também a percepção da fala entra aqui). Barriuso e Hayes-Harb (2018), ao revisar a área de HVPT, descreve que uma questão crítica da área é observar se os resultados dos treinamentos de alta variabilidade produzem resultados “*targetlike*” para a produção da fala. Neste sentido, hipoteticamente, estabelecer um treinamento com o objetivo de inteligibilidade, e não apenas de imitar uma pronúncia, faz-se necessário estabelecer quais são os critérios que constituem uma pronúncia como inteligível. Lee e Hwang (2016b) endereça a questão da mudança de paradigma de um alvo de pronúncia *native like* para uma pronúncia inteligível, e cita o Lingua Franca Core de Jenkins (2000, 2002) como uma possível luz neste sentido, e sugerem um “*gradient of learnability*”, que extrapola as categorias fixas de nativos versus não-nativos. A própria Jenkins (2000) indica que alguns sons da língua inglesa podem não ser tão prejudiciais à inteligibilidade como outros, como é o caso do par de fricativas [θ] [ð], formas fracas no estresse lexical e algumas características do *connected speech*, como lugar de assimilação (como *get way* sendo pronunciado /gerawei/ em vez de /getawei/ e um inglês estadunidense. Cabe notar, porém, que é importante considerar as pesquisas sobre o tema, como Albuquerque (2019), sobre o que é considerado essencial para a inteligibilidade.

Uma outra possibilidade no sentido da Língua Franca e do inglês como língua global é manter os *inputs* de falantes nativos, porém realizar testes de generalização para a percepção falantes não nativos. É possível, também, realizar testes de generalização de produção com a avaliação de ouvintes não-nativos, demarcando a inteligibilidade das produções antes e depois do treinamento.

A pesquisa de Gauer (2024), descrita brevemente em 2.2, que foi publicada meses após o término do processo e elaboração da RS, é um exemplo de uma pesquisa que implementou diversas recomendações presentes para autores da área de TP: uma pesquisa que contou com grupos experimentais, tarefas de generalização para produção da fala, teste de retenção 10 semanas após o final do período e treinamento, um número razoável de sessões (cinco), um número razoável de participantes (48 no total), e, por ser uma pesquisa realizada ao longo de anos, contou com análises quantitativas (na escala macro, considerando as médias dos grupos) e qualitativas (observando com mais detalhes casos específicos de participantes). Ueda e Hashimoto (2019), em uma pesquisa que treinou 24 participantes japoneses para três pares de consoantes da língua inglesa, apontaram a falta de pesquisas de TP em contextos escolares. Gauer (2024) também segue o exemplo de Ueda e Hashimoto (2019) e trabalha com estudantes de ensino regular, o que pode ser muito bem-vindo para a área de TP.

Ademais, pesquisas como a de Ueda e Hashimoto (2019) ~~que~~, mesmo encontrando vários resultados significativos, reconhecem suas limitações e apontam caminhos para que novas pesquisas possam ser mais eficazes, são muito bem-vindas. Estes autores, em apenas 13 páginas, colaboram no sentido de fornecer *insights* no que concerne à quantidade de treinamento (uma vez que julgaram que seis treinamentos de 10 minutos cada, dispostos ao longo de seis semanas, pode não ser suficiente para generalização para novos estímulos), colaboram com a discussão com modelos perceptuais, uma vez que evocam preceitos do PAM para discutir seus dados, trazem um trabalho de Metanálise para tentar explicar maiores dificuldades e menos eficácias dos TP com certos tipos de consoantes, além de discutirem a importância do *feedback*. Reconhecer as

limitações do trabalho e sempre apontar para possíveis caminhos faz o trabalho, os autores, e toda comunidade de TP crescerem.

Apesar de muitas pesquisas terem citado McClelland *et al.* (2002) como fonte principal no que concerne à adoção do *feedback* como benéfico para o aumento na acurácia de respostas provenientes de um TP, seria interessante haver mais pesquisas que isolassem essa variável para verificar se realmente este é o caso. Em McClelland *et al.* (2002), foi observado que, para japoneses aprendizes de inglês e treinados para o contraste /r-l/, o grupo que recebeu o treinamento sem *feedback* não demonstrou resultados estatisticamente significativos no pós-teste para melhor precisão de identificação. Os autores afirmam, na discussão do artigo, que os testes sem a presença de *feedback* resultaram em menos efeitos de aumento na precisão da percepção e menos efeitos de generalização para outro par de palavra (como *road/load*). Como alguém que já passou por um treinamento perceptual, novamente como evidência anedótica e sem referência bibliográfica, foi muito frustrante realizar um treinamento sem *feedback*. Era impossível saber se você estava acertando mais alternativas ou errando mais, e a incerteza perdurava durante toda a sessão de treinamento. Como participante, a ausência de *feedback* foi, de longe, a experiência mais desconfortável com o treinamento perceptual que tive. Como já mencionado, Quin *et al.* (2021) buscaram comparar os treinamentos de alta e baixa variabilidade, uma vez que os treinamentos de alta variabilidade se tornaram a tendência a partir de 1993 e, por quase duas décadas, imperou de forma inquestionável como o modo de se fazer pesquisas em TP (ou seja, HVPT). Não se trata aqui de uma rebeldia infundada, e sim em verificar se os pressupostos que embasam centenas de trabalhos ao longo das décadas realmente são replicáveis e retornam com os mesmos resultados. No caso da pesquisa de Quin *et al.* (2021), nativos de mandarim foram treinados para discriminar contrastes tonais de cantonês (baixo-alto e baixo-médio), recebendo *feedback*, e com grupos experimentais que foram treinados com HVPT e outros com LVPT. Os autores concluíram que a alta variabilidade de falantes no *input* tendia a promover respostas neurais pré-atentivas dos aprendizes de nível básico de cantonês, mas não nos níveis de respostas neurais tardias.

Uma outra possível vereda cogitada para a presente pesquisa foi estabelecer um trabalho inovador que juntaria a Revisão Sistemática com uma narrativa da Historiografia da Linguística (HL). Uma vez que a RS estava feita e foi possível observar o escopo enorme do que estava à frente para ser destrinchado, esta ideia foi deixada para escanteio. Desafio, no entanto, futuros linguistas a aceitarem a empreitada e inovarem, narrando a história de uma área com os pressupostos da HL.

Batista e Bastos (2020), em um livro que aborda questões da HL, relatam que as pesquisas na área se iniciaram nos meados dos anos de 1990, quando o interesse pelo aspecto social e histórico dos acontecimentos linguísticos estava alto. Os autores citam importantes pioneiros da pesquisa como Lourad Koerate e Pierre Swiggers. Como objetivos principais da área, Batista e Bastos (2020) destacam descrever, analisar e interpretar a história do conhecimento sobre linguagem ao longo do tempo. A área transita em uma linha tênue entre a objetividade e a subjetividade, uma vez que há sempre uma lente pela qual os fatos históricos são observados. De maneira interessante, a área se coloca contra a exclusão de “textos menores” (em detrimento das grandes obras principais), o que, tematicamente, é algo que aconteceu na presente tese, uma vez que mesmo TP com poucas sessões ou participantes foram incluídos na RS, uma vez que fazem parte do todo. Batista (2020), ao descrever os métodos, destaca algumas questões importantes da HL:

- A construção interpretativa (e não descrita de forma crônica) das ideias linguísticas;
- A polissemia da palavra “história”, que pode ser entendida através de diversas definições (como um conjunto de fatos situados num espaço e tempo);
- A importância do conceito de historiografia e o escrever sobre a história;
- O paralelo com o conceito de Foucault de arqueologia do saber, uma vez que o trabalho do historiador pode se assemelhar ao de um arqueólogo.

A intersecção de um trabalho de metanálise ou de revisão sistemática ou integrativa com uma área como a HL pode ser extremamente interessante e inovadora, e, já que isto não foi feito aqui, fica o convite para um futuro mestrando ou doutorando se aventurar.

Uma das críticas tecidas ao presente trabalho, em uma apresentação prévia à qualificação realizada no Encontro Intermediário do GT de Fonética e Fonologia da Associação Nacional de Pós-graduação e Pesquisa em Letras e Linguística, foi que talvez este trabalho houvesse ficado “democrático demais” e abarcado pesquisas em fase iniciais, exploratórias e, portanto, que deveriam ter ficado de fora. A resposta para isto é simples: o objetivo foi mapear a área como ela é, sendo as pesquisas mais ou menos robustas, os treinamentos com duas, vinte ou cem sessões. Creio que seja importante levantar estatisticamente a diferença de efeito dos treinamentos em trabalhos com poucas ou muitas sessões, mas este é um trabalho que precisará ser feito no futuro. Acredito que, destes 2744 trabalhos que foram filtrados em 117 no final, foi possível separar muitos trabalhos robustos e de grande impacto para a área.

Como conclusão às considerações finais, fico feliz de ter aproveitado a oportunidade de pesquisar extensamente a bibliografia de uma área tão experimental. Obviamente são as pesquisas e experimentos que produzem mudanças e conhecimentos novos, contudo é possível perceber que muitos treinamentos focam essencialmente nos experimentos e acabam deixando de lado a discussão de teorias, comparação com trabalhos prévios ou buscam uma continuação ou testes de generalização ou retenção para ver os eventos sobre o futuro da área.

Como questionado por um membro da banca de qualificação acerca de se os trabalhos realmente testam as teorias, a resposta para a maioria é categoricamente um “não” sonoro. A maioria das pesquisas que aprofunda pelo menos um pouco as teorias (o que já consiste na absoluta minoria de trabalhos) apenas utiliza os pressupostos de concepções de língua e de modelos teóricos como base para a análise, sem testar os modelos ou concepções. É preciso inovar,

mas com critérios e pesquisas mais robustas: mais sessões, *inputs* não-nativos, metodologias consolidadas, com novas L1 e L2s e, sempre que possível, implementação de treinamentos audiovisuais, testes de retenção e generalização. É possível ir além dos *designs* tradicionais de pesquisas de TP e testar novas possibilidades. Fico muito satisfeito com o trabalho de investigação realizado nesta tese e com a contribuição que possa existir para o futuro do treinamento. Que, novamente, poderia ter uma nomenclatura aglutinadora, para não ser uma área com diversos nomes: Treinamento Perceptual de Sons Não-Nativos (nossa escolha), Treinamento Perceptual de sons de L2, Treinamento perceptual, perceptivo, fonético, quiçá Treinamento-Fonético- Perceptual, para este modelo que foi descrito neste parágrafo e em muitas outras pesquisas abarcadas pela RS.

REFERÊNCIAS

ABRAMSON, Arthur S.; LISKER, Leigh. Voice-timing perception in Spanish word-initial stops. **Journal of Phonetics**, v. 1, n. 1, p. 1-8, 1973.

AKAHANE-YAMADA, Reiko; TOHKURA, T.; BRADLOW, A.R.; PISONI, D.B. Does training in speech perception modify speech production?. In: **Proceeding of Fourth International Conference on Spoken Language Processing**. ICSLP'96. IEEE, 1996. p. 606-609.

ALBUQUERQUE, Jeniffer I. A de. Caminhos dinâmicos em inteligibilidade e compreensibilidade de línguas adicionais: um estudo longitudinal com dados de fala de haitianos aprendizes de português brasileiro. **Universidade Federal do Rio Grande do Sul**. 2019.

ALIAGA-GARCÍA, Cristina; MORA, Joan C.; CERVÍÑO-POVEDANO, Eva. L2 speech learning in adulthood and phonological short-term memory. **Poznań Studies in Contemporary Linguistics**, v. 47, n. 1, p. 1, 2011.

ALIAGA-GARCÍA, Cristina; MORA, Joan C. Perception and production of oral stops by Catalan/Spanish learners of English: A phonetic training experiment. **R. Monroy & A. Sánchez (eds.)**, v. 25, p. 9-15, 2008. *097 ¹⁵

ALIAGA-GARCÍA, Cristina; MORA, Joan C. Assessing the effects of phonetic training on L2 sound perception and production. **Recent research in second language phonetics/phonology: Perception and production**, v. 231, 2009. *056

ALSHANGITI, Wafaa; EVANS, Bronwen G. Comparing the efficiency of vowel production training in immersion and non-immersion settings for Arabic learners of English. In: **ICPhS**. 2015.

¹⁵ As referências com um asterisco seguido de um número (*001 até *117, ordem de organização nos arquivos da tese) pertencem à Revisão Sistemática.

ALSHANGITI, Wafaa; EVANS, Bronwen G. Investigating the domain-specificity of phonetic training for second-language learning: Comparing the effects of production and perception training on the acquisition of English vowels by Arabic learners of English. In: **The Proceedings of the International Seminar for Speech Production**. 2014. *087

ALVES, Mariane Antero. Estudo dos parâmetros acústicos relacionados à produção das plosivas do português brasileiro na fala adulta: análise acústico-quantitativa. Tese de doutorado. **Universidade Federal de Santa Catarina**, 2015.

ALVES, Ubiratã Kickhöfel; KAMPFF, Felipe Rodrigues. Efeitos de longo prazo do treinamento perceptual na percepção e produção das plosivas iniciais surdas do inglês por estudantes brasileiros: implicações para o ensino de pronúncia. **Ilha do Desterro**, v. 72, p. 375-399, 2020. *001

ALVES, Ubiratã Kickhöfel; LUCHINI, Pedro Luis. Effects of perceptual training on the identification and production of word-initial voiceless stops by Argentinean learners of English. **Ilha do Desterro**, v. 70, p. 15-32, 2017. *014

ALVES, Ubiratã Kickhöfel; LUCHINI, Pedro Luis. ¿Entrenamiento perceptivo o instrucción explícita? Percepción y producción de los patrones de voice onset time iniciales del inglés (le) por estudiantes brasileños. **Forma y Función**, v. 33, n. 2, p. 135-165, 2020. *003

ALVES, Ubiratã Kickhöfel; VIEIRA, Felipe Guedes Moreira. O treinamento perceptual no desenvolvimento dos padrões de Voice Onset Time do inglês (L2) por um aprendiz argentino. **BELT-Brazilian English Language Teaching Journal**, v. 13, n. 1, p. e42967-e42967, 2022. *002

ALVES, Ubiratã Kickhöfel; ZIMMER, Márcia Cristina. Percepção e produção dos padrões de VOT do inglês por aprendizes brasileiros: o papel de múltiplas pistas acústicas sob uma perspectiva dinâmica. **Alfa: Revista de Linguística** (São José do Rio Preto), v. 59, p. 157-180, 2015.

ALVES, Ubiratã Kickhöfel. Percepção e produção dos sons de línguas não nativas: fundamentos teóricos e questões de investigação no contexto brasileiro. Editora Zouk. 2024.

AVELLO, Pilar. L2 phonological development in speech production during study abroad. 2014. Tese de Doutorado. **Universitat Pompeu Fabra**.

BARBOSA, Isabela Vieira; DA SILVA FISTAROL, Caique Fernando; HEINZLE, Márcia Regina Selpa. CONCEPÇÕES SOBRE BILINGUISMO: UMA METANÁLISE DE PRODUÇÕES CIENTÍFICAS NAS ÁREAS DE EDUCAÇÃO E LINGUAGEM. **Revista Signos**, v. 44, n. 1, 2023.

BARRIUSO, Taylor Anne; HAYES-HARB, Rachel. High Variability Phonetic Training as a Bridge from Research to Practice. **CATESOL Journal**, v. 30, n. 1, p. 177-194, 2018.

BATISTA, Ronaldo de Oliveira; BASTOS, Neusa Barbosa. Historiografia da Linguística e o ensino de língua como objeto de análise: considerações metodológicas. Questões em historiografia da linguística: homenagem a Cristina Altman. São Paulo: **Pá de Palavra**, p. 53-76, 2020.

BATISTA, Ronaldo de Oliveira; BASTOS, Neusa Barbosa. Questões em historiografia da linguística: homenagem a Cristina Altman. São Paulo: **Pá de Palavra**, 2020.

BENZIES, Yolanda Joy Calvo. Spanish EFL university students' views on the teaching of pronunciation: **A survey-based study. University of Reading language studies working papers**, v. 5, p. 41-49, 2013.

BEREJUK, Denise. GOMES, Maria Lúcia de Castro. Análise acústico-experimental forense das medidas de duração das fases de oclusão e soltura de /p, t, k/ em amostras com e sem uso de disfarce. **No prelo**.

BERNARDO-SILVA, Otavio Augusto Rodrigues. Análise acústica da consoante /p/ por falantes bilíngues de inglês americano e português brasileiro. Dissertação de Mestrado. **Universidade Federal do Paraná**, 2021

BERTI, Larissa Cristina; MARINO, Viviane Cristina de Castro. Contraste fônico encoberto entre /t/ e /k/: um estudo de caso de normalidade e de transtorno fonológico. **Revista CEFAC**, v. 13, p. 866-876, 2011.

BETTONI-TECHIO, Melissa; KOERICH, Rosana Denise. Perceptual training on the pronunciation of /s/-clusters in Brazilian Portuguese/English interphonology. RAUBER, Andrea. S.; BAPTISTA, Barbara O.; WATKINS, Michael A. In: **New Sounds**, p. 53-66, 2007. *022

BETTONI, Melissa. Effects of a Computerized Perceptual Training Task on Non-native Pronunciation. **ICBL2013 – International Conference on Interactive Computer aided Blended Learning**. 2013. *043

BETTONI, Melissa. Immediate and long-term effects of a perceptual training in the production of word-initial/s/-clusters in Brazilian Portuguese/English interphonology. **Cadernos de Estudos Linguísticos**, v. 64, p. e022039-e022039, 2022. *035

BISSIRI, Maria Paola; PFITZINGER, Hartmut R.; TILLMANN, Hans G. Lexical stress training of German compounds for Italian speakers by means of resynthesis and emphasis. In: **Proceedings of the 11th Australian International Conference on Speech Science & Technology**. University of Auckland New Zealand, 2006. p. 24-29.

BOERSMA Paul; WEENIK, David. **Praat: doing phonetics by computer** [Computer program]. Versão 6.4.27, disponível em: <http://www.praat.org/>, 2001.

BORGES, Jorge Luis. A biblioteca de Babel. **Editora Prólogos**, 1944.

BOWEN, Caroline. Give me five. a broad-based approach to phonological therapy. In: Communications ahead. Dunedin: New Zealand Speech-Language Therapists' **Association Biennial Conference Proceedings**. 1998.

BRADLOW, Ann R.; PISONI, D.B.; AKAHANE-YAMADA, R.; TOHKURA, Y.I. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of

perceptual learning on speech production. **The Journal of the Acoustical Society of America**, v. 101, n. 4, p. 2299-2310, 1997. *032

BRADLOW, Ann R.; HANSEN, J.; ZAMPINI, M. Training non-native language sound patterns. **Phonology and second language acquisition**, p. 287-308, 2008.

BRAWERMAN-ALBINI, Andressa; KLUGE, Denise Cristina.; SILVA, Adelaide Hercília P.; ALBINI, L. C.P. Perceptual training effects on the acquisition of English stress by Brazilian learners. **International Journal of English Linguistics**, v. 7, n. 6, p. 1-15, 2017. *018

BURLESON, Deborah F. Improving Intelligibility of Non-Native Speech with Computer-Assisted Phonological Training. **IULC Working Papers**, v. 7, n. 1, 2007.

CALLAN, Daniel E; TAJIMA, K.; CALLAN, A. M., KUBO, R.; MASAKI, S.; AKAHANE-YAMADA, R. Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. **Neuroimage**, v. 19, n. 1, p. 113-124, 2003.

CARDOSO, Teresa Margarida Loureiro. Interação verbal em aula de línguas: meta-análise da investigação portuguesa entre 1982-2002. 2007. Tese de Doutorado. **Universidade de Aveiro** (Portugal).

CARLET, Angélica Foresti. L2 perception and production of English consonants and vowels by Catalan speakers: The effects of attention and training task in a cross-training study. Tese de Doutorado. **Barcelona**, 2017.

CARLET, Angelica; CEBRIAN, Juli; GAVALDÀ, N.; GORBA, C. Does metalinguistic knowledge about the L2 enhance the effectiveness of L2 perceptual training?. In: **Propuestas en fonética experimental: enfoques metodológicos y nuevas tecnologías**. 2022. p. 31-35.

CARLET, Angelica; CEBRIAN, Juli. Assessing the effect of perceptual training on L2 vowel identification, generalization and long-term effects. **A sound approach to language matters**, p. 91-119, 2019. *012

CARLET, Angélica; CEBRIAN, Juli. The roles of task, segment type, and attention in L2 perceptual training. **Applied Psycholinguistics**, v. 43, n. 2, p. 271-299, 2022. *016

CEBRIAN, Juli; GAVALDÀ, N.; GORBA, C.. Perceptual training affects l2 perception but not cross-linguistic similarity. In: **Proceedings of the 19th International Congress of Phonetic Sciences**. 2019. p. 929-933. *019

CHANDRA SEKARAN, B.; SAMPATH, P. D.; WONG, P. C. M. Individual variability in cue-weighting and lexical tone learning. **The Journal of the Acoustical Society of America**, v. 128, n. 1, p. 456-465, 2010. *110

CHAVES, Idalena Oliveira. Panorama dos estudos das vogais pretônicas no Português do Brasil: meta-análise das pesquisas desenvolvidas de 1980 a 2012. **Vozes Dos Vales: Publicações Acadêmicas**, 2013.

CHEN, Yan. Perception of tonal contrasts: high-variability perceptual training with iconic orthographic representations. In: **Sixteenth Australasian International Conference on Speech Science and Technology**. p. 225. 2016. *049

CHENG, Bing; ZHANG, X; FAN, S; ZHANG, Y. The role of temporal acoustic exaggeration in high variability phonetic training: **A behavioral and ERP study**. **Frontiers in psychology**, v. 10, p. 1178, 2019. *076

CIBELLI, Emily. Articulatory and perceptual cues to non-native phoneme perception: Cross-modal training for early learners. **Second Language Research**, v. 38, n. 1, p. 117-147, 2022. *050

CIBELLI, Emily. Training non-native consonant production with perceptual and articulatory cues. **Phonetica**, v. 77, n. 1, p. 1-28, 2019. *054

CORREIA, Susana.; DOS SANTOS RATO, A. A.; FERNANDES, J. D.; GE, Y.; KACHLICKA, M.; SAITO, K.; REBUSCHAT, P.. Effects of implicit perceptual training and cognitive aptitude on the perception and production of non-native contrasts. **OSF.IO preprints**. 2024. *041

DA LUZ AZEVEDO, Wanessa Jéssica Dinelly; DE ANDRADE BONFIM, Lucília Maria Goulart. As variações linguísticas no Brasil: revisão integrativa da literatura. **Caderno Intersaberes**, v. 8, n. 16, 2019.

DA SILVA, Elizabeth Corrêa; JUNIOR, João João Batista Bottentuit. Multiletramentos e práticas com tecnologias digitais: uma revisão sistemática da literatura em dissertações de mestrado. **REVISTA INTERSABERES**, v. 12, n. 26, p. 263-268, 2017.

DAHMEN, S.; GRICE, M.; ROESSIG, S. Prosodic and Segmental Aspects of Pronunciation Training and Their Effects on L2. *Languages (Basel)*, v. 8, n. 1, p. 74, 2023. *112

D'APOLITO, I. Sonia.; Sisinni, B.; Grimaldi, M.; Fivela, B. GPerceptual and ultrasound articulatory training effects on English L2 vowels production by Italian learners. **International Journal of Cognitive and Language Sciences**, v. 11, n. 8, p. 2174-2181, 2017. *047

DE OLIVEIRA ROCHA, Carolina Macedo; FONSECA, Aline Alves; WEISS, Denise Barros. Treinamento fonético e aperfeiçoamento da percepção e da produção oral por falantes estrangeiros de Português como L2. **Veredas-Revista de Estudos Linguísticos**, v. 24, n. 3, p. 322-343, 2020. *004

DIAS, Filipe.; TELES, N.; KARIME, P; GROHMANN, R. Memes. Uma Meta-análise: proposta a um estudo sobre as reflexões acadêmicas do tema. In: **Anais... XXXVIII Congresso Brasileiro de Ciências da Comunicação-Rio de Janeiro-RJ- 4a.** 2015.

DONG, Hanyu.; CLAYARDS, M., BROWN, H.; WONNACOTT, E. The effects of high versus low talker variability and individual aptitude on phonetic training of Mandarin lexical tones. **PeerJ**, v. 7, p. e7191, 2019. *080

DONNELLAN, Mark; LINTON, Kane; JENSEN, James C. Improving the Perception of English Vowels with High Variability Phonetic Training. **Kwansei Gakuin University Humanities Review**, v. 23, p. 175-182, 2019. *096

FENG, Xiaol.; SUN, Y.; ZHANG, J.; XIE, Y. A Study on the long-term retention effects of Japanese C2L learners to distinguish Mandarin Chinese Tone 2 and Tone 3 after perceptual training. In: **The 9th International Symposium on Chinese Spoken Language Processing**. IEEE, 2014. p. 599-603. *039

FLEGE, James Emil. Chinese subjects' perception of the word-final English/t-/ /d/contrast: Performance before and after training. **The Journal of the Acoustical Society of America**, v. 86, n. 5, p. 1684-1697, 1989.

FLEGE, James Emil. Two procedures for training a novel second language phonetic contrast. **Applied Psycholinguistics**, v. 16, n. 4, p. 425-442, 1995. *102

FLEGE, James Emil; BOHN, Ocke-Schwen. The revised speech learning model (SLM-r). Second language speech learning: **Theoretical and empirical progress**, v. 10, n. 9781108886901.002, 2021.

FLEGE, James Emil; WANG, Chipin. Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English/t-/ /d/contrast. **Journal of phonetics**, v. 17, n. 4, p. 299-315, 1989.

FLORENCIO, Hosana Alves. A criação de alvos perceptivos em L2. **Revista Intercâmbio**,v.XLIX:88-112. São Paulo: LAEL/PUCSP. ISSN 2237-759X, 2021. *005

FOUZ-GONZÁLEZ, Jonás; MOMPEAN, Jose A. Exploring the potential of phonetic symbols and keywords as labels for perceptual training. **Studies in Second Language Acquisition**, v. 43, n. 2, p. 297-328, 2021. *011

FOUZ GONZÁLEZ, Jonás. Foreign language pronunciation training with affordable and easily accesible technologies: podcasts, smartphone apps and social networking services. **Proyecto de investigación**, 2015.

FRANKLIN, Amber D.; STOEL-GAMMON, Carol. Using multiple measures to document change in English vowels produced by Japanese, Korean, and Spanish speakers: The case for goodness and intelligibility. **American Journal of Speech-language pathology**, v. 23, n. 4, p. 625-640, 2014.

GAUER, Luana. Percepção e produção do contraste da duração de vogais adjacentes a consoantes plosivas finais do inglês (l2) por brasileiros: o papel do treinamento fonético de alta variabilidade (não) associado à instrução explícita de pronúncia a partir de análises de produto e de processo à luz da teoria dos sistemas dinâmicos complexos (tsdc). Tese de doutorado. **Universidade Federal do Rio Grande do Sul**, 2024.

GHAFFARVAND MOKARI, Payam; WERNER, Stefan. Perceptual Training of Second-Language Vowels: Does Musical Ability Play a Role?. **Journal of psycholinguistic research**, v. 47, p. 95-112, 2018. *017

GIANNAKOPOULOU, Anastasia.; BROWN, H.; CLAYARDS, M.; WONNACOTT, E. High or low? Comparing high and low-variability phonetic training in adult and child second language learners. **PeerJ**, v. 5, p. e3209, 2017. *063

GOLESTANI, N.; ZATORRE, R. J. Individual differences in the acquisition of second language phonology. **Brain and Language**, v. 109, n. 2, p. 55-67, 2009. *108

GÓMEZ LACABEX, Esther; GARCÍA LECUMBERRI, M. L.; COOKE, Martin. Training and generalization effects of English vowel reduction for Spanish listeners. Recent research in second language phonetics/phonology: **Perception and production**, p. 32-42, 2009.

GONZÁLEZ-BUENO, M.; SHAW, D. J. M. The Impact of Perception Training on ELL Spelling: Preventing L1 Phonetic Transfer. **Journal of Language Teaching and Research**, v. 2, n. 6, p. 1193, 2011. *115

GONG, JIAN. YU, Y.; BELLAMY, W.; WANG, F.; JI, X. Effect of Perceptual Training with Noise on Chinese Learners' English Consonant Reception Thresholds. In: **2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)**. IEEE, 2021. p. 1087-1091. *025

GRALIŃSKA-BRAWATA, Anna. Language Experience and Phonetic Training as Factors Influencing Timing Organisation in Polish Learners of English. **Research in Language**, v. 12, n. 2, p. 185-198, 2014. *084

GRENON, Isabelle; KUBOTA, Mikio; SHEPPARD, Chris. The creation of a new vowel category by adult learners after adaptive phonetic training. **Journal of Phonetics**, v. 72, p. 17-34, 2019. *064

HARDISON, Debra M. Second-language spoken word identification: Effects of perceptual training, visual cues, and phonetic environment. **Applied Psycholinguistics**, v. 26, n. 4, p. 579-596, 2005. *008

HAZAN, Valerie.; SENNEMA, A., IBA, M; FAULKNER, A. Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. **Speech communication**, v. 47, n. 3, p. 360-378, 2005. *007

HAZAN, Valerie; KIM, Yoon Hyun. Can we predict who will benefit from computer-based phonetic training?. In: **Second Language Studies: Acquisition, Learning, Education and Technology**. 2010. *090

HAZAN, Valerie; SENNEMA, Anke. The impact of visual training on the perception and production of a non-native phonetic contrast. In: **AVSP**. 2007. p. 24. *099

HERD, Wendy; JONGMAN, Allard; SERENO, Joan. Perceptual and production training of intervocalic/d, r, r/in American English learners of Spanish. **The Journal of the Acoustical Society of America**, v. 133, n. 6, p. 4247-4255, 2013. *048

HIRATA, Yukari. Effects of sentence-vs. word-level perceptual training on the acquisition of Japanese durational contrasts by English native speakers. In: **Proceedings of the 14th International Congress of Phonetic Sciences**. 1999. p. 1413-1416. *027

HIRATA, Yukari; WHITEHURST, Elizabeth; CULLINGS, Emily. Training native English speakers to identify Japanese vowel length contrast with sentences at varied speaking rates. **The Journal of the Acoustical Society of America**, v. 121, n. 6, p. 3837-3845, 2007.

HONG, Soonhyun. Effects of Perceptual Training on the Identification of American English Monophthongal Vowels by Korean Listeners. **현대문법연구**, v. 95, p. 55-80, 2017. *046

HU, Wei et al. Shifting perceptual weights in L2 vowel identification after training. **PloS one**, v. 11, n. 9, p. e0162876, 2016. *051

HUENSCH, Amanda. Perceptual phonetic training improves production in larger discourse contexts. **Journal of Second Language Pronunciation**, v. 2, n. 2, p. 183-207, 2016. *020

HUENSCH, Amanda R. The perception and production of palatal codas by Korean L2 learners of English. **University of Illinois at Urbana-Champaign**, 2013.

HUENSCH, Amanda; TREMBLAY, Annie. Effects of perceptual phonetic training on the perception and production of second language syllable structure. **Journal of Phonetics**, v. 52, p. 105-120, 2015. *061

HWANG, Hyosung. Perceptual training on Korean obstruents for Vietnamese learners. **Phonetics and Speech Sciences**, v. 15, n. 4, p. 17-26, 2023. *028

ICARD ISASA, A. Preliminary Results on the Development of the Perception of Spanish /e/ and /ei/ by Heritage Learners vs. L2 Learners of Spanish in the Classroom. **Languages (Basel)**, v. 7, n. 1, p. 7, 2022. *107

INCEOGLU, Solène. Audiovisual and auditory-only perceptual training: Effects on the pronunciation of French nasal vowels. In: **PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING CONFERENCE** (ISSN 2380-9566). 2014. p. 104. *026

IVERSON, Paul; EVANS, Bronwen G. Auditory training of English vowels for first-language speakers of Spanish and German. In: **Proceedings of the 16th international congress of phonetic sciences**. Saarbrücken, Germany: Universität des Saarlandes, 2007. p. 1625-1628.

IVERSON, Paul; EVANS, Bronwen G. Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. **The Journal of the Acoustical Society of America**, v. 126, n. 2, p. 866-877, 2009. *111

IVERSON, Paul; HAZAN, Valerie; BANNISTER, Kerry. Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English/r/-/l/to Japanese adults. **The Journal of the Acoustical Society of America**, v. 118, n. 5, p. 3267-3278, 2005. *062

JAMIESON, Donald G.; MOROSAN, David E. Training non-native speech contrasts in adults: **Acquisition of the English/ð/-/θ/contrast by francophones. Perception & psychophysics**, v. 40, n. 4, p. 205-215, 1986.

JEROTIJEVIC TISMA, Danica. The Effect of Phonetic Training on Serbian EFL Learners' Perception and Production Accuracy. **Radovi filozofskog fakulteta**, 2012. *085

JEROTIJEVIĆ TIŠMA, Danica. Can audio-visual training equally affect phonemic and phonetic contrasts? An example of /z/ fricative production. **Nasleje** 48 , p317-335. 2021. *086

JESUS, Silvanaide Costa de et al. O uso de aplicativos no ensino da Língua Inglesa: uma revisão sistemática da literatura. **Universidade Federal da Paraíba**, 2020.

JING, G. A. O. An Auditory Perspective to How Phonetic Training Affects Listening. 2019. **4th International Conference on Education and Social Development** . *083

KUBO, R.; AKAHANE-YAMADA, R. Influence of aging on perceptual learning of English phonetic contrasts by native speakers of Japanese. **Acoustical Science and Technology**, v. 27, n. 1, p. 59-61, 2006. *116

LACABEX, Esther Gómez; LECUMBERRI, María Luisa García; COOKE, Martin. Identification of the contrast full vowelschwa: Training effects and generalization to a new perceptual context. **Ilha do Desterro: A Journal of English Language, Literatures in English and Cultural Studies**, n. 55, p. 173-196, 2008. *044

LEE, Ho-Young; HWANG, Hyosung. Gradient of learnability in teaching English pronunciation to Korean learners. **The Journal of the Acoustical Society of America**, v. 139, n. 4, p. 1859-1872, 2016a.

LEE, Ho-Young; HWANG, Hyosung; YOON, Eunkyung. Effects of perceptual training on the perception of Korean boundary tones by Chinese learners. **The Journal of the Acoustical Society of America**, v. 140, n. 4, p. 3341-3341, 2016b. *030

LEONG, Christine Xiang Ru.; PRICE, J. M.; PITCHFORD, N. J.; VAN HEUVEN, W. J. HIGH variability phonetic training in adaptive adverse conditions is rapid, effective, and sustained. **PloS one**, v. 13, n. 10, p. e0204888, 2018. *059

LI, Bei; YANG, Yike; CHEN, Si. Evaluating the Effects of Perceptual Training on Application of Mandarin Tone Sandhi Rules by English Speakers. In: **The 2018 International Conference on Bilingual Learning and Teaching**. 2018. p. 134-140. *029

LI, Yanping.; BEST, C. T., CAO, C.; ZHANG, J. Hybrid perceptual training to facilitate the learning of nasal final contrasts by highly proficient Japanese learners of Mandarin. In: **Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)**, 5-9 August 2019, Melbourne, Australia. 2019. p. 2567-2571. *038

LIN, Susan; CYCHOSZ, M.; SHEN, A.; CIBELLI, E. The effects of phonetic training and visual feedback on novel contrast production. In: **Proceedings of the 19th International Congress of Phonetic Sciences**, Melbourne, Australia. 2019. p. 899-903. *065

LOGAN, John S. Methodological issues in training listeners to perceive non-native phonemes. Speech Perception and Linguistic Experience. **Issues in Cross-Language Research**, 1995.

LOGAN, John S.; LIVELY, Scott E.; PISONI, David B. Training Japanese listeners to identify English/r/and/l: A first report. **The Journal of the Acoustical Society of America**, v. 89, n. 2, p. 874-886, 1991.

LENGERIS, Angelos; HAZAN, Valerie. The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. **The Journal of the Acoustical Society of America**, v. 128, n. 6, p. 3757-3768, 2010. *072

LENGERIS, Angelos; NICOLAIDIS, Katerina. Effect of phonetic training on the perception of English consonants by Greek speakers in quiet and noise. In: **Proceedings of Meetings on Acoustics**. AIP Publishing, 2014. *094

MATTOS, Paulo de Carvalho. Tipos de revisão de literatura. **UNESP**. São Paulo, v. 2, 2015.

MCCANDLESS, Andrew. PHONETIC TRAINING OF L2 SPANISH/u. 2021. **University of Toronto** *069

MELNIK, Gerda Ana; PEPERKAMP, Sharon. High-Variability Phonetic Training enhances second language lexical processing: evidence from online training of French learners of English. *Bilingualism: Language and Cognition*, v. 24, n. 3, p. 497-506, 2021. *066

MIDORI, I. B. A. Perceptual training and the production of English consonants by Japanese learners. *Language and Culture: The Journal of the Institute for Language and Culture*, v. 9, p. 29-40, 2005. *031

MILAN, Polliana. Efeitos do treinamento perceptual na percepção e produção dos heterotônicos por aprendizes brasileiros de espanhol. Tese de doutorado. **Universidade Federal do Paraná**, 2019.

MILAN, Pollianna; KLUGE, Denise Cristina. Treinamento perceptual. **KUPSKE, Felipe Flores; ALVES, Ubiratã K.; LIMA JR., Ronaldo M** *Investigando os sons das línguas não nativas: uma introdução*. Campinas: Editora da Abralín, p. 205-234, 2021.

MILDNER, Vesna; TOMIĆ, Diana. Effects of phonetic speech training on the pronunciation of vowels in a foreign language. 16th **ICPhS**, p. 1665-1668, 2007. *103

MORA, Joan C.; ORTEGA, M., Mora-Plaza, I.; ALIAGA-GARCÍA, Cristina. Training the pronunciation of L2 vowels under different conditions: the use of non-lexical materials and masking noise. **Phonetica**, v. 79, n. 1, p. 1-43, 2022. *105

MORA, Joan C. Assessing L2 vowel production gains after high-variability phonetic training: acoustic measurements vs. perceptual judgements. In: Proc. 3rd International Symposium on Applied Phonetics (**ISAPh 2021**). 2021. p. 9-18. *055

MUELLER, T. H.; NIEDZIELSKI, H. The Influence of Discrimination Training on Pronunciation. **The Modern Language Journal (Boulder, Colo.)**, v. 52, n. 7, p. 410-416, 1968. *117

NAITO, Yuka. The effects of musical experience and training on the perceptual learning of Japanese pitch accent by Italian native speakers. **Training**, v. 1, n. 20, p. 20. *053

NISHIDA, Gustavo. Sobre teorias de percepção da fala. Tese de doutorado. **Universidade Federal do Paraná, 2012.**

NUNES, Ana Margarida Belém; DA CÔRTE, Bruno Camacho. The Impact of Phonetic Training in the Perception and Poduction of European Portuguese Stops by Mandarin Speakers. **E-Revista de Estudos Interculturais**, n. 9, Vol. 2, 2021. *071

OLSON, Daniel J. Feature acquisition in second language phonetic development: Evidence from phonetic training. **Language Learning**, v. 69, n. 2, p. 366-404, 2019. *060

OTHMAN, Diakhshan Y.; MUHAMMED, Himdad A. The impact of phonetic training course on pronunciation performance of kurdish efl university students. **University of Diyala Digital Repository**. 2018. *093

PAIVA, Vera Lucia M. O uso da tecnologia no ensino de línguas estrangeiras: breve retrospectiva histórica. Olhares sobre tecnologias digitais: linguagens,

ensino, formação e prática docente. **Coleção: Novas Perspectivas em Linguística Aplicada**, v. 44, p. 21-34, 2008.

PAIVA, Vera Lucia M. Tecnologias digitais no ensino de línguas: passado, presente e futuro. **Revista da ABRALIN**, 2019.

PEDERSON, Eric; GUION-ANDERSON, Susan. Orienting attention during phonetic training facilitates learning. **The Journal of the Acoustical Society of America**, v. 127, n. 2, p. EL54-EL59, 2010. *068

PEROZZO, Reiner Vinicius. Sobre as esferas cognitiva, acústico-articulatória e realista indireta da percepção fônica não nativa: para além do PAM-L2. Tese de doutorado. **Universidade Federal do Rio Grande do Sul**, 2017.

PISONI, David B. ASLIN, R. N.; PEREY, A. J.; HENNESSY, B. L. Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. **Journal of Experimental Psychology: Human perception and performance**, v. 8, n. 2, p. 297, 1982.

POULSEN, Birgitte; BOHN, Ocke-Schwen; RASMUSSEN, Sidsel. Generalization of training to an untrained phonetic environment: Initial/sz/to final/sz. In: **Proceedings of the International Congress of Phonetic Sciences**. 2023. *077

PREFORS, Amy; DUNBAR, David. Phonetic training makes word learning easier. In: **Proceedings of the Annual Meeting of the Cognitive Science Society**. 2010. *091

PROTOPAPAS, Athanassios; CALHOUN, Barbara. Adaptive phonetic training for second language learners. In: **Proceedings of the 2nd International Workshop on Integrating Speech Technology in Language Learning**. UK: University of Abertay Dundee, 2000. p. 31-38. *092

QIN, Z.; GONG, M.; ZHANG, C. Neural responses in novice learners' perceptual learning and generalization of lexical tones: **The effect of training variability**. **Brain and Language**, v. 223, p. 105029, 2021. *104

QIN, Z.; ZHANG, C. The effect of overnight consolidation in the perceptual learning of non-native tonal contrasts. **PloS One**, v. 14, n. 12, p. e0221498, 2019. *109

RATO, Anabela. Effects of perceptual training on the identification of English vowels by native speakers of European Portuguese. In: **Proceedings of the international symposium on the acquisition of second language speech**. 2014. p. 529-546. *009

RATO, Anabela; CARLET, Angélica. Second language perception of English vowels by Portuguese learners: The effect of stimulus type. **Ilha do Desterro**, v. 73, p. 205-226, 2020.

RAUBER, Andreia. S.; RATO, Anabela.; KLUGE, Denise.; SANTOS, Giane R.; FIGUEIREDO, Marcos. TP-S: Testes/Treino de Percepção – Sons. Disponível em: <http://www.worken.com.br/sistemas/tp-s/>, [Application software], 2012.

REIS, Mara; NOBRE-OLIVEIRA, Denize. Effects of perceptual training on the identification and production of English voiceless plosives aspiration by Brazilian EFL learners. **New Sounds**, p. 398-407, 2007. *013

RYU, Na-Young. Assessing the Effects of Mobile Assisted Phonetic Training on L2 Sound Perception. In: **Society for Information Technology & Teacher Education International Conference**. Association for the Advancement of Computing in Education (AACE), 2021. p. 1310-1315. *095

RYU, Na-Young; KANG, Yoonjung. Web-based high variability phonetic training on L2 coda identification. In: **Proceedings of the 19th International Congress of Phonetic Sciences**. 2019. p. 43-47. *074

SAITO, Kazuya et al.; HANZAWA, K.; PETROVA, K.; KACHLICKA, M.; SUZUKIDA, Y.; TIERNEY, A. Incidental and multimodal high variability phonetic training: Potential, limits, and future directions. **Language Learning**, v. 72, n. 4, p. 1049-1091, 2022. *058

SAVO, Satu; PELTOLA, Maija S. Arabic-speakers learning Finnish vowels: short-term phonetic training supports second language vowel production. **Journal of Language Teaching and Research**, v. 10, n. 1, p. 45, 2019.

UCHIHARA, Takumi; KARAS, Michael; THOMSON, Ron I. Does perceptual high variability phonetic training improve L2 speech production? A meta-analysis of perception-production connection. **Applied Psycholinguistics**, v. 45, n. 4, p. 591-623, 2024.

ANEXOS

1- Planilha de dados da Revisão Sistemática

Guia para a interpretação da planilha:

- Código (organização interna de arquivos);
- L1 (Língua materna, ex: PB = Português brasileiro);
- L2(Língua estrangeira, ex: ING = inglês);
- Segmento: foco da pesquisa;
- Control?: se a pesquisa tem grupo controle, 1= sim, 0=não;
- Control N = número de participantes no grupo controle;
- Num ses = Número de sessões;
- Tempo: duração, em dias, do período experimental;
- Post/ret = tempo, em meses, de quando foi realizado o último teste postergado de retenção;
- HVPT = se a pesquisa trabalha com alta variabilidade;
- Feedback = tipo de feedback da pesquisa;
- Gen = se a pesquisa tem testes de generalização;
- Theory = se a pesquisa cita ou implementa algum tipo de teoria (nominalmente PAM, SLM e NLM);
- Tec: presença de *software* ou *hardware* na pesquisa;
- Dur Min S: duração em minutos das sessões;
- APP extra: *softwares* extras de análise;
- Min/sess: média de minutos por sessões, dividindo a coluna T pela I.
- Type: tipo de tarefa (ID ou DIS).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X			
1	Code	Ano	L1	L2	Segmento	Particip	Control	Control N	Num	ses	Tempo	Host	Res	HVPT	Uso	Feedback	Des	Theory	Tec	Dur	Min	S	APP	extra	min/	sess	Type
2	[1.2]	2020	PB	ING	VOInit	7	1	0	8	3	7	1	HVPT	i/pxn	1b	0			TP	90	praat*	30				ID	
3	[1.3]	2022	ESP	ING	VOInit	1	1	0	24	90			HVPT	i/pxn	1b	0			EAC	168	praat*	7				ID	
4	[1.4]	2020	PB	ING	VOInit	20	1	0	10	3				no f	1b	0			TP	120	praat*	40				ID	
5	[2.2]	2020	SEVERAL	PB	SEVERAL	14	0	0	5	6				i/pxn	1b	0			TP		praat*					ID	
6	[6.2]	2020	PB	ING	/i/ e/ i/ /æ	15	0	0	1	1				@					Praat		EFET					ID & DIS	
7	[9.1]	2003	ING	MAND	TONES	16	1	0	8	14				@	1b	0			WAVES1/ESPS		Bliss					ID	
8	[9.2]	2005	JAP	ING	e/ai/-bi/-fi	101	1	0	27	10	17		HVPT	i/pxn	1b	0			CSLU	400	Baldi (part	40				ID	
9	[9.3]	2005	JAP / KOR	ING	/p/, /t/, /k/	64	1	0	16	15	21			several		0					675		45			ID	
10	[9.5]	2018	PT	ING	/i/-/i/-/i/	19	1	0	11	5				i/pxn+g	1b	0			TP	300		60				ID & DIS	
11	[9.6]	2021	JAP	ING	/i/ /u/	47	1	0	36	10				i/pxn	1b	0			Praat							ID & DIS	
12	[9.8]	2021	SPA	ING	/i/ /æ /a/ /e/	50	1	0	25	4	28		HVPT	i/pxn+g	1b	0			TP	120		30				ID	
13	[9.9]	2018	SPA/CAT	ING	/i/ /æ /a/ /e/	38	1	0	16	10	70			i/pxn	1b	0			TP	150	Cooledit2	15				ID & DIS	
14	[9.10]	2007	BP	ING	VOInit	11	0	0	2					i/pxn	1b	0			TP	18		9				ID & DIS	
15	[9.12]	2018	JAP	ING	VOInit	16	1	0	8	3				1 HVPT	i/pxn	1b	0			TP	120	praat*	40				ID & DIS
16	[9.15]	2018	KOR	ING	/i/-/i/	12	1	0	12	10				HVPT	i/pxn	1b	0									ID	
17	[9.22]	2022	SPA/CAT	ING	/i/ /æ /a/ /e/ + initial fi	0	0	0	5					2 HVPT	i/pxn+g	1b	0			TP	130	Cooledit2	26			ID & DIS	
18	[9.25]	2018	AZERI	ING	14SSBEvc	28	1	0	10	5	17			HVPT	Yes		0			TP	250	praat*	50				ID & DIS
19	[9.27]	2017	PB	ING	Stress	30	1	0	10	1	23			i/pxn+g	1b	0			MS Access							ID	
20	[9.28]	2018	SPA/CAT	ING	s/æ, A, i, j, s, e, o/ p	0	0	0	10					4 HVPT	i/pxn+g	1b	0				180		30			ID & DIS	
21	[9.29]	2018	KOR	ING	singleten	12	1	0	12	6				i/pxn	1b	0					180		20				ID
22	[9.30]	2005	SEVERAL	MAND	TONES	18	0	0	6	21				i/pxn	1b	0			Custom							ID	
23	[9.32]	2007	PB	ING	word-initi	1	1	0	1	6	28			i/pxn	1b	0					180	praat*	30				ID
24	[9.34]	2018	CANT	ING	English Hi	35	0	0	20	10				HVPT	i/pxn+g	1b	0									ID	
25	[9.35]	2021	CANT/MAJ	ING	/p, t, k, b, c	70	1	0	65	20	35			1 HVPT	i/pxn+g	1b	0			* Exemplo	400		20				ID
26	[9.37]	2021	MAND	ING	English	18	1	0	18	4				several		0										ID	
27	[9.38]	2018	ING	FRE	French na	40	1	0	20	6				HVPT	Yes		0			matlab	180		30				ID
28	[9.40]	1999	ING	JAP	Duration	23	1	0	7	10	35			i/pxn	1b	0										ID	
29	[9.42]	2023	Viet	KOR	word-initi	75	1	0	30	5	14			i/pxn	1b	0				perceive							ID
30	[9.44]	2018	ING	MAND	tones	10	1	0	6	1				i/pxn	1b	0			R Core Tea	150		30				ID	
31	[9.45]	2018	KOR	ING	/i/-/i/-/i/	40	1	0	20	20	28			HVPT	i/pxn	1b	0			e-prime							ID
32	[9.51]	2005	JAP	ING	/i/ /i/	25	0	0	1					HVPT	Yes		0			PESAA	176		8.8				ID
33	[9.52]	1997	JAP	ING	/i/ /i/	11	1	0	12	45	24			HVPT	Yes	1b	0									ID	
34	[9.53]	2018	JAP	ING	h/-/i/-/i/	24	1	0	25	6	42			HVPT	several	1b	0				60		10			ID	
35	[9.55]	2018	KOR	ING	duration j	20	1	0	9	20	5			i/pxn	1b	0					500		25			ID	
36	[9.57]	2020	JAP	ING	initial /i/-/i/	15	1	0	8	6	1			i/pxn	1b	0					315	praat	52.5			ID & DIS	
37	[9.64]	2001	ING	MAND	tones	16	1	0	8	14				76 HVPT	i/pxn	1b	0			Bliss	22	100	WAVES1/i	0		ID	
38	[9.66]	2018	CANT	ING	/i/ and /i/	29	1	0	10	10	5			HVPT	i/pxn+g	1b	0									ID	
39	[9.67]	2018	JAP	MAND	nasal final	8	1	0	8	4	4			HVPT	i/pxn	1b	0			E-prime			0			ID	
40	[9.69]	2018	JAP	MAND	TONES	13	0	0	6	4				6 HVPT	@	1b	0				240	praat	40			ID	
41	[9.83]	2018	JAP	MAND	TONES	22	0	0	6	6				HVPT	Yes		0				320		53,3333			ID & DIS	
42	[9.86]	2024	ENG	PT	/i/-/i/-/i/	130	1	0	65	2	2			no	1b	0				The Gorilla	10		5			DIS	
43	[9.89]	2023	Malay	HAUSA	/b/ AND /h/	34	1	0	12	3	21			2 HVPT	@	1b	0			Praat	180		60			ID & DIS	
44	[9.98]	2018	BP	ING	Word-initi	23	0	0	9	28	8			i/pxn	1b	0				SPSS						ID & DIS	
45	[9.99]	2008	SPA	ING	initial /i/-/i/	34	1	0	7	12	90			indiv f	1b	0					360		30			ID	
46	[9.101]	2020	KOR	ING	Vocalic Str	54	0	0	8	8				Yes	1b	0					160		20			ID	
47	[9.105]	2018	KOR	ING	Vowels	47	0	0	30	15				2 HVPT	i/pxn	1b	0				120		4			ID	
48	[9.140]	2018	PT	ING	Vowels	6	1	0	6	2				HVPT		0										ID	
49	[9.111]	2018	ING	SPA	/d, t, r, /i/	42	1	0	20	6	17			HVPT	i/pxn	1b	0				150	praat	25			ID	
50	[9.112]	2016	SEVERAL	CANT	TONES	87	0	0		3	3			HVPT	i/pxn	1b	0				180		60			DIS	
51	[9.117]	2022	eng	hindi	VOT	21	0	0		4	15			several		0				Psycholog	120		30			DIS	
52	[9.118]	2018	MAND	ING	vowels	34	0	0		6	6			i/pxn	1b	0					360		60			ID & DIS	
53	[9.124]	2020	SPA/CAT	ING	/i/-/i/-/i/	64	1	0	10	10	35			HVPT	i/pxn	1b	0			DmDr	350	Praat	35			ID & DIS	
54	[9.125]	2021	JAP	ING	pitch	2	0	0		2	3			Yes	1b	0				Gorilla sof	50		25			ID	
55	[9.128]	2019	ENG	HINDI	VOT	19	0	0		8	16			several		0				Psychology Software Tools Inc						ID & DIS	
56	[9.135]	2021	SPA/CAT	ING	/ae/-/i/	105	0	0		4				HVPT	@		0									ID & DIS	
57	[10.1]	2009	SPA/CAT	ING	/p/-/b/-/t/-/k/	18	1	0	11	6	42			HVPT	i/pxn+g	1b	0			EyeSpeak	720		120			ID & DIS	
58	[10.2]	2007	SPA/CAT	ING	/h/-/h/-/i/	18	1	0	11	6	42			HVPT	i/pxn+g	1b	0			EyeSpeak	720		120			ID & DIS	
59	[10.4]	2022	JAP	ING	/ae/ /i/ /i/	58	0	0		6	6			HVPT	Yes		0				180	Core	30			ID	
60	[10.5]	2018	SEVERAL	ING	cons and v	28	0	0		5	5			3 HVPT	i/pxn	1b	0			Software:	60		12			ID & DIS	
61	[10.6]	2019	ING	SPA	VOT	0	0	0		3				visual f		0				Praat		psych, R				DIS	
62	[10.7]	2018	KOR	ING	palatal coc	12	1	0	12	8	10			i/pxn	1b	0				E-prime, i	160		20			ID	
63	[10.8]	2005	JAP	ING	/i/ /i/	63	0	0		10	17			HVPT	i/pxn	1b	0				300	praat	30			ID	
64	[10.9]	2007	GREEK	ING	/i/ /i/	93	0	0		10	25			HVPT	i/pxn	1b	0			Exbuilder						DIS	
65	[10.11]	2019	JAP	ING	/i/-/i/	23	0	0		10	17			HVPT	i/pxn	1b	0					300		30			ID
66	[10.12]	2019	ING	MARATHI	dental vs. retroflex	0	0	0		3				ultrasound f		0				Ultrasonix	SonixTabl					DIS	
67	[10.16]	2021	FRE	ING	Initial /h/	24	0	0		8				4 HVPT	@	1b	0			JSPsych LI	126		15,75			ID & Decis	
68	[10.21]	2022	BP	ING	esyllabic f	4	0	0		1	1			@		0					50		50			ID	
69	[10.27]	2020	ENG	HINDI	cons / vov	42	0	0		3	14			i/pxn+g		0										ID & DIS	
70	[10.28]	2021	ENG	SPA	/i/	1	1	0		1	4			@		0										ID	
71	[10.29]	2021	MAND	ING	/i/-/i/ c	40	1	0	20	7				HVPT	@		0			CHRONO	350	Praat	50			ID	
72	[10.33]	2021	MAND	PT	/p t k b d g/ either initi	0	0	0		5				2 HVPT	i/pxn+g	1b	0			TP 3.1	125		25			ID & DIS</	

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
93	[10.62]	2000	JAP	ING	/r/ /l/	8	0		3.5	2.5			/p x n		0					400		114.286		ID & DIS
94	[10.63]	2018	KURDISG	ING	connecte	50	1	50	16	112			@		0					600	praat	37.5	?	
95	[10.65]	2014	GREEK	ING	7 english c	13	1	9	5				Yes		0				TP	150		30	ID	
96	[10.71]	2021	ENG	KOR	three-way	44	0		3	7			Several		0					60		20	ID	
97	[10.76]	2019	JAP	ING	vowels	34	1	34	14	98		HVPT	@		1a	nim				140		10	ID	
98	[10.77]	2008	SPA/CAT	ING	/p t b d/	9	1	5	6	42			/p x n + g		0					720		120	ID	
99	[10.81]	2008	CANT	ING	English Vc	22	1	23	10	21		HVPT	/p x n + g		0				designed by the reser	Adobe Audition 1.5		15	ID	
100	[10.85]	2007	JAP	ING	/r/ /l/	15	1	11	7	14		@			0				CSLU	280		40	DIS	
101	[10.86]	2019	JAP	ING	Noise	55	0		15	18		HVPT	@		0								ID	
102	[10.89]	2021	FRE	DUTCH	1.Vowels	15	1	15	5	20		1 HVPT	/p x n		0					praat			ID	
103	[10.94]	1995	MAND	ING	/r/ /d/ codi	20	1	10	10	21		2	Several		0								ID & DIS	
104	[10.95]	2007	SEVERAL	CROAT	croat vowel	6	1	4	12.5	7			@		0					Praat and Excel/SPSS			?	
105	[12.5]	2021	MAND	CANT	tones	40	0		1	3		HVPT /lvp	/p x n		0				Paradigm	80		80	ID & DIS	
106	[13.1]	2022	SPA/CAT	ING	/æ/-/a/	48	1	14	4	14		HVPT	/p x n		1a	pam	sim			120	praat	30	ID, DIS, IR	
107	[13.2]	2012	MAND	ING	English vo	48	0		8	21		HVPT	/p x n		1a	sim				140	Praat (	17.5	ID	
108	[13.5]	2022	ING	SPA	/el-/el/	27	1	7	3	-		HVPT	No		1a	pam	sim				praat		DIS	
109	[13.6]	2009	ING	HINDI	Dental-ret	59	0		8	14			/p x n		0								ID & DIS	
110	[13.8]	2019	MAND	CANT	TONES	32	0		1	2			/p x n		0								ID & DIS	
111	[13.9]	2010	ING	MAND	TONES	16	0		9	14			/p x n		0				E-prime		praat		ID & DIS	
112	[13.10]	2009	SPA/GER	ING	Vowels	20	0		5	14		4	Yes		1a	sim				225		45	ID	
113	[13.12]	2023	ITA	GER	VOT	25	1	8	20	20			@		0								ID & DIS	
114	[13.13]	2018	SEVERAL	MAND	TONES	20	0						Several		1a	pam	sim			custom-designed sol	goldwave		ID	
115	[13.14]	2008	ING	JAP	LENGHT i:	20	1	39	10	5			/p x n		0							0	ID	
116	[13.19]	2011	SPA	ING	/d, b/	11	0		20	14			@		1a	pam	sim	nim		300		15	ID	
117	[13.21]	2006	JAP	ING	/r/ and /l/	41	0		37	-			/p x n		0				tape			0	ID	
118	[13.23]	1968	ING	FRE	sentence	63	1	180	25	63			@		0					1500		60	DIS	
119						30,2909		16,2687	9,00885	22,3242		6		36x1bb						244,915	média	33,3556		
120						23		12	6	14		2								#NOME?	moda	30		
121																				72				
122	General	Import				Particip	Control?	ControlN	Num	sess	Tempo	Post/Ret												
123	Convert	as dias na duração																						
124	Contabilizar	todos os grupos experimentais nas estatísticas de grupos experimentais																						
125	Um																							
126																								

2. Pasta com arquivos e 117 PDFs da Revisão Sistemática:

[htTP://drive.google.com/drive/folders/1TUS16_KoOguOioZ3NPdFXK0tEBmhsnW?usp=sharing](https://drive.google.com/drive/folders/1TUS16_KoOguOioZ3NPdFXK0tEBmhsnW?usp=sharing)