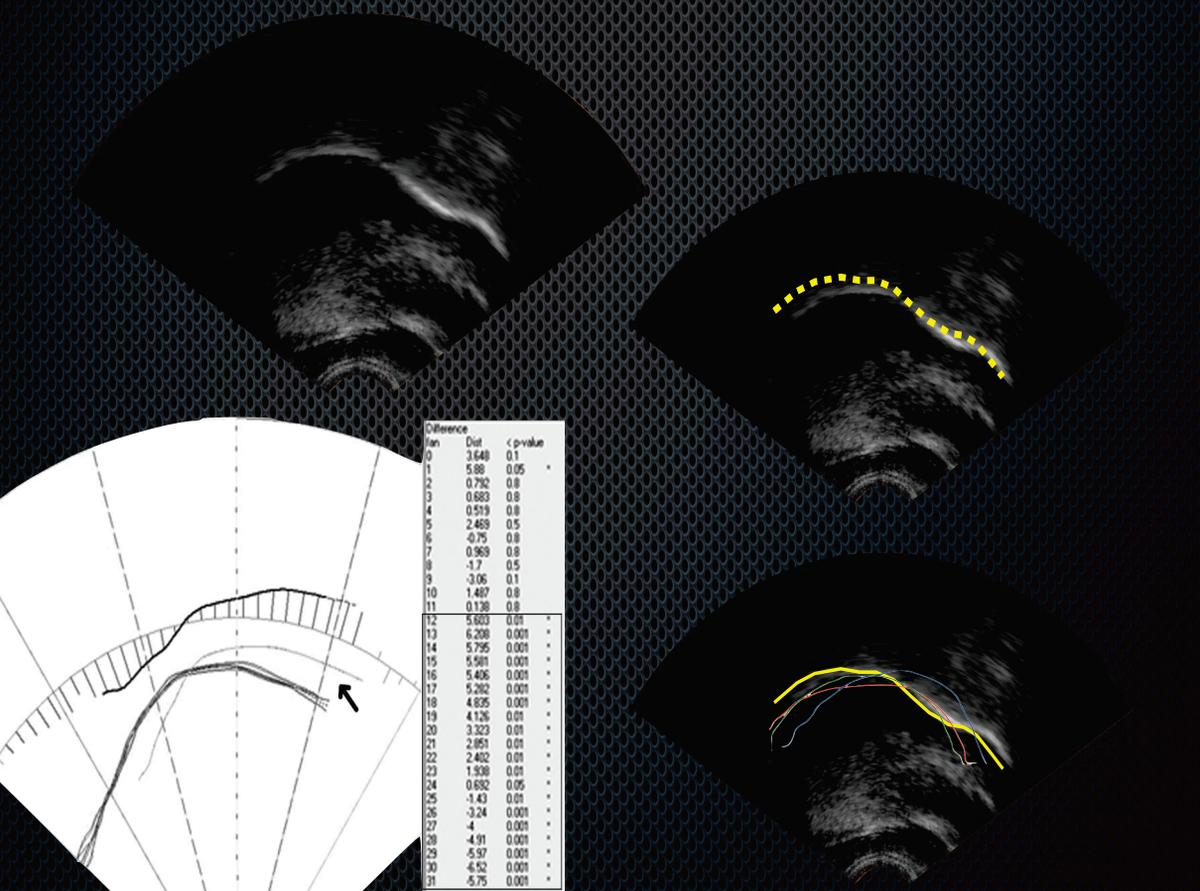


Giovana Ferreira-Gonçalves e Mirian Rose Brum-de-Paula (Orgs)

A ultrassonografia

&

os gestos da fala



A ultrassonografia e os gestos da fala

**Giovana Ferreira-Gonçalves
Mirian Rose Brum-de-Paula**
(Organizadoras)

A ultrassonografia e os gestos da fala

Copyright © Autoras e autores

Todos os direitos garantidos. Qualquer parte desta obra pode ser reproduzida, transmitida ou arquivada desde que levados em conta os direitos das autoras e dos autores.

Giovana Ferreira-Gonçalves; Mirian Rose Brum-de-Paula [Orgs.]

A ultrassonografia e os gestos da fala. São Carlos: Pedro & João Editores, 2021. 288p.

ISBN: 978-65-5869-255-3 [Impresso]
978-65-5869-256-0 [Digital]

1. Ultrassonografia. 2. Aquisição da língua. 3. Língua e linguagem. 4. Gestos da fala. I. Título.

CDD – 410

Capa: Petricor Design

Diagramação: Diany Akiko Lee

Editores: Pedro Amaro de Moura Brito & João Rodrigo de Moura Brito

Parecer e revisão por pares

Os capítulos que compõem esta obra foram submetidos para avaliação e revisados por pares.

Conselho Científico da Pedro & João Editores:

Augusto Ponzio (Bari/Itália); João Wanderley Geraldi (Unicamp/ Brasil); Hélio Márcio Pajeú (UFPE/Brasil); Maria Isabel de Moura (UFSCar/Brasil); Maria da Piedade Resende da Costa (UFSCar/Brasil); Valdemir Miotello (UFSCar/Brasil); Ana Cláudia Bortolozzi (UNESP/ Bauru/Brasil); Mariangela Lima de Almeida (UFES/Brasil); José Kuiava (UNIOESTE/Brasil); Marisol Barenco de Mello (UFF/Brasil); Camila Caracelli Scherma (UFFS/Brasil); Luís Fernando Soares Zuin (USP/Brasil).



Pedro & João Editores
www.pedroejoaoeditores.com.br
13568-878 - São Carlos – SP
2021

SUMÁRIO

A dinâmica do movimento: gestos como objeto de investigação	7
Mirian Rose Brum-de Paula	
Giovana Ferreira-Gonçalves	
Parte Um	
1) Ultrassonografia de língua e o uso do <i>biofeedback</i> visual no tratamento das desordens dos sons da fala	19
Helena Bolli Mota	
2) Ultrassonografia de língua na terapia dos distúrbios dos sons da fala: uma revisão sistemática	31
Isadora Mayer Rosado	
Marieli Barichello Gubiani	
Marileda Barichello Gubiani	
Márcia Keske-Soares	
3) Ultrassonografia de língua na aquisição de língua estrangeira	67
Laís Silva-Garcia	
Misael Krüger Lemes	
Giovana Ferreira-Gonçalves	
4) Ultrassonografia de língua na variação da linguagem	95
Aline Rosinski	
Thalena Evangelista Santos	
Mirian Rose Brum-de-Paula	
5) Parâmetros em análises qualitativas e quantitativas: dados ultrassonográficos de língua	117
Bruna Teixeira-Correa	
Giovana Ferreira-Gonçalves	

Parte Dois

1) Proposta de intervenção com o apoio da ultrassonografia dos movimentos da língua para os distúrbios do som da fala	153
Luciana da Silva Barberena	
Denis Altieri de Oliveira Moraes	
Larissa Cristina Berti	
Márcia Keske-Soares	
2) Ultrassonografia de língua no desvio fonológico: da avaliação à terapia	189
Aline Mara de Oliveira	
Rafaela Duarte Silva	
Izabel Christine Seara	
3) Ultrassonografia e erros de fala	207
Letícia Bitencourt Uberti	
Caroline Rodrigues Portalete	
Denis Altieri de Oliveira Moraes	
Márcia Keske-Soares	
4) Análise ultrassonográfica de fricativas posteriores em onset silábico em diferentes contextos vocálicos	219
Carine Haupt	
Izabel Christine Seara	
5) Configuração Gestual do <i>tap</i> na fala de pelotenses	233
Patrícia Pereira-Melcheque	
Giovana Ferreira-Gonçalves	
Mirian Rose Brum-de-Paula	
Referências	253
Sobre as autoras e os autores	281

A DINÂMICA DO MOVIMENTO: GESTOS COMO OBJETO DE INVESTIGAÇÃO

Mirian Rose Brum-de Paula
Giovana Ferreira-Gonçalves

A tecnologia do ultrassom, empregada em investigações sobre a linguagem, é recente, mas a proposta deste livro não é o de lançá-la ou reafirmá-la como um recurso indispensável capaz de resolver questões de ordem teórica ou prática. Sem dúvida, novas tecnologias podem fazer avançar teorias e métodos, e ideias podem potencializar o uso dos mais variados recursos tecnológicos. O livro e o foco no ultrassom, porém, trazem principalmente à tona inquietudes, a vontade de aprender mais sobre a produção da fala e os mecanismos que a suportam, além de um forte desejo de melhor compreender a singular conquista dos precisos e delicados gestos que arquitetam a realização dos sons de um idioma materno ou de uma língua adicional.

Para poder observar e especular acerca de fenômenos acústico-articulatórios, os autores, dentro de seus respectivos laboratórios e universidades, optaram por uma mudança de posição relativa aos seus objetos de trabalho, para inspecioná-los a partir de outra perspectiva. Compreendem, desse modo, que tal perspectiva não corresponde somente a um campo, a uma teoria ou a uma abordagem outra, mas também à adoção de um ângulo diverso, proporcionado pelo deslocamento e conseqüente troca de posição. É, então, de um inabitual local que a vista se revela, com verdes e frescas nuances de uma paisagem já entrevista. O ultrassom entra em cena nesse contexto.

Na pesquisa internacional, as ciências da fala são alimentadas por investigadores de horizontes muito diversos. Os trabalhos desenvolvidos envolvem equipes compostas por engenheiros, fonoaudiólogos, matemáticos, físicos, filósofos e linguistas, dentre outros profissionais. Além disso, o emprego de ferramentas, físicas e/ou virtuais, é comum porque são, de fato, usuais e porque são, também, compartilhadas. No Brasil, esse tipo de colaboração e prática é menos recorrente, o que explica o pouco emprego de equipamentos de imagem nos trabalhos sobre a linguagem articulada pois, frequentemente, habitam espaços vinculados a

outras áreas do conhecimento, áreas que possuem pouco ou nenhum interesse pela língua, sua estrutura e funcionamento.

No ambiente acadêmico brasileiro das Ciências Humanas, a dificuldade na obtenção de tecnologias mais performáticas do que gravadores e computadores é patente. Mas existem exceções, pois, nos últimos quinze anos, alguns linguistas souberam, por meio de editais da CAPES, do CNPq e de agências de fomento regionais, captar, aos poucos, recursos para equipar seus laboratórios de pesquisa. O ultrassom surgiu, então, como uma boa opção de equipamento para o estudo da fala, pois é pequeno, pouco oneroso (comparado a outras tecnologias), sem riscos para os informantes e fornece imagens em tempo real dos movimentos efetuados pela língua durante a articulação de sons significativos. Permite, pois, que o mais veloz e flexível órgão do corpo humano e suas evoluções possam ser observados e analisados no seu tempo e espaço.

A ideia de trabalhar com essa ferramenta chegou ao Brasil por meio de colegas fonoaudiólogos, que possuíam contato com a pesquisa que a incluía em trabalhos sobre queixas de fala. Por razões óbvias, esses mesmos colegas já haviam se tornado também linguistas em cursos de pós-graduação, realizados nas últimas décadas do século XX.

A tradição relativa aos estudos experimentais ligados à fala sempre promoveu alianças, possibilitando a elaboração de projetos comuns, o compartilhamento de materiais, a criação de novos artefatos. A exemplo dessa tradição, encontram-se no ambiente universitário – em que cada área defende o seu próprio território e suas especificidades – frestas mal vedadas para a realização de trabalhos conjuntos, interdisciplinares e/ou de pesquisas cujos métodos e resultados transitam em *papers*, disciplinas e práticas docentes. Assim sendo, o leitor encontrará, nesta obra, contribuições advindas da linguística e da fonoaudiologia, e verificará que o ultrassom compõe a metodologia dos trabalhos apresentados. Porém, o que de fato une os diferentes pesquisadores é a atenção que voltam para o corpo que fala. Muito do que se produziu sobre a fala excluiu o corpo que a produz. Não somente o corpo, o que o constitui, mas o que o associa à vida, ou seja, o movimento. A atração pelo movimento é, pois, o que verdadeiramente aproxima os autores desta coletânea.

Na introdução de *Método gráfico nas ciências experimentais e principalmente na fisiologia e medicina* (1878, p. iii), Maray destacou a importância do elemento gráfico para auxiliar a ciência a observar o “estado dinâmico” dos corpos, que escapa aos sentidos e à linguagem do

pesquisador. O que o método mostra revelaria, segundo o autor, “a linguagem própria dos fenômenos”. Maray demonstrou o seu apreço pelas manifestações dinâmicas da natureza e a importância de se obter, no papel, transcrições gráficas contendo informações sobre mudanças de temperatura, variações de peso, pressão, volume ou velocidade; sobre o desenvolvimento, duração e ritmo de eventos sucessivos ou que de modo concomitante ocorrem. Médico, fisiologista e inventor, logo percebeu os limites do método gráfico proposto para a “linguagem da vida”. Produziu e aperfeiçoou, então, diversos e diferentes mecanismos para investigar o impalpável, interno e externo de um organismo. Assim, com suas cronofotografias, por exemplo, demonstrou que os movimentos naturais se organizam em segmentos discretos. Para Braun (1983), Maray foi o primeiro a iniciar a incrível e progressiva incursão dentro do invisível.

Atualmente, é ingrata a tarefa de enumerar os diversos equipamentos que podem auxiliar na investigação do “invisível” da fala. Eles representam esforços de muitos indivíduos e a contribuição de vários campos de estudo. Cada equipamento pode revelar um ponto de vista do objeto investigado e o conjunto da aparelhagem existente poderia, talvez, expor toda a complexa engrenagem que envolve as estruturas e os movimentos articulatórios. No entanto, sistemas dinâmicos tão espetaculares, que possam dar acesso ao que internamente ocorre no momento da fala, ainda não estão disponíveis.

De fato, sincronizar alguns materiais (áudio e imagem ultrassonográfica, no caso dos trabalhos aqui dispostos) já é tarefa árdua, requer formação complementar e equipe. Importa ressaltar que, ao longo dos anos, a acumulação de conhecimentos tirou a imagem do papel e a migrou para a tela. Esse feito fornece imagens contendo sequências de estruturas externas ou internas do corpo em movimento ou, ainda, o movimento parado, a postura, prestes a acontecer. Assim, no que tange aos movimentos da superfície da língua, acessível por meio da tecnologia do ultrassom, pode-se indagar e obter bons elementos de resposta acerca de sua posição, alongamento, contração, velocidade, aceleração, ajustes e compensações. Em suma, sobre muitas das ações efetuadas pelos gestos de uma das principais peças do quebra-cabeças da articulação e da coarticulação de segmentos linguísticos.

Este livro contém dez capítulos divididos em duas partes, cada uma com um número idêntico de textos. 23 pesquisadores são os seus autores, vinculados a cinco universidades brasileiras: Universidade Estadual Paulista

Júlio de Mesquita Filho (UNESP/Campus de Marília), Universidade Federal de Pelotas (UFPEL), Universidade Federal de Santa Catarina (UFSC), Universidade Federal de Santa Maria (UFSM) e Universidade Federal do Tocantins (UFT). Nas referidas Instituições, quatro laboratórios viabilizam a realização das pesquisas com a utilização de aparelhos de ultrassom: Laboratório de Análise Acústica e Articulatória (LAAc/UNESP), Laboratório Emergência da Linguagem Oral (LELO/UFPEL), Laboratório de Fala (LabFala/UFSM) e Laboratório de Fonética Aplicada (FONAPLI/UFSC).

A primeira parte é iniciada pelo capítulo *Ultrassonografia de língua e o uso do biofeedback visual no tratamento das desordens dos sons da fala (DSF)*, de Helena Bolli Mota. Segundo a autora, uma fatia importante dos pacientes infantis da fonoaudiologia possui DSF e mesmo que, na maior parte do tempo, não se saibam as causas dessas alterações, elas advêm de problemas cognitivo-linguísticos ou motores. A tais pacientes, pistas auditivas lhes são fornecidas, com o intuito de que compreendam o movimento e o modo de posicionar seus articuladores. A autora realiza, neste capítulo, uma revisão da literatura sobre o *biofeedback* visual na intervenção das DSF, já que nem sempre o *input* auditivo soluciona os problemas apontados pelos pacientes e seus familiares. Mota constata, então, que os estudos reportados detalham a fase da produção da terapia, mas dão pouca ênfase à fase pré-prática das intervenções propostas com o ultrassom. Sugere, conseqüentemente, que haja uma maior atenção a essa fase inicial, já que o uso da ferramenta se mostrou promissor nos estágios iniciais de aquisição linguística. O capítulo termina com uma lista de direções para a pesquisa, fornecida pelos próprios autores consultados.

Na seqüência, o texto *Ultrassonografia de língua na terapia dos distúrbios dos sons da fala: uma revisão sistemática*, de Isadora Mayer Rosado, Marieli Barichello Gubiani, Marileda Barichello Gubiani e Márcia Keske-Soares, apresenta, também, uma revisão da literatura concernente ao *biofeedback* visual ultrassonográfico na terapia dos DSF. Como no texto anterior, as autoras reportam o modo como procederam para encontrar os artigos a serem revisados, apresentando diagramas de fluxo dessa identificação e seleção. Na parte relativa aos achados dessa empreitada, destacam que quase todos os estudos visitados continham as etapas de aquisição, generalização e retenção, ou seja, as três fases necessárias para a aquisição motora do som tratado e que o uso do ultrassom foi eficaz na maior parte das terapias empreendidas. Sugerimos que os dois primeiros capítulos, 1 e 2, sejam lidos num mesmo

momento, pois complementam-se: trazem informações relevantes para profissionais que desejam empregar o ultrassom em terapias de fala e fornecem elementos fundamentais para a construção do desenho metodológico a ser implementado na terapia. Muitos achados metodológicos são também relevantes para pesquisadores que trabalham com a fala típica. Destacamos, neste segundo capítulo, a sessão dedicada às limitações encontradas no uso do ultrassom.

O terceiro capítulo, *Ultrassonografia de língua na aquisição de língua estrangeira*, de Misael Krüger Lemes, Laís Silva-Garcia e Giovana Ferreira-Gonçalves, apresenta uma revisão das principais pesquisas realizadas, no Brasil e no exterior, acerca da aplicabilidade da ultrassonografia no processo de ensino-aprendizagem de uma língua estrangeira (LE). Os autores iniciam o capítulo discorrendo sobre os principais aspectos que viabilizam o uso da ferramenta para fins instrucionais, sendo o principal deles justamente o fato de o aprendiz poder visualizar, em tempo real, os movimentos de língua que ocorrem na produção dos sons da LE. Destacam, ainda, o fato de ser o ultrassom uma ferramenta não invasiva, de fácil portabilidade e custo reduzido. Na sequência, trazem a compilação de seis pesquisas internacionais e de cinco pesquisas brasileiras que utilizaram o US como ferramenta de instrução explícita no processo de ensino-aprendizagem do espanhol, francês e, principalmente, inglês.

O quarto texto, intitulado *Ultrassonografia de língua na variação da linguagem*, de Aline Rosinski, Thalena Evangelista Santos e Mirian Rose Brum-de-Paula, expõe, de modo detalhado, pesquisas sociofonéticas que incorporaram o uso do ultrassom a suas metodologias. As autoras destacam que os dados ultrassonográficos servem, principalmente, para confirmar ou lapidar resultados advindos da análise acústica ou, ainda, mostrar novas facetas do fenômeno investigado. A acústica e a articulação fornecem, então, informações diversas e complementares relativas ao mesmo fenômeno. Porém, para que isso ocorra, adaptações metodológicas são necessárias, pois os equipamentos deixam o ambiente laboratorial para que habitem, temporariamente, uma dada comunidade de fala. Assim, num espaço construído com planejamento, mas sem as condições ideais para a coleta de dados linguísticos, adaptações precisam ser efetuadas implicando, por exemplo, o tempo de registro das produções e o tipo de interação a ser estabelecido com cada informante. Nas considerações finais, reportam aspectos metodológicos do primeiro trabalho brasileiro a empregar o ultrassom

em pesquisa de campo. Trata-se da pesquisa de Rosinski (2019) sobre o português brasileiro como língua de imigração.

Parâmetros em análises qualitativas e quantitativas: dados ultrassonográficos de língua, de Bruna Teixeira-Correa e Giovana Ferreira-Gonçalves, é o capítulo que fecha, enfim, a primeira parte deste livro. As autoras realizam uma revisão teórica acerca dos parâmetros de análise ultrassonográficos que têm sido utilizados nas pesquisas linguísticas. Na primeira seção, abordam os parâmetros qualitativos relativos às porções da língua, como ponta, dorso e raiz, bem como compilam trabalhos que analisam os movimentos da língua de forma mais ampla nos eixos vertical e horizontal. Na segunda seção, parâmetros quantitativos, como *Índice de Excursão do Dorso*, *Índice de Anterioridade* e *Índice de Posição de Restrição da Língua* são elencados. As autoras destacam que a maior parte das análises incluem, no entanto, os dois tipos de parâmetros, ou seja, qualitativos e quantitativos, cujas escolhas, a serem feitas pelos pesquisadores, determinam, sem dúvida, o sucesso da pesquisa que utiliza a ferramenta ultrassonográfica.

A segunda parte abre com uma *Proposta de intervenção com o apoio da ultrassonografia dos movimentos da língua para os distúrbios do som da fala*. O capítulo, de Luciana da Silva Barberena, Denis Altieri de Oliveira Moraes, Larissa Cristina Berti e Márcia Keske-Soares, discorre sobre o uso da ultrassonografia como ferramenta de intervenção em terapia fonoaudiológica para a aquisição do *tap* [r] por duas crianças com distúrbios dos sons da fala. A terapia com o ultrassom se demonstrou exitosa, tanto no que concerne à percepção quanto à produção dos gestos articulatórios que constituem o rótico. As autoras destacam que a ferramenta foi, ainda, capaz de revelar certos gradientes na produção do *tap*, ocultos, em análises de oitiva, pela aparente substituição por [l] ou [j]. Os resultados, portanto, evidenciam o quanto a ultrassonografia é promissora para uma aplicação cada vez mais recorrente nas pesquisas e na clínica fonoaudiológica.

Aline Mara de Oliveira, Rafaela Duarte Silva e Izabel Christine Seara, no segundo capítulo desta parte, intitulado *Ultrassonografia de língua no desvio fonológico: da avaliação à terapia*, trazem à tona um estudo de caso clínico longitudinal contendo um tratamento fonoaudiológico em que o *biofeedback* visual, por meio do ultrassom, se encontra associado à terapia tradicional (SILVA, 2020). Tal relato é o fio condutor de um capítulo que visa à descrição dos resultados da terapia e à identificação das vantagens e

desvantagens da ferramenta ultrassonográfica aplicada à aquisição típica e atípica da linguagem, com foco nas fricativas /s/ e /ʃ/ pré e pós terapia. O último tópico diz respeito a possíveis desdobramentos da pesquisa fonoaudiológica e linguística com o uso do ultrassom, já que a ferramenta se mostrou promissora em processos terapêuticos e tem avançado do ponto de vista tecnológico.

O terceiro capítulo, *Ultrassonografia e erros de fala*, de Letícia Bitencourt Uberti, Caroline Rodrigues Portalete, Denis Altieri de Oliveira Moraes e Márcia Keske-Soares, propõe uma caracterização dos contornos médios de língua das fricativas e líquidas /s/, /z/, /l/ e /r/ de adolescentes e adultos com produção de fala típica e Erros Residuais de Fala (ERF). Os autores reportam que os ERF estão ligados à posição ou ao movimento anormal da língua, o que os torna de difícil diagnóstico e tratamento. Estão, pois, vinculados a características de um órgão de difícil acesso visual, o que justifica o uso do ultrassom. O enfoque nos sons fricativos e líquidos foi motivado pela literatura da área, que os consideram como os que mais causam dificuldades de articulação no falante adulto. Esse dado foi confirmado na pesquisa que empreenderam e apresentam neste capítulo. Uberti, Portalete, Moraes e Keske-Soares ressaltam a importância de se investigar a produção da fala de adolescentes e adultos falantes do português brasileiro, pois a “presença de alterações de fala pode influenciar de maneira negativa a vida desses sujeitos no meio social, educacional e profissional”. Em pesquisas futuras, sugerem o emprego de métodos de análise quantitativa, já que a tecnologia do ultrassom, dentre outras ferramentas de imagem, avança, possibilitando a realização de trabalhos sobre as bases motoras da fala de maneira cada vez mais performática e inovadora.

Análise ultrassonográfica de fricativas posteriores em onset silábico em diferentes contextos vocálicos, de Carine Haupt e Izabel Christine Seara, versa sobre a análise e descrição fonética desses segmentos do Português Brasileiro (PB) por meio do ultrassom. Em 2020, as autoras já haviam proposto uma caracterização acústico-articulatória de /r/ em onset silábico. Para tanto, elegeram as vogais extremas [a, i, u] do triângulo vocálico do PB. Neste capítulo, expandiram os contextos vocálicos, incluindo as sete vogais orais da língua portuguesa em posição tônica, aprofundando, assim, os achados encontrados anteriormente. Objetivaram com essa ampliação de segmentos e análises, identificar “o comportamento das fricativas

anteriores diante de diferentes contextos vocálicos e contribuir na descrição articulatória da categoria dos róticos”. Como o estudo contou com informantes naturais de Florianópolis, Haupt e Seara, nas considerações finais, instigam os seus pares no sentido de envolver participantes das demais regiões brasileiras na realização de trabalhos semelhantes, para que se possa constatar o “papel de variáveis extralinguísticas na escolha da fricativa que ocupa a posição de rótico em contraposição com fatores linguísticos, como o contexto vocálico”.

Por fim, Patrícia Pereira Melcheque, Giovana Ferreira-Gonçalves e Mirian Rose Brum-de-Paula, no texto *Configuração gestual do tap na fala de pelotenses*, visam a caracterizar esse segmento rótico em locutores do dialeto de Pelotas. Para tanto, o analisam em sílabas CV, CCV e CVC e investigam a natureza do elemento vocálico adjacente nas sílabas CCV e CVC. A pesquisa tornou possível tratar padrões acústico-articulatórios dos róticos da fala daquela comunidade. Assim, no que concerne à acústica, o segmento foi identificado como *tap* em sílabas CV e CCV, e como *tap* e vibrante em sílabas CVC. No que diz respeito aos aspectos articulatórios, a análise revelou que é constituído de movimentos complexos presentes em um único gesto, conforme o que propõe Recasens (2016). O trabalho também apoia os achados de Silva, Clemente e Nishida (2006) quanto à natureza do elemento vocálico que acompanha o *tap*: possui a “mesma natureza da vogal nuclear em sílabas CCV, mas comporta-se como uma vogal neutra em sílabas CVV medial e final”.

Não é possível finalizar este texto sem, antes, agradecer a todos os colegas que responderam afirmativamente à ideia da elaboração e escrita desta publicação. Agradecemos, particularmente, o pronto auxílio de Márcia e Helena, da UFSM, parceiras de longa data na promoção de eventos e publicações, bem como de Izabel, da UFSC, com quem temos firmado recentes e profícuas parcerias. Indiscutível também é o nosso agradecimento a colegas que estiveram ou que estão em processo de formação na pós-graduação e que assinam, com antigos ou atuais orientadores, a maior parte dos capítulos deste livro. Eles compõem os elementos humanos indispensáveis para que o motor da paixão pela pesquisa e sua aplicação continue vibrando.

Agradecemos, enfim, o apoio fundamental do CNPq, da CAPES e de outros organismos de fomento, como a FAPERGS e a FAPESP, tanto para a criação, aquisição e manutenção de laboratórios de pesquisa que têm na linguagem o seu principal objeto, quanto para a divulgação do que nesses

ambientes universitários se produz. Esta publicação é fruto desses investimentos, pois foi financiada com verba proveniente da CAPES. Tal recurso chegou ao Laboratório Emergência da Linguagem Oral (LELO) por intermédio do Programa de Pós-Graduação em Letras da Universidade Federal de Pelotas (UFPel). Assim, nesse difícil momento que atravessa a população e as instituições brasileiras, continuamos no combate, tentando preservar laços, afinar parcerias e levar adiante ideias e projetos desenhados antes e durante a pandemia da Covid-19.

PARTE UM

ULTRASSONOGRAFIA DE LÍNGUA E O USO DO BIOFEEDBACK VISUAL NO TRATAMENTO DAS DESORDENS DOS SONS DA FALA

Helena Bolli Mota

As desordens dos sons da fala (DSF) são os distúrbios de comunicação mais comuns na infância. De acordo com alguns autores, sua prevalência varia de 2,3% a 24,6% (LAW et al., 2000; e WREN et al., 2016).

Enquanto algumas crianças com DSF adquirem a fala típica por volta dos 8 anos, outras com essa idade continuam apresentando distorções ou uma alteração de fala persistente (WREN et al. 2016) e aproximadamente 1 a 2% de crianças mais velhas ou jovens adultos terão erros residuais de fala ou DSF persistentes (FLIPSEN, 2015).

Dada a alta prevalência dessa desordem, não é de se surpreender que as crianças com DSF compreendam uma grande proporção de pacientes para os fonoaudiólogos no mundo inteiro (HEGARTY et al. 2018; e McLEOD e BAKER, 2014).

Embora não haja um sistema universalmente aceito para classificar as diferentes apresentações das DSF, é amplamente aceito que as crianças que apresentam essa desordem formam um grupo heterogêneo (WARING e KNIGHT, 2013).

Para algumas crianças a causa da DSF é conhecida, por exemplo, quando está associada a uma perda auditiva, a uma fissura palatina ou a uma síndrome genética. Porém, para a maioria das crianças com DSF, as causas dessas alterações são desconhecidas (SHRIBERG et al 2010). Para essas crianças, em geral, considera-se a raiz da desordem ou uma dificuldade cognitivo-linguística para a aquisição do sistema fonológico da língua ambiente ou uma dificuldade para adquirir o correto plano motor para um som ou sequência de sons particular (DODD et al. 2018; e WARING e KNIGHT, 2013).

As diferentes causas dão origem a diferentes subtipos ou combinações de diferentes subtipos de DSF, tais como atraso ou desordem fonológica, desordem da articulação ou Apraxia de fala infantil (CAS).

No mundo inteiro, a maior parte das abordagens de intervenção usadas pelos fonoaudiólogos para tratar as DSF baseiam-se no input auditivo (BAKER et al 2018; HEGARTY et al. 2018; SUGDEN et al 2018; e BRUMBAUGHT e SMIT, 2013) proporcionando à criança informações auditivas a respeito do som alvo e de suas próprias produções. Nestas abordagens de intervenção, em particular aquelas que têm por objetivo corrigir erros fonéticos, o fonoaudiólogo geralmente dá à criança pistas auditivas relativas a como movimentar e onde colocar seus articuladores (SECORD et al, 2007). No entanto, esse tipo de informação, para muitas crianças, pode ser de difícil compreensão e de difícil aplicação em sua produção dos sons alvo, uma vez que vários articuladores não são completamente visíveis durante a fala.

Para minimizar essas dificuldades, muitos fonoaudiólogos têm explorado, em suas terapias, o uso de diferentes técnicas de imagem como uma ferramenta para proporcionar informação articulatória adicional na forma de biofeedback visual (RUSCELLO, 1995).

O *biofeedback* visual é definido como “o uso de instrumentalização para tornar processos fisiológicos encobertos mais explícitos” (HUANG et al., 2006 I). No campo das DSF, em geral envolve proporcionar informação visual sobre a posição, forma, movimento e/ou lugar dos articuladores, mais comumente da língua. Acredita-se que este tipo de informação é útil para o tratamento das DSF, uma vez que permite tanto ao paciente quanto ao clínico acessar informações até então inacessíveis sobre a articulação do paciente. Essa informação pode ser benéfica para guiar o diagnóstico e o tratamento. Considera-se que o biofeedback visual, quando usado na intervenção, proporciona um conhecimento da performance do indivíduo em relação à natureza de uma articulação alvo. O feedback do conhecimento da performance é importante quando se está adquirindo ou modificando um plano motor existente (MAAS et al., 2008).

Diferentes abordagens que proporcionam biofeedback visual articulatório em tempo real na intervenção das DSF têm sido exploradas pelos clínicos, entre elas a eletropalatografia (EPG), o ultrassom da imagem da língua e a articulografia eletromagnética.

Existem muitos estudos sobre o tratamento usando a EPG, que é uma técnica que mostra o tempo e o ponto de contato da língua no palato. Embora esta abordagem seja efetiva para tratar as DSF, o

equipamento e a confecção personalizada dos palatos tornam os custos muito altos.

Por outro lado, nos últimos anos, o custo dos sistemas de ultrassom tem diminuído e máquinas portáteis já são disponíveis. Além disso a melhora dos métodos de análise tem levado a um aumento de pesquisas e do interesse clínico no uso do ultrassom de imagem de língua no tratamento das DSF.

A imagem dinâmica do movimento da língua proporcionada pelo ultrassom pode ser usada na intervenção tanto pelo terapeuta como pelo paciente para guiar e proporcionar feedback dos movimentos e da posição da língua.

Embora o *feedback* visual por ultrassom seja um procedimento ainda não tão comumente utilizado na clínica fonoaudiológica, a combinação da diminuição do preço e maior número de evidências de pesquisa têm levado a um aumento do interesse clínico por essa abordagem. No entanto, quando o clínico busca, na literatura, essas evidências científicas para o uso do US como ferramenta de terapia o que encontra é uma diversidade nos desenhos dos estudos, nas características dos participantes, nos métodos clínicos utilizados e nos resultados encontrados fazendo com que a interpretação das pesquisas seja difícil.

Uma forma de sintetizar as evidências das pesquisas, com o objetivo de ajudar os fonoaudiólogos em sua prática clínica é a realização de uma revisão sistemática da literatura relevante.

O objetivo deste capítulo, é apresentar os achados de uma revisão sistemática sobre as evidências para o uso do biofeedback visual com ultrassom na intervenção das DSF realizada pelos autores Eleanor Sugden, Susan Lloyd, Jenny Lam e Joanne Cleland, publicada no *International Journal of Language & Communication Disorders* em 2019.

Dada a alta prevalência das DSF nos casos atendidos no mundo inteiro, esta revisão teve como foco principal o uso do biofeedback visual com ultrassom no tratamento das DSF.

As estratégias de busca utilizadas pelos autores seguiram as diretrizes da PRISMA (The Preferred Reporting Items for Systematic Reviews and Meta-Analyses).

Para a identificação dos artigos, foram pesquisadas as seguintes bases de dados: Scopus, MEDLINE, PROQUEST, LLBA, CINAHL, speech-BITE, ASHA online jornal site e Cochrane Library utilizando-se

os termos: intervenção, terapia, tratamento, ultrassom, biofeedback, fala, articulação, fonologia, apraxia, dispraxia sendo incluídos artigos publicados até 2017.

Inicialmente foi realizada uma busca por título e resumo feita pelos dois primeiros autores usando o software Covidence. Conflitos em relação a esses aspectos foram resolvidos por meio de discussões entre os dois autores.

As cópias completas dos artigos foram avaliadas de acordo com os seguintes critérios de inclusão: artigos revisados por pares publicados em 2017 ou antes; artigos disponíveis em inglês; artigos que apresentassem uma investigação sobre o uso do biofeedback visual com ultrassom em intervenção; a intervenção deveria ter sido aplicada em crianças ou adultos com DSF não adquiridas. Não houve critérios de exclusão.

Após a seleção dos artigos que se adequaram aos critérios de inclusão, os autores realizaram a análise dos mesmos a partir da extração de relevantes aspectos, bem como detalhes da publicação, desenho do estudo, características dos participantes (incluindo número, idade, tipo de DSF, presença de outros transtornos concomitantes, história prévia de intervenção), detalhes do tratamento (incluindo procedimentos, outras intervenções na mesma sessão, orientação do transdutor, exercícios para casa), medidas de resultados utilizadas, análises realizadas, informações sobre como a terapia foi realizada, e intensidade do tratamento.

No processo de extração dos dados apenas as informações explícitas no artigo foram extraídas.

Os dados foram analisados descritivamente não sendo possível a realização de meta-análise devido à heterogeneidade dos desenhos de pesquisa e das medidas de resultado usadas em cada artigo.

Os autores também analisaram os artigos considerando o nível de evidência de cada um, o risco de viés na avaliação e como foi feita a descrição da intervenção.

O nível de evidência dos artigos foi determinado de acordo com a hierarquia de nível de evidencia da ASHA (2004) que atribui um nível de alto (I) a baixo (IV) com base no desenho do estudo.

Os artigos foram avaliados de acordo com o risco de viés usando a ferramenta PEDroP (Perdices et al. 2009) para estudos de grupo ou

a escala RoBiN.T (TATE et al., 2015) para pesquisas de caso único ou estudos de caso.

Todos os estudos incluídos foram também classificados usando-se o Template para Descrição e Réplica da Intervenção (TIDieR) (HOFFMANN et al. 2014). Trata-se de um checklist que considera os protocolos e procedimentos de intervenção descritos considerando a facilidade de réplica dos estudos de intervenção. Dois autores completaram essas avaliações e foram usadas medidas de concordância interjuízes.

Um total de 3529 artigos foram encontrados nas bases de dados e mais outros três adicionados a partir de outras estratégias de busca. Após a remoção dos artigos duplicados restaram 2128 artigos. Desses, 66 textos completos foram analisados, sendo que apenas 29 estudos se adequaram aos critérios de inclusão, descritos em 28 artigos.

Os estudos foram realizados nos seguintes países: Estados Unidos (n= 14; 48,3%), Canadá (n=12; 41,4%), Escócia (n=2; 6,9%), Austrália (n=1; 3,4%) e foram publicados em 14 periódicos diferentes, sendo que a maior parte deles foi publicada na *Clinical Linguistics and Phonetics* (n=10, 34,5%).

Os desenhos de pesquisa envolveram pesquisa experimental de caso único n=13 (44,8%), série de casos n=8 (27,6%), estudo de caso n=6 (20,7), experimento controlado randomizado n=1 (3,4%) e grupo quase-experimental n=1 (3,4%).

Na análise do nível de evidência, a maioria dos estudos apresentaram nível III (n=15; 51,7%), 13 estudos (44,8%) tiveram nível II b de evidência e apenas 1 estudo teve nível I b de evidência.

A maioria dos estudos foi classificada como estudos de eficácia (n=23; 79,3%), com 5 estudos examinando a efetividade do biofeedback com ultrassom (17,2%) e um investigando os resultados da intervenção a longo prazo (3,4%).

O número de participantes dos estudos variou de 1 (n=6; 20,7%) a 13 (n=1; 3,4%) participantes, com uma média de 4,1. A idade dos mesmos variou de 4 a 27 anos, com média de idade de aproximadamente 11 anos.

Considerando todos os artigos, 103 indivíduos receberam tratamento com uso de biofeedback visual com ultrassom.

Os principais subtipos de DSF relatados nos estudos foram erros residuais de fala (n=7), apraxia de fala infantil (n=5) e disartria (n= 2). Os outros estudos não identificaram o subtipo de DSF.

Na maioria dos estudos, os participantes haviam recebido terapia prévia para seus erros de fala (n=25 estudos; 86,2%).

Analisando os detalhes da intervenção, os autores verificaram que o *biofeedback* visual com ultrassom foi usado, em geral, como um componente da intervenção. A utilização de prática de produção baseada na articulação sem o uso de ultrassom foi mencionada em 19 estudos (65,5%).

Nesses 19 estudos, o uso do *biofeedback* visual com ultrassom foi administrado de diferentes maneiras: somente em uma fase de pré-prática da intervenção, em períodos alternados dentro da sessão, de uma forma gradualmente decrescente determinada pelo progresso dos participantes ou podendo ser usado em algumas, mas não em todas, as sessões de terapia. Outros estudos (n=5; 17,2%) usaram outros tipos de *biofeedback* além do ultrassom, tais como a eletropalatografia (EPG). Cinco estudos usaram o treinamento auditivo ou perceptual além da intervenção com foco na produção.

No que diz respeito aos alvos de intervenção, a análise dos estudos evidenciou que dos 103 participantes, 83 receberam terapia para róticos (80,6%), 17 sujeitos receberam tratamento para as fricativas (16,5%), 7 participantes tiveram como alvo as velares (6,8%), “sequências envolvendo sons linguais” foram alvo para 6 participantes (5,8%), 4 participantes receberam terapia para vogais (3,9%), /l/ foi alvo para 4 participantes (3,9%). Os sons /n/ e /t/ foram tratados (um para cada participante). Um estudo não especificou os alvos da intervenção. 15 participantes (14,6%) receberam intervenção para mais de um alvo.

Quanto à orientação do transdutor, 15 estudos (51,7%) usaram tanto imagens sagitais como coronais na intervenção, 4 estudos (13,8%) usaram apenas a visão sagital, mas nenhum estudo usou apenas a visão coronal. Um estudo (3,4%) relatou o uso de uma visão transversa e 10 estudos (34,5%) não descreveram a orientação das imagens usadas na intervenção.

Sobre a estabilização do transdutor, constatou-se que apenas um estudo relatou o uso do capacete, 8 estudos (27,6%) relataram que o transdutor foi seguro pela mão (portanto, não estabilizado), 2 estudos (6,9%) relataram que os participantes apoiaram suas cabeças em um

transdutor estabilizado e 18 estudos (62,1%) não relataram nada sobre este aspecto.

Analisando o tipo de entrega do serviço e a intensidade da intervenção, observou-se que, em geral, o tratamento foi realizado individualmente (n=16; 55,2%). Dois estudos (6,9%) referiram a combinação de atendimento em grupo e individual e 11 estudos não referiram a forma de entrega do tratamento (37,9%).

Em relação aos locais de tratamento, várias possibilidades apareceram, tais como: clínica de universidade, escola do participante, casa do participante e centro de pesquisa. Quatro estudos relataram que o atendimento foi realizado em mais de um local e 18 estudos não mencionaram onde o atendimento foi realizado.

Em 21 estudos (72,4%) a terapia foi realizada por um fonoaudiólogo ou por um estudante de fonoaudiologia (n=4; 13,8%) ou por ambos (n=1) e 3 estudos não mencionaram quem realizou a terapia (10,3%).

Em relação aos exercícios para casa, 8 estudos mencionaram o envio de exercícios para prática em casa (sem o uso do ultrassom e com a supervisão dos pais). Dois estudos deixaram explícito que não solicitaram prática em casa e 19 estudos não mencionaram nada sobre esse aspecto.

Quanto à intensidade da intervenção, a maioria dos estudos relatou que a mesma foi realizada uma vez por semana (n=11) ou 2 vezes por semana (n=9). 4 estudos não mencionaram a frequência das sessões.

Em geral as sessões foram de 60 minutos (14 estudos), mas variaram sua duração de 20 minutos a duas horas.

Dos estudos que relataram a dose da intervenção, o número de tentativas de produção (trials) variou de 60 a 366 por sessão na fase de prática da terapia. 18 estudos não referiram nada sobre esse aspecto.

Em relação ao número total de sessões realizadas, variou entre uma a 18 sessões de terapia, sendo o número mais comum 14 sessões (31,0% dos estudos). A terapia foi realizada em um período entre 1 a 22 semanas, sendo que 12 estudos não mencionaram a duração total do tratamento realizado.

Na análise dos resultados da intervenção 19 estudos relataram resultados positivos para todos os participantes (65,5%) após a terapia com uso do biofeedback com ultrassom, 10 estudos mostraram resultados mistos para os participantes com alguns indivíduos respondendo bem à intervenção e outros não.

Vários estudos relataram baixos níveis de generalização para palavras não tratadas, enquanto outros estudos relataram generalização extensiva para alguns participantes, mas não para outros.

Quanto às medidas dos resultados pós tratamento, a maioria dos estudos incluiu um julgamento perceptual de correção do som alvo por outros experts (n=19). Onze estudos analisaram a correção da produção com base na transcrição fonética, 5 estudos incluíram medidas de análise acústica, 3 estudos usaram descrições qualitativas das imagens de ultrassom antes e depois da intervenção, 2 estudos realizaram uma análise quantitativa das imagens de ultrassom, um estudo reportou o conhecimento dos participantes sobre os componentes linguais da articulação alvo e um estudo fez uma descrição geral das mudanças articulatórias após a intervenção. Somente um estudo considerou os resultados de longa duração após a intervenção com o uso do ultrassom.

Os estudos mais antigos explorando o uso de ultrassom na terapia, em geral, não faziam uso de medidas estatísticas para medir os resultados enquanto estudos mais recentes passaram a usar essas medidas. No entanto, tamanhos de efeito para diferentes aspectos da intervenção foram relatados (por exemplo: por alvo, por participante ou por fase de intervenção) o que limita a comparação entre os estudos. De um modo geral, os resultados e conclusões dos estudos sugerem que a intervenção com uso de biofeedback por ultrassom pode ser efetiva para a aquisição de estágios de aprendizado motor para alguns indivíduos com DSF.

A análise dos indicadores de qualidade, seja com o PEDro-P, com a escala RoBiN.T ou com as avaliações pelo TiDieR mostraram que os estudos foram melhor avaliados em relação à validade externa do que à interna. Por exemplo, a maioria dos estudos tinha alguma informação sobre as medidas de generalização, sobre as variáveis dependentes e independentes e sobre medidas de linha de base. Em relação à validade interna, em geral os estudos tiveram escores pobres no RoBiN.T quanto a itens como desenho da pesquisa, randomização e amostragem. Houve uma tendência de estudos publicados mais recentemente terem melhor escores em relação a esses aspectos do que estudos mais antigos. As avaliações usando a checklist TiDieR mostraram que, em geral, os estudos não relataram detalhes suficientes sobre o tratamento realizado (apenas 20,7% fizeram isso),

quando e quanto de tratamento foi realizado, onde a intervenção foi feita, como foi entregue e que materiais foram utilizados.

Quando os resultados de todos os estudos são considerados em conjunto, observa-se que o uso do biofeedback visual com ultrassom na terapia pode facilitar a aquisição de diferentes alvos linguais para alguns indivíduos com DSF, mas que isso nem sempre leva à generalização. Muitos estudos apresentaram baixos níveis de evidência e apresentaram informações insuficientes que permitam sua implementação e sua réplica com níveis altos de fidelidade. Combinados, esses aspectos têm implicações para pesquisas futuras e para os fonoaudiólogos que desejam utilizar esse tipo de intervenção emergente em suas práticas clínicas.

Uma variedade de sons alvos foi tratada com o uso do ultrassom na terapia. A maioria dos estudos incluiu o tratamento dos róticos. O uso do ultrassom como feedback visual foi efetivo como um procedimento adjunto a outros tipos de tratamento mais tradicionais. Um estudo que comparou a terapia com ultrassom e a terapia tradicional mostrou um benefício claro para a terapia com ultrassom nos estágios iniciais da intervenção quando comparada com o modelo tradicional.

No entanto, ambas as abordagens tiveram ganhos equivalentes no final do tratamento. Enquanto muitos estudos relataram resultados positivos para todos os participantes, 10 estudos mostraram que houve pacientes que responderam e aqueles que não responderam positivamente à terapia com o uso do ultrassom. O uso do ultrassom na terapia foi eficiente nos estágios iniciais da intervenção, se mostrando importante para o aprendizado motor. Porém, vários estudos relataram que houve pouca generalização dessas habilidades para palavras, posições e sons não-tratados.

Quando se analisa a possibilidade de implementação e réplica das pesquisas, a maioria dos estudos forneceu detalhes suficientes sobre como a intervenção foi administrada para cada participante. Isso reflete o desenho de pesquisa predominante: estudos de caso único ou séries de casos, que permitem intervenções individualizadas. Por outro lado, de um modo geral, os artigos continham informações pobres sobre a entrega do serviço e a intensidade da intervenção, em particular sobre onde e quando a intervenção foi feita e quanto de tratamento foi realizado. Dada a importância da intensidade para uma intervenção efetiva para as DSF, a falta de informação sobre esse

aspecto torna difícil examinar adequadamente as contribuições da intensidade para as respostas à intervenção demonstradas pelos participantes. Isso tem implicações para a implementação de intervenções efetivas com o uso do ultrassom na prática clínica e na replicação de estudos de pesquisa.

Ficou claro nesta revisão que, enquanto a maioria dos estudos apresentam informações detalhadas sobre a fase de produção da terapia, pouco estudos especificam detalhes sobre a fase de pré-prática. A pré-prática é importante para assegurar a estimulabilidade dos alvos e, visto que o uso do ultrassom pode ser benéfico nos estágios iniciais de aquisição, informações adequadas sobre este estágio da intervenção é crucial para a implementação e réplica.

A fase pré-prática inclui instruções verbais e visuais que ajudam os participantes a compreender o que é necessário para a correta produção dos movimentos alvo. Os estudos, em geral, não incluem informações sobre essas instruções, que são muito importantes para a implementação da terapia. Mais informações sobre as melhores estratégias para ensinar (dar pistas) produções corretas são necessárias na literatura

As contribuições trazidas por esta revisão sistemática da literatura são inúmeras, porém os próprios autores citam algumas limitações do estudo, tais como: não incluir pesquisas relatando o uso do ultrassom em outras populações (adultos com dispraxia de fala, adultos com glossectomia, transtornos de fala adquiridos); considerar apenas artigos revisados por pares (não foram usados trabalhos apresentados em congressos ou teses) e o fato de só considerar estudos conduzidos em inglês com participantes monolíngues ou bilingues falantes de inglês.

Nesse sentido, os autores sugerem como direções para pesquisas futuras, além dos aspectos apontados como limitações, que sejam feitos estudos em larga escala, com grupos maiores de participantes, estudos com desenhos de pesquisa mais bem estruturados, buscando altos níveis de evidência, estudos que explorem características individuais dos participantes associadas a respostas positivas ao tratamento com uso do ultrassom (tais como estimulabilidade pré-tratamento, motivação, idade, habilidades cognitivas, tipo e gravidade das DSF).

Considerações finais

Neste capítulo procurou-se relatar os achados descritos em uma revisão sistemática das evidências sobre o uso do biofeedback visual com ultrassom no tratamento das DSF de origem desenvolvimental. Um total de 29 estudos publicados em 28 artigos foram revisados. A maioria dos estudos apresentou baixos níveis de evidência, no entanto, os mais recentes apresentaram desenhos de pesquisa mais robustos. Os resultados mostraram que o biofeedback visual com ultrassom pode ser usado como parte da intervenção de diferentes subtipos de DSF e para uma variedade de sons linguais. Em particular, os resultados indicaram que o uso de ultrassom na terapia pode ser efetivo para alguns indivíduos quando usado nos estágios iniciais do aprendizado motor, mas pode ser menos efetivo para promover a generalização a sons e palavras não tratados. Pesquisas futuras de alto nível devem explorar os ingredientes ativos da intervenção e a efetividade da abordagem em contextos clínicos reais de modo que os fonoaudiólogos possam oferecer o melhor tratamento a seus pacientes.

ULTRASSONOGRRAFIA DE LÍNGUA NA TERAPIA DOS DISTÚRBIOS DOS SONS DA FALA: UMA REVISÃO SISTEMÁTICA

Isadora Mayer Rosado
Marieli Barichello Gubiani
Marileda Barichello Gubiani
Márcia Keske-Soares

Introdução

Este capítulo apresenta uma pesquisa de revisão sistemática da literatura sobre o *biofeedback visual ultrassonográfico* na terapia dos distúrbios dos sons da fala (DSF). Os DSF são muito prevalentes em crianças e adolescentes, e o diagnóstico diferencial destes transtornos, bem como um tratamento eficaz é fundamental. Na terapia de fala, o uso de recursos instrumentais vem sendo bastante estudado e utilizado na prática clínica.

A linguagem oral se refere ao uso convencional de palavras faladas e tem como principal objetivo a comunicação humana. Esta é elemento essencial para a socialização e integração na comunidade (RUBEN, 2000). O uso da linguagem envolve aquisição e domínio de diversos aspectos linguísticos–comunicativos, dentre os quais está inserida a fala, denominada como uma habilidade comunicativa.

Durante o desenvolvimento linguístico, algumas crianças apresentam desvios ou atrasos no que diz respeito a aquisição fonológica, passando a realizar estratégias de reparo como forma de aproximar-se ao som alvo padrão, surgindo assim os chamados DSF (GHISLENI et al., 2010).

Os DSF são decorrentes de uma variedade de etiologias que acarretam em prejuízo nos diferentes níveis de percepção e produção da fala. Quando se tem o nível linguístico-fonológico prejudicado, caracterizado por omissões e/ou substituições de fonemas, os quais permanecem após os 5 anos, tem-se o chamado desvio fonológico (MOTA et al., 2007). Quando essas alterações ainda se fazem presentes em sujeitos com idade superior aos 8-9 anos de idade, têm-se os Erros Residuais de Fala (ERF) (PRESTON; EDWARDS, 2007 e SHRIBERG et al.,

1997). E os DSF podem também ser resultado da aquisição de sons típicos de fala, porém com distorções durante a produção (FLIPSEN, 2015), ou estar associados à apraxia de fala infantil (AFI).

Quando o comprometimento ocorre no nível de programação motora da fala, tem-se a chamada AFI, a qual se caracteriza por uma considerável dificuldade na produção da fala e acurácia fonética, havendo lentidão, intermitência e variabilidade durante a produção da fala, sendo a terapia geralmente mais lenta (STRAND et al., 2013; e SHRIBERG, 2006).

Entende-se, assim, que os DSF têm uma relação de impacto direto sobre a vida social do sujeito e, também, sobre o sucesso acadêmico e ocupacional (MCKINNON et al., 2007; e KEEGSTRA et al., 2007), portanto, os DSF são reconhecidos como importantes questões de saúde pública a serem tratadas (SOMEFUN et al., 2006).

A fim de reparar estas dificuldades, foram criados modelos de terapia fonológica (MOTA et al., 2007) com o objetivo de conduzir o procedimento terapêutico a ser adotado. Porém, para um diagnóstico preciso e um delineamento terapêutico adequado, é necessário que se tenham avaliações cada vez mais objetivas, para que haja uma maior precisão sobre o nível de produção alterado, e assim, uma melhor intervenção terapêutica.

A utilização de avanços tecnológicos tem contribuído para a área da saúde. Na fonoaudiologia, os estudos relacionados à terapia com recursos instrumentais têm sido crescentes, permitindo melhores condições para avaliação e obtenção de medidas mais objetivas (BALL et al., 2001; BERTI; MARINO, 2008; BARBERENA et al., 2014) por meio da utilização de *biofeedbacks* para modificação do padrão de produção da fala (MANTIE-KOZLOWSKI; PITT, 2013). Crianças com DSF têm sido contempladas com tratamento terapêuticos que visam abordagens com uso de *biofeedback* (MELO et al., 2016).

A ultrassonografia de fala (USG) funciona como uma ferramenta de *biofeedback* visual. É um instrumento seguro, prático e não invasivo (ZHARKOVA et al., 2011) e vem sendo utilizado em pesquisas relacionadas à fala (RUBEN, 2000; SHRIBERG, 2006; MCKINNON et al., 2007; KEEGSTRA et al., 2007; SOMEFUN et al., 2006; BALL et al., 2001; BERTI; MARINO, 2008; BARBERENA et al., 2014; MANTIE-KOZLOWSKI; PITT, 2013; MELO et al., 2016; ZHARKOVA et al., 2011; PRESTON; LEAMAN, 2014; ADLER-BOCK et al., 2007; BYUN et al., 2014; e PRESTON

et al., 2013). Além disso, também pode ser utilizado na prática clínica, complementando outras terapias, especialmente em casos de sujeitos com alterações de fala que não obtiveram resultados positivos com o tratamento tradicional.

A USG fornece a representação visual dos movimentos da língua envolvidos durante a produção de um determinado som em tempo real, permitindo que o paciente visualize a produção correta do som estimulado e tente se aproximar ao gesto de língua ensinado pelo terapeuta (CAGLIARI, 2009).

A utilização da USG se baseia no princípio da aprendizagem motora, pois entende-se que quanto maior for o número de tentativas durante a prática, conseqüentemente haverá maior retenção da informação e uma possível generalização do padrão aprendido (BYUN et al., 2016). Ainda, acredita-se que são necessárias três fases para a aprendizagem: a aquisição, retenção e generalização das produções (MAAS et al., 2008; SCHMIDT e LEE, 1999).

Entende-se assim, que podem ser diversos os prejuízos gerados pelos DSF e que o *biofeedback* visual ultrassonográfico (BV-U) é um recurso que vem sendo utilizado em pesquisas com o intuito de se obter uma terapia mais eficiente. Buscou-se nessa revisão de literatura, responder aos seguintes questionamentos: Quantos estudos com o uso do BV-U abordaram a terapia nos DSF? Quais as contribuições geradas pelo BV-U nos DSF?

Para responder a essas questões, esse estudo buscou apresentar uma revisão sistemática da literatura sobre o uso do *biofeedback* visual ultrassonográfico (BV-U) na terapia dos Distúrbios dos Sons da Fala (DSF).

Estratégia de pesquisa

A estratégia de busca utilizada segue as diretrizes *Preferred Reporting Items for Systematic Reviews e Meta-Analyzes* (PRISMA) (MOHER et al., 2006). A seguir, são apresentados os procedimentos realizados nas quatro fases consideradas para a busca e seleção dos artigos: identificação, *screening*, elegibilidade e inclusão.

Nesta primeira fase, o objetivo foi identificar as evidências de estudos relacionados ao uso do BV-U nos DSF. Para obtenção de uma confiabilidade adequada, dois pesquisadores conduziram de maneira

independente pesquisas nos bancos de dados PubMed, Scopus e Web of Science durante os meses de outubro e novembro de 2018.

Para realizar a busca em cada uma das bases citadas, foram utilizados os descritores: “*intervention*”, “*therapy*”, “*treatment*”, “*ultrasound*”, “*visual feedback*”, “*biofeedback*”, “*speech sound disorders*”, “*phonological disorders*”, “*articulation disorders*”, “*childhood apraxia of speech*”, “*residual speech sound errors*”, “*speech sound distortions*”.

As combinações dos descritores ocorreram da seguinte forma: foram realizados três constructos, sendo o primeiro relacionado à intervenção, com os descritores “*intervention OR therapy OR treatment*”. O segundo relacionado ao uso do instrumento de USG, com os descritores “*ultrasound OR visual feedback OR biofeedback*” e o terceiro relacionado às patologias de fala, com os termos “*speech sound disorders OR phonological disorders OR articulation disorders OR childhood apraxia of speech OR residual speech sound errors OR speech sound distortion*”.

Em seguida, em cada um dos mecanismos de busca, a associação entre os três construtos foi incorporada, desta vez utilizando o operador booleano AND. O campo de pesquisa em todos os casos foi limitado ao “título ou resumo”, “língua inglesa”; além de “artigos e artigos no prelo”.

Nesta segunda fase o objetivo foi revisar a primeira versão dos estudos encontrados na busca inicial, com o intuito de identificar aqueles potencialmente elegíveis para uma análise de texto completa.

Para isso, os artigos encontrados nas três bases de dados foram inseridos em uma planilha de dados. Nesse momento, artigos duplicados foram eliminados. Considerando a amplitude da busca inicial, os resultados foram refinados de acordo com as informações apresentadas no título, resumo e palavras-chave.

A partir disso, os estudos que: (1) não eram empíricos; (2) não envolviam terapia com BV-U; (3) os participantes apresentavam outras desordens de linguagem que não o DSF; e que (4) o idioma não era a língua inglesa, foram excluídos desse estudo.

Como na pesquisa inicial, a seleção dos artigos foi realizada de forma independente por dois dos autores. Cada um revisou os títulos e resumos previamente selecionados e obteve o texto completo dos

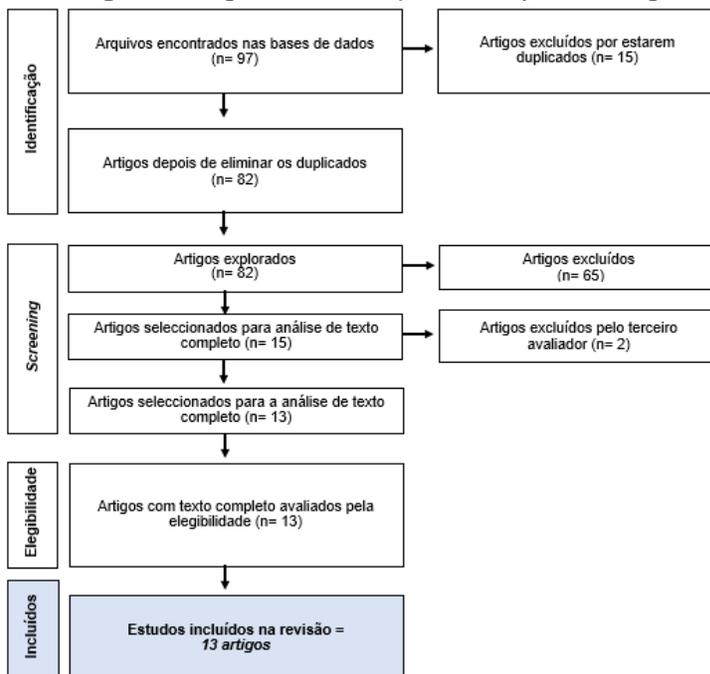
artigos considerados importantes para a pesquisa, tomando como referência os critérios supracitados.

Após, os artigos que não obtiveram concordância entre os dois primeiros avaliadores, passaram pela avaliação de um terceiro avaliador-autor. O terceiro avaliador, após sua leitura decidiu pela permanência ou não do manuscrito.

Critérios de elegibilidade

O objetivo desta terceira fase foi fazer uma seleção dos artigos de texto completo que foram incluídos na revisão final. Para a seleção dos estudos disponíveis nas bases de dados, não foi realizada uma restrição quanto ao ano de publicação de cada estudo, devendo somente obedecer aos critérios: ser um estudo empírico, envolver terapia com BV-U, os participantes deveriam apresentar DSF, o idioma deveria ser o inglês, os artigos deveriam estar publicados ou no prelo, disponíveis na versão online.

Figura 1 - Diagrama de fluxo para a seleção dos artigos



Após a seleção dos artigos incluídos na revisão foram extraídos e analisados os dados relevantes de cada pesquisa, os quais foram codificados em uma tabela e um quadro. Esse estudo de revisão sistemática não fornece um protocolo disponível na web, tendo sido realizado a apresentação dos estudos em dois quadros sendo um com a identificação dos artigos e o outro para o entendimento dos procedimentos terapêuticos realizados em cada estudo. Os quadros estão apresentados na seção de resultados. O fluxograma (Figura 1) representa o fluxo de seleção dos artigos desta revisão.

Análise dos dados

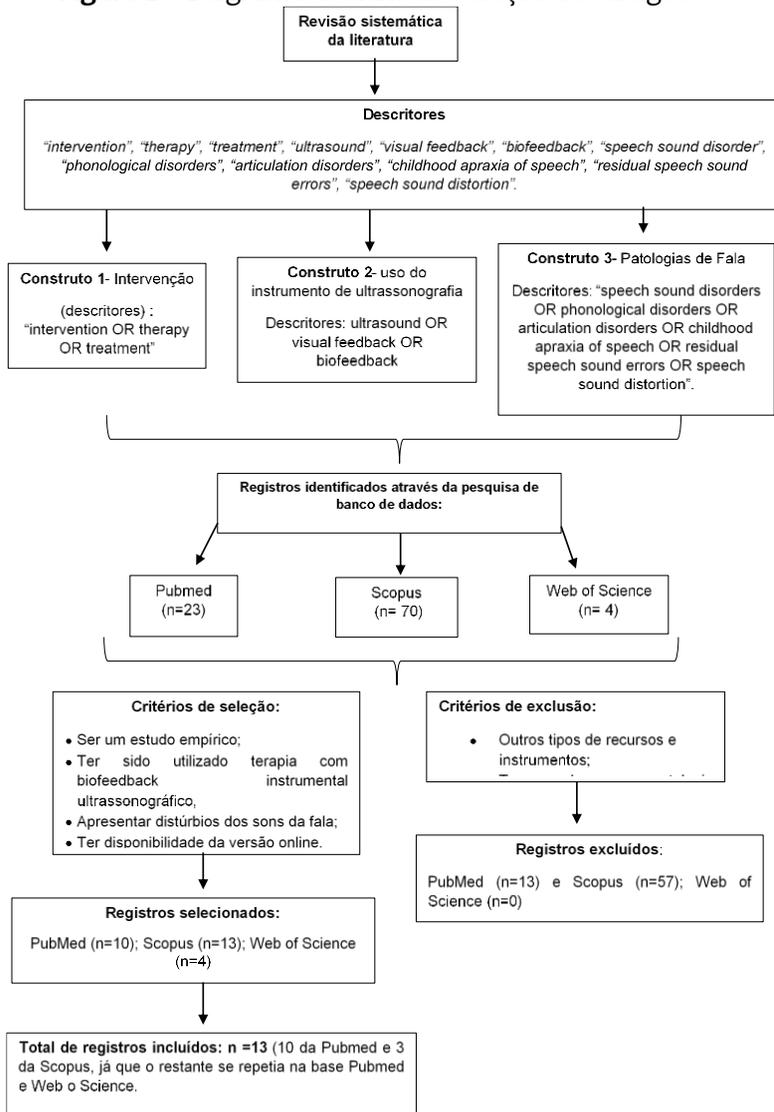
Na fase de identificação, a busca inicial realizada pelos pesquisadores, resultou em um total de 97 resumos (PubMed = 23, Scopus = 70, Web of Science = 4). Houve 100% de concordância entre os pesquisadores na busca. Dos 97 resumos que a pesquisa inicial mostrou, artigos duplicados e artigos que não se adequavam aos critérios de inclusão foram eliminados (n = 84), restando um total de 13 artigos a serem explorados.

Na fase de triagem, dos 97 estudos explorados, 84 foram *excluídos* de acordo com os critérios definidos nessa fase (PubMed = 10, Scopus = 67, Web of Science = 4). Desta forma, foram estabelecidos 13 artigos para a revisão do texto completo. Entre os pesquisadores, houve concordância na seleção de 13 artigos (100%).

Na fase de elegibilidade, o texto completo de 13 artigos foi então avaliado. Destes, todos preenchiam os critérios de inclusão do estudo na revisão final.

Para indicar o processo de seleção dos estudos, foi realizado um diagrama de fluxo (Figura 2).

Figura 2 – Diagrama de fluxo de seleção dos artigos



A revisão identificou 13 artigos que atendiam aos critérios de busca e elegibilidade. O texto completo de todos os estudos selecionados foi revisado pelo primeiro autor. Para facilitar a análise, as informações coletadas foram dispostas em dois quadros. O Quadro 1 apresenta os estudos encontrados, e são informados os seguintes

itens: (1) autor e ano, (2) título do artigo, (3) tipo de DSF, (4) objetivo, (5) tipo do estudo; (6) amostra, (7) idade e (8) escala Pedro P. No Quadro 2 foram inseridos os aspectos relevantes aos procedimentos terapêuticos realizados nestes estudos, e os resultados obtidos em cada estudo, incluindo-se os itens: (1) número de sessões; (2) tempo da sessão; (3) uso da USG; (4) som-alvo; (5) outros auxílios; e (6) resultados.

Quadro 1 – Identificação dos estudos envolvendo terapia USG em DSF

Estudo	AUTOR/ ANO	DSF	Título	Objetivo	Tipo de Estudo	Amostra (n)	Idade (anos)	Escala Pedro P
E1	Preston, Brick, Landi (2013)	Apraxia de fala infantil	Ultrasound biofeedback treatment for persisting childhood apraxia of speech	Avaliar a eficácia de um tratamento com biofeedback ultrassonográfico em crianças com erros residuais de fala associados à apraxia da fala na infância	Estudo de caso não randomizad o	6	9 a 15	Não aplicável
E2	Preston, et al. (2014)	Erros residuais de fala	Ultrasound visual feedback treatment and practice variability for residual speech sound errors.	Testar a eficácia de um tratamento com biofeedback ultrassonográfico em participantes com erros residuais de fala e investigar se a adição de pistas prosódicas facilita a aprendizagem do som alvo	Estudo de caso não randomizad o	8	10 a 13	Não aplicável
E3	Byun, Hitchcoc k, Swartz (2014)	Erros residuais de fala	Retroflex Versus Bunched in Treatment for Rhotic Misarticulation: Evidence From Ultrasound Biofeedback Intervention	Documentar a eficácia do tratamento de biofeedback por USG para a má articulação do rotacismo inglês norte-americano em crianças.	Estudo de caso não randomizad o	1ºE= 4 2ºE= 4	6 a 10 7 a 15	Não aplicável

E4	Cleland, Scobbie, Wrench (2015)	Erros residuais de fala	Using ultrasound visual biofeedback to treat persistent primary speech sound disorders	Utilizar o biofeedback USG em crianças que não responderam à tratamentos anteriores. Além disso, utilizar como modelo terapêutico, vídeos de falantes típicos de idade correspondente.	Estudo em série	7	6 a 11	Não aplicável
E5	Sjolie, Leece, Preston (2016)	Erros residuais de fala	Acquisition, retention, and generalization of rhotics with and without ultrasound visual feedback.	Fornecer uma comparação preliminar entre participantes da terapia de fala com e sem feedback ultrassonográfico para os róticos pós-vocálicos (/r/)	Estudo randomizado	4	7 a 9	Aplicável
E6	Preston, Leece, Maas (2016)	Apraxia de fala infantil	Intensive Treatment with Ultrasound Visual Feedback for Speech Sound Errors in Childhood Apraxia.	Verificar a eficácia de um programa intensivo de terapia de duas semanas, incluindo biofeedback ultrassonográfico	Não randomizado	3	10 a 14	Não aplicável
E7	Bressmann, et al. (2016)	Erros residuais de fala	Perceptual, durational and tongue displacement measures following articulation therapy for rhotic sound errors.	Desenvolver e testar medidas quantitativas que possam descrever as alterações motoras resultantes da terapia de articulação para /r/.	Não randomizado	6	7 a 10	Não aplicável

E8	Preston, et al. (2016)	Apraxia de fala infantil	Limited acquisition and generalization of rhotics with ultrasound visual feedback in childhood apraxia	Verificar se o biofeedback ultrassonográfico da língua facilitaria a aquisição de /ɹ/; se levaria à generalização e a influência da prosódia.	Não randomizado	3	10 a 13	Não aplicável
E9	Preston, Leece (2017)	Erros residuais de fala	Intensive Treatment for Persisting Rhotic Distortion: A Case Series.	Verificar mudanças na precisão dos róticos durante terapia intensiva de 1 semana em adolescentes e adultos jovens com erros residuais de fala que não obtiveram resultados em terapias anteriores.	Estudo de caso não randomizado	4	13 a 22 anos	Não aplicável
E10	Preston; Leece; Maas (2017)	Erros residuais de fala	Motor-based treatment with and without ultrasound feedback for residual speech-sound errors	Comparar a generalização de róticos /ɹ / tratados com e sem ultrassom em crianças com erros residuais de fala que já haviam realizado tratamentos anteriores sem sucesso.	Estudo de caso não randomizado	12	10 a 16	Não aplicável
E11	Preston et al. (2017)	Apraxia de fala infantil	Variable Practice to Enhance Speech Learning in Ultrasound Biofeedback Treatment for Childhood Apraxia of Speech: A Single Case Experimental Study	Avaliar o efeito da variação prosódica na generalização do som da fala durante a terapia com biofeedback ultrassonográfico em crianças com apraxia.	Estudo de caso não randomizado	6	8 a 16	Não aplicável

E12	Preston, et al. (2018)	Erros residuais de fala	Tongue shapes for rhotics in school-age children with and without residual speech errors.	Caracterizar as formas da língua nas produções corretas e distorcidas de róticos em crianças escolares falantes do inglês americano com e sem erros residuais e observar mudanças nas formas da língua naquelas com erros residuais após um breve período de tratamento.	Estudo de caso não randomizado	30 (14 fala típica e 16 erros residuais de fala)	9 a 16	Não aplicável
E13	Preston, et al. (2018)	Erros residuais de fala	Treatment for Residual Rhotic Errors With High- and Low-Frequency Ultrasound Visual Feedback: A Single-Case Experimental Design	Explorar como a frequência com que o biofeedback visual de ultrassom (BV-U) é fornecido durante a fonoterapia afeta a aprendizagem do som da fala.	Estudo de caso não randomizado	12		Não aplicável

Quadro 2. Procedimentos terapêuticos realizados em cada estudo.

Estudos	Procedimentos terapêuticos					Resultados
	N° Sessões	Tempo de sessão	Uso do USG	Som-alvo	Outros auxílios	
E1	18 sessões	1 hora	Foram tratadas 2 sequências-alvo por sessão, organizadas da seguinte maneira: 15 minutos para biofeedback por USG para uma sequência, 20 minutos com terapia convencional, 15 minutos para biofeedback por USG com outra sequência e 10 minutos para lista de sondagem.	Róticos, encontros consonantais com /r/, com /s/ e //.	-----	Foi verificada a aquisição, generalização e retenção. Todos adquiriram precisão do som alvo em pelo menos 2 sequências de sons tratados. A porcentagem de fonemas corrigidos aumentou significativamente desde o pré-tratamento até o seguimento de 2 meses. Para os 6 participantes, o critério de desempenho foi atingido para 23 das 31 sequências tratadas em uma média de 5 sessões. Alguns participantes conseguiram generalizar o som aprendido para sequências não tratadas quando eram foneticamente semelhantes às sequências tratadas. Houve retenção na maioria dos participantes, sendo os ganhos mantidos 2 meses após o final do tratamento.
E2	14 sessões	1 hora	O som alvo foi tratado em dois contextos: com USG aliado à sinalização prosódica por 7 sessões; e com USG sem	/ɹ/ e /s/	Sinalização prosódica	Foi verificada a aquisição, generalização e retenção dos róticos. Para a maioria dos participantes, houve um aumento na precisão dos

			<p>sinalização prosódica por 7 sessões.</p> <p>- 6 a 8 minutos iniciais para sondagens, seguidos dos períodos de produção: A, B, C, D, durante 13 minutos cada. Os períodos A e C foram com feedback ultrassonográfico e B e D, sem esse feedback.</p>			<p>contextos sonoros alvo no nível da palavra, independentemente de haver ou não indicações de prosódia. Foi observada generalização entre onset simples e complexo e para o nível de sentenças. Houve evidências de retenção do aprendizado para 7 dos 8 participantes, visto nas sondagens feitas 2 meses após o tratamento.</p>
E3	16 sessões	30 minutos	<p>Foram 2 sessões introdutórias para interpretação de imagens de USG. As sessões seguintes foram marcadas por 3-5 minutos iniciais de pré-prática, para a criança aproximar-se ao som-alvo com base em imagens de USG. Os minutos restantes foram para a produção de 40 tentativas do som em contexto silábico, somente com auxílio verbal do terapeuta.</p> <p>1ºEstudo= 4 participantes orientados para corresponder a um alvo em forma de língua; 2ºEstudo= 4 participantes receberam sugestões</p>	Róticos	-----	<p>Foi verificada apenas a aquisição e generalização.</p> <p>1ºEstudo: Houve apenas efeitos mínimos de tratamento.</p> <p>2ºEstudo: Todos os participantes obtiveram generalização, visto por meio de uma melhoria na produção dos róticos em palavras não tratadas, produzidas sem biofeedback, demonstrando que a terapia com USG pode ser uma intervenção muito eficaz para crianças com erros róticos persistentes. Além disso, na comparação qualitativa dos Estudos 1 e 2 sugere que o tratamento para o inglês norte-americano deve incluir oportunidades para explorar diferentes formas de língua, para</p>

			individualizadas visando elicitar a forma da língua mais facilitadora para a obtenção de uma percepção maior para melhor produção dos róticos.			encontrar a variante mais facilitadora para cada falante individual.
E4	12 sessões	1 hora	Cada sessão foi dividida em: 30 minutos com apoio do USG e 30 minutos sem esse apoio, além de serem indicados exercícios para casa, feitos 1 ou 2 vezes na semana.	Frontalização de velar (P1) Frontalização de alveolar (P2) Processos idiossincráticos (P3)	-----	Foi avaliada a aquisição, generalização e retenção dos sons alvo. <u>Processo de frontalização de velar /r/:</u> 3 dos 4 participantes conseguiram produzir o som alvo durante e após a terapia. Apenas um participante não conseguiu adquirir o som. <u>Processo de frontalização pós alveolar /S/→/s/:</u> O som-alvo (S) foi totalmente adquirido no pós terapia. <u>Processo idiossincrático:</u> O participante apresentava as plosivas alveolares [t], [d] de forma variável (como [k], [g], [s], [ʃ], [r]), sendo estimulável apenas uma vez no pré terapia para o som [d] mas não pra [t], tendo sucesso de 100% de produção no pós terapia. Após a intervenção, todos obtiveram progressos significativos em segmentos-alvo, evidenciados tanto por medidas perceptuais quanto por mudanças na forma da língua.

E5	14 sessões	40 minutos	Foram realizadas 7 sessões com biofeedback USG e 7 sem esse recurso, sorteadas em ordem aleatória. Cada sessão incluiu três períodos de 13 minutos de terapia (A, B, C). O tratamento com USG e o tratamento com sem USG foram idênticos no alvo, métodos de seleção, duração do tratamento, horário de prática e frequência de feedback. A diferença entre as condições foi o uso do USG por 26 min nas sessões que tinham o tratamento por USG (13 min nos Períodos A e C), aumentando a especificidade do feedback de desempenho.	Róticos (/j/ pós vocálico)	-----	Foi avaliada a aquisição, generalização e a retenção dos sons alvo (após 1 mês sem atendimento fonoaudiológico). Dois participantes não mostraram evidências de aquisição, retenção ou generalização de róticos (<5%). Dos dois que mostraram evidência de aquisição, um se destacou pela significativa melhora nas sessões ultrassonografia. Não houve diferença entre as condições de tratamento na generalização ou retenção de róticos. Conclui-se que para algumas crianças, a aquisição pode ser facilitada pelo biofeedback visual de USG, mas que nessa pesquisa ele não inibiu nem facilitou a retenção ou a generalização de róticos.
E6	10 sessões de terapia intensiva (2 semanas)	(2,5 horas por dia)	Cada hora abordou um som em uma posição silábica: onset (simples e complexo) e rima, com 30 minutos de intervalo; 6 a 10 minutos iniciais foram para treino da percepção auditiva. O restante para o treinamento de produção. A	-Sons tratados: /j/, /s/ e /ʃ/. -Para o participante A: /j/ em onset e /j/ rima; - Para o B: /j/ onset, /j/ rima; - Para o C: /s/	-----	Foi avaliada a aquisição, generalização e a retenção dos sons alvo Todos os participantes mostraram sinais de aquisição dos sons em que haviam trocas. Os resultados de generalização e retenção foram mistos. Um participante mostrou generalização e retenção de sons que

			<p>produção foi dividida em quatro períodos de 12 minutos (A, B, C e D). O período A e C incluíram o apoio ultrassonográfico e os períodos B e D não.</p>	<p>onset e /tʃ/ rima.</p>		<p>foram tratados; um mostrou generalização, mas retenção limitada; e o terceiro não mostrou evidências de generalização ou retenção. Características individuais que podem facilitar a generalização são discutidas. Programas de tratamento intensivo a curto prazo usando biofeedback por ultrassom podem resultar na aquisição de padrões motores mais precisos e melhor articulação de sons previamente em erro, com níveis variados de generalização e retenção.</p>
E7	10 sessões	1 hora	<p>O som alvo foi abordado em posição inicial, medial e final, além de encontros consonantais;</p> <p>Dos 6 participantes, 4 tiveram suas sessões suplementadas com 10 min de biofeedback USG.</p> <p>Foram avaliadas as impressões perceptivas de sentenças com /ɹ/ e para avaliar se a terapia (com ou sem biofeedback por USG) tinha um impacto na duração das sílabas VV, foi medido o</p>	<p>Róticos /ɹ/.</p>	-----	<p>Foi avaliada a aquisição, generalização e a retenção dos sons alvo.</p> <p>Perceptivamente, os ouvintes perceberam melhora do /ɹ/ nas avaliações intermediárias e finais. As durações das sílabas VV no ponto médio da terapia foram maiores. As medidas de duração e deslocamento de língua, que aumentou na sessão final, ilustraram como a terapia de articulação afetou o movimento da língua e pode ser útil na terapia de articulação. Todos os participantes adquiriram os sons-alvo, mas a</p>

			tempo total a produção de cada V1V.			generalização e retenção foi mista. Um participante mostrou generalização e retenção de sons que foram tratados; um mostrou generalização, mas retenção limitada; e o terceiro não mostrou evidências de generalização ou retenção. Tratamentos intensivos a curto prazo usando biofeedback por ultrassom podem resultar na aquisição de padrões motores mais precisos e melhor articulação de sons previamente em erro, com níveis variados de generalização e retenção.
E8	14 sessões	1hora	Tratamento foi dividido em 2 fases: Fase I (sessões 1-7) abordou uma posição (onset inicial ou coda) junto às condições de “com prosódia” ou “sem prosódia”. A Fase II (sessão 8-14) ocorreu da mesma forma, porém com o / J/ na outra posição; Cada sessão foi dividida em quatro períodos de 13 minutos (A, B, C, D). A e C incluíram o apoio ultrassonográfico e os	Róticos (/J/)	Sinalização prosódica	Foi avaliada somente a aquisição e generalização dos sons-alvo. Houve uma aquisição limitada em relação ao som abordado, visto que duas das três crianças conseguiram adquirir o som, porém não generalizaram. Um dos participantes que teve boa evolução recebeu a sinalização prosódica apenas na Fase II, não sendo possível determinar se esse recurso teve influência na facilitação da aprendizagem motora. O estudo indicou que o biofeedback de ultrassonografia pode não facilitar a rápida aquisição ou generalização

			períodos B e D não. O estímulo prosódico foi incluído em uma fase selecionada aleatoriamente para cada participante.			para os róticos em determinadas crianças com apraxia de fala, sendo necessária uma maior investigação por meio de outros estudos.
E9	1 semana (14 horas de terapia intensiva.	2 horas no período da manhã e 2 horas à tarde	O som alvo foi abordado em duas posições: onset inicial (simples e complexo) e em rima (núcleo e coda). As sessões 1-7 incluíram 6 a 8 minutos de julgamentos perceptuais auditivos das gravações corretas e distorcidas / ɹ / . Cada sessão incluiu quatro períodos de 12 minutos (A, B, C, D). A e C tiveram biofeedback ultrassonográfico e os períodos B e D não.	Róticos (/ɹ/)	-----	Foi avaliada somente a aquisição do som-alvo. Todos apresentaram melhora mensurável e houve diferença na precisão do som da fala. Nos quatro participantes, a precisão rótica no nível da palavra melhorou de 35% para 83%. No nível da sentença, essa precisão aumentou de 11% antes do tratamento para 66% após o tratamento em 1 semana. A promessa de um programa de tratamento intensivo que inclui treinamento baseado em motor, biofeedback e percepção auditiva é ilustrado pelas apresentações de casos em que melhorias substanciais na precisão do som da fala foram observadas.
E10	14 sessões	1 hora	O tratamento foi dividido em duas fases: fase I - som-alvo em uma posição silábica; fase II: em outra posição silábica. Uma fase contou com o	Róticos /ɹ /	-----	Foi avaliada apenas a aquisição e generalização do som alvo Em média, ambos os tratamentos (com e sem USG) geraram um aumento de 30% na precisão das

			biofeedback ultrassonográfico e a outra não. Para isso, foram sorteados os sons alvo individualmente para cada fase do tratamento, tendo cada uma delas, 7 sessões. As sessões foram divididas em períodos de 13 minutos (A e C com Ultrassom e B e D sem apoio do ultrassom).			palavras não tratadas em 7 sessões, visto que 10 dos 12 participantes apresentaram progressos, ou seja, generalizaram. No entanto, a variabilidade na resposta sugeriu que algumas crianças mostraram uma resposta preferencial a um tratamento em detrimento de outro, algumas responderam bem a ambas as intervenções e algumas responderam minimamente a ambas as intervenções. Ambas as abordagens podem ser opções viáveis para algumas crianças, necessitando ainda de estudos para determinar quais crianças são as melhores candidatas para intervenções com e sem biofeedback visual de USG.
E11	14 sessões	1 hora	Foram tratados dois sons alvo com biofeedback USG: um com variação prosódica e um sem variação prosódica. Cada alvo foi tratado por metade da sessão, sendo sorteados quanto à condição de “com variação prosódica” e “sem variação prosódica”. Foram designados 8 minutos para	/ɹ/ e /s/	Sinalização prosódica	Foi avaliada a aquisição, generalização e retenção do som-alvo. Todos os participantes mostraram maior mudança nas sondagens de generalização para alvos tratados sob a condição de variação prosódica. O aumento médio nos escores de generalização foi de 38% na condição de variação prosódica em

			treino de percepção, seguidos de 20 para a produção (10 minutos com o USG e 10 minutos sem USG).			comparação com 31% sem a variação. Em geral, os quatro participantes obtiveram retenção das habilidades. Entende-se que o biofeedback por USG pode facilitar a aprendizagem do som da fala, que pode ser aprimorada pelo tratamento de sons com variação prosódica explícita.
E12	14 horas	Não relatado	A pesquisa foi realizada com dois grupos: participantes com fala típica e participantes com erros residuais de fala (ERF). Nos participantes com ERF foi realizada terapia com feedback visual de USG. Foram 7 horas de terapia com USG, depois a criança era avaliada e realizava mais 7 horas de terapia também com USG, a fim de verificar as mudanças na forma de língua após o tratamento e se essas correspondiam à precisão percebida.	Róticos (/l/)	-----	Foi avaliada somente a aquisição (precisão) do som-alvo. As crianças com fala típica produziram róticos com maior precisão e com formas de língua mais complexas do que as crianças com erros residuais de fala. Após 7 h de terapia fonoaudiológica para o grupo de ERF, houve um aumento tanto na precisão percebida dos róticos quanto nos valores de medida de complexidade do movimento de língua. Percebeu-se que a curvatura da língua na primeira sessão para esse grupo foi de menor modulação, tendo formas menos complexas quando comparadas ao grupo de fala típica. Nas sessões média e pós tratamento, também para esse grupo, houve um aumento significativo na modulação de língua em onset inicial, mas sem

						efeito significativo para onset complexo. Esse aumento de complexidade de língua foi correlacionado positivamente à classificação perceptivo-auditiva sobre a precisão dos róticos, com o onset complexo sendo julgado como mais precisos.
E13	16 sessões	1hora	Foram especificadas duas condições de tratamento: uma condição apresentou 8 horas de biofeedback visual de USG de alta frequência (AF) (89% dos ensaios), enquanto a outra incluiu 8 horas de biofeedback visual de USG de baixa frequência (BF) (44% dos ensaios). A ordem das condições de tratamento foi contrabalançada entre os participantes. O progresso foi rastreado medindo a generalização em /ɹ/ em palavras não tratadas.	Róticos (/ɹ/) vocálico	-----	Foi avaliada somente a aquisição e generalização do som alvo. Após a 1ª fase de tratamento, os participantes que receberam a condição AF superaram aqueles que receberam BF. No final das duas fases de tratamento, as comparações intra participantes mostraram variabilidade entre os resultados individuais nas duas condições. No entanto, uma análise em nível de grupo sugeriu que os participantes cuja ordem de tratamento era AF-BF, obtiveram ganhos maiores do que aqueles cuja ordem de tratamento foi BF- AF. Isso sugere que o biofeedback da USG pode ser benéfico nos estágios iniciais de aquisição de novos alvos de fala.

Resultados

A partir dos achados dessa revisão sistemática, destaca-se que a maioria dos estudos (E1, E2, E4, E5, E6, E7, E11) buscou verificar, ao longo do processo terapêutico de diferentes DSF, a ocorrência das três fases necessárias para a aprendizagem motora de um determinado som (SJOLIE et al., 2016). O restante dos estudos buscou entender se havia ocorrido as duas primeiras fases do aprendizado motor (aquisição e generalização), sendo visto que desses, quatro (E3, E8, E10, E13) se propuseram a verificar a ocorrência de ambos e apenas dois (E9, E12) verificaram somente a ocorrência da primeira fase (aquisição).

Em relação ao tipo de DSF, destaca-se que 31% (n=4) dos artigos envolviam a AFI, e 69% (n=9) envolviam os ERF. Dentre o total dos 13 estudos, 77% (n=10) apresentaram resultados condizentes com melhora na precisão, generalização e retenção do som-alvo, conforme o objetivo de verificação proposto nas terapias com BV-U. Isso demonstra que para a maioria dos participantes, esse tipo de terapia é eficaz, já que muitos já haviam realizado terapia convencional anteriormente e não haviam obtido melhoras significativas.

Notou-se, no entanto, que em 23% (n=3) dos estudos (E5, E6 e E8), os resultados diferiram da maioria.

No E5, que tratou ERF e verificou as três fases para aprendizagem motora do som (aquisição, generalização e retenção), apenas metade dos sujeitos conseguiram adquirir o som-alvo, não ocorrendo a generalização do som aprendido em palavras não tratadas e nem a retenção, verificada meses após o final do tratamento. No E6, também foram verificadas as três fases da aprendizagem motora do som tratado, porém, para casos de AFI. Notou-se que todos os sujeitos conseguiram adquirir o som-alvo durante as sessões, porém, os resultados verificados na generalização de palavras não tratadas e na retenção do aprendizado foram mistos.

O E8, que avaliou duas fases da aprendizagem motora (aquisição e generalização) no tratamento de AFI, dois dos três sujeitos tiveram aquisição do som-alvo durante as sessões, porém nenhum obteve generalizações para palavras não tratadas ao final das 14 sessões de tratamento ou no pós-tratamento.

É importante destacar que nesses estudos em que não houve uma melhora significativa nas fases de generalização e retenção do som alvo, os participantes foram capazes de realizar o gesto motor de língua para o som treinado durante o uso do BV-U, mas não foram capazes de generalizar para palavras não treinadas quando o recurso da USG era retirado, demonstrando assim, que podem ter se tornado dependentes do apoio que lhe foi oferecido, sem terem tido a necessidade de memorização dos gestos de língua quando sem a visualização por meio da USG (HODGES e FRANKS, 2001).

O som-alvo tratado pode ter sido um fator contribuinte para essa dificuldade, visto que os três estudos envolviam o rótico /r/ - som este, de complexidade maior, por requerer uma movimentação coordenada e simultânea tanto de ponta quanto de dorso de língua, sendo um fator de dificuldade tanto para casos de AFI quanto para casos de ERF (GICK et al., 2007).

Além disso, nos três estudos citados, envolvendo dificuldades nas fases de generalização e retenção, dois eram de AFI, o que pode ter contribuído para tais resultados, visto que essas dificuldades já foram descritas em outros estudos (BALLARD et al., 2010; MAAS et al., 2012), os quais referiram que é necessário um maior número de sessões para auxiliar no processo de transição da aquisição para aprendizagem, já que há uma maior dificuldade na programação e coordenação de movimentos refinados nesses casos (MAAS et al., 2014).

Com base na proporção dos achados dessa revisão, observa-se que houve um maior benefício da terapia com BV-U em casos de ERF quando comparados aos casos de AFI, já que do total de quatro estudos encontrados de AFI, (E1, E6, E8 e E11), 50% (n=2) (E6 e E8) apresentaram dificuldades na generalização e retenção das habilidades, apesar de terem conseguido adquirir o som-alvo durante as sessões. Acredita-se que isso se atribua à dificuldade inerente ao distúrbio, que embora com uma maior estimulação visual, fornecida pelo instrumento de USG, as dificuldades na realização motora do gesto ainda continuam existindo, principalmente quando esse recurso é retirado. Talvez se o tempo de tratamento fosse ampliado para esses casos, pudesse surgir um maior efeito a longo prazo (PRESTON; LEAMAN, 2014; BYUN et al., 2014).

a) Estrutura das sessões

Apesar de haver variações em relação ao período do tratamento e ao momento do uso do instrumento de USG, em todos os estudos foram realizadas sondagens ao longo do tratamento. Em todos os estudos foram feitas avaliações pré-terapia; durante a metade do tratamento; após a última sessão do tratamento e após dois meses do término do tratamento, sendo esta última não realizada nos estudos que não se propuseram a analisar a ocorrência da retenção (E9, E12, E3, E8, E10, E13).

O tempo de duração das terapias variou entre 10 a 18 sessões ($\mu=14$ sessões). O tempo para a duração de cada atendimento na maioria dos estudos foi equivalente a 60 minutos, exceto para os E6 e E9, que realizaram terapia intensiva. No E6, a terapia teve duração equivalente a duas horas por dia, durante dez dias. No E9 a terapia teve duração de 4 horas, sendo duas no período da manhã e duas no período da tarde, totalizando 14 horas de terapia, dispostos em apenas 5 dias.

Em seis estudos a estrutura da terapia ocorreu de forma semelhante (E2, E5, E6, E8, E9 e E10), incluindo um treinamento inicial de percepção e posteriormente, o treinamento da produção com o instrumento de USG, com uma subdivisão em 4 períodos de produção com duração de aproximadamente 12 minutos cada (A, B, C, D) dentro da mesma sessão. O período A e C incluíam o biofeedback por USG e os períodos B e D incluíam a prática sem o instrumento. Os demais estudos variaram em relação a estrutura da sessão, mas em todos sempre existiram momentos com e sem o BV-U, com o intuito de que a criança não necessitasse deste recurso em tempo integral, para favorecer o aprendizado do gesto de língua (SECORD et al., 2007; SHRIBERG, 1975). É importante que seja feita a distinção entre o desempenho durante a aquisição do som alvo e a aprendizagem desse som, que se dará posteriormente, quando o sujeito generalizar o gesto para sons não-treinados e conseguir reter essa informação ao longo do tempo (SCHMIDT; LEE, 1999; GUENTHER et al., 1999).

b) Som alvo

Observou-se que todos os estudos envolveram a investigação da produção dos róticos /ɹ/ do Inglês norte americano. Não foram

encontrados estudos com falantes do Português Brasileiro. Estes costumam ser abordados em pesquisas porque é uma alteração comumente encontrada em casos de AFI e ERF, por requerer uma maior complexidade articulatória, com uma retroflexão de língua e maior habilidade motora fina, sendo por isso, um som que muitas vezes leva mais tempo a ser adquirido (GUENTHER et al., 1999; MCGOWAN et al., 2004; SMIT et al., 1990). Em todos esses estudos, envolvendo esse som-alvo, notou-se melhoras significativas na aquisição, generalização e retenção. Os únicos estudos que abrangeram outros sons além dos róticos, foram: E1, E2, E4, E6, E11.

No E1, tratamento para AFI, foram tratadas oito sequências-alvo de um som por criança, envolvendo sons linguais (os sons róticos e os encontros consonantais com /r/, /s/ e /l/). Toda vez que fosse atingido 80% de precisão para uma sequência alvo tratada em duas sondagens consecutivas, trocava-se a sequência, não sendo abordadas mais que duas por sessão. A USG foi utilizada nos 15 minutos iniciais em uma determinada sequência, depois foram 10 minutos para uma atividade lúdica e mais 15 minutos para novamente o uso do biofeedback da USG em outra sequência (como o /kl/). Os 10 minutos finais foram para aplicação de lista de sondagem. Observou-se que todos adquiriram precisão do som alvo em pelo menos duas sequências de sons tratados e que alguns participantes conseguiram generalizar o som aprendido para sequências não tratadas quando eram foneticamente semelhantes às sequências tratadas. Além disso, houve retenção do aprendizado na maioria dos participantes.

No E2, o tratamento abrangeu os sons /r/ e /s/ em casos de ERF, com um dos sons trabalhado na primeira fase do tratamento (primeiras 7 sessões) e o outro na segunda fase (as 7 sessões seguintes), variando a ordem para cada participante, conforme sorteio, além da condição “com e sem prosódia” que também foi abordada e descrita aqui, na sessão de “outros auxílios”. Destaca-se que em relação aos sons descritos, houve melhorias em uma ou em ambas as fases da intervenção para a maioria dos participantes, havendo um aumento na precisão dos contextos sonoros alvo no nível da palavra, independentemente de haver ou não indicações de prosódia. Foi observada generalização entre onset simples e complexo e para o nível de sentenças. Além disso, houve evidências de retenção do aprendizado, para 7 dos 8 participantes, visto nas sondagens feitas

dois meses após o tratamento, concluindo-se assim, que um programa de tratamento por USG pode facilitar o aprendizado dos sons da fala em indivíduos com ERF.

No E4, tratamento para ERF, foram abordados os processos de frontalização de velar /r/, frontalização pós alveolar (/S/→[s]) e os processos idiossincráticos (/t/ e /d/ eram produzidos de forma variável, como /k/, /g/, /s/, /S/, /r/). Foi observado que para o primeiro processo, três dos quatro sujeitos tiveram como resultados uma ausência de produção velar na pré terapia, conseguindo produzi-las na sondagem durante a terapia, porém sem generalizar. Durante o pós-terapia, o sujeito atingiu 100% de produções corretas. Apenas um sujeito não conseguiu adquirir o som-alvo, mantendo-se com percentual correspondente a 0% no pré-terapia, sendo isso associado a um problema de atenção diagnosticado.

Para o segundo processo, no pré terapia o sujeito produziu apenas 30% do som-alvo corretamente. No pós-terapia estava totalmente adquirido o som /S/ e para o terceiro processo, o sujeito apresentava as plosivas alveolares [t], [d] de forma variável ([k], [g], [s], [S], [r], sendo estimulável apenas uma vez no pré terapia para o som [d] mas não pra [t], tendo sucesso de 100% de produção no pós-terapia. Notou-se que independente do processo abordado, todas as crianças obtiveram progressos significativos com o tratamento por BV-U.

O E6, que envolveu terapia em casos de AFI, investigou a produção dos sons /r/, /s/ e /tS/. Os resultados sugeriram que todos os participantes mostraram sinais de aquisição dos sons em que haviam trocas. Os resultados de generalização e retenção foram parciais, conforme citado anteriormente. Um participante mostrou generalização e retenção de sons que foram tratados; um mostrou generalização, mas retenção limitada; e o terceiro não mostrou evidências de generalização ou retenção.

O E11, que envolveu terapia em casos de AFI, investigou a produção dos sons /j/ e /s/. Ambos tratados sob condições diferentes em cada sessão. Nos primeiros 30 minutos foi abordado um som-alvo e nos 30 minutos seguintes, o outro som alvo. Para cada participante, um som alvo foi sorteado a receber ou não a variação prosódica, sendo assim, se um som-alvo havia sido sorteado como sendo “com variação prosódica”, o segundo era “sem variação prosódica”. A estrutura da sessão era: sondagem do som-alvo seguida de 8 minutos para treino

de percepção auditiva, 20 minutos para o treino da produção (10 minutos com USG e 10 minutos sem USG) sob a condição de “com prosódia”, por exemplo. Nos 30 minutos seguintes, era abordado o outro som, sob o mesmo esquema.

Todos os participantes mostraram maior mudança nas sondagens de generalização para alvos tratados sob a condição de variação prosódica. O aumento médio nos escores de generalização foi de 38% na condição de variação prosódica em comparação com 31% sem a variação. Em geral, os quatro participantes obtiveram retenção das habilidades, entendendo-se assim, que o biofeedback por USG pode facilitar a aprendizagem do som da fala e que esta pode ser aprimorada pelo tratamento de sons com variação prosódica explícita.

c) Outros auxílios terapêuticos

Além dessa variação em relação ao som alvo abordado, destaca-se que houve também uma variação em relação a um fator de “possível auxílio” para aquisição do som alvo, que foi a pesquisa com BV-U aliado à investigação da influência da prosódia como recurso facilitador, abordado em três estudos, sendo um relacionado a casos de ERF (E2) e os outros dois a ocorrência de AFI (E8 e E11). Destaca-se que desses três estudos existentes até o momento, pode-se concluir que somente o último (E11) trouxe a prosódia como um fator a ser correlacionado aos demais dados, demonstrando que a prosódia parece influenciar no resultado terapêutico. Nos outros estudos, as estruturas das terapias foram organizadas de forma que não se pôde concluir que as melhorias obtidas foram devido ao auxílio da prosódia. Sendo assim, observou-se alguns aspectos que serão comentados a seguir.

No E2, tratamento para /t/ e /s/ em casos de ERF, o som-alvo foi tratado conforme a estrutura “ABCD”, recebendo auxílio da USG nos momentos A e C. Para cada participante, nas primeiras 7 sessões, foi abordado um contexto sonoro em uma das duas condições do tratamento: com ou sem auxílio prosódico, conforme sorteio. Nas 7 sessões seguintes, foi abordado um contexto sonoro diferente, na outra condição de auxílio (com ou sem sinalização prosódica). Por meio de sondagens, foi realizado um monitoramento das três fases essenciais para o aprendizado, observando-se que houve tanto a aquisição, quanto a generalização e retenção dos sons-alvo.

Em relação à sinalização prosódica, esta parece não ter fornecido uma adição robusta ao tratamento de *biofeedback* por USG, visto que o estudo não foi projetado para isolar um único fator responsável pelos ganhos de tratamento, havendo uma variabilidade nos resultados. Alguns participantes apresentaram respostas igualmente fortes a longo prazo tanto para a condição prosódica quanto para a não prosódica (por exemplo, participantes 86 e 87); outros apresentaram resposta mínima a ambas as condições (por exemplo, participante 88); alguns tiveram o benefício mais claro da condição prosódica sobre a condição não prosódica (por exemplo, participante 89); e alguns participantes apresentaram maior aprendizado sem a sinalização prosódica (por exemplo, participante 91).

No E8, tratamento para róticos (/l/) em casos de AFI, também foi realizado seguindo a estrutura “ABCD”, com a utilização do feedback da USG nos períodos A e C, com presença da prosódia conforme sorteio para cada contexto sonoro. Uma das duas crianças apresentou evolução, adquirindo o som alvo, porém sem generalizar, mesmo recebendo a sinalização prosódica. Entretanto, essa sinalização foi na fase da metade do tratamento (sessão 8 a 14), não sendo possível afirmar que esse recurso teve influência para a facilitação da aprendizagem motora.

No E11, que abordou casos de AFI, foram trabalhados dois sons: um com sinalização prosódica e o outro não, ambos com feedback da USG. Notou-se que houve um aumento médio nos escores de generalização, passando de 38% na condição de “com sinalização prosódica” em comparação com 31% sem essa sinalização. Desta forma, todos os participantes mostraram maior mudança nas sondagens de generalização para alvos tratados sob a condição de variação prosódica.

Segundo estudo de PRESTON et al. (2017), o *biofeedback* visual da USG de língua pode facilitar mudanças no movimento desta para sons de fala direcionados ou sequências de som. No entanto, segundo a teoria da aprendizagem motora baseada em esquemas, a prática de tarefas de fala variável, que exijam a integração dos comandos motores que governam os aspectos segmentais e prosódicos podem acelerar esse aprendizado motor (MAAS et al., 2008; SCHMIDT; LEE, 2011), já que há evidências em tarefas motoras sem fala de que a generalização pode ser melhor facilitada com prática variável em

comparação com a prática constante (HALL; MAGILL, 1995; WULF; SCHMIDT, 1997).

Entende-se assim, que se faz necessário maiores pesquisas para investigar se a influência da prosódia como um recurso capaz de influenciar na aprendizagem, principalmente em casos de AFI (E8, E11), podendo ser usado assim como nos estudos descritos, juntamente ao uso do BV-U, para intensificar e facilitar o aprendizado de sons-alvo.

d) Limitações

Durante a análise, percebeu-se que 8 estudos apontaram algumas limitações (E1, E2, E5, E7, E9, E10, E12 e E13).

No E1, tratamento de AFI, os autores apontaram que os dados mostraram que houve melhora rápida do som-alvo (exceto algumas sequências), no entanto nem todos os efeitos de generalização puderam ser previstos, já que alguns alvos não tratados não tiveram melhorias e nem uma precisão estável e baixa, o que seria preferível, já que se trata de um design de base múltipla, sendo ainda observadas algumas linhas de base crescentes, devido aos efeitos de generalização.

A escolha dos sons-alvo, também foi questionada, visto que não foi utilizado nenhum padrão sistemático para a seleção, sendo esta, baseada nas transcrições fonéticas e julgamento clínico.

Notou-se também que o BV-U é mais eficaz para a produção em sequências curtas (sílabas e palavras). Não há evidências, segundo o estudo, que a generalização ocorra para sequências longas (fala encadeada).

Ainda sobre o E1, o estudo indica que não foi utilizada uma pontuação cega durante as sondagens. A forma de condução do atendimento também parece ter influência, já que 74% das sessões foram feitas por estudantes de fonoaudiologia, sendo possível que os maiores efeitos tenham sido observados quando os terapeutas foram clínicos experientes. Outro aspecto relevante foi que o estudo incluiu graus diferentes de AFI (leves ou moderados), notando-se que o caso mais severo teve um pior desempenho. Com relação a esses resultados pode-se inferir que o biofeedback com ultrassom parece ser mais eficaz em casos de AFI mais leves.

Notou-se também que a função cognitiva de atenção, nesse caso a atenção especialmente ao biofeedback parece ser muito importante para bons resultados. Um dos sujeitos da pesquisa apresentava diagnóstico de Transtorno do déficit de Atenção com Hiperatividade (TDAH), e não teve um bom desempenho quando comparado aos demais. Dessa forma, isso pode ter sido considerado um viés do estudo, uma vez que as crianças deveriam ser comparadas com as mesmas comorbidades associadas (ou de AFI somente).

No E2, sobre o tratamento de AFI, foi destacado que não se pode afirmar que os efeitos obtidos pelo biofeedback da USG de fala tenham ocorrido exclusivamente pelo conhecimento de performance, mas sim por um conjunto de princípios, como a estrutura do protocolo, que tem uma crescente complexidade de tarefas. É possível que se ocorresse uma mudança neste protocolo, os resultados fossem diferentes. Destacou-se ainda que o uso da USG resultou em melhorias, no entanto, neste estudo não foi realizada comparação com outros métodos.

O E5 teve como desenho dois grupos de participantes, recebendo ambos tratamentos (com e sem BV-U), acontecendo nesse caso à randomização dentro de um bloco único, para evitar efeitos cumulativos em um dos tipos de tratamento. No entanto, os autores destacam que não se pôde controlar os efeitos cumulativos que ocorreram durante o tratamento como um todo, sendo essa uma limitação do estudo.

O pressuposto nesse caso era que a generalização fosse atribuída à sessão de tratamento anterior, mas é possível que os efeitos combinados de todos os tratamentos anteriores possam ter influenciado para essa generalização. Além disso, não se pôde tirar conclusões sobre a eficácia do tratamento apenas na terapia com USG ou apenas no tratamento sem USG ao final desse estudo. Ainda, conforme a teoria da aprendizagem motora, é provável que o biofeedback visual proporcione benefícios durante os primeiros estágios de prática, quando a criança ainda não se tornou dependente desse recurso, porém, nessa pesquisa não se pode dizer sobre um ponto durante o processo de aprendizagem em que a visualização das imagens foi mais benéfico ou mais prejudicial.

O E7 teve limitações relacionadas ao desenho do estudo e procedimentos terapêuticos. Uma das limitações indicada pelos

autores foi a ausência da realização de uma linha de base entre as sessões de avaliações, a fim de descartar qualquer melhora espontânea possível. Outro aspecto importante foi que o estudo baseou-se numa abordagem motora fonética diferente da abordagem motora descrita em outro estudo, que permitiria sessões diferenciadas, consistentes e comparáveis.

Além disso, o estudo destacou que nem sempre as imagens ultrassonográficas relacionadas aos encontros consonantais e a retroflexão, de /ɺ/ foram claramente discerníveis para cada participante, pois o contexto fonético parece ter influenciado, principalmente nas vogais adjacentes, não tendo sido tomado esse cuidado na escolha das palavras.

Como o padrão de erro não foi detalhado, não foi possível descartar se o caractere variável de /ɺ/ também pode ter distorcido a análise perceptual, principalmente em relação às medidas instrumentais quantitativas. Os itens de nível de palavra e sentença que foram usados para a avaliação perceptiva foram diferentes das sílabas de VɺV que foram usadas para as medidas de duração, deslocamento e variabilidade. As razões para essa discrepância eram práticas. O movimento para /ɺ/ poderia ser mais precisamente identificado em uma posição medial, especialmente quando a produção era afetada por um erro sonoro. No entanto, é importante entender que as medidas perceptivas e instrumentais foram baseadas em diferentes produções e que sua comparabilidade pode ser limitada.

A tarefa de julgamento perceptual algumas vezes exigia que o ouvinte resumisse suas avaliações de duas realizações de [ɺ] em um único julgamento. A tarefa de categorização perceptiva também exigiu que os ouvintes classificassem realizações distorcidas de [ɺ] como [w] ou [l]. Observou-se que os ouvintes não usavam a categoria [l] com muita frequência. Os ouvintes também poderiam escolher a categoria alternativa de "outros" se suas percepções não se encaixassem nas seleções fornecidas. A categoria foi vagamente definida e os itens classificados como "outros" não foram analisados com mais detalhes. Finalmente, a medida de deslocamento é um quantificador que ainda requer um refinamento em pesquisas futuras, mas permitiu comparar um conjunto muito heterogêneo de superfícies de língua e medi-las independente de seus tamanhos e alturas.

O E9 apresentou bons resultados, porém os autores indicaram que não é possível determinar se isso ocorreu especificamente pelo uso do BV-U, já que envolve também o trabalho de percepção auditiva, por exemplo. O progresso observado nos participantes foi medido por ouvintes experientes em fala cegados para a avaliação, porém não é possível afirmar em qual parte do tratamento houve maior precisão, já que pode acontecer a generalização, tendo em vista que foi feito um monitoramento do pré-tratamento ao pós tratamento, sendo também impossível comparar a condição de “sem tratamento” e “com tratamento”.

O E10 descreveu como limitação o fato de metade dos participantes terem sido tratados por sete sessões com USG antes de serem tratados na condição de “sem USG” (embora tenham sido abordadas diferentes possibilidades em cada fase do tratamento).

Para esses participantes, a exposição inicial ao feedback visual ultrassonográfico pode ter melhorado a compreensão dos movimentos linguais para /ɹ/ e com isso influenciado nos momentos de prática sem o ultrassom. Notou-se também que a melhora na precisão do som alvo em determinada posição silábica influenciou na precisão de outra posição, diferindo da suposição de independência relativa no tratamento dos sons alvo.

O E12 destacou a importância de reconhecer uma associação entre a complexidade da forma de língua medida pelo número de inflexões na curvatura do contorno de língua (medida NINFL – *Number of Inflections*) e a percepção dos ouvintes sobre a precisão rótica, mas destacou que essa correspondência não é direta entre as medidas quantitativas e as classificações dos ouvintes.

Foi indicado que apesar de o estudo ter demonstrado uma relação de um para um, isso pode se atribuir a uma variedade de fatores, incluindo o ambiente fonético, que pode utilizar formas de língua menos complexas; preferência individual em formas de língua para a produção dos róticos; complexidade parassagital não capturada por USG sagital mediana, configuração labial ou falta de estabilização da cabeça, limitando a precisão das imagens.

Destaca-se ainda que a medida NINFL é aplicável às imagens sagitais da língua, mas provavelmente existem outras importantes exigências articulatórias para a precisão que não são capturadas por essa medida, como a elevação das margens laterais da língua e o

posicionamento dos lábios. Assim, a quantificação de imagens sagitais sozinhas pode não refletir totalmente os requisitos articulatórios necessários para atingir uma qualidade rótica aceitável.

O E13 apontou que foi utilizado o BV-U de alta frequência (*High Frequency - HF*) em 88% dos ensaios e o de baixa frequência (*Low Frequency - LF*) em 44% das vezes, mas que poderiam ter sido obtidos resultados diferentes se a proporção do uso desse instrumento tivesse sido definida de forma diferentes para as duas condições.

Além disso, o estudo incluiu apenas crianças que foram submetidas à ordem: HF-LF ou LF-HF para permitir que a exposição total ao ultrassom fosse semelhante após as 16 sessões, porém, nenhuma criança recebeu exclusivamente o tratamento de alta frequência ou exclusivamente o de baixa frequência.

Ainda, embora HF-LF tenha sido a opção mais eficaz neste estudo, é possível que a vantagem observada para HF sobre LF ao longo de oito sessões possa ser aumentada se as crianças submeterem-se à ordem de HF-HF para todas as 16 sessões.

Outra questão apontada pelos autores, refere-se à dificuldade de adaptação desse estudo na prática, visto que apesar de todos os participantes começarem o tratamento no mesmo nível, tendo o mesmo critério para avanço em níveis mais elevados, o tempo gasto praticando em um determinado nível variou entre os participantes devido a características individuais). Assim, é ressaltada a importância de estudos com uma amostra maior, para identificar possíveis fatores que possam influenciar nas respostas às intervenções com biofeedback.

Como limitação desta revisão sistemática, destaca-se que apenas um artigo era randomizado, sendo ainda necessário um número maior de pesquisas envolvendo a terapia de fala nos casos de ERF e AFI juntamente ao uso da USG de fala, a fim de gerar novas evidências e contribuir com os resultados já destacados.

Recente estudo de revisão sistemática (SUDGEN et al., 2019) concluiu que as evidências científicas do uso da BV-U ainda estão em desenvolvimento, pois a maioria utiliza amostras pequenas (estudo de casos). A BV-U pode ser um complemento para indivíduos com ERF que não lograram êxito em terapias anteriores. Os resultados evidenciam que são necessárias mais pesquisas de alta qualidade.

Considerações finais

O uso da USG de fala como um recurso de biofeedback terapêutico em casos de DSF mostrou-se de grande importância para a maioria dos estudos. Evidencia-se que o uso da USG permitiu além da aquisição, que houvesse o aprendizado motor, identificado pela capacidade da maioria dos sujeitos de generalizar o som aprendido para palavras que não foram trabalhadas e reter esse conhecimento meses do término do tratamento. No entanto, ainda são necessários mais estudos envolvendo terapia com ERF e AFI, abordando outros sons alvo, a fim de determinar se esse tipo de intervenção gera o mesmo resultado positivo.

ULTRASSONOGRRAFIA DE LÍNGUA NA AQUISIÇÃO DE LÍNGUA ESTRANGEIRA

Laís Silva-Garcia
Misael Krüger Lemes
Giovana Ferreira-Gonçalves

Introdução

O desenvolvimento tecnológico é um dos fatores determinantes para o avanço das investigações em todas as áreas do conhecimento. As ciências biológicas, por exemplo, utilizam a tecnologia para decodificar e recompor uma estrutura genética em tempo hábil; a física para analisar os limites do espaço e do tempo; a computação para criar aparelhos cada vez mais potentes, rápidos e com ampla capacidade de armazenamento. Por sua vez, a linguística, mais especificamente, a fonética e a fonologia, utilizam o desenvolvimento tecnológico para acessar o que antes era invisível aos olhos, como os movimentos da língua durante a fala (BRUM-DE-PAULA e FERREIRA-GONÇALVES, 2012). A ferramenta ultrassonográfica (US) pode, assim, contribuir de forma ímpar nesse sentido.

São diversos os benefícios trazidos pela utilização da ultrassonografia em estudos da fala, por exemplo, a possibilidade de análises dos gestos articulatórios envolvidos na produção dos sons, captados em tempo real, de forma dinâmica. Por meio dessas análises, é possível destacar as similaridades e diferenças entre os sons em diferentes línguas e discorrer sobre sua complexidade. Também é possível utilizar a ferramenta como facilitadora para o processo de aquisição de línguas estrangeiras e em terapias de fala. A utilização de uma nova ferramenta metodológica promove novas perspectivas de análise, sendo capaz de auxiliar na reestruturação de concepções vigentes na área.

A literatura acadêmica aponta que o US tem sido utilizado em estudos de produção da fala desde os anos 60 e 70 (GICK, 2002). A aplicação dessa ferramenta tem se mostrado eficiente, especialmente por oferecer benefícios como: “ser segura, não-invasiva, cada vez mais

acessível, portátil, design amigável, amplamente disponível e intuitiva” (BIRD e GiCK, 2018, p. 05). Estudos que investigaram o US aplicado a terapias de fala (BERNHARDT 2003; 2005; ADLER-BOCK, 2007; e MELO *et al*, 2016) mostraram-se eficientes no ensino de sons desafiadores em contextos clínicos. Com base nos métodos de tradição clínica, nas duas últimas décadas, o US passa a ser, também, empregado no ensino de línguas estrangeiras, como ferramenta de *feedback* visual.

Este capítulo, organizado em cinco seções, dedica-se, pois, a discorrer sobre a aplicabilidade da ultrassonografia no processo de ensino-aprendizagem de uma língua estrangeira. Na sequência a esta Introdução, serão apresentados os principais aspectos que tornam o uso da ferramenta como promissor no rápido ganho articulatório por parte dos aprendizes. Nas seções 3 e 4, uma revisão teórica sobre os estudos que empregam o ultrassom em pesquisas sobre a aquisição de línguas estrangeiras realizados em âmbito internacional e no Brasil, respectivamente. Por fim, as considerações finais acerca do tema.

O potencial instrutivo da ultrassonografia no ensino de línguas estrangeiras

Considerando-se o papel da pronúncia no êxito da aquisição de uma língua estrangeira (LE) ou segunda língua (L2)¹, importa salientar, primeiramente, que o ensino dessa habilidade, durante longo período, foi ignorado no campo dos estudos em Linguística Aplicada, acarretando na ausência de diretrizes adequadas para a abordagem da pronúncia na aula de LE (DERWING e MUNRO, 2005). Nesse sentido, o ensino de pronúncia é, de modo frequente, percebido como desafiador por professores da língua, além de ser custoso para estudantes, especialmente na fase adulta.

No que tange às abordagens gerais do ensino e aprendizagem de pronúncia de uma LE, Wilson (2014) explica que duas são as principais: (a) abordagens intuitivo-imitativas, que envolvem fundamentalmente práticas de imitação, e (b) abordagens linguístico-analíticas, que passam a utilizar determinadas ferramentas, tais como o alfabeto fonético internacional (IPA). Nessas abordagens, apesar de o aprendiz contar com o *feedback* proprioceptivo – ou seja, a capacidade de reconhecer o

¹ Neste capítulo, não serão feitas distinções entre os termos “língua estrangeira” e “segunda língua”.

movimento de sua própria língua –, Wilson e Gick (2006) salientam que, a depender do segmento, aprendizes podem ter menor *feedback* proprioceptivo, como no caso de vogais e consoantes líquidas.

Além disso, Wilson (2014) chama a atenção para o fato de que aprendizes podem não ser capazes de compreender instruções dadas verbalmente pelos professores acerca de como mover suas línguas. Assim, o US emerge como uma ferramenta capaz de preencher essa lacuna, ao proporcionar aos alunos a visualização de seus movimentos linguais, em tempo real, e, dessa forma, conforme sinalizado por Wilson e Gick (2006), possibilitar uma melhor percepção sobre ajustes necessários nos movimentos da língua a fim de completar, com êxito, determinada tarefa articulatória.

A ultrassonografia aplicada à instrução de pronúncia de LE: estudos internacionais

O presente tópico versa sobre estudos, realizados em âmbito internacional, da aplicação do US ao ensino/treinamento² de pronúncia de sons na aquisição de uma LE.

Meadows (2007) conduziu um estudo exploratório no Laboratório Fonológico de Imagem do Arizona (APIL) com o objetivo de investigar a aplicação de US à instrução de L2. Em sua investigação, o autor selecionou três estudantes universitários que estudavam japonês como LE, sendo as vogais longas, consoantes nasais e *flaps*, os segmentos alvos.

Três questionamentos foram propostos: “o US pode contribuir com o desenvolvimento articulatório da L2? Quais protocolos de instrução se mostram mais eficientes? Como se dá a resposta dos participantes em relação ao uso da ferramenta?” (MEADOWS, 2007, p. 18)

As sessões de instrução incluíram: instrução convencional e instrução mediada pelo US. Quatro sessões foram executadas, e cada uma delas estruturada conforme o Quadro 1, que ilustra a distribuição dos conteúdos em cada uma das instruções realizadas.

² Os termos “treinamento” e “instrução” serão utilizados de forma intercambiável aqui.

Quadro 1 – sessões de instrução explícita

Sessão 01	Sessão 02	Sessão 03	Sessão 04
*US e fonética – introdução	*Vogal longa [o:]	*Nasais (cont.) *Flaps *Flaps palatalizados	*Nasais (cont.) *Flaps *Flaps palatalizados
*Vogal longa [e:]	*Nasais		

Fonte: adaptado de Meadows (2007, p. 19).

Para a realização das atividades de instrução, contou-se com a contribuição de um instrutor falante nativo de japonês e com o auxílio do pesquisador. Na primeira metade da sessão, realizava-se a instrução convencional e, no segundo momento, partia-se para a abordagem com o US. Na etapa do US, o instrutor era responsável por produzir os sons em estudo lançando mão da ferramenta, a fim de modelar as articulações alvos, e, nesse momento, o pesquisador tecia comentários sobre pontos pertinentes das imagens. Utilizaram-se, também, vídeos ultrassonográficos de movimentos de língua como fonte adicional de *input*.

Na etapa final, os participantes tinham a oportunidade de praticar a produção dos sons alvos utilizando o aparelho de US, e, também nesse estágio, cada aprendiz tinha a oportunidade de receber feedback individualizado do instrutor acerca de suas produções da L2.

Os resultados constituíram-se de análises qualitativa e quantitativa. Para a realização da análise qualitativa, foram consideradas as anotações feitas pelo pesquisador durante as sessões de instrução e as respostas dadas pelos alunos por meio de questionários aplicados ao término de cada sessão.

A análise qualitativa sinalizou aspectos positivos e negativos das práticas de instrução. As imagens dos movimentos de língua realizados pelo instrutor na articulação dos sons e apresentados aos participantes foram consideradas um dos aspectos positivos da instrução. Além disso, a interação entre os sujeitos, o instrutor e o US foi destacada como a melhor etapa da sessão de instrução.

Por outro lado, a leitura de frases contendo os sons alvos pelos aprendizes e a exibição de vídeos ultrassonográficos foram reportadas como práticas pouco proveitosas. Em relação ao papel do pesquisador,

apesar de não mencionado pelos participantes, entende-se que sua contribuição foi fundamental na interpretação das imagens ultrassonográficas e na detecção de problemas referentes à pronúncia.

No que tange aos dados quantitativos, cada participante foi submetido a coletas de dados, que ocorreram antes e depois das sessões de instrução. Cada sujeito produziu, por coleta, 32 palavras apresentadas de forma isolada, que se repetiram igualmente em todas as coletas. Os dados foram avaliados por um painel de juízes, composto por três falantes nativos de japonês. Cada produção foi classificada em uma escala de 1 a 4, de acordo com maior ou menor proximidade da fala nativa. A atribuição do valor 1 indicaria maior aproximação à fala nativa, ao passo que o valor 4 representaria fala ininteligível. Em seguida, esses valores foram convertidos em porcentagens para quantificação.

Embora a média de concordância dos juízes tenha sido abaixo do esperado, a classificação atribuída nas avaliações se mostrou consistente, estando S2 melhor avaliado, seguido de S3 em uma posição intermediária, e, S1 inferior aos demais. O autor destaca, ainda, que os três sujeitos obtiveram ganhos importantes pós-instrução, uma vez que S3 obteve 7,3% de melhoria, S2 obteve 6% e S1 obteve 5%. A análise estatística, no entanto, revelou que esses resultados não foram significativos. Para o autor, uma das razões para esse fato pode estar no baixo número de sujeitos, o que reduz o poder estatístico dos cálculos.

Em síntese, tendo por base as observações e as respostas dos participantes aos questionários, o US parece contribuir com o aprimoramento da pronúncia da L2. Contudo, considerando-se que a abordagem deste estudo incluiu tanto as instruções convencionais, quanto as instruções via US, não foi possível, segundo o autor, determinar qual dos dois tipos de instrução favoreceu, predominantemente, o desenvolvimento dos segmentos produzidos pelos aprendizes. Logo, grupos experimentais e grupos-controle são necessários para que o impacto da ferramenta *per se* seja examinado sem intervenção de práticas de instrução convencional.

Posteriormente, em 2008, uma pesquisa piloto conduzida por Gick *et al* (2008) examinou a aplicação do US no treinamento das consoantes aproximantes /l/ e /ɹ/ do inglês por três estudantes japoneses de Linguística, recém chegados na América do Norte.

Os aprendizes participaram de um único encontro, com duração de aproximadamente uma hora, que consistiu de: (i) uma coleta de dados inicial; (ii) sessão de treinamento e (iii) coleta de dados pós-treinamento. Para a gravação, os estímulos escolhidos encontravam-se em posições diferentes dentro das palavras – início, meio e fim –, em seis contextos vocálicos variados. Os participantes deveriam produzir os segmentos alvos inseridos na frase veículo: “See X be”, de forma randomizada, totalizando-se dez repetições.

A etapa de treinamento, com duração de 30 minutos, foi constituída de duas partes principais: (i) os participantes eram apresentados às suas melhores e mais problemáticas produções (gravadas na coleta de dados com o US), e (ii) deveriam compará-las em termos de aspectos gerais e específicos dos movimentos da língua. Em suma, a coleta de dados inicial revelou que todos os participantes foram capazes de produzir os sons /l/ e /ɺ/ em, pelo menos, alguns contextos fonéticos. Após a sessão de treinamento, no entanto, os três sujeitos foram capazes de produzir a aproximante alvo até mesmo nos contextos mais desafiadores.

Também sobre as aproximantes do inglês, Tsui (2012), em sua tese, explorou a aplicação de US na aquisição dos sons /l/ e /ɺ/ por seis falantes nativos de japonês, com idades entre 19 e 28 anos, sendo quatro participantes do sexo feminino e dois participantes do sexo masculino. As principais perguntas de pesquisa delineadas tiveram por objetivo averiguar se esses indivíduos, após submetidos a treinamento com US, seriam capazes de produzir acuradamente os segmentos alvos em contextos diversificados, e se as produções com potencial mudança assemelhar-se-iam, em termos acústicos e articulatórios, àquelas produzidas por falantes nativos.

Inicialmente, duas semanas antes do início dos treinamentos, foi realizada uma coleta de dados pré-teste, com a produção de 44 palavras, dentre elas 24 com /ɺ/ e 20 com /l/. Foram produzidos três tokens por palavra, mas, para a análise, foi considerada somente a terceira repetição. Após, ocorreram quatro sessões de instrução por um período de duas semanas consecutivas e com duração de 45 minutos cada.

Cada sessão foi estruturada com as seguintes etapas: (i) discussão sobre os sistemas fonológicos do inglês e do japonês; (ii) descrição do US; (iii) orientação sobre o uso da ferramenta; (iv) ensino dos

componentes da língua na produção dos segmentos /l/ e /ɫ/; (v) práticas dos aprendizes com o US e (vi) práticas a serem realizadas em casa³. Duas semanas após a última sessão de instrução, a coleta de dados pós-teste foi realizada.

Para a análise dos dados, houve, em um primeiro momento, julgamentos por dois juízes falantes nativos de inglês. Suas respostas foram comparadas e compuseram um cálculo de confiabilidade que revelou uma porcentagem de 77% para as produções de palavras com o alvo /l/ e de 80% para as produções de palavras com o alvo /ɫ/. Esses resultados evidenciaram produções mais acuradas para ambos os segmentos no pós-teste. Em relação às produções de /ɫ/, ambos os juízes avaliaram 52% das palavras produzidas como mais acuradas no pós-teste e 15% igualmente acuradas no pré e no pós-teste. Para as produções de /l/, 62% das produções foram julgadas como mais acuradas no pós-teste e 12% igualmente acuradas nos dois testes.

As análises acústicas e de imagens ultrassonográficas consideraram palavras em que ocorreu maior aprimoramento após as sessões de treinamento. A Figura 1 mostra imagens de *frames* ultrassonográficos das produções de /ɫ/, por um mesmo sujeito, no pré-teste (à esquerda) e no pós-teste (à direita).

Figura 1 – Imagens ultrassonográficas de /ɫ/ em “read”



Fonte: adaptado de Tsui (2012, p. 58).

³ Cabe salientar que as práticas realizadas em casa eram, evidentemente, executadas sem o emprego de US, e consistiram, fundamentalmente, na produção oral de listas de palavras.

Comparando-se as imagens pré-treinamento e pós-treinamento, evidencia-se que, após ser submetido à instrução, o participante passou a produzir o segmento /ɹ/ com elevação da lâmina da língua e sua curvatura para trás. A análise acústica corroborou a inspeção visual, uma vez que revelou diminuição dos valores de F3 no pós-teste, o que aproxima essas frequências daquelas tipicamente encontradas na fala de nativos.

As produções de /l/ que obtiveram aprimoramento pós-instrução também apresentaram queda em F3, o que as aproxima, igualmente, dos alvos nativos. Em termos gerais, os ganhos foram mais evidentes nas produções de /l/ do que nas produções de /ɹ/, em todos as posições de palavras, exceto nos encontros consonantais de final de palavra.

Os resultados apresentados por Tsui (2012), apesar de algumas limitações do experimento, tais como a não existência de um grupo controle, apontam ganhos nas produções dos segmentos supracitados pelos aprendizes desde a primeira sessão de treinamento, além do aperfeiçoamento gradual ao longo das sessões posteriores. Os ganhos mantiveram-se, ainda, duas semanas após o término das instruções. Os participantes, por meio de um questionário, responderam positivamente em relação ao uso da ferramenta ultrassonográfica para o treinamento de fala.

Tateishi (2013), por sua vez, investigou se o treinamento da produção com o US contribuiria com a produção e a percepção⁴ de contrastes de sons não nativos, ou seja, das consoantes /l/ e /ɹ/ do inglês. Os sujeitos da pesquisa foram dez falantes nativos de japonês – quatro do sexo masculino, e seis do sexo feminino –, com idades entre 18 e 30 anos, aprendizes de inglês. Em relação à produção, dois questionamentos foram propostos: “o treinamento da produção com o US conduz a melhorias na produção dos sons supracitados em termos de F2 e F3? O treinamento com o US melhora a inteligibilidade das produção desses sons do inglês por aprendizes japoneses?” (TATEISHI, 2013, p. 23)

A estrutura dos experimentos, cuja execução ocorreu em uma cabine acústica no Laboratório de Fonética da Universidade de Calgary, encontra-se disposta no Quadro 2.

⁴ Aspectos perceptuais não serão aqui abordados, considerando-se que o foco deste capítulo está voltado a aspectos relativos à produção. Para mais detalhes, ver Tateishi (2013).

Quadro 2 – Etapas de realização dos experimentos

Dia I	Dia II	Dia III	Dia IV	Dia V	Dia VI
1. Teste de percepção 2. Gravação das produções	Treinamento de produção				1. Treinamento de produção 2. Teste de percepção 3. Gravação das produções

Fonte: adaptado de Tateishi (2013, p. 26).

O Quadro 2 evidencia a estrutura dos experimentos, que compreenderam coletas de dados pré-treinamento e pós-treinamento. Cada coleta integrava 40 estímulos: 20 pares mínimos de palavras e 20 palavras monossilábicas. Os estímulos e os procedimentos adotados eram os mesmos para as coletas pré e pós-treinamento.

Em relação às sessões de treinamento com o US, foram realizadas cinco sessões, de duração aproximada de 30 minutos cada, conduzidas individualmente com cada participante. Cabe ressaltar que, durante o treinamento, os aprendizes manuseavam o transdutor do US. Para o treinamento, 14 alvos foram selecionados e distribuídos em três estágios, conforme o Quadro 3.

Quadro 3 – Estágios das sessões de US

Estágio I	Estágio II	Estágio III
/ʌ/ e /l/ isolados	/ri/, /ru/, /ræ/ /li/, /lu/, /læ/	reek, room, rack leak, loom, lack

Fonte: adaptado de Tateishi (2013, p. 30).

É possível observar que, dentre os 14 alvos, no primeiro estágio, os segmentos eram apresentados isoladamente; no segundo estágio, eram inseridos em sílabas CV e, por último, em pares mínimos de palavras. A fim de obterem-se imagens com movimentos de língua que servissem de modelo, os alvos também foram gravados por um falante nativo de inglês, e apresentados aos aprendizes nas sessões de treinamento. Em cada sessão, produções dos aprendizes eram gravadas e posteriormente apresentadas para fins de discussão, bem como para possibilitar comparações com as imagens da fala nativa.

No que se refere à análise dos dados, foram consideradas as médias de F2 e F3 do ponto estacionário. Essas medidas foram normalizadas por meio da fórmula de z-score de Lobanov. A inteligibilidade das produções foi avaliada por três juízes falantes nativos de inglês.

A análise acústica revelou maior aperfeiçoamento das produções de /l/ do que das produções de /ɹ/, corroborando os achados de Tsui (2012). Em linhas gerais, apesar de F3 não ter apresentado mudanças significativas para as produções de /l/ e /ɹ/, suas médias, após o treinamento, aproximam-se daquelas exibidas pelos falantes nativos. Em relação a F2, não houve, igualmente, resultados significativos entre as produções pré e pós-treinamento.

A avaliação dos juízes revelou que o treinamento não contribuiu significativamente com uma maior inteligibilidade das produções dos aprendizes, embora a não-significância possa estar ligada ao excesso de variabilidade entre as respostas dos juízes, considerando-se os valores de desvio padrão.

Outro estudo, conduzido por Pillot-Loiseau, Antolík e Kamiyama (2013), investigou o emprego do US como ferramenta de *feedback* visual por japoneses aprendizes de Francês, tendo como ênfase o contraste entre os segmentos /y/ – /u/. Foram selecionadas quatro participantes japonesas, com idades entre 31 e 33 anos, e uma falante nativa de língua francesa, cujos dados foram gravados para comporem a análise. Dois sujeitos constituíram o grupo experimental (AK e CS), que receberam treinamento com US, e dois sujeitos constituíram o grupo controle, que não receberam treinamento com a ferramenta.

As participantes foram submetidas a três sessões de instrução, com duração de 45 minutos cada. As vogais eram exercitadas de quatro formas: (i) isoladamente; (ii) em palavras não existentes; (iii) em palavras reais e (iv) em frases curtas. Em relação às coletas de dados, os sujeitos experimentais tiveram suas produções gravadas antes do início dos treinamentos (pré-teste), uma semana após a última sessão (pós-teste) e dois meses depois do pós-teste (retenção).

Os dados acústicos e articulatórios foram gravados simultaneamente e sincronizados por meio do *Software AAA*. Apesar de o *corpus* ter sido constituído de contextos e estímulos diversificados, foram apresentados pelos autores somente os resultados referentes às vogais isoladas (produzidas com dez

repetições). A análise acústica considerou os valores de F1, F2, F3 e F4, extraídos de três pontos da vogal – 25%, 50% e 75%.

Os resultados do grupo controle mostram que as distâncias euclidianas entre [y] e [u] aumentaram pouco para os aprendizes, mas, no pós-teste, os contornos de língua indicam certa diferenciação entre essas vogais que, no pré-teste, mostravam-se sobrepostas no espaço acústico.

O grupo experimental, por sua vez, apresentou evolução na produção do contraste [y] – [u]. As vogais no pré-teste quase sobrepunham-se para ambos os sujeitos, mas, no pós-teste e no teste de retenção, passaram a apresentar distâncias euclidianas consideráveis. Além disso, no pré-teste, AK e CS articulavam as vogais quase que no mesmo ponto da cavidade oral, porém, no pós-teste, passou a haver maiores diferenças entre os contornos da língua para as referidas produções, e essas diferenças foram ainda maiores no teste de retenção, reforçando, dessa forma, ganhos consistentes e duradouros.

Posteriormente, White *et al* (2017) investigaram a aplicação do US no ensino da consoante lateral do inglês, em sua variante classificada como *dark*, por falantes nativos de cantonês, aprendizes de inglês. Foram selecionados oito participantes – estudantes de Hong Kong, sendo quatro do sexo feminino e quatro do sexo masculino, com idades entre 15 e 16 anos.

Primeiramente, foi realizada uma coleta de dados pré-teste, em que os sujeitos deveriam realizar a leitura de frases curtas contendo [ɫ] em final de palavras, em três contextos: /i:/, /ɔ/ e /u:/. As frases foram as seguintes: (i) *The meal heater was broken*; (ii) *The hall model was very nice* e (iii) *The pool food cannot be eaten*. Cada frase foi lida por cinco vezes. Após a coleta pré-teste, uma sessão de treinamento, com duração de 30 minutos, foi conduzida, com instruções sobre como articular o segmento [ɫ]. Para tanto, os aprendizes foram divididos em dois grupos de quatro integrantes cada: grupo 1 (G1) e grupo 2 (G2). Os grupos receberam tarefas articulatórias idênticas, com a diferença de que G1 recebeu *feedback* visual com o US, o mesmo não ocorrendo com G2.

Após a sessão de instrução, nova coleta de dados foi efetuada, seguindo-se as mesmas diretrizes do pré-teste, ou seja, a leitura das frases, com cinco repetições. A análise dos dados articulatórios foi realizada por meio do software *Articulate Assistant Advanced* (AAA) e as *splines* dos formatos das línguas comparadas via SS-ANOVA (GU,

2002). Além da análise articulatória, houve, também, julgamento das produções pelos autores do trabalho.

Os resultados do pré-teste evidenciaram que, à exceção de um participante de G2, os demais sujeitos articularam o segmento com gesto posterior de língua, ou seja, vocalizado, porém, sem um gesto anterior subsequente, e, portanto, mais semelhante a uma vogal posterior do que a uma lateral velarizada. A partir do pós-teste, três (dentre os quatro) participantes do G1 adicionaram um gesto anterior de língua às suas produções, o que corrobora o impacto positivo da ferramenta na aquisição da lateral pós-vocálica do inglês. Em relação a G2, não houve diferenças relevantes entre as produções pré-teste e pós-teste,

Em síntese, a análise articulatória e o julgamento dos juízes apontam maiores ganhos para o grupo que contou com a inserção de US no treinamento. Apesar de ambos os grupos apresentarem melhorias após a sessão de treinamento, SS-ANOVA mostrou diferenças significativas mais nítidas entre o grupo que recebeu treinamento com US.

O Quadro 4 condensa as principais informações sobre os estudos supracitados.

Quadro 4 – Síntese dos estudos de US aplicado ao ensino de pronúncia de L2 no exterior

Autor(es)	Ano	Segmentos	Resultados
Meadows	2007	[e:] e [o:] do inglês Nasais do inglês Flaps do inglês	Os participantes obtiveram ganhos articulatórios importantes pós-instrução, embora esses resultados não tenham sido significativos em termos estatísticos.
Gick et al	2008	/l/ e /ɹ/ do inglês	Após o treinamento, os três sujeitos foram capazes de produzir as aproximantes alvos.

Tsui	2012	/l/ e /ɹ/ do inglês	A avaliação dos juízes evidenciou produções mais acuradas para ambos os segmentos no pós-teste. A análise acústica exibiu resultados positivos: queda de F3 para /l/ e /ɹ/. Ganhos articulatórios foram mais evidentes para /l/.
Tateishi	2013	/l/ e /ɹ/ do inglês	Ganhos articulatórios mais evidentes para /l/ do que para /ɹ/; apesar de F3 não ter apresentado mudança significativa para /l/ e /ɹ/, suas médias aproximam-se dos valores dos nativos.
Pillot-Loiseau et al	2013	/y/ e /u/ do francês	Após o treinamento com US, as vogais passaram a apresentar distâncias euclidianas consideráveis, havendo também maiores diferenças entre os contornos da língua.
White et al	2017	Lateral (<i>dark</i>) do inglês [ɫ]	A partir do pós-teste, três (dentre os quatro) participantes do G1 – que recebeu feedback com US –, adicionaram um gesto anterior às suas produções

Com base no Quadro 4, é possível constatar que o uso do US é promissor no ensino de pronúncia de uma L2. Sabe-se, no entanto, que ainda são poucas as pesquisas que abordam essa temática, além de apresentarem limitações, tais como: baixo número de informantes, pouca recorrência de grupo controle, dentre outros aspectos.

Também é notável que os segmentos selecionados nas pesquisas são, de forma recorrente, os róticos e as laterais, com poucas exceções. Nesse sentido, estudos que investiguem a aquisição de outros segmentos vocálicos e consonantais são fundamentais para solidificar a relevância da aplicabilidade do US no ensino de línguas estrangeiras.

A ultrassonografia aplicada à instrução de pronúncia de LE: estudos brasileiros

De acordo com Brum-de-Paula e Ferreira-Gonçalves (2013), estudos da fala aliados ao uso da ultrassonografia de língua são ainda incipientes no Brasil. Atualmente, o emprego da ferramenta em investigações linguísticas é crescente. Sua contribuição tem papel positivo nas análises acerca da aquisição do português brasileiro como língua materna – Vassoler (2016), Barbieri (2019), Barberena et al (2020) –, bem como acerca da aquisição de línguas estrangeiras – Dias-Cavaleiro (2016), Teixeira-Corrêa, Ferreira-Gonçalves e Brum-de-Paula (2017) e Silva (2018). No entanto, apesar de contribuir para a análise e descrição de fenômenos linguísticos, são ainda mais restritas as pesquisas que investigam a utilização do US como ferramenta facilitadora no processo de aquisição de LE.

Os trabalhos desenvolvidos no Brasil que utilizaram o US aliado ao ensino e aprendizagem de LE o empregaram em sessões de instrução explícita (IE). A primeira pesquisa foi desenvolvida em Iniciação Científica (IC) por Pereira e Ferreira-Gonçalves (2016) no Laboratório Emergência da Linguagem Oral (LELO) da Universidade Federal de Pelotas (UFPel). Este estudo compõe uma série de quatro trabalhos vinculados ao projeto “A ultrassonografia e o ensino de línguas”⁵, desenvolvido no LELO. No referido trabalho, os autores verificam a aplicabilidade de sessões de IE via US na aquisição do rótico retroflexo do inglês. A publicação dos resultados pode ser constatada em Ferreira-Gonçalves, Pereira e Lemes (2019).

Para realizar a pesquisa, contou-se com a participação de dois informantes do sexo feminino, uma aprendiz de inglês de nível básico e outra de nível avançado. A informante de nível básico nunca havia recebido instrução formal de inglês, enquanto a informante de nível avançado era formanda do curso de graduação em Letras – Português

⁵ Este projeto foi apoiado pelo CNPq por meio do processo nº 312273/2015-0.

e Inglês da UFPEl, além de ter frequentado por quatro anos escolas especializadas em ensino de inglês.

As coletas de dados do trabalho foram divididas em três etapas: (i) pré-teste; (ii) pós-teste I e; (iii) pós-teste II. Em todas as fases das coletas, o mesmo instrumento foi aplicado. As palavras foram selecionadas a partir da posição do segmento rótico na sílaba e na palavra – onset inicial, onset medial, onset complexo, coda medial e coda final; da tonicidade – posição tônica – e da vogal adjacente – seguido pela vogal /a/. Ainda, para facilitar as análises acústicas, o segmento se encontrava entre consoantes plosivas surdas. No Quadro 5, dispõe-se a lista de palavras utilizadas por Ferreira-Gonçalves, Pereira e Lemes (2019).

Quadro 5 – Palavras utilizadas no instrumento de coleta de dados de Ferreira-Gonçalves

Posição na palavra	Inglês	Português
onset inicial	rabbit	-
onset medial	interaction	barata
onset complexo	traffic	trave
coda medial	apartamento	carta
coda final	car	mar

Fonte: Pereira e Lemes (2019)

As palavras eram produzidas em frase-veículo no idioma do item lexical referido: “I say ___ to you”, para palavras em inglês, e “Digo ___ para você”, para palavras em português. Cada item lexical era repetido cinco vezes. O pré-teste foi realizado antes de qualquer sessão de instrução explícita com o US, a fim de possibilitar comparações com os resultados obtidos posteriormente.

As coletas acústicas e articulatórias foram realizadas em uma cabine de isolamento acústico e ocorreram simultaneamente. Os equipamentos utilizados foram: um gravador digital, modelo Zoom H4N; um aparelho de US Mindray DP-6600, com sonda endocavitária – 65EC10EA – acoplada; um capacete estabilizador de movimentos da sonda, projetado pela *Articulate Instruments*; um sincronizador de áudio e imagem *Sync BrightUp*, modelo SBU 1.0; uma placa de vídeo, responsável pela ligação entre o US e o computador, e o *software Articulate Assistant Advanced (AAA)*, versão 2.14, para coleta e análise dos dados articulatórios.

Ao total, foram realizadas duas sessões de IE, apenas com a informante de nível básico, uma vez por semana. As coletas pós-teste foram realizadas após as sessões de IE. O aparelho de US utilizado para esta fase do estudo foi o *Chison Eco1-Vet*, com sonda micro-convexa MC6-A acoplada.

As sessões de IE foram divididas em três etapas. Na primeira, o professor apresentava e explicava, por meio de imagens ultrassonográficas, os movimentos articulatórios envolvidos na produção do segmento retroflexo do inglês.

O som-alvo era produzido pelo professor isoladamente, de forma audível – para ênfase no *input* sonoro e visual – e de modo silencioso – para direcionar a concentração do aluno apenas para a visualização dos gestos articulatórios envolvidos na produção do retroflexo. O rótico também foi produzido em um contexto de dez palavras previamente selecionadas para cada IE. O professor repetia as palavras e a produção do segmento isolado por três vezes.

Na segunda etapa, o aprendiz realizava os exercícios apresentados anteriormente pelo professor. Neste momento, o aprendiz treinava a produção do segmento retroflexo de forma isolada, de forma audível e silenciosa, e também em contexto das palavras apresentadas na primeira etapa. O aluno, assim como o professor, repetia três vezes a produção dos segmentos, tanto em contexto de palavras quanto de forma isolada. Por fim, a terceira etapa correspondia igualmente aos passos realizados na primeira etapa da sessão de IE.

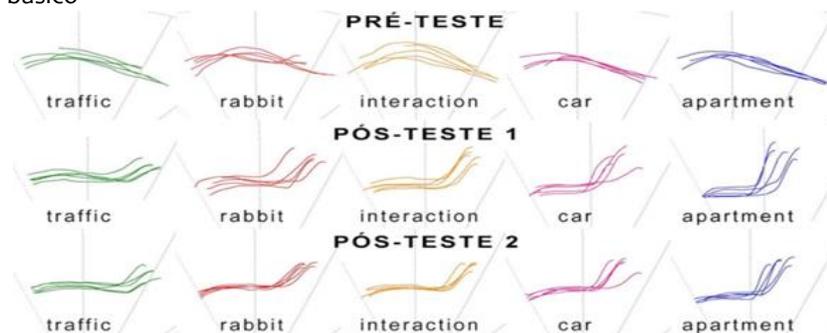
Para as análises acústicas, foram considerados os valores de duração relativa e absoluta de cada segmento, a fim de comparar os resultados do pré-teste com os dos pós-testes I e II. De acordo com os autores, houve um expressivo aumento dos valores de duração do rótico, tanto relativa quanto absoluta, em todos os contextos silábicos.

Um exemplo importante percorrido por Ferreira-Gonçalves, Pereira e Lemes é o do rótico em *onset* medial nos pares *barata* e *interaction*: na coleta pré-teste os valores de duração relativa eram muito próximos, 3,55% e 3,01%, respectivamente. Já no pós-teste I, essa diferença aumenta consideravelmente, com 5,21% para *barata* e 10,66% para *interaction*. Os resultados acústicos apontam um efeito positivo da IE via US na aquisição do rótico retroflexo.

No caso das análises ultrassonográficas, foi selecionado o *frame* de maior magnitude do gesto articulatório do retroflexo de cada *token*

a fim de traçar o contorno da borda da língua por meio do *software* AAA. Os autores indicam que os resultados da informante de nível básico corroboram com os dados acústicos, uma vez que evidenciaram um importante ganho articulatório em todos os contextos silábicos. A Figura 2 apresenta a disposição das *splines* da língua em todas as coletas do estudo.

Figura 2 – Splines das produções do rótico retroflexo da informante de nível básico



Fonte: FERREIRA-GONÇALVES, PEREIRA e LEMES (2019, p. 15)

Os pesquisadores apontam que, a partir das sessões de IE, a informante evidencia o encurvamento da ponta da língua característico da variante retroflexa. Esse movimento era inexistente no pré-teste, como mostram as *splines* na Figura 2.

Por fim, o trabalho de Ferreira-Gonçalves, Pereira e Lemes (2019) indica um papel positivo da ferramenta ultrassonográfica no processo de aquisição do rótico retroflexo do inglês.

O segundo trabalho vinculado ao projeto “A ultrassonografia e o ensino de línguas” foi desenvolvido por Duarte e Ferreira-Gonçalves (2018) e também contribui para a investigação relativa ao papel da instrução explícita via US na aquisição do rótico retroflexo do inglês.

O estudo de Duarte e Ferreira-Gonçalves (2018) contou com a participação de quatro sujeitos do sexo feminino, aprendizes de Inglês como Língua Estrangeira (ILE) de nível básico e intermediário. As informantes eram graduandas do curso de Letras-Português e Inglês da UFPel, sendo três naturais de distintas regiões do Rio Grande do Sul (RS) e uma natural do interior de São Paulo (SP). A faixa etária dos sujeitos era de 18 a 30 anos, sendo que as informantes de nível básico

não tinham recebido instrução formal de LE, à exceção do ensino fundamental e médio.

As coletas de dados das autoras incluíram quatro etapas: (i) pré-teste; (ii) pós-teste I; (iii) pós-teste II e (iv) teste de retenção. Diferentemente do estudo de Ferreira-Gonçalves, Pereira e Lemes (2019), o trabalho de Duarte e Ferreira-Gonçalves (2018) contou com realização de três sessões de IE e duas coletas pós-teste, uma após a primeira sessão de IE e outra ao fim da última.

No que concerne à elaboração do instrumento de coleta, a produção das palavras ocorreu de forma isolada, sendo cada item lexical repetido por cinco vezes de forma aleatória. Já na coleta de retenção, as palavras foram repetidas por três vezes. As autoras optaram pela produção isolada das palavras de forma a garantir uma maior sincronização entre o dado acústico e o articulatório (TURTON, 2017).

As palavras utilizadas nas sessões de IE, realizadas com intervalo de uma semana, e nas coletas de dados estão dispostas nos Quadros 6, 7 e 8, respectivamente.

Quadro 6 – Palavras utilizadas no instrumento de coleta de pré e pós-teste.

Posição na palavra	Inglês	Português
onset inicial	rabbit, rat	-
onset medial	arrive, arrest	barata
onset complexo	practice, press	prato
coda medial	market, parking	carpa
coda final	car, far	mar

Fonte: Duarte e Ferreira-Gonçalves (2018)

Quadro 7 – Palavras utilizadas no instrumento de coleta do teste de retenção

Posição na palavra	Inglês	Português
onset inicial	rabbit, rat racket, rapid	rato rapa
onset medial	arrive, arrest carafe, garage	barata barata
onset complexo	practice, press crab, crash	prato praga
coda medial	market, parking overcoat underpants	carpa farpa
coda final	car, far, jar, war	mar, lar

Fonte: Duarte e Ferreira-Gonçalves (2018)

Quadro 8 – Palavras utilizadas durante as sessões de IE

Posição na palavra	Sessão 1	Sessão 2	Sessão 3
onset inicial	raght, race	rule, rescue	rose, reach
onset medial	configuration harass	ironic arena	direction erase
onset complexo	grave cry	drive trace	problem frame
coda medial	universal tutle	emergency internal	important commercial
coda final	winter, your	wear, fur	poor, flower

Fonte: Duarte e Ferreira-Gonçalves (2018)

Assim como Ferreira-Gonçalves, Pereira e Lemes (2019), os resultados obtidos por Duarte e Ferreira-Gonçalves (op.cit) comprovam a eficácia da ferramenta ultrassonográfica na aquisição do rótico retroflexo do inglês. As autoras relatam que, a partir das sessões de IE, informantes que não realizavam a produção do segmento na coleta pré-teste passam a produzi-lo. Além disso, os dados ainda confirmam o aprimoramento gestual no caso de informantes que já produziam o segmento, uma vez que a duração do segmento aumenta, componente característico do segmento.

Outro trabalho vinculado ao projeto “A ultrassonografia e o ensino de línguas” é Silva-Garcia e Ferreira-Gonçalves (2019). O estudo das autoras versa sobre a aquisição da lateral pós-vocálica do espanhol por aprendizes brasileiros. A pesquisa contou com a participação de cinco informantes, dois alunos do primeiro e três do sétimo semestres do curso de Letras-Português e Espanhol da UFPel. Todos do sexo feminino, naturais de Pelotas/RS e com instrução formal de língua espanhola apenas durante o ensino fundamental e médio. De acordo com as autoras, a seleção de sujeitos de diferentes níveis de ensino é uma estratégia para verificar a eficácia da ferramenta em alunos com baixo e alto nível de instrução da língua espanhola.

As coletas de dados foram realizadas em quatro etapas em uma cabine de isolamento acústico alocada no LELO: (i) pré-teste; (ii) pós-teste I; (iii) pós-teste II e (iv) teste de retenção. Ainda, foi realizada uma coleta controle com a produção de palavras em língua portuguesa. A coleta pré-teste foi realizada uma semana antes da primeira sessão de IE sem que fosse mencionado o elemento alvo da pesquisa. O pós-teste I ocorreu após a primeira sessão de IE; o pós-teste II, após a

última. No total, foram realizadas três sessões de IE, também com intervalos de uma semana. O teste de retenção e a coleta controle ocorreram no mesmo encontro, 30 dias após a última sessão de IE.

O pré-teste foi realizado para que as produções das informantes servissem como referência para a posterior comparação com as coletas realizadas após as sessões de IE via US. Os dois pós-testes viabilizaram, assim, acompanhar o desenvolvimento das participantes mediante às intervenções de IE com US. Já a coleta de retenção teve por objetivo observar se os possíveis avanços obtidos permaneceriam internalizados, evidenciando a sedimentação dos movimentos adquiridos durante as aulas de IE. Por fim, a coleta controle serviu para comparar os movimentos referente a lateral pós-vocálica do espanhol e da língua materna (PB).

A fim de realizar comparações precisas entre as produções das informantes realizadas no pré-teste e nos pós-testes, foi mantido o mesmo instrumento de coleta.

O método de acesso aos itens lexicais selecionado pelas autoras foi o de nomeação de imagens. Esta escolha permite que não haja a interferência da grafia na produção dos sujeitos – no caso de o estímulo ser a palavra escrita. Conforme os critérios de seleção de palavras, o segmento pós-vocálico deveria estar em coda silábica, em posição medial ou final, em sílaba tônica e átona, e ser antecedido pelas cinco vogais do espanhol, ou seja, [i, e, a, o, u]. Em todas as etapas das coletas, as palavras eram repetidas seis vezes, três com o transdutor de US em posição sagital e outras três com o transdutor em posição coronal. Nos Quadros 9, 10 e 11, apresentam-se as palavras utilizadas pelas autoras nos três instrumentos de coleta:

Quadro 9 – Lista de palavras utilizadas no pré-teste e nos pós-testes

Contextos	/a/	/e/	/i/	/o/	/u/
fim de sílaba/tônica	talco	celta	pilcha	colcha	pulso
fim de sílaba/átona	altar	felpuda	filtrar	olfato	cultura
fim de palavra/tônica	astral	papel	brasil	caracol	azul
fim de palavra/átona		cóctel	fértil	apóstol	cónsul

Fonte: Silva-Garcia e Ferreira-Gonçalves (2019)

Quadro 10 – Lista de palavras utilizadas no teste de retenção

Contextos	/a/	/e/	/i/	/o/	/u/
fim de sílaba/tônica	salsa talco	felpa celta	filtro pilcha	escolta colcha	último pulso
fim de sílaba/átona	altura altar	delgado felpuda	silvestres filtrar	soltero olfato	escultor cultura
fim de palavra/tônica	total astral	pastel papel	fusil brasil	girassol caracol	caracul azul
fim de palavra/átona	-	cárcel cóctel	fácil fértil	futebol apóstol	cônsul

Fonte: Silva-Garcia e Ferreira-Gonçalves (2019)

Quadro 11 – Lista de palavras utilizadas na coleta-controle

Contextos	/a/	/e/	/i/	/o/	/u/
fim de sílaba/tônica	salsa	felpa	filtro	escolta	último
fim de sílaba/átona	altura	delgado	silvestres	solteiro	escultor
fim de palavra/tônica	total	pastel	fusil	girassol	caracul
fim de palavra/átona		coquetel	fácil	futebol	cônsul

Fonte: Silva-Garcia e Ferreira-Gonçalves (2019)

As sessões de IE de Silva-Garcia e Ferreira-Gonçalves (2019) foram executadas seguindo a metodologia desenvolvida por Duarte e Ferreira-Gonçalves (2018) e Ferreira-Gonçalves, Pereira e Lemes (2019), contendo três etapas: “(i) explicação articulatória do segmento lateral pós-vocálico do espanhol, por meio de US, por parte do pesquisador; (ii) realização de exercícios articulatórios com produções em tempo real, via US, pelo aprendiz; (iii) repetição das explicações discorridas pelo pesquisador na primeira etapa” (SILVA-GARCIA; e FERREIRA-GONÇALVES, 2019, p.72).

Figura 3 – Espaço de realização das sessões de IE



Fonte: Silva-Garcia e Ferreira-Gonçalves (2019, p.73)

As análises acústicas foram realizadas considerando os valores duracionais do segmento lateral pós-vocálico do espanhol, bem como a diferença dos valores formânticos de F1 e F2. Segundo Recasens *et al* (1995), quanto maior o distanciamento dos valores entre F2 e F1, mais alveolar é a produção de /l/; quanto menor, presume-se uma produção mais velarizada.

A partir de uma análise parcial das médias dos resultados acústicos de duas informantes, uma do 1º e outra do 7º semestres, as autoras constataram que, após as instruções explícitas, as aprendizes obtiveram resultados positivos no que concerne ao aumento da diferença de F2-F1, indiciando uma produção mais alveolar de /l/, ou seja, mais próxima das configurações da lateral pós-vocálica do espanhol.

A partir de uma análise parcial, Silva-Garcia e Ferreira-Gonçalves (2019) concluem que as sessões de IE via US possibilitaram maior apropriação e automatização dos gestos de /l/ pós-vocálico do espanhol pelas informantes, as quais já realizavam os movimentos articulatórios referentes à lateral alveolar da língua-alvo na coleta pré-teste. Os movimentos articulatórios que já eram realizados se tornaram mais acurados nas produções das aprendizes do 1º e do 7º semestres, evidenciando o caráter positivo da ferramenta aliada à realização das sessões de IE. Atualmente, uma continuidade do referido estudo é desenvolvida com maior aprofundamento, por meio de análises acústica e articulatória dos dados, na dissertação de Mestrado de Silva-Garcia (2022, no prelo).

O último trabalho elaborado a partir do projeto “A ultrassonografia e o ensino de línguas”, no LELO, foi desenvolvido por Teixeira e Ferreira-Gonçalves (2020). O trabalho das autoras investiga

a aquisição da consoante lateral pós-vocálica do inglês através de sessões de IE via US, por aprendizes brasileiros de Inglês como Língua Estrangeira. A proposta das autoras decorre da percepção de que alunos de ILE apresentam dificuldades na produção deste segmento em inglês, uma vez que aplicam na LE os padrões da LM, o que resulta, em muitos dialetos do PB, a produção da lateral pós-vocálica do inglês como vocalizada /w/. Na língua inglesa, /l/ pós-vocálico é produzido, predominantemente, de forma velarizada, conhecido como dark-l [ɫ].

As etapas do estudo de Teixeira e Ferreira-Gonçalves (2019) consistiram na realização de: (i) coleta pré-teste; (ii) três sessões de IE e (iv) uma coleta pós-teste, após a primeira sessão de IE. Duas outras etapas previstas, ou seja, a realização de um pós-teste II, após a última sessão de instrução explícita, e a aplicação de um teste de retenção não foram concluídas, por conta da suspensão das atividades presenciais em decorrência da pandemia de COVID-19.

A seleção dos sujeitos da pesquisa ocorreu a partir da aplicação de um questionário aos alunos do curso de Letras-Português e Inglês da UFPel. Assim, participaram quatro informantes do sexo feminino naturais da cidade de Pelotas/RS, duas graduandas do segundo semestre, sem contato anterior com instrução formal de inglês, e duas graduandas do oitavo semestre.

O instrumento de coleta de dados da pesquisa foi proposto a partir de Scoobie e Wrench (2003). Sendo assim, palavras com movimentos pré-linguais adjacentes à lateral foram excluídas, já que tais movimentos podem interferir na articulação do som-alvo. Contextos de segmentos pré-labiais sonoros também foram excluídos para que o segmento lateral fosse visualizado de forma ainda mais clara nas análises acústicas. Foram, então, selecionadas palavras em que a lateral pós-vocálica estivesse em contexto de pré-pausa e de pré-labiais surdas – /p/ e /f/ –, precedida pelas vogais altas /i/ e /u/ e pela vogal baixa /a/.

As coletas de dados ocorreram em cabine de isolamento acústico. Em cada coleta, as palavras foram repetidas dez vezes, cinco em posição sagital e cinco em posição coronal, totalizando 120 produções por coleta. A coleta pré-teste foi realizada sem nenhuma instrução prévia acerca do elemento analisado e, em todas as etapas das coletas, os itens lexicais eram apresentados ao informante de forma aleatória.

As sessões de IE foram realizadas em duplas compostas por uma informante do 2º e outra do 8º semestres da graduação. Os encontros

duraram aproximadamente 50 min e ocorreram uma vez por semana durante três semanas, totalizando três sessões de IE. Durante as sessões, as informantes receberam informações sobre os movimentos referente a lateral pós-vocálica do inglês, por meio de vídeos e de movimentos realizados pela pesquisadora no US. Em seguida, as aprendizes praticaram os movimentos do segmento de forma isolada e também em contexto de palavras, visualizando-os pelo US e recebendo *feedback* instantâneo da pesquisadora.

Em relação aos resultados, Teixeira e Ferreira-Gonçalves (2019) dispõe das análises acústicas, as quais foram realizadas verificando a duração do segmento lateral, bem como a diferença entre os valores de F2-F1. A análise da duração da lateral é importante porque pode fornecer pistas sobre o grau de velarização do segmento – quanto maior a duração, maior o grau de velarização de /l/ (TEIXEIRA e FERREIRA-GONÇALVES, 2019, p. 3). Os valores de F2-F1, como já mencionado no trabalho de Silva-Garcia e Ferreira-Gonçalves (2019), estão também relacionados com o grau de velarização: quanto menor a diferença entre essas frequências formânticas, mais velarizado é o segmento.

Os resultados ainda parciais do trabalho indicam uma produção menos vocalizada da lateral para as informantes do 2º semestre, quando comparados o pré-teste, com 48,7% de produções vocalizadas, e o pós-teste, com apenas 21,4% de produções vocalizadas. Em posição de coda medial, as autoras apontam diferenças ainda mais expressivas, uma vez que, no pré-teste, 87,5% das produções das aprendizes foram vocalizadas, enquanto, no pós-teste, a produção vocalizada ocorreu em apenas 15,3% dos dados.

Sobre os valores de F2 e F1, os dados das informantes do 2º semestre indicam um maior grau de velarização, já que a diferença F2-F1 reduziu de 557,8(Hz) no pré-teste, para 377,8(Hz) nos dados de pós-teste. A média de F2 também diminuiu após a primeira sessão de IE, de 984,8(Hz) no pré-teste para 880,2(Hz) no pós-teste. Considerando que as alterações do valor de F2 estão relacionadas aos movimentos horizontais da língua, logo, sons mais velares requerem movimentos mais posteriores de língua, os quais apresentam valores de F2 mais baixos.

Os resultados de pré-teste das aprendizes do 8º semestre também indicaram a produção da lateral do inglês de forma vocalizada. A porcentagem de produções de segmentos vocalizados nesta etapa foi de 56%. No entanto, após a sessão de IE, o valor das

produções reduziu para 21,6%, sinalizando um maior domínio dos gestos referentes a [t] do inglês. O aumento do grau de velarização da lateral também foi verificado: (i) pela elevação exponencial da média de duração do segmento, o qual era produzido em 131,7(ms) no pré-teste e passou para 201,2(ms) na coleta pós-teste, e (ii) pela diminuição da diferença entre F2-F1, cuja média reduziu de 507(Hz) no pré-teste para 442,9(Hz) no pós-teste.

Dispostos os resultados de Teixeira e Ferreira-Gonçalves (2019), evidencia-se o papel positivo da IE via US na produção de um som articulatoriamente complexo, como o dark-l (TEIXEIRA e FERREIRA-GONÇALVES, 2019, p.4). Por meio dos resultados, é possível verificar que a ferramenta contribui com a formação acadêmica dos aprendizes, uma vez que, enquanto futuros professores de língua inglesa, poderão alcançar a acurácia fonético-fonológica necessária para a profissão mais rapidamente.

Ainda no âmbito de trabalhos que investigam a ferramenta ultrassonográfica aplicada ao ensino de LE no Brasil, destaca-se a primeira dissertação de Mestrado que versa sobre o tema no Brasil, desenvolvida por Lemes (2021, no prelo), cujos resultados parciais já podem ser encontrados em Lemes e Ferreira-Gonçalves (2020). Em seu estudo, os autores investigam os efeitos de sessões de IE via US na aquisição das vogais altas – [i] e [ɪ] – do inglês por aprendizes brasileiros acadêmicos do curso de licenciatura em Letras-Português e Inglês da UFPel.

Lemes e Ferreira-Gonçalves (2020) apontam que as vogais [i] e [ɪ] geralmente são distinguidas a partir de suas características duracionais, sendo [i] uma vogal longa e [ɪ] uma vogal curta. No entanto, a revisão da literatura destaca que as diferenças fundamentais entre estes segmentos estão no padrão articulatorio de cada um. Nesse sentido, a pesquisa dos autores reitera a importância da análise e da aprendizagem dos movimentos característicos e determinantes de cada vogal, já que trata-se de um componente discriminatório entre os sons.

Para realizar o estudo, Lemes e Ferreira-Gonçalves (2020) selecionaram seis participantes, dividindo-as em três grupos, a saber: (i) grupo experimental formado por quatro estudantes universitárias do curso de Letras – Português e Inglês, três do período inicial e uma do período final do curso; (ii) uma falante monolíngue de PB e (iii) uma falante nativa de Inglês Americano (LEMES e FERREIRA-GONÇALVES, 2020, p.1).

Assim como nas pesquisas supracitadas, as coletas de dados realizadas com o grupo experimental de informantes ocorreram em três etapas: (i) pré-teste, anterior aos encontros com sessões de IE; (ii) pós-teste I, após a primeira sessão de IE; (iii) pós-teste II, ao fim da última sessão de IE. Ao todo, foram realizadas três sessões de IE, uma vez por semana. Igualmente aos estudos anteriores, as coletas foram realizadas em cabine de isolamento acústico, localizada no LELO, da UFPel. A realização de um teste de retenção também não pôde ocorrer por conta da suspensão das atividades acadêmicas, em decorrência da pandemia de COVID-19.

O instrumento de coleta elaborado pelos autores foi constituído a partir da seleção de vinte e quatro palavras com as vogais [i-I] e, também, contendo as vogais [e] e [Λ]. Em todas as coletas, cada palavra foi repetida seis vezes, totalizando cento e cinquenta e seis *tokens*.

As sessões de IE foram realizadas a fim de que as participantes recebessem informações acerca do detalhamento fonético-fonológico de cada vogal e, também, para que pudessem exercitar, por meio de atividades de produção oral via US, os movimentos articulatorios apreendidos no momento da instrução. Todas as sessões foram ministradas pelo pesquisador, que também é professor de inglês.

Os resultados acústicos preliminares, a respeito de uma informante do período inicial do curso de graduação, revelaram efeitos expressivos nos valores de F1 dos segmentos alvos, uma vez que apontam uma diminuição, no caso da vogal tensa [i] – de 403(Hz) no pré-teste para 364(Hz) no pós-teste II –, e uma elevação, no caso da vogal frouxa [ɪ] – de 482(Hz) no pré-teste para 535(Hz) no pós-teste II. Em relação à análise de F2, os autores afirmam não haver alterações expressivas dos valores formânticos.

Nesse sentido, apesar de discorrer acerca dos resultados de apenas uma participante do estudo, Lemes e Ferreira-Gonçalves (2020) verificaram a eficácia da ferramenta ultrassonográfica no que diz respeito ao aperfeiçoamento das vogais altas anteriores [i] e [ɪ] do inglês. De acordo com os autores, o US promove ganhos relativos à aceleração do aperfeiçoamento fonético-fonológico da língua-alvo.

No Quadro 12, estão dispostos a série de trabalhos aqui reportados.

Quadro 12 – Síntese dos estudos de US aplicado ao ensino de pronúncia de L2 no Brasil

Autor(es)	Ano	Segmento	Resultados
Duarte e Ferreira-Gonçalves	2018	rótico retroflexo do inglês - [ɹ]	(i) produção do segmento anteriormente não realizado; (ii) aprimoramento dos gestos já executados.
Ferreira-Gonçalves, Pereira e Lemes	2019	rótico retroflexo do inglês - [ɹ]	(i) aquisição do gesto de encurvamento característico de [ɹ].
Silva-Garcia e Ferreira-Gonçalves	2019	lateral alveolar pós-vocálica do espanhol - [l]	(i) automatização dos movimentos de [l] devido; (ii) aprimoramento do gesto alveolar de [l] devido ao aumento da diferença entre F2-F1.
Teixeira e Ferreira-Gonçalves	2019	lateral velar pós-vocálica do inglês - [ɫ]	(i) redução de produções vocalizadas; (ii) aprimoramento do gesto velar do dark-l.
Lemes e Ferreira-Gonçalves	2020	Vogais altas anteriores do inglês - [i] e [ɪ]	(i) produções das vogais [i] e [ɪ] mais próximas das formas alvos.

Como pode ser constatado, no Brasil, os estudos nesta área surgiram há cerca de cinco anos, a partir dos investimentos realizados por agências de fomento – CNPq, CAPES e FAPERGS – no Laboratório Emergência da Linguagem Oral, o que viabilizou a aquisição de equipamentos portáteis de US para a realização de atividades de instrução explícita.

Os resultados obtidos até o momento coadunam-se com os achados dos estudos internacionais, os quais indicam o papel benéfico do US na aquisição precoce e no aprimoramento gestual dos segmentos da L2. No entanto, assim como as pesquisas internacionais, os estudos do Brasil também apresentam algumas limitações.

Um aspecto importante a ser destacado está no fato de que todos os trabalhos apresentam, como etapa metodológica, a coleta de dados articulatórios. No entanto, apenas um, até o momento, dispõe seus

resultados em publicação. A futura análise das imagens articulatórias obtidas poderá, sem dúvida, promover a utilização do US em estudos acerca do ensino e aprendizagem de LE no Brasil.

Salienta-se, também, a ênfase dada, assim como nos estudos internacionais, às classes das soantes. O Quadro 13 exhibe as consoantes [l], [ɫ], [ɮ] e as vogais [i] e [ɪ] como os segmentos-alvo das análises. A exploração de segmentos que constituem outras classes de sons pode, além de sedimentar estudos de base articulatória via US, trazer explicações acerca de fenômenos fonético-fonológicos que permeiam a aquisição de uma língua estrangeira. A aplicabilidade da ferramenta no processo de aquisição de outros segmentos proporciona, enfim, uma maior facilidade e um menor tempo dispendido no que concerne ao aprimoramento fonético-fonológico.

Considerações Finais

As pesquisas acerca da aplicação do US ao ensino de línguas são, sem dúvida, inovadoras. Os trabalhos já realizados evidenciam o quanto a ferramenta é promissora para uma compreensão mais adequada sobre os fenômenos fonéticos-fonológicos presentes na aquisição de L2 e, principalmente, para o sucesso na obtenção de uma pronúncia acurada, aspecto relevante aos futuros professores de línguas.

O avanço na utilização de US aplicado ao ensino de línguas depende, essencialmente, do refinamento metodológico que advirá dos estudos pioneiros. É importante, portanto, aumentar o número de sujeitos experimentais, criar um grupo-controle e ampliar o *corpus* das coletas, dando, assim, mais robustez para as análises.

No Brasil e no exterior, são demasiadamente novas as pesquisas que averiguam a eficácia da ferramenta ultrassonográfica como componente facilitador para aquisição de LE. Revela-se, todavia, o papel positivo do aparelho para a aquisição de segmentos por vezes complexos aos aprendizes brasileiros. Os estudos confirmam, via análise acústica e articulatória, a aquisição ou o aperfeiçoamento gestual dos segmentos-alvo. Trata-se, portanto, de uma ferramenta promissora não só para o desenvolvimento de novas pesquisas, como, de forma mais prática, um instrumento a ser diretamente utilizado no ensino e aprendizagem de LE no país.

ULTRASSONOGRRAFIA DE LÍNGUA NA VARIAÇÃO DA LINGUAGEM

Aline Rosinski
Thalena Evangelista Santos
Mirian Rose Brum-de-Paula

Pesquisas sobre variação linguística têm sido alimentadas por um conjunto de enunciados produzidos por um número expressivo de informantes, o que visa à robustez dos resultados obtidos. Assim, embora a constituição de *corpora* orais exija tempo e conhecimento metodológico específico, amostras amplas são empreendidas com poucos entraves tecnológicos, visto que podem ser adquiridas com gravadores portáteis de uso comum, desde que sejam de boa qualidade. O obstáculo maior na obtenção desse tipo de amostra está mais relacionado ao *paradoxo do observador* do que a questões que compreendem a manipulação dos acessórios usados para seu registro. A aquisição de dados de fala por meio do ultrassom, ao contrário, não é um recurso de fácil manejo para o entrevistador, ou de cômoda utilização para o entrevistado. A técnica pode, de fato, apresentar características não muito vantajosas para a pesquisa sociolinguística. Assim sendo, em um primeiro momento, tender-se-ia a descartar o uso dessa ferramenta para esse tipo de estudo, já que exige transporte adequado, formação para o seu emprego, realização de entrevistas mais longas, compreensão do participante quanto ao emprego de sonda, capacete estabilizador e gel, bem como uma maior tolerância do informante relativa a possíveis desconfortos causados pelo uso desses elementos durante a entrevista. Consequentemente, essas peculiaridades do ultrassom reduzem a possibilidade de que integrem pesquisas que requerem amostras amplas. Encontra-se, nessas singularidades, um empecilho para seu uso na investigação sociolinguística? Empregá-lo contribuiria para mostrar facetas não reveladas por meio do dado acústico?

Nas investigações sobre a fala, observa-se que diversos trabalhos incorporaram o ultrassom aos seus procedimentos metodológicos. A ferramenta tem sido empregada por linguistas, fonoaudiólogos e engenheiros — especialistas do sinal da fala — em pesquisas

experimentais que envolvem um número pequeno de sujeitos, principalmente (RECASENS, 2004; GICK ET AL, 2008; ZHARKOVA ET AL, 2012, TEIXEIRA-CORREA, FERREIRA-GONÇALVES e BRUM-DE-PAULA, 2017; BARBIERI, 2019). Pode-se também empregá-lo quando se busca descrever a variabilidade da fala de um grupo de indivíduos? De que modo uma ferramenta que, por várias razões, não se adapta à aquisição de dados de um grande número de sujeitos pode servir à descrição da variação linguística?

Para melhor refletir sobre o uso da ultrassonografia com esse objetivo, será posto em foco, inicialmente, o campo intitulado sociofonética que, segundo Hay e Drager (2007), integra a sociolinguística e se debruça sobre a distribuição das realizações das formas variáveis e sua avaliação social.

A sociofonética iniciou sua trajetória de modo tímido: o termo foi empregado por Deshaies-Lafontaine, em 1974, em sua tese de doutorado. Nela, a autora propôs uma descrição da língua francesa de Trois Rivières, uma comunidade quebequense. Depois disso, pesquisadores como Thomas (1982) e Léon e Cichocki (1989), por exemplo, também intitularam desse modo os estudos que empreenderam. Mas as pesquisas sociofonéticas precederam sua denominação. Baranowski (2013) identificou em Labov, Yaeger e Steiner (1972) a primeira investigação sociofonética propriamente dita. Mas do que tratam as investigações em sociofonética?

Retomando Foulkes et al (2010), Candea e Trimaille (2015) destacam que esses estudos visam a "dar conta de mecanismos ligados à aprendizagem, à memorização, à avaliação, [a dar conta] do processamento cognitivo de tudo o que constitui a estruturação social da variação" (p.15). Boa parte dessas pesquisas — que se identificam como sociofonéticas — têm, na análise acústica, o seu trunfo. No entanto, essa perspectiva, que conquistou a academia (sob forma de disciplinas e cursos), figura em publicações temáticas (como em FOULKES, SCOBIE & WATT, 2010; BAYLEY R., CAMERON R. e LUCAS C., 2013; CANDEA, M. e TRIMAILLE, C., 2015 ou, ainda, BRESCANCINI e FOULKES, 2017) e manuais (a exemplo de DI PAOLO e YAEGER-DROR, 2010 e THOMAS, 2011), aos poucos, tornou-se uma frente investigativa que alia o interesse na variação da fala, essencial à sociolinguística, à sua captação por meio de instrumentos diversos, alicerces da fonética experimental. Essas ferramentas, quando utilizadas de modo

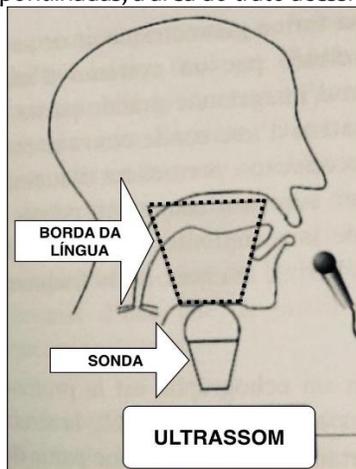
concorrente, fornecem informações diversas e complementares sobre o mesmo fenômeno linguístico, como destaca-se nas seções que compõem este capítulo.

Com o intuito de desenvolver esse tópico, que diz respeito à relação existente entre as análises acústicas e articulatórias em estudos sociofonéticos, na seção seguinte, reportam-se trabalhos considerados como sendo de referência. Na sequência, destaca-se o trabalho de Rosinski (2019), a primeira pesquisa sociofonética de campo a fazer uso do ultrassom no Brasil. Enfim, nas considerações finais, apresenta-se seu emprego para além do ambiente de laboratório.

O ultrassom e a variação linguística

A ferramenta ultrassonográfica tem sido empregada nos estudos do trato vocal e, particularmente, da produção da fala desde a sua aplicação clínica regular, entre as décadas de 1960 e 1970. O aparelho permite a obtenção de imagens dinâmicas atinentes à superfície/borda da língua, como é possível observar na Figura 1.

Figura 1 – Entre linhas pontilhadas, a área do trato acessível ao feixe ultrassônico



Fonte: (adaptada de HUEBER e DENBY, 2009)

Com o seu uso, é possível aferir medidas linguais e investigar interações entre vogais e consoantes (GICK, 2002). De acordo com Gick, associado à análise acústica, o ultrassom permite o acesso a imagens de

dados articulatórios anteriormente inacessíveis para os pesquisadores da área. Na presente seção, apresentam-se, então, alguns estudos sociofonéticos de referência para que se possa elaborar uma ideia mais clara sobre esses dados e sobre os procedimentos adotados com o uso dessa ferramenta em pesquisas da fala.

Lawson, Stuart-Smith e Scobbie (2008) investigaram aspectos articulatórios a respeito da não realização do /r/ em final de sílaba no inglês escocês. A partir desse estudo, os pesquisadores visaram a um melhor entendimento dos processos articulatórios associados à ausência dessa produção e, além disso, a revelar os ganhos do emprego da ferramenta ultrassonográfica em trabalhos de cunho sociolinguístico. Dentre as suas vantagens, destacaram o fato (i) de essa tecnologia ser segura, de manejo relativamente simples, possível de ser transportado e (ii) de as imagens ultrassonográficas permitirem a visualização de movimentos linguais de maneira dinâmica. A ferramenta pode gerar imagens de quase todo o comprimento da língua (da raiz à ponta da língua), aspecto importante para o estudo da duração e da sincronização dos gestos articulatórios. Um estudo piloto, desenvolvido por Scobbie e Stuart-Smith, originou essa investigação. Ao observarem dados acústicos, constataram que as características geralmente associadas à produção do /r/ escocês em posição pós-vocálica — concernente ao aumento de F₂ e queda de F₃ — não foram identificadas no espectrograma. As imagens ultrassonográficas evidenciaram que a elevação frontal da língua ocorria na fase final do vozeamento. Logo após a fonação, a parte da frente da língua alcançava sua máxima elevação. Lawson, Stuart-Smith e Scobbie (2008) destacaram que o levantamento da língua — inferido quando os formantes se dissipavam — tendia a ser percebido auditivamente e, de fato, os ouvintes reportavam ter a impressão de que havia algo que soava de maneira similar ao /r/. De acordo com Lawson, Stuart-Smith e Scobbie (2008), esse fenômeno pode ser caracterizado como uma variação diastrática, tendendo a estar presente em contextos informais de fala. Os autores manifestaram que outros estudos sobre a fala informal deveriam ser desenvolvidos. Para tanto, defenderam a ideia do desenvolvimento de aparelhos de ultrassom menores e do aprimoramento de seus acessórios (sonda e estabilizador de cabeça, por exemplo) para que trabalhos de caráter

sociolinguístico pudessem ser efetuados em trabalhos de campo. Na sequência, mais detalhes sobre esse trabalho.

O estudo contou com a participação de 14 jovens do sexo masculino, estudantes do ensino médio, com faixa etária entre 12 e 13 anos, residentes em Livingston, localizada em West Lothian. O fato de a escola receber estudantes de áreas carentes da cidade permitiu o nivelamento socioeconômico dos participantes da pesquisa.

As gravações ocorreram no ambiente escolar, em local silencioso, e contou com dois tipos de gravação: uma conversa informal e a leitura de uma lista de palavras.

A coleta foi efetuada em dois momentos. No primeiro, registrou-se o áudio de todos os participantes. No segundo, simultaneamente, foram gravados dados acústicos e articulatórios (com o uso do aparelho de ultrassom) de dez locutores. Os outros quatro participantes compunham o grupo controle e tiveram somente os dados sonoros obtidos. Os dados ultrassonográficos foram coletados de apenas 1 informante de cada dupla. Todos participantes, entretanto, foram induzidos a crer que os dados — sonoros e ultrassonográficos — estavam sendo registrados.

O aparelho usado foi um ultrassom portátil – modelo *Concept M6*. Duas sondas com ângulos convexos de 100° e 120° foram empregadas. Os participantes que tiveram os dados de imagens ultrassonográficas coletados eram orientados a usar um estabilizador de cabeça — constituído de alumínio leve —, visando à manutenção da posição da sonda em um local fixo sob o queixo e possibilitando uma relativa liberdade de movimento durante a coleta. Esse tipo de *capacete* foi especialmente desenvolvido por Alan Wrench, diretor administrativo da *Articulate Instruments Ltd*, empresa que cria *softwares* e outros acessórios destinados à pesquisa da fala articulada.

A análise das imagens ultrassonográficas foi realizada por meio do *software AAA – Articulate Assistant Advanced*, concebido por essa mesma empresa. No que concerne aos resultados obtidos, durante a fase 1, compilaram-se 1283 *tokens* relativos ao /r/ pós-vocálico. Segundo Lawson, Stuart-Smith e Scobbie (2008), a partir da análise qualitativa dos dados, verificou-se que o levantamento da ponta da língua ocorria com atraso. A análise quantitativa revelou que a tendência à não produção do rótico em posição final de sílaba tônica existia. Esses resultados indicaram que a não realização do /r/ no inglês

escocês estava provavelmente relacionada a um atraso na execução do gesto, gerado pelo alongamento. Os autores destacaram que o fato tendia a ocorrer em contextos específicos: em sílabas átonas e em posição final de enunciado. Sílabas átonas mais curtas podem obliterar /r/ em posição final de sílaba. Em contrapartida, as sílabas tônicas alongadas que emergem no final dos enunciados podem gerar a separação do gesto de retração da raiz da língua e o gesto relativo ao levantamento da ponta da língua para a realização de /r/. Dessa maneira, o gesto anterior é articulado após o vozeamento, deixando-o menos perceptível ou, até mesmo, inaudível. Por fim, Lawson, Stuart-Smith e Scobbie (2008) afirmam que o emprego do ultrassom permitiu identificar que há, de fato, um atraso no gesto de elevação da língua em numerosas pronúncias supostamente não róticas.

De Decker e Nicz (2012) empregaram a ferramenta ultrassonográfica em investigação sobre aspectos acústicos e articulatórios envolvendo o segmento /æ/ do inglês americano, caracterizado como tenso. De acordo com os autores, a distinção entre as variantes de /æ/ está principalmente ligada a posições adotadas pela língua, que são geradas pelo avanço ou elevação de sua raiz. Os pesquisadores reportaram, então, que a percepção de [æ] como tenso pode estar fundamentalmente associada a duas fontes articulatórias: de um gesto de avanço/ levantamento da língua ou da nasalização. Destacam, ainda, que pode haver uma combinação dos aspectos elencados. De Decker e Nicz (2012) indagaram sobre a possibilidade de haver variações entre locutores na articulação do [æ] tenso, e se um locutor que produz o segmento [æ] tenso em numerosos contextos fonológicos realiza todas essas variantes da mesma forma.

Motivados pelos questionamentos supracitados, De Decker e Nicz (2012) desenvolveram um estudo que contou com a participação de quatro sujeitos, dois do sexo feminino e dois do sexo masculino, falantes nativos de inglês americano. Todos tinham aproximadamente 20 anos de idade e eram de Nova Jersey.

O instrumento continha quatro palavras principais – *pat*, *pad*, *pass* e *pan* – e oito palavras distratoras, que deveriam compor a frase veículo *Say ____ very loudly* (que significa *Diga ____ muito alto*). Os doze itens lexicais tinham a estrutura CVC. O material foi apresentado de forma aleatória por meio de um monitor de computador. Os dados acústicos e articulatórios foram coletados concomitantemente. Como na pesquisa de Lawson, Stuart-Smith e Scobbie (2008), os informantes

usaram um estabilizador de cabeça no momento da coleta. As imagens referentes ao plano médio sagital foram obtidas por meio do uso de um ultrassom *Sonosite Titan*, com uma taxa de 30 *frames* por segundo. Os áudios foram registrados por um microfone *Audio Technica*. Para realizar a análise articulatória, as imagens obtidas contaram com o emprego do sistema *EdgeTrak*¹. A análise acústica dos dados foi realizada com o *software* PRAAT. Foram medidos os dois primeiros formantes a partir do ponto médio do segmento. Segundo De Decker e Nicz (2012), apesar de todos os locutores serem provenientes de uma mesma região de Nova Jersey, as produções sugeriam a existência de dois padrões diferentes relacionados ao segmento /æ/. Essas produções caracterizavam-se como nasal clássica e nasal complexa.

Dois locutores realizaram o som nasal clássico, com uma evidente distinção entre /æ/ pré nasal e pré oral. Em relação às produções do sistema nasal complexo, os investigadores observaram que os sujeitos realizaram produções diferentes entre a vogal /æ/, presente em *pan*, e /æ/, constante em *pass*, *pad* e *pat*. No entanto, destacaram que as vogais dessas palavras não se sobrepunham no espaço acústico. Ainda sobre as especificidades dessas produções, salientaram que o sexo dos locutores pode ter atuado na articulação dos segmentos, uma vez que os participantes cujas produções se alinhavam ao sistema nasal clássico eram do sexo feminino, enquanto que os informantes do sexo masculino adotaram o sistema nasal complexo.

Os dados articulatórios revelaram que os locutores possuíam produções variadas da vogal /æ/ tensa e /æ/ não tensa, mas que elas eram mais evidentes nas informantes, pois apresentavam importantes diferenças acústicas no contexto pré nasal. Contudo, essas distinções não exibiam a mesma base articulatória. Enquanto a articulação de uma das informantes, em contexto pré nasal, tinha a parte frontal da língua mais elevada, a articulação de outra informante era de outro modo realizada, visto que sua língua não se encontrava tão elevada, nem tão anteriorizada. De Decker e Nicz (2012) propuseram, então, que as diferenças acústicas existentes entre a produção de *pan* e as demais palavras realizadas por essa locutora não podiam estar associadas a uma simples alteração no gesto de língua, mas que realizava produções nasalizadas para criar a distinção acústica, fazendo com que as vias auditivas identificassem o tensionamento.

¹ Para mais detalhes, ver Yi *et al.* (2005).

Os pesquisadores salientaram que o ultrassom poderia ser empregado em pesquisas futuras para investigar essa hipótese, associando-o a medidas de fluxo aéreo nasal. Os resultados indicaram que era necessário dar continuidade à investigação do referido segmento, com o objetivo de identificar se a variação era diatópica ou diastrática e como está relacionada a aspectos diacrônicos da língua.

Mielke (2013) também utilizou a ferramenta ultrassonográfica, a pesquisa versa sobre a influência do rótico na produção das vogais /ø/, /œ/ e /œ̃/ do francês canadense. Segundo o autor, alguns canadenses locutores de francês articulam essas vogais de forma semelhante ao [ɹ] da língua inglesa. O fato é exemplificado por meio dos itens lexicais *heureux*, *docteur* e *commun*, realizados como [ɹɛ], [dɔkɹœɹ] e [kɔmɹ]. O contato entre o inglês e o francês, línguas oficiais do Canadá, poderia estar atuando na realização dessas produções. Mielke (2013) salientou, entretanto, que os locutores nativos não identificavam as diferenças existentes entre as produções róticas e não róticas. Produções róticas não são tão evidentes para locutores nativos de francês canadense, aspecto que é perceptível para locutores da língua inglesa que não têm contato com o francês².

A coleta dos dados articulatórios foi realizada com 27 sujeitos, entretanto, os dados usados na análise foram de 23 participantes, sendo que 16 eram do sexo feminino e sete do sexo masculino. A faixa etária dos locutores variou entre 18 e 38 anos.

Os informantes foram orientados a ler uma lista de palavras em que constavam os sons alvo /ø/, /œ/, /œ̃/. Cada palavra foi proferida três vezes: duas vezes inseridas na frase veículo *Je dis ____* (que significa *Eu digo ____*) e uma vez na frase veículo *Je dis ____ encore* (que significa *Eu digo ____ de novo*). A coleta foi realizada numa cabine com isolamento acústico e os dados articulatórios foram obtidos com o emprego de um ultrassom modelo t3000. Os dados acústicos foram armazenados a uma taxa de amostragem de 44,1 kHz e 16 bits, com a utilização de um microfone *Shure MX412/C Gooseneck Condenser*. Também foram obtidos registros faciais dos participantes com o uso de uma câmera de vídeo. Para executar a correção no movimento de cabeça, foi empregado o sistema *Palatron*³.

² Para maior detalhamento, ver Lamontagne e Mielke (2013).

³ Para esclarecimentos, consultar Mielke et al. (2005).

A partir da análise dos dados, Mielke (2013) reportou que a relação entre o valor baixo de F3 e o abaixamento do dorso da língua em produções não róticas indicava “um caminho gradual” (p. 148) da vogal para o /ɹ/ *bunched*. De acordo com o autor, a vogal rótica podia ter emergido por meio do aumento gradual dos valores diminutos de F3 nas vogais arredondadas anteriores, que podia, também, estar associada ao F2 baixo. Inicialmente, a qualidade rótica foi observada em palavras que apresentavam o segmento /ɹ/ em posição final de palavra. A referida vogal apresentava evidências de que existia uma correlação entre o grau de *bunching* e o grau de qualidade rótica, tendo relação, também, com uma intensa queda no valor de F2. A qualidade rótica podia estar ligada à retração da raiz da língua e sua retroflexão. Em contrapartida, os dados concernentes às vogais /œ/ e /œ̃/ não indicaram esses aspectos de forma contundente. Os dados articulatórios também não revelaram que a retroflexão poderia surgir paulatinamente.

Hirkham e Wormald (2015) desenvolveram um estudo acerca das características acústicas e articulatórias presentes na variação dos segmentos líquidos produzidos por locutores ingleses e asiáticos que empregavam o dialeto inglês falado em Bradford, cidade localizada no distrito de West Yorkshire, no norte da Inglaterra. Segundo os autores, essa região era um bom local para a realização do estudo, visto que apresentava um número expressivo de asiático-britânicos. Partiram, então, da hipótese de que esses locutores tendiam a produzir o segmento consonantal /l/ com valores de F2 mais elevados, com uma constrição efetuada em uma região mais anterior do trato vocal do que a realizada pelos locutores ingleses.

A pesquisa contou com a participação de nove informantes do sexo feminino (cinco paquistanesas e quatro inglesas), locutoras nativas de inglês com idades que variavam entre 21 e 36 anos. Todas eram, desde o nascimento, residentes de Bradford e falantes da variedade dessa mesma cidade. As participantes paquistanesas eram bilíngues (inglês/ panjabi – mirpuri) e as participantes inglesas monolíngues.

O instrumento aplicado continha 28 itens lexicais com a presença de /l/ ou /r/ em início de palavra (como em *lead*, por exemplo), em meio de palavra em sílaba pretônica, (como em *believe*), em meio de palavra em sílaba postônica (como em *belly*), em borda de morfemas (como em *filing*) e em posição final de palavra (como em *peel*). Tais itens lexicais foram apresentados às participantes em meio a 54 palavras

distratoras. Para realizar o estudo, os autores elencaram pares mínimos como, por exemplo *lead/reed*, em posição de início de palavra, e *belly/berry*, em posição de meio de palavra. Os estímulos foram repetidos por três vezes.

A coleta dos dados articulatórios ocorreu com um ultrassom modelo *Mindray DP-2200* e o *software AAA – Articulate Assistant Advanced* – em um centro comunitário localizado em Bradford. Para a obtenção de imagens de um mesmo plano médio sagital do trato, foi utilizado o estabilizador de cabeça concebido pela *Articulate Instruments Ltd*. Os dados acústicos foram obtidos com um microfone *headset Beyerdynamic Opus 55*, conectado a um pré-amplificador *Grace m101* em uma interface de áudio *Sound Devices USBPre2*.

Hirkham e Wormald (2015) constataram que as participantes asiáticas tendiam a executar a constrição – para a articulação dos segmentos consonantais /l/ e /r/ – em uma região mais anterior do trato vocal, realizada com a ponta da língua ou com sua parte anterior. As participantes inglesas, entretanto, comumente articulavam essas consoantes em uma região mais posterior, com uma retração do dorso da língua.

Westerbeg (2016), por sua vez, realizou um trabalho relacionado aos aspectos articulatórios, acústicos e sociofonéticos da vogal /i:/ do idioma sueco, variante conhecida como *Viby-i*. A pesquisa objetivava “preencher uma lacuna na literatura” (WESTERBEG, 2015, p.5) no que se refere à descrição fonológica e articulatória dessa vogal. O estudo contou com a participação de treze locutores – dez do sexo feminino e três do sexo masculino. Sete participantes eram naturais de Estocolmo, enquanto dois eram de Gotemburgo, outros dois de Varberga, um de Ionecopinga e um de Katrineholm. A faixa etária dos participantes incluiu indivíduos de 18 a 27 anos. Os locutores eram fluentes em língua inglesa e quase todos empregaram outras línguas em casa quando crianças (como inglês, francês, finlandês ou lingala). No momento em que foi realizada a coleta dos dados, os participantes residiam na Escócia, com períodos que variavam de cinco meses a oito anos. Todos permaneciam fluentes em sueco, embora pudessem ter alguns problemas para recordar certos itens lexicais dessa língua.

Em relação à coleta dos dados acústicos e articulatórios, Westerbeg (2016) reportou que foram gravados com o uso do *software AAA – Articulate Assistant Advanced*. Como nos outros trabalhos aqui reportados, os registros foram sincronizados de

maneira automática ao longo das gravações. As coletas foram realizadas na *Queen Margaret University*, em duas salas com isolamento acústico denominadas *sala de controle* e *sala de gravação*. Os dados articulatórios foram obtidos com o uso de um ultrassom modelo *Ultrasonix Sonix RP*, por meio de uma sonda microconvexa. Para a obtenção dos registros de áudio, utilizou-se um microfone *Audio Technica*. Duas coletas ocorreram: uma de fala controlada e outra de fala espontânea. Além disso, os participantes responderam a um questionário, com o objetivo de verificar, na sociobiografia de cada um, aspectos não necessariamente linguísticos que poderiam estar relacionados às suas produções.

No que tange aos resultados, Westerbeg (2016) verificou que a característica articulatória mais evidente do segmento /i:/, quando comparado a /e:/, era que o corpo da língua assumia uma posição mais baixa, em altura semelhante a /u:/, fato observado em todos os locutores. A língua encontrava-se, assim, posicionada em uma região mais anterior do trato vocal sempre que atingia o seu ponto mais alto. Tal resultado apresentava-se oposto ao que indicava o baixo valor de F2. A autora informou que /i:/ é a vogal mais anterior para todos os participantes. Comparada a outras vogais, na articulação de /i:/, a ponta da língua tendia a estar mais elevada e mais anterior. Contudo, a parte posterior da língua ficava mais retraída para /i:/ do que para /e:/. Segundo a autora, essa retração podia estar associada ao baixo valor de F2.

Westerbeg (2016) reportou que foram identificadas duas maneiras diferentes de articular /i:/: alguns locutores produziram o segmento *bunched*⁴ enquanto outros o realizaram como um *bunched* duplo. Há também casos em que alguns participantes articularam produções intermediárias. A autora constatou que, ao comparar os locutores provenientes das regiões urbanas e metropolitanas, os metropolitanos costumavam articular /i:/ com valor de F1 elevado e F2 mais baixo, havendo, também, um grau considerável de abaixamento e posteriorização da língua. Comparando outras variáveis, como sexo e situação socioeconômica, não foi identificado nenhum padrão evidente. De acordo com a autora, essas variáveis não atuaram significativamente. No entanto, as participantes do sexo feminino tenderam a usar uma versão mais robusta da vogal /i:/.

⁴ O termo *bunched* é mantido no transcrito do texto, tendo em vista a possível perda de sentido quando da realização de sua versão para o português.

Hwang, Charles e Lulich (2019) elaboraram um estudo relacionado às características articulatórias e variação das laterais do coreano para melhor compreender a produção desses segmentos. A língua coreana apresenta somente um fonema líquido /l/, que tem dois alofones: [r] e [l]. Apesar de já existir uma vasta discussão sobre a representação fonológica da consoante líquida do coreano, os autores destacaram que muito ainda precisava ser investigado acerca de seus aspectos fonéticos. No que tange à sua descrição articulatória, principalmente, pouco havia a respeito. Os autores destacaram que a lateral [l] apresenta dois gestos principais: (i) uma constrição formada pela ponta da língua e (ii) elevação do corpo da língua (palatalização). Juntamente com a elevação, também observaram um gesto na parte da frente de sua raiz. Ao comparar [r] e [l], evidenciaram que o levantamento da ponta da língua era maior para [l].

Para melhor identificar os aspectos articulatórios envolvidos na produção da lateral, Hwang, Charles e Lulich (2019) coletaram dados provenientes de cinco locutoras nativas de coreano *Standard* (Seul). As informantes da pesquisa nasceram e cresceram na capital da Coreia do Sul. Por razões acadêmicas, contudo, no momento da coleta dos dados, as participantes residiam nos Estados Unidos, e estavam no país há períodos que variavam de 3 meses a 4 anos. Sobre o inglês empregado pelo grupo, as participantes tinham nível de proficiência entre intermediário e avançado. Todas começaram a estudar a língua inglesa depois dos dez anos de idade.

Os estímulos foram selecionados com o objetivo de conter o segmento-alvo em posição final de palavra, em itens lexicais monossilábicos. Os pesquisadores delimitaram como contexto vocálico precedente /i, a, ʌ, o, u/. Para gerar imagens volumétricas tridimensionais em formato de cunha, os dados foram obtidos com o uso de um ultrassom Philips 3D/4D EPIQ 7G, com um transdutor digital 3D/4D Philips xMatrix x6-1. O transdutor foi posicionado abaixo da mandíbula das participantes por meio do uso do estabilizador *Articulate Instruments Ltd*, como em outros trabalhos aqui citados.

Além desse procedimento, outro modo foi empregado para a aquisição dos dados articulatórios: a impressão do palato, elaborada com alginato dental⁵. Essa impressão tridimensional foi digitalizada com o uso do scanner *NextEngine Desktop 3D Scanner*. Os dados

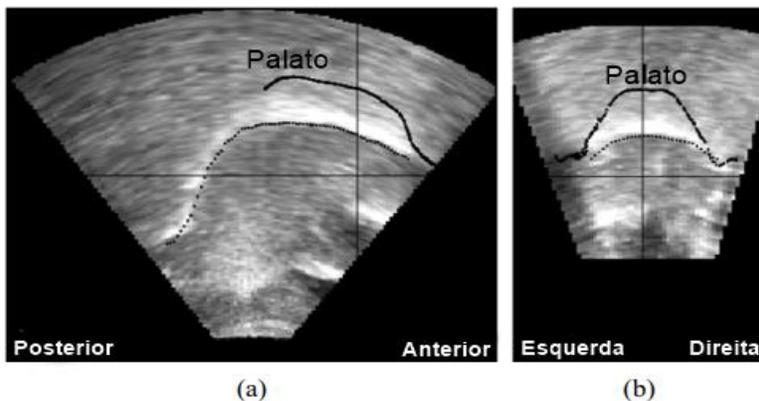
⁵ Para mais detalhes, ver Ladefoged (2003).

provenientes da ferramenta ultrassonográfica foram armazenados de acordo com o método proposto por Charles e Lulich (2018). Os áudios foram registrados a partir da utilização de um microfone SHURE KSM32. Os dados acústicos foram gravados com uma taxa de amostragem de 48 kHz e 16 bits de quantização.

No que tange à análise qualitativa dos dados articulatórios, Hwang, Charles e Lulich (2019) reportaram que foi realizada de acordo com uma *toolbox*⁶ customizada via MATLAB. O sistema *toolbox* possibilita a combinação de dados ultrassonográficos, impressões do palato e dados acústicos. A partir da inspeção visual, a superfície da língua foi identificada e segmentada manualmente em cada quinto plano sagital e décimo plano coronal.

Segundo Hwang, Charles e Lulich (2019), os dados de impressão tridimensional e os dados do ultrassom foram empregados com o objetivo de identificar como o palato e a língua estavam alinhados. A análise qualitativa foi embasada na observação de dois planos bidimensionais que se transpassam, e outro tridimensional. No que concerne ao plano bidimensional, o palato é indicado por uma linha sólida, enquanto a superfície da língua é evidenciada por meio de uma linha pontilhada, como pode ser observado na Figura 2.

Figura 2 – Dois planos bidimensionais: (a) médio sagital e (b) coronal

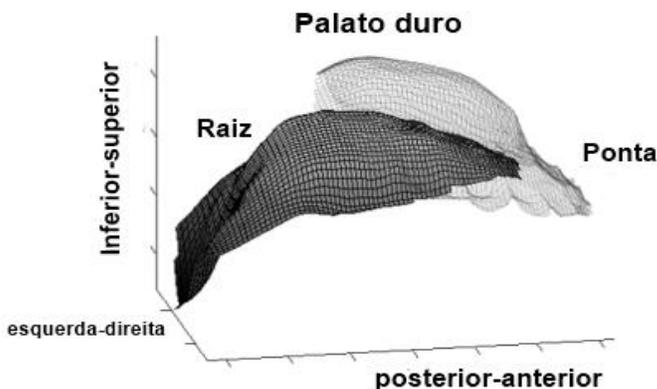


Fonte: (adaptado de Hwang, Charles e Lulich, 2019, p. 21).

⁶ Nesse caso, a *toolbox* compreende um conjunto de funções desenvolvidas no ambiente de computação técnica do MATLAB. Disponível em: https://www.mathworks.com/?s_tid=gn_logo. Acesso em: 21 jan 2021.

Ao observar a Figura 2, é possível visualizar duas linhas – uma na horizontal e outra na vertical – que se cruzam. Elas delimitam a interseção entre os dois planos. A Figura 3, a seguir, ilustra essa interseção de forma tridimensional.

Figura 3 – Imagem tridimensional da articulação da lateral lamino-alveolar precedida pela vogal /i/



Fonte: (adaptado de Hwang, Charles e Lulich, 2019, p. 21).

No que se refere à análise quantitativa dos dados articulatórios, foram consideradas distâncias anatômicas concernentes à altura do corpo da língua e à parte da frente da raiz da língua. Os dados foram aferidos a partir de uma imagem do plano médio sagital de cada token.

As medidas correspondentes à altura do corpo da língua foram mensuradas por meio da distância vertical do músculo genioglosso até o ponto mais elevado da língua, em centímetros. A aferição da parte frontal da língua foi realizada considerando a distância do músculo genioglosso até a raiz da língua, em centímetros, a partir de um ângulo de 45°. Foi considerado o ponto médio das produções de /l/ e cada estímulo foi produzido duas vezes.

Quanto aos resultados do estudo, Hwang, Charles e Lulich (2019) evidenciaram que cada participante lançou mão de estratégias próprias para a produção da lateral [l] em posição de final de palavra. Dois participantes, contudo, realizaram gestos de língua similares para a articulação do segmento alvo. As produções tenderam a ser produzidas com um bloqueio localizado apenas em um lado da cavidade oral, de maneira dissimétrica, havendo uma sutil abertura no

lado oposto do molar mais posterior até os incisivos. Um dos participantes realizou a constrição no lado esquerdo, enquanto o outro a articulou no lado direito. Hwang, Charles e Lulich (2019) reportam que os dois locutores realizaram gestos de língua de maneira muito parecida em contextos vocálicos distintos, havendo pouca variação na altura da língua, fato que já era esperado considerando determinados contextos vocálicos. Os resultados sugeriram que, ao articular a lateral, a língua se encontra em uma posição mais próxima da empregada para a articulação da vogal /i/ do que da vogal /a/.

Houve, ainda, um participante que realizou uma produção ápico-dental. Contrastando com as produções dos outros participantes, mencionados anteriormente, a constrição realizada foi menor. Os autores destacaram que a vogal adjacente anterior ao segmento gerou alterações sutis na altura do corpo da língua. Outro participante articulou o /l/ como um som retroflexo, enrolando a ponta da língua, promovendo o contato entre o palato e a superfície sublaminal. Essa especificidade na articulação foi identificada nos contextos vocálicos /u, ʌ, o, ɑ/. Além disso, também foi reportada uma produção lamino-alveolar, com uma constrição assimétrica do lado direito. Enfim, os resultados do estudo sugeriram que a variação está condicionada a produções peculiares de cada informante, algo que ainda não constava na literatura da área, evidenciando que não há um único padrão entre os locutores.

O ultrassom e a variação linguística no Brasil

Descrever fonético-fonologicamente a variação de uma comunidade/grupo de fala, utilizando imagens da articulação dos sons produzidos pelos seus membros é uma realidade recente e ainda bastante restrita no Brasil. Estudos amparados por uma metodologia sociofonética vêm descrevendo os dados de variação por meio da observação de aspectos acústicos, principalmente. Assim, Mileski (2013, 2017), por exemplo, propôs uma descrição de aspectos articulatórios, descrevendo as vogais médias de sujeitos bilíngues, falantes de português e polonês como língua de imigração. Bilharva-da-Silva (2019) apresentou padrões articulatórios para as vogais suarabáticas que ocorrem nos encontros [Cr] e [rC] do português de bilíngues português-pomerano. Em ambos estudos, a articulação dos

segmentos, inferida por meio de parâmetros acústicos, baseia-se, pois, em indícios deixados pela acústica.

A contribuição de pistas articulatórias mais sutis, advindas do ultrassom (ou outra ferramenta que possa fornecê-las) pode contribuir para uma maior robustez dos resultados obtidos na investigação empreendida. Aspectos finos da articulação de um som, em muitos casos, são responsáveis por peculiaridades da fala de uma comunidade — que podem ser identificadas nos mais variados âmbitos, como geográfico, econômico e cultural.

Rosinski (2019) e Ferreira-Goncalves e Rosinski (2020), por exemplo, buscaram ampliar os métodos que possibilitam descrever a variação dos sons do português brasileiro dentro de um grupo linguístico específico. Os estudos analisaram dados relativos à produção da consoante líquida lateral do português — o segmento // —, realizada em posição pós-vocálica, na fala de sujeitos bilíngues português/polonês como língua de imigração. Nesses trabalhos, as autoras descrevem, por meio de indícios acústico-articulatórios, os níveis de velarização da lateral, seguindo Sproat e Fujimura (1993). Segmentos com maior velarização são produzidos na parte mais posterior do trato articulatório, enquanto sons com menor nível de velarização são realizados na parte mais anterior, podendo ser caracterizados como alveolares. A caracterização de // no polonês é indicada pela literatura como sendo predominantemente alveolar⁷. Na língua polonesa, desse modo, o segmento realiza-se de forma distinta ao que é tido como padrão no português brasileiro, isto é, na forma vocalizada⁸ (realização posterior). Assim, também foi objetivo dos trabalhos detectar possíveis influências da língua de imigração na configuração da lateral produzida no português por sujeitos bilíngues.

A comunidade de fala cujos dados foram observados localiza-se no interior de Dom Feliciano-RS. Trata-se de uma comunidade rural desse município, fundada por imigrantes poloneses entre o final do século XIX e início do século XX. Além de outros aspectos culturais, prevaleceu na comunidade o uso do polonês como língua de imigração, ainda que o uso da língua esteja, atualmente, restrito a algumas famílias de descendentes e, nestas, a apenas alguns de seus membros.

⁷ Para mais informações, ver Szreder (2013) e Swan (2002).

⁸ Para detalhamento de caracterização, ver literatura de descrição da produção de // no português brasileiro, como Collischonn e Quednau (2009).

O número total de sujeitos que forneceram dados de fala para a realização dos estudos foi doze. Tal grupo incluiu seis sujeitos bilíngues e seis monolíngues (falantes apenas de português brasileiro). A participação de sujeitos monolíngues foi programada com o intuito de comparar os padrões de realização da lateral pós-vocálica na fala de indivíduos que utilizam a língua de imigração com as produções de falantes que dominam somente o português. Todos os sujeitos eram do sexo feminino e habitavam, desde o nascimento, a comunidade observada. As idades dos sujeitos ocuparam uma faixa etária de quinze a 59 anos. Os doze sujeitos tiveram produções contendo /l/ registradas por meio de um gravador digital modelo Zoom H4n. As coletas com a utilização do ultrassom ocorreram com um número reduzido de sujeitos: seis informantes fizeram parte desse subgrupo. Novamente, o grupo dividiu-se entre bilíngues e monolíngues com três representantes de cada grupo.

No tocante aos materiais elaborados e utilizados para a captação dos dados, empregou-se instrumentos de nomeação de imagens. O grupo monolíngue produziu o segmento lateral em português, enquanto o grupo bilíngue, ao qual foram apresentados dois instrumentos, produziu palavras em polonês e em português. As palavras foram realizadas seis vezes individualmente pelos sujeitos. Na produção oral, sem o uso do ultrassom, os vocábulos foram produzidos, em português, inseridos na frase veículo *digo _____ para você* e, em polonês, na frase veículo *“mowię _____ ponownie”* (NEWLIN-ŁUKOWICZ, 2012), que significa *“digo _____ de novo”*. No registro efetuado de forma simultânea com gravador digital e ultrassom, as palavras foram produzidas isoladamente, também repetidas por seis vezes.

Os resultados indicaram uma caracterização diferente da lateral pós-vocálica para sujeitos bilíngues em relação aos monolíngues. Utilizando-se como parâmetros de caracterização acústica os valores de F1 e F2 e os valores de diferença entre os dois formantes ($F_2 - F_1$), medidos no ponto médio da porção estável da lateral, identificaram-se realizações menos velarizadas de /l/ para os bilíngues. Tais produções indicaram uma articulação mais anterior para a maioria dos sujeitos. Este resultado destaca-se para aqueles cujo convívio diário é apenas familiar, ou seja, ocorre no espaço em que primordialmente se utiliza o polonês, o lar, já que se dedicam à agricultura e trabalham com seus

familiares. Acusticamente, as produções dos monolíngues apresentaram uma diferença menor na média de $F_2 - F_1$, evidenciando uma produção mais posterior de /l/ pós-vocálico.

As autoras apresentaram também uma descrição articulatória dos dados de apenas um dos sujeitos, caracterizado como bilíngue. Constatou-se que os padrões observados por meio da caracterização gestual ajustavam-se aos resultados obtidos por meio da observação acústica. A informante cujos dados foram analisados articulatoriamente apresentou caracterização variável para a lateral pós-vocálica: em alguns casos, sua produção foi mais anterior e, em outros, mais posterior, tal como pode ser visto, ao analisar-se dado a dado, o valor da diferença entre o primeiro e o segundo formantes.

A configuração articulatória de /l/ para os dados que apresentaram um valor mais baixo para $F_2 - F_1$ demonstrou um recuo do dorso e um abaixamento do corpo da língua, indicando uma posteriorização do articulador. Em contrapartida, produções de /l/ cujo valor dessa diferença foi mais elevado indicaram, articulatoriamente, elevação do dorso e da parte anterior da língua, o que denota anterioridade. Isso pôde ser visto nas produções do vocábulo *culpa*, por exemplo. Tomando-se três realizações seguidas da palavra, identificou-se, na primeira, uma diferença F_2-F_1 de 212 Hz. Observou-se, então, a elevação do articulador como consequência do movimento de dorso e da parte anterior da língua para frente. Na segunda e terceira produções, os valores de diferença F_2-F_1 foram de 1221Hz e 1415Hz, respectivamente, o que revelou claro direcionamento do dorso para a parte de trás do trato articulatório.

Por meio das imagens ultrassonográficas de /l/ produzido no polonês de imigração, as autoras constataram uma provável influência da língua de imigração sobre a forma como o segmento é produzido em português. A configuração articulatória observada nas produções do polonês acompanhou a que foi identificado para a lateral mais anterior do português: elevação do corpo e da região anterior da língua. Dessa forma, os indícios obtidos por meio das imagens ultrassonográficas permitiram admitir uma possível atuação da língua de imigração na forma como se caracteriza a lateral pós-vocálica no português da comunidade investigada.

Por fim, esses resultados, proporcionados pelo registro ultrassonográfico dos movimentos de língua, iluminam os dados a partir

de um outro ponto de vista, confirmando — ou não — o que a análise acústica revela. Dessa forma, fenômenos linguísticos analisados a partir de um olhar mais panorâmico dos dados, isto é, de médias gerais obtidas pela análise acústica, são complementares a uma observação pontualmente vigiada — em procedimentos que incluem seleção e monitoramento de dados expostos a um efeito lupa. O ultrassom é, pois, uma ferramenta bem-vinda para descrever a variação segmental da fala de um grupo, ainda que não dê conta, pelo menos até o momento, de coletas de um número expressivo de sujeitos.

Para além do laboratório: considerações finais

O estudo de Rosinski (2019), seguindo a metodologia sociofonética de captação de dados de fala, realizou o primeiro experimento de campo do Laboratório Emergência da Linguagem Oral (LELO) com o uso do ultrassom. A coleta demandou remodelações ligadas a adaptações dos materiais empregados para o registro dos dados (pois deveriam ser transportados e utilizados em ambiente não laboratorial, na comunidade de fala selecionada), ao tempo de captação das produções e ao tipo de interação a ser estabelecida com cada informante (que, no caso, contou com participantes de idade avançada). A proposta da metodologia sociofonética inclui justamente isso: obter dados com o emprego dos mesmos instrumentos utilizados dentro do laboratório, porém, com aplicação no ambiente de convívio dos informantes. Mas, além dos cuidados mais técnicos, a naturalidade dos dados também é algo a ser preservado (THOMAS, 2011). Por isso, coletar os dados no local em que os informantes habitam, tendo como entrevistadores integrantes da própria comunidade, pode proporcionar interações mais próximas e, conseqüentemente, uma realização mais natural dos sons que o pesquisador objetiva adquirir.

Como os materiais deviam ser transportados até a comunidade de fala, para um espaço provisório para sua instalação, os registros ocorreram em uma única data. Assim, foi necessário também contar com a disponibilidade e pontualidade dos informantes, o que foi obtido com sucesso.

O equipamento ultrassonográfico empregado foi o modelo *Mindray DP 6600*, um equipamento portátil. A coleta, a organização e a montagem das ferramentas ocorreram da seguinte forma: i)

instalação do equipamento de ultrassom, com acoplamento da sonda transdutora modelo 65EC10EA micro-convexa, ii) montagem do computador desktop contendo o *software* AAA (*Articulate Assistant Advanced*), por meio do qual foram apresentados os instrumentos de nomeação e captadas as imagens para análise, e sua conexão ao ultrassom, iii) instalação do sincronizador de áudio e vídeo Sync BrightUp modelo SBU 1.0, interligando-o ao computador e ao ultrassom e iv) disposição de um gravador digital modelo Zoom H4n, para captar os áudios das produções registradas. Foi também disponibilizado um capacete estabilizador de cabeça, desenvolvido pela *Articulate Instruments Ltd*, para ser utilizado pelos informantes a fim de manter fixa, sob a mandíbula dos sujeitos, a sonda transdutora. Esta foi a primeira etapa da organização da coleta.

A segunda etapa compreendeu a recepção e ambientalização dos sujeitos. Antes do início de cada coleta, ao informante, eram apresentados os equipamentos que seriam utilizados e, em seguida, as imagens a serem nomeadas no momento da captação dos dados. Após, foram realizados a colocação e os ajustes relativos ao capacete estabilizador, para proporcionar conforto durante a coleta e garantir uma firme conexão entre a sonda e o capacete. Após o cumprimento dessas etapas, acessórios e informante estavam prontos para o procedimento. Na Figura 4, pode-se observar a disposição dos materiais empregados no início de uma das coletas de dados via ultrassom.

Figura 4 – Imagem de arquivo



Fonte: (cedida por Aline Rosinski).

A captação de imagens articulatórias por meio do ultrassom e sua posterior análise necessitam de uma boa imagem da borda/superfície da língua. Logo, para a obtenção de uma imagem nítida, ajustes precisaram ser realizados no equipamento⁹, tais como o ganho, a potência acústica, a profundidade e a posição do foco. Ainda, por conta do curto espaço de tempo, as coletas foram realizadas quase de modo sucessivo, exigindo uma reorganização ágil e atenta na recomposição dos materiais antes de cada registro. Também, entre uma coleta e outra, os dados armazenados no gravador foram salvos em outro dispositivo, o que visou a garantir que não houvessem perdas de arquivos.

Para o desenvolvimento de pesquisa de campo, muito se alerta sobre o estabelecimento de um ambiente confortável para os informantes, o que é, de fato, fundamental, mas deve-se, igualmente, levar em conta o fato de os pesquisadores encontrarem-se em um local improvisado, que precisa ser acolhedor para o desenvolvimento das etapas planejadas. Trata-se de um ambiente que deve acolher, mas que não deixa de ter características de laboratório. Essa situação é desafiadora pois, caso haja qualquer imprevisto que danifique a captação das produções acústicas ou articulatórias, o estudo pode ficar comprometido. Uma coleta programada fora do laboratório tende a ser um momento irreplicável, já que depende de muitos delicados procedimentos, como o deslocamento do maquinário, a disponibilidade de espaço para a coleta dentro da comunidade, o agendamento e comparecimento dos informantes em janela de pouca extensão de tempo, como aqui apresentado.

Enfim, o dado articulatório obtido por meio do ultrassom e o mesmo dado adquirido por meio do gravador, principalmente quando é possível sincronizar o áudio e a imagem, fornecem perspectivas diferentes de um mesmo fenômeno linguístico. Os registros podem, então, como constatou-se nos trabalhos reportados, ser analisados por meio de *softwares* específicos, como o PRAAT e o AAA (há outros programas que também podem realizar essas análises, mas esses são os mais empregados por serem bastante amigáveis e intuitivos). Assim, os resultados logrados com o ultrassom atuam como uma lupa em relação às amostras acústicas, corroborando, descartando ou

⁹ Para mais detalhes sobre os ajustes aplicados ao equipamento de ultrassom, ver Ferreira-Gonçalves e Brum-de-Paula (2013)

desvendando detalhes da articulação inferidos por meio de oscilogramas e espectrogramas. Com o acesso a imagens dinâmicas da realização de um segmento, padrões articulatórios deduzidos através da análise acústica dos dados podem ser, de fato, observados. A ultrassonografia visa, então, a revelar e/ou confirmar finos detalhes articulatórios que podem advir de um condicionamento social. Pode, portanto, auxiliar na descrição da fala de um grupo de informantes.

A descrição da variação linguística de uma comunidade deve incluir a imersão no espaço onde as formas linguísticas são realizadas. É esse movimento que pode levar a conhecer a língua efetivamente utilizada nas comunidades de fala investigadas. Essa diretiva, inaugurada por Labov, continua vigorando, mesmo quando surgem novos obstáculos entre entrevistados e entrevistadores. O tempo ensina que empecilhos tendem a ser demovidos. Assim, gravadores, computadores e, mais recentemente, aparelhos de ultrassom, ao longo do tempo, por conta de avanços tecnológicos e novas demandas, ficaram mais flexíveis, performáticos e de pequenas dimensões. Lawson, Stuart-Smith e Scobbie, no estudo de 2008, já sentiam a necessidade de uma disponibilização de aparelhos de ultrassom menores para o estudo da fala informal. Tais aparelhos "de bolso" já foram desenvolvidos, servindo, notadamente, para auxiliar em terapias de fala. Os dispositivos tecnológicos, não reduzidos a gravadores e ultrassons unicamente, promovem novos roteiros e rotinas para que coletas e análises de dados ocorram sem ou com poucos entraves. Eles possibilitam, ainda, que, na medida em que passam a habitar os laboratórios universitários, compreendam-se melhor a linguagem e seus gestos.

PARÂMETROS EM ANÁLISES QUALITATIVA E QUANTITATIVA: DADOS ULTRASSONOGRÁFICOS DE LÍNGUA¹

Bruna Teixeira Correa
Giovana Ferreira Gonçalves

Introdução

Atualmente, com o avanço das ferramentas tecnológicas, vários são os aparelhos e técnicas que permitem a observação direta da articulação do trato vocal, como a ressonância magnética (MRI), o raio-x, a fotonasografia, a fibroscopia, a articulografia eletromagnética (EMA) e o ultrassom (US). Quando comparados, são recorrentes as desvantagens constatadas em relação ao ultrassom, o que faz com que pesquisadores lancem mão das imagens ultrassonográficas para fins de pesquisas nas áreas da linguística e da fonoaudiologia (BARBARENA *et al.*, 2014; BARBIERI, 2019; BERTI, 2010; BRESSMAN, 2011; STOLAR, GICK, 2013; MÉNARD *et al.*, 2012; TEIXEIRA-CORREA, FERREIRA-GONÇALVES, BRUM-DE-PAULA, 2017; e ZHARKOVA, 2013). O custo, o caráter invasivo, a dificuldade de acesso, a portabilidade e, inclusive, a segurança em relação a outras tecnologias levam muitos estudiosos da área a aderirem a esse método inovador, não invasivo e de maior naturalidade na produção dos sons, tendo em vista que não é necessário inserir artefatos no aparelho fonador.

A observação da articulação em tempo real trouxe muitos ganhos em relação a outros métodos de análise dos dados da fala, como a inspeção acústica e de oitiva, os quais apresentam limitações. No primeiro, por meio da observação de oscilogramas e espectrogramas, podemos fazer inferências acerca do funcionamento do trato vocal para a produção de um determinado som. No segundo, utiliza-se um sentido humano que, muitas vezes, pode ser bastante impreciso e passível de interferências: a audição. Por conta disso, demanda, a fim de obter um grau de confiabilidade necessário, outros ajustes

¹ Pesquisa apoiada pelo Conselho Nacional de Desenvolvimento em Pesquisa, processo 314022/2018-0.

metodológicos, como uma maior quantidade de dados coletados e a inclusão de juízes para a descrição dos dados.

Fazer uso da ferramenta ultrassonográfica para a investigação dos dados de fala, especialmente se unida a outras técnicas, a depender do objetivo da pesquisa, pode, de fato, trazer contribuições relevantes às análises fonético-fonológicas. Com o uso do ultrassom, é possível investigar contrastes fônicos de diversas línguas e caracterizar os movimentos linguais desses segmentos (JESUS, SHADLE, 2002; JONGMAN *et al.*, 1998; SILVA, 2017), promover a emergência de gestos articulatórios no processo de aquisição de línguas estrangeiras, por meio da realização de atividades de instrução explícita (PILLOT-LOISEAU *et al.*, 2013; TATEISHI, 2013; WHITE *et al.*, 2017), contribuir com os estudos em sociofonética (LAWSON *et al.*, 2011; SCOBIE *et al.*, 2008), observar a maturação do controle da língua na aquisição de língua materna e nas terapias de fala (NOIRAY *et al.*, 2018; e ZHARKOVA, HEWLETT, HARDCASTLE, 2011), entre muitas outras opções de linhas de trabalho.

Várias são as decisões metodológicas que influenciam direta ou indiretamente na análise dos dados. Uma delas é a definição do melhor plano para a observação do objeto de estudo, isto é, se sagital (separa as metades esquerda e direita da borda da língua) ou coronal (separa as metades anterior e posterior), cujas imagens representam “reconstruções de fatias ou cortes virtuais bidimensionais do trato vocal” (FERREIRA-GONÇALVES; e BRUM-DE-PAULA, 2013, p. 72). Outro ponto a ser destacado é a definição de qual tipo de sonda (linear, micro convexa, convexa) será utilizado para a captação das imagens, especialmente considerando o som a ser investigado e a faixa-etária dos informantes, já que essa escolha está relacionada à qualidade das imagens obtidas. Ainda, o pesquisador deve fazer uso de equipamento adequado para a imobilização dos movimentos de cabeça em relação à sonda, de forma a viabilizar a posterior análise quantitativa dos dados.

A análise das imagens ultrassonográficas da borda da língua demanda o uso de *softwares* específicos e o consequente estabelecimento de parâmetros para, assim, ser possível fazer comparações, caracterizações, distinções e categorizações. Sem a definição de parâmetros, sejam de base qualitativa e/ou quantitativa, o objeto de estudo não pode ser aferido e muito menos contrastado.

Considerando as análises quantitativa e qualitativa, há diversas possibilidades para o tratamento dos dados ultrassonográficos. Incorporadas ao *software* AAA² (*Articulate Assistant Advanced*), por exemplo, existem as opções de (i) realização do contorno automático da borda da língua – o que auxilia sobremaneira no tempo despendido pelo pesquisador para a realização dos trabalhos – ou do contorno manual; (ii) observação de *frames* quadro a quadro ou em movimento; (iii) plotagem dos contornos das línguas em gráficos, para melhor visualização e posterior comparação; (iv) realização de cálculos estatísticos na busca de pontos de diferenças entre bordas de línguas analisadas, dentre muitas outras possibilidades.

Como é possível observar, são muitos os detalhes e os ajustes metodológicos para que uma análise de imagens ultrassonográficas seja robusta e possa, então, ser validada. Ajustes nas ferramentas, no modo de conduzir a pesquisa, na preparação do informante e, especialmente, na escolha dos parâmetros quantitativos e qualitativos para a investigação de determinado fenômeno fonético-fonológico são de extrema importância.

O objetivo deste capítulo é, portanto, fazer uma revisão teórica acerca dos parâmetros de análise mais utilizados em trabalhos que utilizam o ultrassom como ferramenta metodológica em pesquisas linguísticas. Para isso, primeiramente, abordaremos os parâmetros qualitativos de análise, discorrendo sobre as porções da língua, sua curvatura e deslocamento e, em seguida, falaremos sobre os parâmetros quantitativos, como índice de anterioridade, grau de curvatura, entre outros.

Parâmetros qualitativos

Muitos são os parâmetros utilizados em uma análise qualitativa das imagens ultrassonográficas da borda da língua, a depender do objetivo do trabalho e do objeto de estudo. Normalmente, para esse tipo de análise, o número de informantes é pequeno, pois é feita uma investigação detalhada da produção – considerando seus aspectos articulatórios e acústicos³. Por meio desse tipo de análise e,

² *Software* criado pela *Articulate Instruments* (<http://www.articulateinstruments.com/aaa/>).

³ Importante reportar que, para a realização das análises ultrassonográficas, o pesquisador tem por base, primeiramente, o dado acústico, pois é com a análise de

consequentemente, da escolha de parâmetros, as imagens ultrassonográficas são amplamente descritas, comparadas e explicadas. O índice de anterioridade da língua, a ponta para cima ou para baixo, uma porção da língua (anterior, média ou posterior) retraída, a elevação de dorso, a postura mais achatada, o local e o grau de constrição de uma porção específica da língua são alguns exemplos de parâmetros qualitativos encontrados na literatura. A seguir, reportaremos alguns, especificando a sua função e trazendo exemplos de pesquisas que deles fizeram uso. Salientamos, porém, que as pesquisas exemplificadas geralmente fazem uso de mais de um parâmetro, incluindo também os quantitativos, no entanto, focaremos, sempre, naquele correspondente à seção.

Ponta da língua

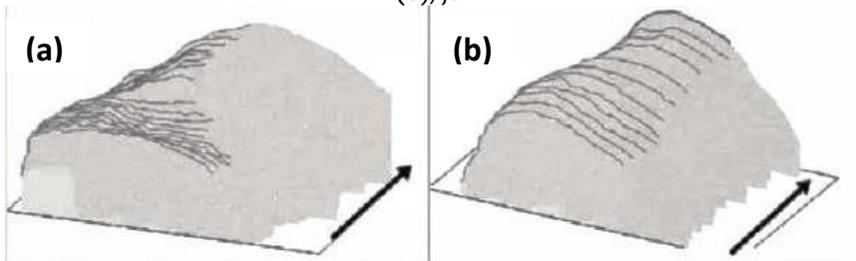
O movimento da ponta da língua é o que viabiliza a produção de alguns segmentos, como a fricativa dental surda do inglês, [θ] – encontrada, por exemplo, nas palavras *death* e *three* – e o retroflexo [ɻ] do português brasileiro – encontrado em palavras como *mar* e *carta*, característico, dentre outras, da variedade mineira e de cidades do interior paulista. Para a produção de [θ], a ponta da língua é articulada ao encontro dos dentes; já no que concerne ao retroflexo, a ponta da língua faz uma curvatura e se direciona para parte posterior da boca. Apesar de haver variações no movimento realizado, se para cima, para baixo, para frente ou para trás, a ponta da língua é de suma importância para a produção de alguns sons, como os exemplificados, e, portanto, é utilizada como parâmetro para análises qualitativas.

Lawson *et al.* (2011) fizeram um estudo sociofonético acerca do /r/ do inglês escocês e observaram que sutis diferenças articulatórias podem indicar casos de estratificação social. Assim, enquanto falantes de classe média utilizavam articulações *bunched*⁴ para a produção do rótico, os da classe trabalhadora utilizavam a ponta da língua e a sua parte frontal.

oscilogramas e espectrogramas que são selecionados os *frames* para a análise. Observa-se, assim, a extrema relevância da perfeita sincronização que deve ocorrer entre som e imagem.

⁴ A referência do termo em inglês foi mantida considerando a dificuldade de correspondência de sentido para a tradução em português.

Figura 1 – Exemplos de seqüências de contornos de língua para palavras terminadas em /ar/ – em um período de tempo de 30 ms. Em (a), *par* e, em (b), *far*



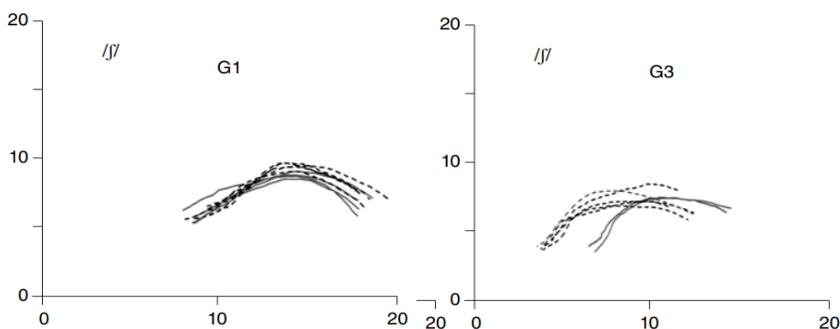
Fonte: (adaptada de LAWSON et al., 2011, p. 260).

Como é possível observar na Figura 1, ao final da flecha, que corresponde à produção da consoante final da palavra, diferentes gestos são ativados para a realização do mesmo segmento. Em (a), temos a ponta da língua sendo ativada e levada para cima; já em (b), toda a porção frontal da língua, incluindo a ponta, faz esse movimento. Segundo os autores, os falantes estão pouco cientes desses movimentos articulatórios; no entanto, sugerem que essa sistematização possa ser, em um futuro, socialmente ordenada. A observação de um conjunto de movimentos, incluindo a elevação da ponta da língua, possibilitou, aos autores, um estudo sociofonético, cujos resultados indicaram correlação com a classe social e o gênero: a classe de trabalhadores e os homens utilizam mais a ponta da língua durante a produção do /r/.

Outro estudo que fez uso desse parâmetro foi o de Wertzner, Francisco e Neves (2014), porém, observaram o abaixamento da ponta da língua. Esses autores fizeram uma descrição do contorno da língua na produção dos alvos fricativos /s/ e /ʃ/ em três grupos de sujeitos com e sem transtornos⁵ fonológicos. Observaram que, enquanto o grupo sem transtorno fonológico produzia a palatal com a ponta da língua para baixo, em direção ao soalho da boca, e o dorso elevado, o grupo com transtorno fonológico produzia com a língua achatada na boca, com uma pequena elevação do dorso e sem o abaixamento da ponta da língua, conforme pode ser observado na Figura 2.

⁵ O termo *transtorno fonológico* é utilizado em acordo com o texto dos autores.

Figura 2 – Contornos de língua para a produção de /ʃ/ pelo G1 (sem transtorno fonológico, à esquerda da Figura) e pelo G2 (com transtorno fonológico, à direita da Figura). Linhas tracejadas = meninos; Linhas contínuas = meninas. À esquerda de cada imagem, está a raiz da língua e, à direita, a ponta da língua



Fonte: (adaptado de WERTZNER et al., 2014, p. 250).

Os pesquisadores perceberam que o G2, grupo com transtornos fonológicos, não diferenciava os gestos articulatórios para /s/ e /ʃ/, pois realizavam os mesmos movimentos para ambos os sons. Ainda, esse conjunto de gestos era diferente daquele utilizado pelos informantes de G1, sem transtorno fonológico. Essas observações trouxeram contribuições relevantes à investigação, até então perceptivo-auditiva, do fonoaudiólogo, propiciando, dessa forma, evidências para o diagnóstico do transtorno fonológico por meio de imagens articulatórias em tempo real.

Em Barberena, Soares e Berti (2014), o movimento da ponta de língua também é observado, mas para a produção das líquidas /r/ e /l/. O objetivo do trabalho foi investigar os dados de natureza articulatória de cinco falantes típicos, cuja língua materna é o português brasileiro, a partir da produção de consoantes líquidas e da coarticulação dessas consoantes quando em diferentes contextos vocálicos, como em barata/salada, ferida/falida, charuto/cabeludo. Com isso, poderiam estabelecer padrões para os falantes típicos do PB e, assim, confirmar ou não os achados obtidos por meio de análises acústicas, presentes na literatura, os quais indicam dois movimentos articulatórios para as líquidas: um de ponta e um de dorso, concomitantemente. Para isso, escolheram, como *frame* de análise, aquele que evidenciava a maior elevação da ponta da língua, conforme pode ser observado na Figura 3.

Figura 3 – Frames relativos às produções de *charuto* (à esquerda) e *cabeludo* (à direita) por um único sujeito, que expressam a elevação máxima da ponta da língua (flechas) para as consoantes /r/ e /l/



Fonte: (adaptado de BARBERENA et al., 2014, p. 342)

As autoras observaram, em seus dados, os mesmos resultados apresentados pela literatura (SILVA, 2002; e SILVA et al, 2006), embora por outros métodos de análise. Viram que, por haver gestos duplos, provavelmente haja mais dificuldade na coordenação desses movimentos e que, para a sua aquisição, talvez sejam necessárias “habilidades coordenadas e refinadas de movimentos dissociados e simultâneos da língua” (BARBERENA et al., 2014, p. 343).

Em relação à ponta de língua, observaram que o grau de constrição foi maior para /r/ do que para /l/. A lateral, portanto, demonstra menor complexidade de coordenação, ainda mais quando constatado que a produção do *tap* para o alvo /r/ necessita de uma “coordenação simultânea entre o movimento balístico da ponta da língua e a constrição em direção à faringe” (BARBERENA et al., 2014, p. 343). Foi observada, também, uma maior constrição na ponta da língua quando o *tap* foi produzido diante da vogal alta e anterior, o que pode indicar o papel facilitador dessa vogal.

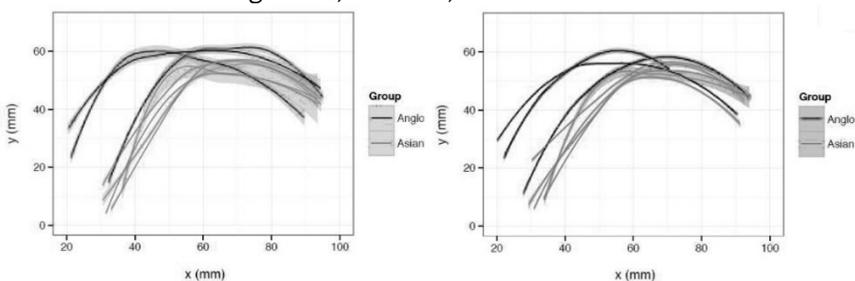
Os três estudos exemplificados investigaram, entre outros movimentos, aqueles relativos à ponta da língua: levantamento, abaixamento, local e grau de constrição. A partir da observação desses movimentos, estabeleceram padrão, confirmaram diagnóstico e reconheceram estratificação social no nível articulatório. A ponta da língua, portanto, é um parâmetro promissor para análise de imagens ultrassonográficas da borda da língua.

Dorso da língua

Assim como a ponta, o dorso da língua desempenha papel importante na articulação de alguns segmentos, como as consoantes velares, cuja produção depende da elevação do dorso da língua, o qual entra em contato com o palato, como em [k] e [g]. Assim como ocorre com a ponta da língua, sua movimentação pode ser variada: para cima, para baixo, mais para frente, mais para trás.

Um dos estudos que observou a movimentação do dorso como parâmetro qualitativo foi o de Hirkham e Wormald (2015). Os autores observaram as características acústicas e articulatórias encontradas na variação dos segmentos líquidos produzidos por locutores ingleses e asiáticos, cujo dialeto é o inglês falado em Bradford, na Inglaterra. A pesquisa contou com cinco informantes do sexo feminino que produziam pares mínimos do tipo *lead/reed*. Em seus resultados, constataram que as informantes asiáticas fazem uso da região anterior do trato para a produção das consoantes líquidas /l/ e /r/, enquanto as participantes inglesas a parte posterior e o dorso da língua, cujo movimento é de retração, conforme pode ser observado na Figura 4.

Figura 4 – Contorno das bordas das línguas para (a) o ponto médio de /l/ em *lead* e para (b) o ponto médio de /r/ em *berry*. Em preto, os dados das inglesas e, em cinza, os das asiáticas

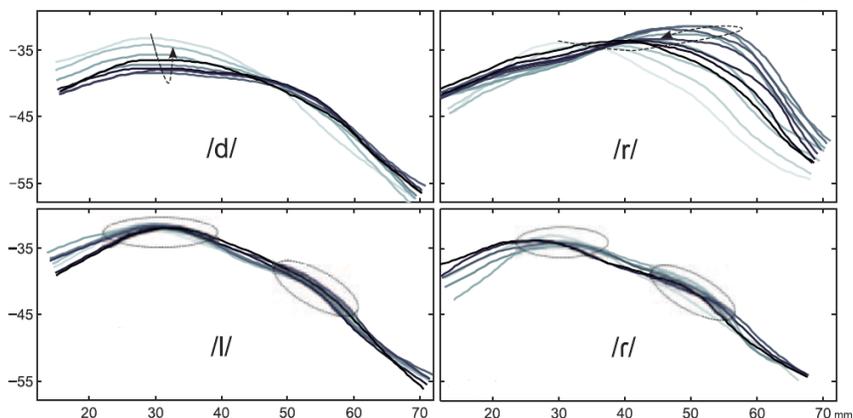


Fonte: (adaptado de HIRKHAM e WORMALD, 2015, p. 3-4).

Como é possível constatar, as falantes inglesas, ao contrário das asiáticas, apresentam maior elevação do dorso e também maior retração. Conforme os autores, a ausência desses movimentos nos dados das asiáticas é resultante de um padrão apresentado por falantes asiáticos em processo de aquisição de uma variante do inglês.

O trabalho de Proctor (2011) também utilizou o mesmo parâmetro de análise. O autor busca descobrir um aspecto específico que possa caracterizar, de forma universal, os segmentos líquidos, como ocorre com outras classes de sons. Para isso, Proctor (*op.cit*) comparou /d/ com as líquidas laterais e não laterais do espanhol, observando o comportamento do movimento articulatorio do dorso. Dessa forma, seria possível verificar o que os segmentos apresentavam de semelhanças e de diferenças. O seu experimento contou com a participação de cinco falantes nativos do espanhol da América Latina, os quais produziram palavras como *ere/erre/ele/hede*.

Figura 5 – Contornos das bordas das línguas da obstruinte e das líquidas do espanhol em contexto intervocálico frontal, à esquerda parte frontal da língua

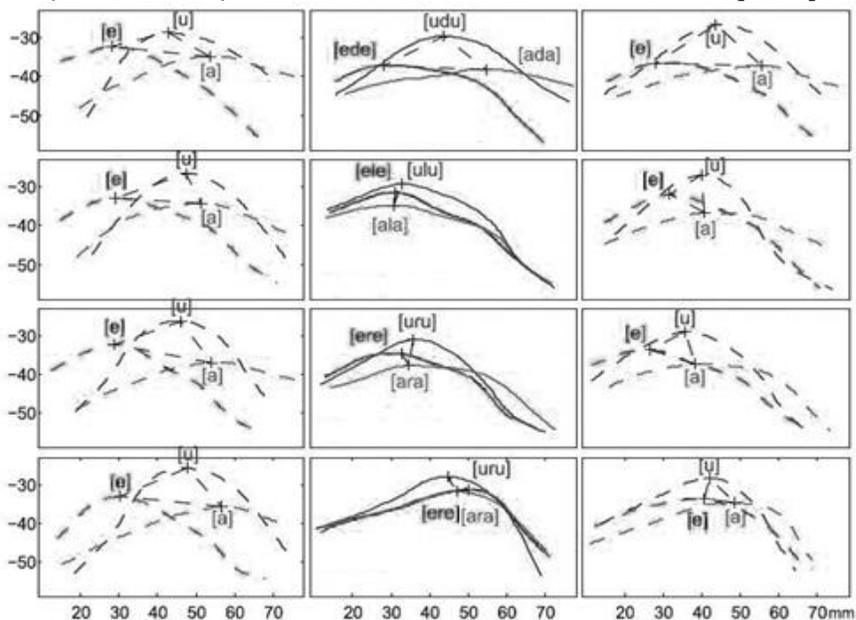


Fonte: (adaptada de PROCTOR, 2011, p. 458)

Como é possível observar, por meio da Figura 5, para o alvo /r/, o dorso da língua é direcionado para a parte de trás da boca, realizando movimento elevado e posterior, que é, justamente, o sentido contrário da direção das vogais escolhidas como contexto anterior e posterior. A lateral e o *tap*, da mesma forma, não têm movimento de língua em direção à parte anterior do trato, o que indica que o gesto é de fato característico da consoante e não tem influência do contexto vocálico. O dorso para esses dois sons se mantém em estabilidade.

A fim de investigar, em maior detalhamento, a influência do contexto nos movimentos articulatorios das líquidas, o autor as inseriu nos contextos [e_e], [a_a] e [u_u], observando, outra vez, a movimentação dorsal.

Figura 6 – Contornos das bordas das línguas na produção da obstruente /d/ e das líquidas /l, r, r/, respectivamente, nos três contextos intervocálicos [e, a, u]



Fonte: (adaptada de PROCTOR, 2011, p. 459).

Em relação à obstruente, na primeira linha, o autor observou posições dorsais influenciadas pelo contexto intervocálico; já para as líquidas, esse efeito de coarticulação foi menos expressivo. Os dados revelaram que há certa variação na articulação dorsal das líquidas, com papel de itens lexicais e de diferentes falantes. Destacam-se também as seguintes constatações: (i) o alvo dorsal da lateral é mais anterior do que o de /r/; o alvo dorsal do /r/ é mais posterior e tipicamente mais elevado do que o da lateral; (iii) o alvo dorsal do /r/ se encontra entre o da lateral e do tap. Por fim, o autor concluiu que as consoantes líquidas do espanhol compartilham da propriedade comum de um componente dorsal na realização da sua produção e que esse movimento articulatorio as diferencia da obstruente /d/, produzida apenas com um gesto coronal.

O parâmetro qualitativo do movimento do dorso da língua, como exemplificado, auxilia na caracterização de vogais orais, na descrição das produções provenientes do contato com outra língua – como no caso da comparação entre o inglês das asiáticas e das inglesas – e

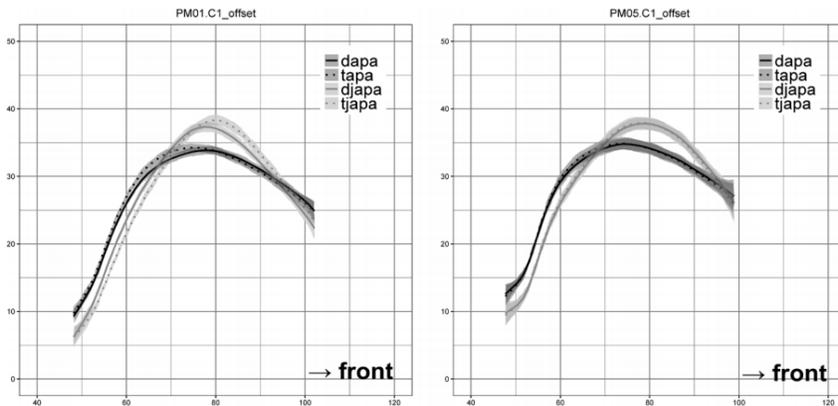
também na distinção entre consoantes líquidas, e entre líquidas e outras classes. Ao que tudo indica, é, portanto, um parâmetro relevante para investigações dessa ordem, a depender do segmento ou da classe em questão.

Raiz da língua

Ao contrário dos parâmetros já abordados, a visualização da raiz da língua, nas imagens do ultrassom, não é tão clara. Isso se deve à sombra do osso hioide, que gera um escurecimento na parte posterior da língua. Talvez esse seja o motivo de o número de estudos qualitativos que tem como parâmetro essa porção da língua ser bastante menor em comparação àqueles que observaram a ponta e/ou o dorso. Em consequência, são recorrentes as decisões metodológicas que devem ser tomadas para evitar quaisquer tipos de problemas na observação não tão clara dessa porção da língua, como a escolha entre o traçado automático ou manual, o descarte de *frames* e o cuidado na análise de pontos de significância que sejam encontrados na região da raiz. Por conta disso, a maioria dos estudos que observa essa porção da língua é de natureza quantitativa, para maior confiabilidade. Apesar disso, os movimentos da raiz, para frente e para trás, são de extrema importância para muitos segmentos. Falaremos sobre dois estudos que fizeram uso desse parâmetro.

O primeiro deles é o de Matsui e Kochetov (2018), sobre o fenômeno de palatalização no russo. Os autores investigaram o posicionamento da raiz da língua em plosivas sonoras e surdas, que possuem palatalização contrastiva /t, d, tʲ, dʲ/, e observaram se esse movimento teria papel em sua distinção. Para isso, analisaram as imagens ultrassonográficas das produções de cinco falantes nativos de russo.

Figura 7 – Contornos de bordas de línguas das consoantes plosivas /t, d, tʃ, dʃ/ produzidas por dois informantes do estudo

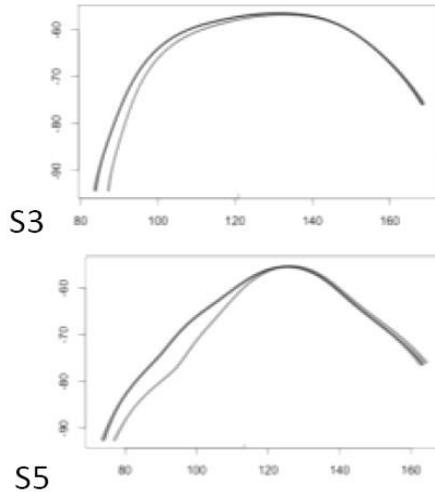


Fonte: (adaptada de MATSUI, KOCHETOV, 2018, p. 88).

Segundo os autores, os informantes distinguem as consoantes plosivas palatalizadas e não palatalizadas, entre outros movimentos, pelo avanço da raiz da língua, que é maior para as palatalizadas, conforme pode ser visto pelos contornos em cinza, do que para as não palatalizadas. A raiz da língua, no entanto, não se mostrou um parâmetro relevante para distinguir plosivas surdas de sonoras, ao contrário do que algumas pesquisas vinham demonstrando, relativas a outras línguas, em que a raiz é movimentada para atender aos requisitos aerodinâmicos da sonorização, conforme veremos a seguir.

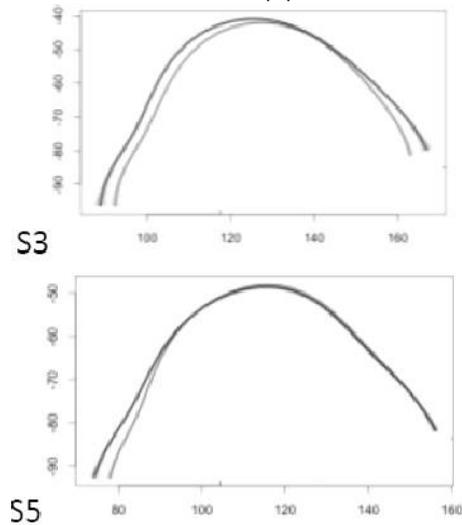
No estudo de Ahn (2015), foi investigado se havia, no inglês americano, diferenças nos movimentos articulatórios da raiz quando produzida uma plosiva surda ou sonora. Para isso, coletou dados de oito informantes americanos, os quais produziram as plosivas surdas e sonoras (labial, velar e alveolar) em posição inicial de enunciado, como em *pine trees* e *tie shoes*.

Figura 8 – Médias dos contornos das bordas das línguas de dois informantes para as plosivas alveolares. Em cinza, a plosiva sonora /d/ e, em preto, a plosiva surda /t/



Fonte: (adaptado de AHN, 2015, p. 8)

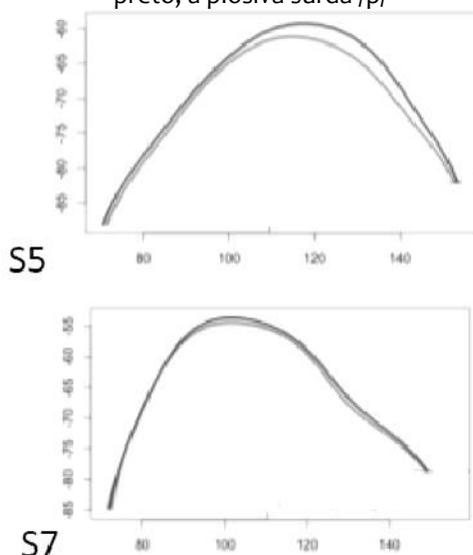
Figura 9 – Médias dos contornos das bordas das línguas de dois informantes para as plosivas velares. Em cinza, a plosiva sonora /g/ e, em preto, a plosiva surda /k/



Fonte: (adaptado de AHN, 2015, p. 9)

Os resultados apontaram que, na posição inicial do enunciado, havia uma distinção clara entre plosivas surdas e sonoras, em relação ao posicionamento da raiz da língua, para as alveolares e velares. Como é possível observar pelas Figuras 8 e 9, a raiz da língua avança em maior grau quando uma plosiva alveolar/velar sonora é produzida. No caso de S5, nas plosivas velares, essa é a única diferença encontrada entre a surda e a sonora, estando sobreposto o restante do traçado das outras porções da língua.

Figura 10 – Médias dos contornos das bordas das línguas de dois informantes para as plosivas labiais. Em cinza, a plosiva sonora /b/ e, em preto, a plosiva surda /p/



Fonte: (adaptado de AHN, 2015, p. 9)

Tal diferença, no entanto, não foi encontrada entre as plosivas labiais. Como é possível observar nos exemplos trazidos na Figura 10, as raízes das línguas estão praticamente sobrepostas, não apresentando diferenças para esse movimento. No caso de S7, quase não há distinção entre o par surdo e sonoro; já S5 apresenta diferença no dorso e na ponta da língua, não sendo, portanto, a raiz da língua relevante para o estabelecimento de diferenças entre /b/ e /p/.

Como já mencionado, são poucos os trabalhos que utilizam o ultrassom para observar, qualitativamente, o movimento da raiz da

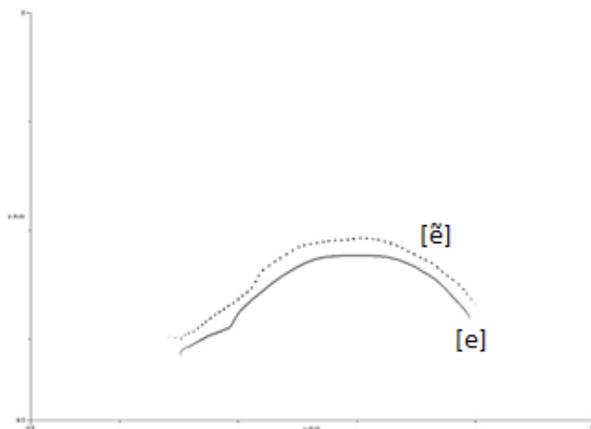
língua. Na maior parte dos estudos, há recorrência de análises estatísticas para uma maior confiabilidade dos dados advindos de uma região cuja visibilidade é restrita nas imagens ultrassonográficas, mesmo que a sonda seja posicionada na parte mais posterior do trato oral. Ainda assim, trouxemos um recorte de dois estudos que, primeiramente, fazem uma descrição de cunho qualitativo, embora depois apresentem resultados com base na estatística inferencial. Como ficou claro, a raiz da língua, assim como outros parâmetros, também é relevante para a distinção de alguns segmentos e de suas características, como a palatalização e o vozeamento.

Eixos vertical e horizontal

Alguns estudos, ao contrário dos já citados, fazem observações menos específicas acerca dos movimentos de língua, investigando somente o seu deslocamento nos eixos vertical e/ou horizontal, isto é, se a posição da língua para produzir determinado som é mais alta, mais baixa, mais anterior ou mais posterior, incluindo, então, todas as suas partes, ponta, dorso e raiz.

Teixeira-Correa *et al.* (2017) observaram a aquisição das vogais nasais francesas por aprendizes brasileiras via análise de imagens ultrassonográficas. Para isso, utilizaram também os dados de dois grupos controle: o da nativa monolíngue do português brasileiro e o da nativa do francês, a fim de estabelecer comparações com os dados das aprendizes. Em vez de lançarem mão de uma ferramenta metodológica que desse acesso à visualização do véu palatino – tendo em vista que é um dos principais articuladores envolvidos na realização de uma vogal nasal –, como a ressonância magnética, por exemplo, optaram por observar se havia, nos movimentos de língua, diferenças específicas entre uma vogal oral e uma vogal nasal, no português brasileiro e no francês.

Figura 11 – Médias dos contornos das bordas de língua do ponto médio das vogais [ɛ̃] (linha pontilhada) e [e] (linha sólida) produzidas pela nativa do PB

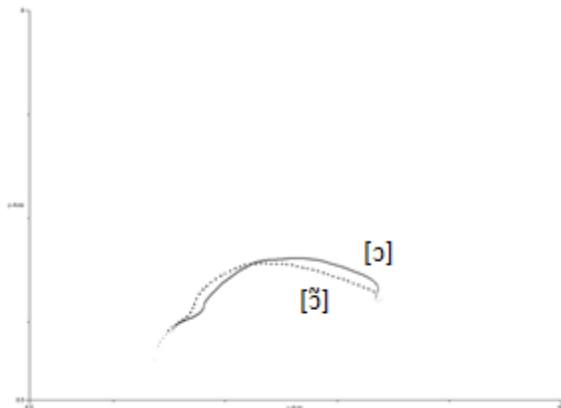


Fonte: (adaptada de TEIXEIRA-CORREA *et al.*, 2017, p. 271).

As autoras observaram, pelos dados da nativa do PB, que, para todos os pares de vogais oral–nasal, inclusive naqueles diferentes em altura, como [ɛ] – [ɛ̃], e exceto para [o]-[õ], o eixo vertical era determinante para a observação da nasalidade no português brasileiro. Como é possível averiguar pela Figura 11, a linha pontilhada, referente à vogal nasal, está completamente mais elevada do que a linha sólida da vogal oral. A língua, isto é, sua ponta, dorso e raiz, se desloca com maior elevação no eixo vertical.

No entanto, para o francês, as autoras observaram que o eixo horizontal era determinante na diferenciação vogal oral – vogal nasal, pois, para todos os pares de vogais, a produção da nativa evidenciava o recuo da língua para a parte posterior do trato, conforme pode ser visto na Figura 12, pela linha pontilhada.

Figura 12 – Médias dos contornos das bordas de língua do ponto médio das vogais [õ] (linha pontilhada) e [o] (linha sólida) produzidas pela nativa do francês

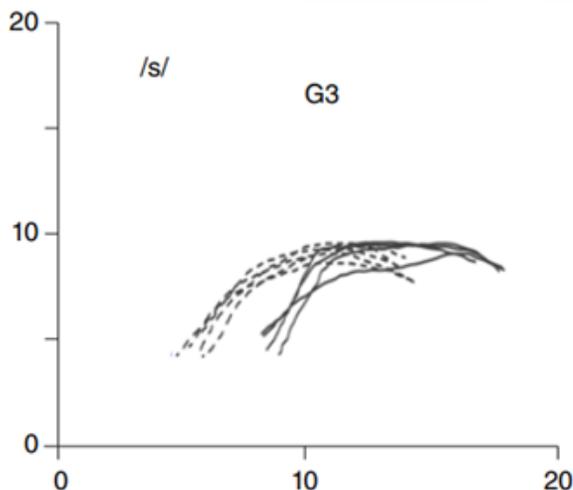


Fonte: (adaptada de TEIXEIRA-CORREA et al., 2017, p. 290).

As autoras também observaram que, algumas das aprendizes de francês, embora não conseguissem diferenciar as nasais do francês entre si, utilizavam a mesma estratégia da nativa para diferenciar uma vogal oral de uma nasal na língua francesa, ou seja, o eixo horizontal como determinante.

Outro estudo que tratou da importância dos eixos de deslocamento da língua, vertical e horizontal, foi o já citado Wertzner, Francisco e Neves (2014). Na descrição de /s/ do Grupo 3 (duas crianças apresentando o processo fonológico de anteriorização da palatal), observaram que os sujeitos utilizavam movimentos distintos para a produção da fricativa, conforme pode ser constatado na Figura 13.

Figura 13 – Contornos da borda de língua da produção de [s] de uma menina – linha contínua – e de um menino – linha tracejada



Fonte: (adaptado de WERTZNER, FRANCISCO, NEVES, 2014, p. 250).

Por meio da Figura 13, podemos ver que, na tentativa de produzir a fricativa surda alveolar, a língua da menina foi direcionada mais para frente no eixo horizontal, isto é, em direção à parte anterior da boca, enquanto a do menino fez o movimento contrário. Além disso, tanto o menino quanto a menina posicionaram a língua de forma mais alta no eixo vertical, diferentemente das crianças dos outros grupos do estudo. O mesmo aconteceu, em relação ao eixo vertical, para o alvo /ʃ/.

Como foi possível constatar, são vários os parâmetros qualitativos, encontrados na literatura, utilizados nas análises de imagens ultrassonográficas. Poderíamos ter reportado ainda outros nesta seção, porém, consideramos que as divisões da língua em porções, as quais têm relação direta com a produção de muitos segmentos, e um parâmetro mais geral, relacionado somente à movimentação da língua como um todo nos eixos x e y, constituem alicerces básicos para compreendermos o que deve ser considerado quando estamos diante de uma imagem, em tempo real, desse importante órgão para a produção da fala.

Parâmetros quantitativos

Como forma de mensurar e estabelecer valores de referência advindos das imagens do ultrassom, muitos pesquisadores fazem uso de parâmetros quantitativos para a análise dos dados. Por meio desses parâmetros, é possível realizar a comparação de resultados, sua verificação e validação estatística. Sendo assim, é necessário obter uma maior quantidade de dados – no comparativo à pesquisa qualitativa –, o que nem sempre é uma tarefa fácil, considerando as especificidades de uma coleta articulatória.

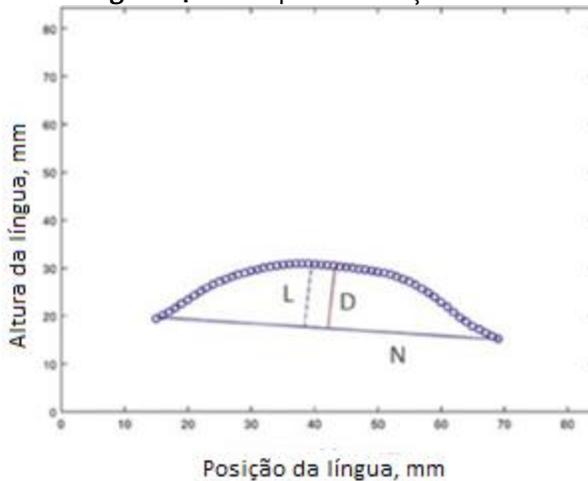
Nas pesquisas relacionadas ao ultrassom, na área de linguística, podemos encontrar diversos parâmetros quantitativos empregados para análise dos movimentos de língua: índice de concavidade, índice de anterioridade, grau de curvatura, diferença entre picos, distância entre curvas, medida angular, entre outros. Assim como ocorre com os parâmetros qualitativos, esses são escolhidos pelo pesquisador, levando em consideração o objeto de estudo em questão. Nesta seção, reportaremos alguns desses parâmetros, de forma a descortinar os critérios de análise e os índices utilizados. Novamente, salientamos o fato de que os trabalhos geralmente fazem uso de mais de um parâmetro, mas focaremos naquele relativo à seção em que se encontra descrito.

Índice de Excursão do Dorso (DEI)

O Índice de Excursão do Dorso é uma medida que quantifica a amplitude do dorso da língua em direção ao palato durante a articulação de determinado som. Bahgban *et al.* (2020) fizeram uso desse parâmetro em seu estudo com quatro crianças típicas falantes de persa como língua materna, com média de 7.4 anos, investigando o comportamento do movimento lingual para a produção dos alvos /p, t, k, f, s/ quando em contexto vocálico de /a/ e /i/ em sílaba CV. A hipótese dos autores era a de que os valores de DEI seriam significativamente diferentes de acordo com o contexto vocálico em que se encontrava a consoante obstruinte. Para obter a medida do DEI, alguns procedimentos precisam ser realizados: (a) desenhar uma linha reta (N) entre o ponto mais anterior e o ponto mais posterior da curva da língua; (b) especificar o ponto médio de (N); (c) traçar linhas perpendiculares a (N) até quaisquer pontos da curva da

língua; (d) desenhar uma linha vertical que cruze o ponto médio de (N), a qual é chamada de (D), representando o grau de amplitude do dorso na parte da frente e de trás da língua e (e) calcular o DEI a partir da divisão da razão de (D) por (N). A Figura 14 exemplifica o cálculo de DEI e de outro parâmetro utilizado pelo estudo, (L), o qual pode ser, neste momento, desconsiderado.

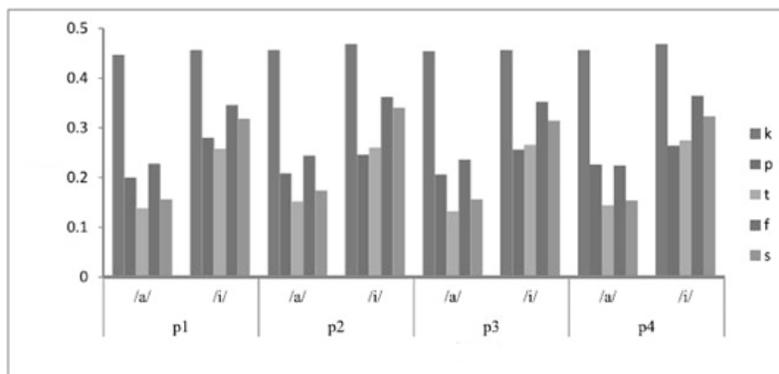
Figura 14 – Exemplo de medição de DEI



Fonte: (adaptada de BAHGBAN et al., 2020, p. 2).

Os autores observaram, em sua pesquisa, que, para todas as cinco consoantes investigadas, o valor médio de DEI foi significativamente maior em contexto vocálico de /i/ quando comparado ao de /a/, para todos os participantes, conforme pode ser observado na Figura 15.

Figura 15 – Valores médios de DEI de todos os informantes para as consoantes em contexto de /a/ e /i/



Fonte: (adaptada de BAHGBAN et al., 2020, p. 4).

O valor de DEI também colaborou para diferenciar as consoantes entre si, apresentando valores significativos para tal distinção, especialmente entre consoantes alveolares e plosivas velares, cujo DEI é maior, o que se mostrou consistente para todos os participantes. Além disso, as consoantes /p/ e /f/ apresentaram medidas muito distintas das demais. Segundo os autores, isso se deve ao fato de que, para a produção de consoantes labiais, a língua não precisa assumir uma postura especial, fazendo com que adote a posição da língua das vogais vizinhas e, portanto, apresente bastante variação.

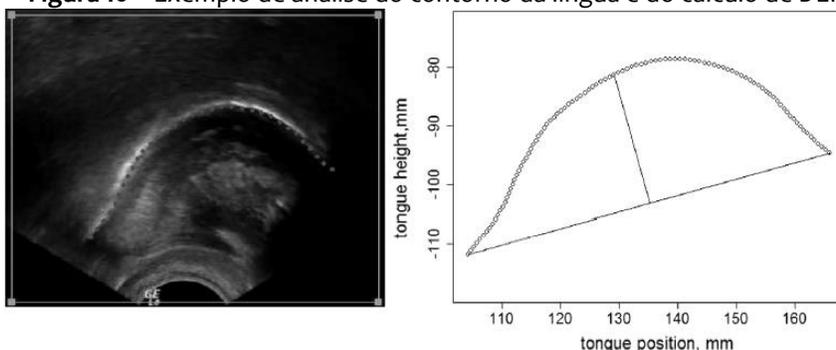
Também foi possível observar, por meio desse parâmetro quantitativo, que não houve efeito significativo para a consoante velar /k/. Segundo os autores, para a produção das consoantes não velares, o dorso tem algum espaço para ir em direção ao palato quando em contexto de uma vogal alta, nesse caso, /i/; já para uma vogal baixa, ocorre o contrário. Assim, o fato de DEI, para /k/, não apresentar diferenças, considerando-se os contextos vocálicos, corrobora ser uma medida válida da amplitude do dorso da língua, pois o dorso atinge o palato, independentemente do contexto vocálico na articulação de /k/.

Byun et al. (2016) foi outro estudo que fez uso do DEI, tendo como foco de investigação os contrastes encobertos das plosivas velares e alveolares na fala infantil. Para os autores, a aquisição dos sons da fala é um processo gradual e, portanto, seguidamente podem ser observados contrastes encobertos na fala de crianças que estão

adquirindo uma língua. A fim de investigar esses contrastes, lançaram mão da ferramenta do ultrassom, a qual possibilitaria a observação de diferentes formatos de língua durante a produção de um som percebido como neutralizado. Os autores salientaram que esse parâmetro só pode ser utilizado para imagens que mostram toda a língua, da sombra do osso hioide à sombra da mandíbula. Enfatizaram, ainda, que, quanto maior o DEI, maior a atividade do dorso da língua durante a produção de um som.

A pesquisa contou com a participação de duas crianças diagnosticadas com *anteriorização velar*, isto é, ao invés de produzirem um som alvo articulado na parte de trás da boca, produziam outro na parte anterior do trato. Participaram do estudo mais duas crianças que produziam o contraste velar-alveolar de forma acurada. Para isso, fizeram uso das palavras *cup-tub* e *key-T*, na posição inicial, e *back-bat* e *bug-mud*, na posição final.

Figura 16 – Exemplo de análise do contorno da língua e do cálculo de DEI



Fonte : (adaptado de BYUN *et al.*, 2016, p.24)

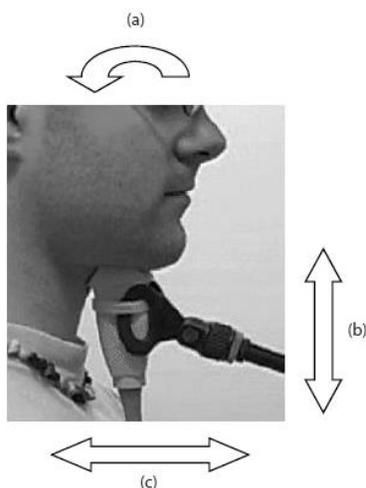
As plosivas velares e alveolares produzidas pelas duas crianças com produção acurada diferiram em relação à medida do ultrassom (DEI); as plosivas velares e alveolares produzidas por uma das crianças com anteriorização velar, no entanto, não apresentaram diferenças no valor de DEI. Alguns dados de Max, um dos informantes, foram salientados pelos autores, pois apresentavam contrastes encobertos, os quais foram observados por meio do DEI e por meio de inspeção acústica. Nesse sentido, os autores concluíram que o DEI não é uma medida mais sensível do que as medidas acústicas para a observação de contrastes encobertos, embora seja possível utilizá-lo como recurso de análise.

O Índice de Excursão do Dorso é um parâmetro quantitativo de análise de imagens ultrassonográficas que cumpre com o seu papel de quantificar a amplitude da língua em direção ao palato. Assim como ocorre com os outros parâmetros, nem sempre essa medida vai ser a mais adequada para observar determinado objeto de estudo. A sua melhor contraposição e comparação está com aqueles sons que, para serem produzidos, têm a movimentação do dorso bastante ativa, como as vogais altas e as consoantes velares.

Grau e posição de curvatura

O grau e a posição de curvatura da língua aparecem como parâmetro em diversos trabalhos, seja para observar um segmento em específico, a coarticulação, ou para avaliar uma questão metodológica, como as que exemplificaremos a seguir. O estudo de Ménard *et al.* (2012) tem, dentre os seus objetivos, avaliar a validade de diversas métricas desenvolvidas para caracterizar as formas da língua, como o grau e a posição de curvatura, por meio de imagens do ultrassom, sem fazer uso de nenhum sistema de estabilização de cabeça, e através de diferentes posições da sonda: deslocamento horizontal, vertical e em rotação, conforme Figura 17, a seguir.

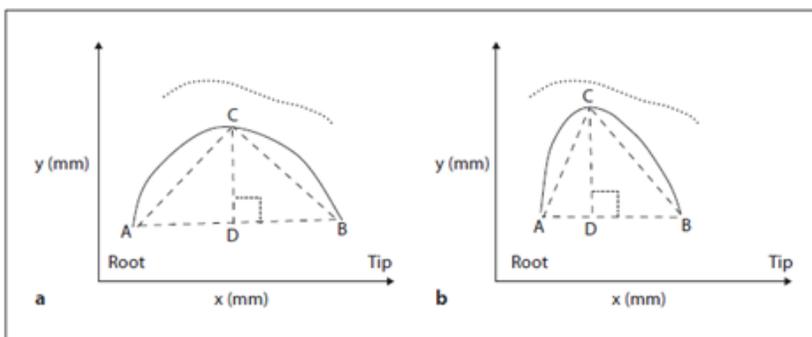
Figura 17 – Simulação dos movimentos da sonda do ultrassom: (a) rotação, (b) movimento vertical e (c) movimento horizontal



Fonte : (adaptada de MÉNARD *et al.*, 2012, p. 64).

Os autores buscavam observar, portanto, se, para esses dois parâmetros, grau e posição de curvatura, haveria diferença de resultados, a depender da posição da sonda, para as vogais orais do francês /a, e, i, o, u, y/. Para isso, os dados foram comparados a imagens de formas de línguas sintetizadas por meio de um modelo articulatório, no qual a posição do palato e de vários articuladores é reconhecida. Dessa forma, segundo os autores, é possível determinar quais parâmetros têm relação com o deslocamento da sonda e quais são diferenças reais de formato de língua.

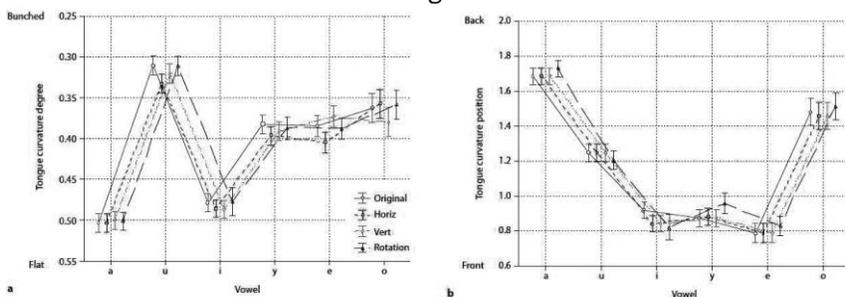
Figura 18 – Representação esquemática do contorno da língua (linha sólida), do palato (linha pontilhada) e do triângulo para extrair parâmetros



Fonte: (MÉNARD *et al.*, 2012, p. 64).

Para a realização dos cálculos, as bordas das línguas receberam os seus contornos e, em seguida, foram remodeladas em um triângulo. O primeiro e o último ponto de intersecção com as linhas foram ligados e considerados como a base do triângulo (A-B). O pico do triângulo (C) foi o ponto mais alto do contorno da língua em relação à base do triângulo. Medidas angulares de A, B e C foram, então, extraídas; medidas de curvatura da língua e de posição de curvatura da língua foram determinadas por meio dos pontos A-D. A curvatura da língua é definida como a proporção da distância C-D sobre a distância A-B, já a posição da curvatura da língua é definida como a razão da distância A-D sobre a distância D-B.

Figura 19 – Efeitos dos diferentes movimentos da sonda em valores médios do (a) grau de curvatura da língua e da (b) posição de curvatura da língua

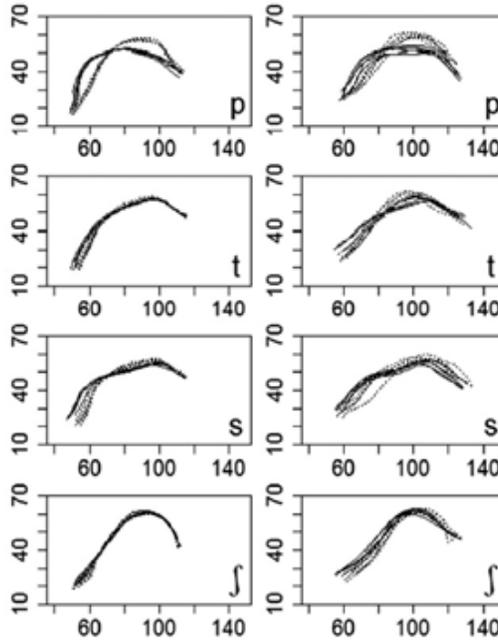


Fonte: (adaptada de MÉNARD *et al.*, 2012, p. 69).

Os resultados apontaram *p* significativo para a influência da vogal no grau de curvatura da língua, aumentando significativamente de /a/ para /i/, /y, e/, /o/ e, por fim, /u/. Da mesma forma, para a posição da curvatura da língua, foi encontrado valor significativo para *p* em relação à influência da vogal, diferindo entre todos os segmentos vocálicos, exceto para /y, i/. Não foi observado efeito do movimento da sonda na curvatura da língua e na sua posição, indicando que esse parâmetro não é significativo, independentemente do tipo de deslocamento da probe.

Outro estudo que observou os parâmetros grau e posição de curvatura da língua, concomitantemente, foi o de Zharkova *et al.* (2015). Seu objetivo era analisar os dados de ultrassom de dez falantes adolescentes, coletados com e sem estabilização de cabeça, produzindo os mesmos estímulos. A hipótese dos autores era a de que, tendo em vista que a posição estabilizada da cabeça provavelmente restringiria o movimento da mandíbula em comparação com o não uso do estabilizador, as localizações absolutas de grau e de posição das curvas durante a estabilização seriam menos variadas do que durante a não estabilização. Para isso, os informantes produziram as consoantes do inglês /ʃ/, /s/, /t/ e /p/ em dois contextos vocálicos distintos, /a/ e /i/, totalizando 800 dados analisados. Além do contorno automático da borda da língua possibilitado pelo software utilizado, os cálculos utilizados para a definição do grau e da posição de curvatura são os mesmos de Ménard *et al.* (2012), apresentados anteriormente na Figura 18.

Figura 20 – Exemplo de contornos de bordas da língua durante a produção de um informante (em contexto de /a/ = linhas sólidas; em contexto de /i/ = linhas pontilhadas; coluna à esquerda: com estabilização de cabeça; coluna à direita: sem estabilização de cabeça).



Fonte: (adaptada de ZHARKOVA *et al*, 2015).

Os resultados do trabalho, em relação à coarticulação da vogal sobre a consoante, apontam que o grau de curvatura pode ser utilizado para comparar formas de língua relativas às produções de consoantes alveolares entre diferentes contextos vocálicos e entre grupos de falantes com condições de estabilização de cabeça distintas. O grau de curvatura também pode ser utilizado para análises similares de /ʃ/. Nesse caso, na condição de cabeça estabilizada, apenas o efeito da vogal foi significativo.

Em relação às diferenças específicas de consoantes, foi encontrado resultado significativo para a curvatura na condição estabilizada somente para a consoante fricativa pós-alveolar. Sem a estabilização da cabeça, no entanto, é possível que a diferença significativa na curvatura tenha aparecido devido a menor pressão da sonda portátil no queixo. Tendo em vista a posição da cabeça, /p/ foi a única consoante com efeito

significativo, com valores maiores para a condição estabilizada, mostrando que, para as bilabiais, é necessário muito cuidado para a escolha de parâmetros a serem avaliados, levando em consideração a ocorrência ou não de movimento da sonda. Por fim, os autores concluíram que é possível obter informações confiáveis sobre os movimentos coarticulatórios da língua, especificamente sobre o grau e a posição de curvatura, sem a estabilização da cabeça.

Essas duas medidas, juntas, constituem um parâmetro capaz de fornecer, ao pesquisador, uma visão holística do movimento da língua, pois deixa claro onde exatamente está ocorrendo a curvatura, se no dorso, na ponta ou na raiz – o que já classifica muitos segmentos –, e o grau dessa curvatura – o que diferencia determinados sons, evidenciando também a influência da coarticulação.

Índice de Anterioridade

O índice de anterioridade (IA), segunda Silva *et al.* (2017, p. 3), “é a multiplicação dos valores extraídos em milímetros (mm) para cada ponto mensurado da língua em um número ponderado correspondente”. Valores mais altos de IA indicam que essa parte da língua apresentou maior movimentação quando comparada às outras partes, como média e posterior, por exemplo.

Silva *et al.* (2017) observaram, em seu trabalho, se esse parâmetro, entre outros, poderia depreender a diferença de modo e ponto de articulação entre 14 fonemas consonantais do português brasileiro (/l/, /ʎ/, /r/, /R/, /n/, /ɲ/, /t/, /d/, /k/, /g/, /s/, /ʃ/, /z/ e /ʒ/). Para isso, a pesquisa contou com 20 participantes de produção típica da fala, os quais produziam as consoantes em contexto de /a/, a partir da criação de logatomas monossílabos, como *ala* e *aga*, totalizando 840 dados analisados.

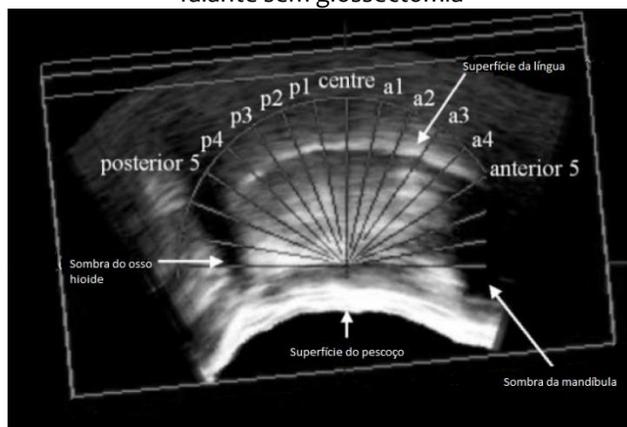
Para cada produção, foi selecionado o ponto máximo de constrição da língua, de forma que toda a sua extensão pudesse ser observada; após, foi realizado o desenho manual da borda da língua e, por fim, os valores obtidos a partir da multiplicação em milímetros (mm) passaram por testes estatísticos. O índice de anterioridade, nesse trabalho, não se mostrou eficaz para diferenciar ponto e modo de articulação dessas 14 consoantes analisadas, apresentando valores altos de *p*: 0,88 para ponto e 0,70 para modo. Esperava-se que o IA pudesse diferenciar os fonemas produzidos pela porção anterior da

língua – como as alveolares e as pré-palatais – daqueles produzidos pela porção média e posterior – como as velares e palatais.

Conforme as autoras, algumas razões podem estar associadas a não significância desse parâmetro em sua análise: (i) em relação ao ponto de articulação, deve-se considerar o posicionamento da língua no trato vocal, pois o espaço existente entre a ponta da língua e os alvéolos é menor do que o existente entre as outras porções da língua e o palato duro. Assim, a porção anterior sempre terá valores e deslocamentos menores, por questões anatômicas. O mesmo raciocínio pode ser considerado para o modo de articulação, pois uma obstrução total ou parcial na região velar, por exemplo, sempre terá valor maior de altura do que uma obstrução na região dos alvéolos; (ii) o fato de o referido índice não ser uma medida relativa, ou seja, questões como o tamanho da cavidade oral dos indivíduos ou, ainda, as possíveis movimentações de cabeça, podem ter interferido nos resultados.

Outra pesquisa que fez uso do IA como parâmetro quantitativo para análise de imagens ultrassonográficas foi o de Bressman *et al.* (2005). O objetivo do trabalho era avaliar, a partir de um ultrassom tridimensional, a protusão, o sulco e a simetria da língua durante a produção de um som da fala produzido por doze falantes sem glossectomia e por um informante que passaria por uma glossectomia parcial, isto é, cirurgia que remove a língua ou parte dela. Os autores queriam avaliar qual parâmetro seria determinante para um bom resultado de fala após um procedimento cirúrgico desse tipo. Para isso, observaram produções dessa informante quatro semanas antes da cirurgia e após o procedimento. Os dados continham fonemas do inglês, a saber: /a, i, u, s, l, n, ʃ, ɲ/. Os sons vocálicos foram utilizados porque representam os extremos das articulações linguais para as vogais; as consoantes porque apresentam gestos típicos para os sons da fala em inglês, como retroflexão, elevação alveolar e fechamento velar.

Figura 21 – Grade de medição da borda da língua de /j/, produzido por um falante sem glossectomia



Fonte: (adaptada de BRESSMAN et al., 2005, p. 577)

Os pesquisadores fizeram uso do *software* FreeScan, o qual projeta visualizações coronal, sagital e horizontal do volume tridimensional da língua, todos advindos do conjunto original de imagens coronais. No modo de exibição do *software*, esses três planos são exibidos simultaneamente e um marcador indica a sua intersecção, a qual foi usada como ponto de ancoragem para as medições. A fim de obtê-lo, foi identificado, pela imagem coronal, o corte mediano paralelo ao septo lingual. No corte sagital, o marcador de intersecção foi colocado na borda superior entre as sombras da mandíbula e do osso hioide. A partir desse ponto médio de ancoragem sagital, foram identificados os planos direito e esquerdo. Esses planos eram paralelos ao ponto médio da imagem sagital e, então, foram colocados entre o septo lingual e os lados da língua pela imagem coronal.

Usando esses pontos de ancoragem anatômicos intrínsecos, segundo os autores, é possível medir e fatorar a rotação da mandíbula, além de viabilizar a medição das modificações da língua. Após, uma grade de medição concêntrica com linhas de medição espaçadas foi sobreposta às imagens, a qual contou com 15 medições (anteriores, posteriores e central), cinco de cada plano (coronal, sagital e horizontal). Dessa forma, foi observado, entre outros parâmetros, o índice de anterioridade da língua.

Os autores encontraram índices baixos de anterioridade para /a/ e mais altos para /i/, estando /u/ com valores entre as duas vogais

anteriormente citadas. Tais resultados já eram esperados, pois as vogais /a/ e /i/ demandam um deslocamento mais extremo do corpo da língua. Para a informante que passaria por glossectomia, antes da cirurgia, foi encontrado padrão muito semelhante ao dos outros participantes, exceto para o valor de anterioridade de /u/, bem menor do que para /a/. Os valores obtidos no pós-operatório foram semelhantes aos obtidos no pré-operatório, porém, um pouco superiores e comprovadamente significativos.

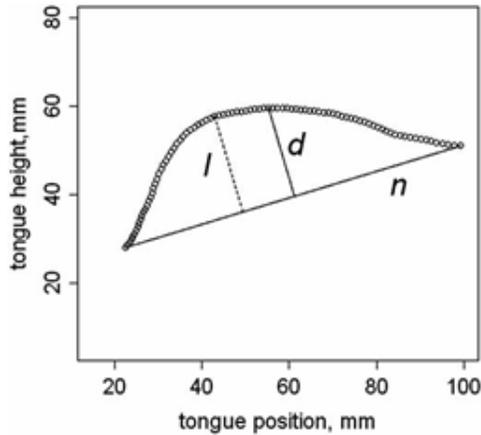
Segundo os autores, o índice de anterioridade foi informativo, mas não tanto quanto outros parâmetros utilizados no estudo. A pesquisa concluiu que a cirurgia não afetou a capacidade de a informante projetar a sua língua, pois apresentou valores significativamente diferentes para o índice de anterioridade após a cirurgia. Concluiu, também, que o ultrassom tridimensional contribuiu com o detalhamento e a avaliação concisa da forma da língua e dos parâmetros influentes para a realização da fala.

Índice de Posição de Restrição da Língua

O TCPI representa o lugar máximo de excursão da língua, é o Índice de Posição de Restrição da Língua. Tem valor igual a zero quando, ao longo do contorno da língua, a parte mais agrupada (*bunched*) está na região do dorso; tem valores positivos quando o local de excursão/amplitude máxima está situado na parte mais anterior e valores negativos quando na parte mais posterior. As vogais frontais, por exemplo, provavelmente tenham valores de TCPI positivos, enquanto as vogais posteriores, ao contrário, valores negativos.

Zharkova (2013) utilizou o parâmetro quantitativo TCPI para analisar os dados ultrassonográficos de seis adultos falantes de inglês escocês *standard*, sem alterações na fala, produzindo os sons consonantais [p, t, k, s, l, r, f], seguidos das vogais [a] e [i], a fim de observar a coarticulação advinda do segmento vocálico seguinte. Esperava-se valores significativamente maiores de TCPI para todas as consoantes em contexto de [i] quando comparado ao de [a]. O objetivo do trabalho era distinguir as articulações, realizadas para a produção dos sons em questão, com e sem excursão do dorso. O total de dados analisados foi de 419.

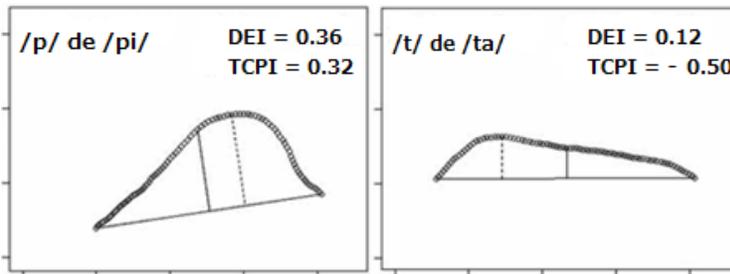
Figura 22 – Curva da língua, ilustrando cálculo de TCPI



Fonte: (adaptado de ZHARKOVA, 2013, p. 487)

O denominador, para o TCPI, foi igual a $n/2$. O numerador, segunda a autora, é a distância entre a perpendicular mais longa de n para a curva da língua, isto é, a linha tracejada l na Figura 22, e a perpendicular para a curva da língua traçada a partir do ponto médio de n , ou seja, a linha contínua d . Nos casos em que a linha l cruzou o contorno da língua na área posterior àquela cruzada pela linha d , o TCPI foi arbitrariamente atribuído a um valor negativo; quando a linha l cruzou a curva da língua mais à frente do que a linha d , o TCPI foi positivo.

Figura 23 – Exemplo de TCPI positivo e negativo, respectivamente



Fonte: (adaptada de ZHARKOVA, 2013, p. 488)

Os resultados interfalantes apontaram que o valor de TCPI foi maior quando a consoante estava em contexto de $[i]$ em comparação ao de $[a]$,

conforme esperado. O mesmo padrão foi observado para os falantes individualmente, exceto para um em específico, em relação à produção da consoante [k]. Sendo assim, portanto, a vogal teve efeito significativo. O estudo cumpriu com o seu objetivo de estabelecer valores para adultos típicos a partir de um índice do formato da língua. Os valores de TCPI formaram resultados altamente consistentes entre os falantes. Os padrões observados nos seis falantes demonstram que o TCPI pode medir mudanças na forma da língua e fazer inferências sobre o comportamento do seu dorso. Segundo a autora, esse índice pode ser usado para estudar a coarticulação lingual e para quantificar mudanças na forma da língua dependendo do grau de anterioridade das vogais.

O TCPI é um parâmetro ainda pouco usado pela literatura da área, porém, como pôde ser observado, bastante promissor em relação à movimentação do dorso da língua. Assim como constatado acerca dos parâmetros qualitativos, são vários os parâmetros quantitativos utilizados nas análises ultrassonográficas – índice de concavidade, índice de assimetria, medida angular, índice de modificação de curvatura, entre outros–, porém, tentamos trazer, nesta seção, uma revisão de alguns deles, os quais, certamente, continuarão sendo explorados em novas pesquisas da área.

Considerações finais

A língua, enquanto hidróstato muscular, isto é, órgão composto principalmente por músculos, tem, como principais características: (a) a constância de volume, ou seja, qualquer diminuição em uma dimensão causará, conseqüentemente, aumento compensatório em pelo menos uma outra dimensão e (b) maior variedade de movimentos quando comparada a outros sistemas, como o musculoesquelético, por exemplo (KIER e SMITH, 1985). Essas duas características têm relação direta com o que observamos nas diferentes línguas faladas pelo mundo: cada uma, a sua maneira, faz uso da incompressibilidade e do alto número de movimentos possíveis conforme sua necessidade fonético-fonológica, seja para diferenciar sons, coarticulá-los ou para expressar alguma gradiente.

Com o emprego do ultrassom nas pesquisas linguísticas, foi possível analisar em tempo real os diversos movimentos do hidróstato lingual – um dos órgãos mais importantes para a produção da fala –,

em suas diferentes porções. São inegáveis os ganhos que essa tecnologia inovadora trouxe para a área da linguística e também da fonoaudiologia, colaborando com ensino de línguas estrangeiras, com inspeções socioarticulatórias, possibilitando diagnósticos clínicos, entre muitas outras aplicações.

O trabalho com imagens ultrassonográficas, no entanto, exige do pesquisador alguns domínios, como o próprio manuseio do equipamento, o conhecimento dos diversos comandos e ajustes das imagens, o cuidado com a sonda e com todos os aspectos metodológicos que interferem na análise dos dados, incluindo a escolha do *software* e dos parâmetros de análise.

Nesse sentido, a escolha dos parâmetros selecionados para a análise de determinado som, via imagens de ultrassom, é de fato ponto central na realização exitosa da pesquisa. Quais são as principais características articulatórias do segmento a ser estudado? O que o diferencia em relação a outros segmentos dentro da mesma classe? Em que contextos linguísticos o encontramos? Para responder à pergunta de pesquisa, é interessante a escolha de parâmetros qualitativos? Quantitativos? Ou ambos? Enfim, são várias as questões que precisam estar claras para o pesquisador que utiliza a ultrassonografia para a análise de fala. Se bem respondidas e articuladas, são grandes as chances de uma pesquisa inovadora e com relevantes contribuições para a literatura da área.

PARTE DOIS

PROPOSTA DE INTERVENÇÃO COM APOIO DA ULTRASSONOGRRAFIA DOS MOVIMENTOS DA LÍNGUA PARA OS DISTÚRBIOS DOS SONS DA FALA

Luciana da Silva Barberena
Denis Altieri de Oliveira Moraes
Larissa Cristina Berti
Marcia Keske-Soares

Este capítulo apresenta um estudo que utiliza a ultrassonografia (USG) dos movimentos da língua como ferramenta para a intervenção com o som [r] *tap*. A produção do ‘r’ na posição inicial da sílaba dentro da palavra no Português Brasileiro (PB) é bastante diversificada, sendo que na região central do sul do Brasil ocorre a produção do [r] *tap* nessa posição da palavra.

No PB, as consoantes líquidas são as que mais dificuldades oferecem às crianças com distúrbios dos sons da fala (DSF). Os DSF são definidos, segundo a *American Speech-Language and Hearing Association - ASHA* (2014), como alterações que afetam a percepção, produção motora e/ou a representação fonológica dos sons da fala. Neste capítulo serão abordados especificamente o desvio fonológico (LAMPRECHT et al. 2004), que são quadros de dificuldade na organização do sistema fonológico de sua comunidade linguística.

GICK et al. (2008) referem que os sons de ‘r’ no Inglês apresentam maior complexidade na aquisição por envolverem duplos gestos de língua quando analisados em imagens de USG. No Inglês Americano, as posturas de língua para ‘r’ são frequentemente divididas em duas categorias principais: retroflexo e *bunched*. No retroflexo a ponta da língua levanta e enrola um pouco para trás, na vizinhança do rebordo alveolar. No *bunched* a ponta da língua diminui enquanto o corpo da língua levanta para aproximar-se do palato duro, sendo esse o som que se assemelha à produção do [r] *tap* do Português Brasileiro (PB).

Avaliações instrumentais abordam aspectos importantes sobre a inabilidade na produção da fala. Os avanços das tecnologias disponíveis permitiram propor um modelo dinâmico de tratamento da fala, capaz de interpretar os gestos articulatórios e a multimodalidade

dos mesmos para a sua caracterização. Os modelos dinâmicos procuram desvendar a inexistente ruptura entre a fonética e a fonologia e interpretam fenômenos na fala que revelam a existência de estados intermediários/gradientes, que podem ser esclarecidos por instrumentos, como a USG da fala (ALBANO, 2012, BALL; KENT, 1997; BROWMAN; GOLDSTEIN, 1992; GOLDSTEIN; BYRD; SALTZMAN, 2006 KENT; READ, 2002; e SCOBIE, 1998).

Browman; Goldstein (1992) afirmam que os gestos articulatórios são unidades básicas de contraste fonológico, caracterizações abstratas de eventos articulatórios com tempo intrínseco. Dessa forma, o gesto articulatório representa tanto eventos abstratos quanto físicos, sendo considerada uma unidade pré-linguística, a criança utiliza os gestos inicialmente mais simples no balbúcio e, gradativamente, até que sejam alcançados os alvos adultos, aprende a coordenar e sincronizar os diferentes gestos entre si. Para alguns gestos, de aquisição mais tardia, é necessária maior habilidade para diferenciar, por exemplo, dois gestos de língua em um único articulador anatomicamente acoplado, como a língua, tal como na produção de sons de 'r'. KLEIN et al. (2013) referem que devido às características complexas que envolvem os sons de 'r', estudos internacionais têm recorrido a metodologias que utilizem instrumentos na análise desses sons, tais como a USG, tanto na avaliação como no processo de intervenção.

Stone (2005) descreve os detalhes técnicos e princípios da USG. A ultrassonografia cria imagens de tecido do corpo usando ondas sonoras de ultra alta frequência, revelando distintas formas da língua e movimentos de articulações alveolares, velares e para vogais. Estudos têm investigado os movimentos da língua pela USG durante a fala com os sons de 'r' em crianças e adultos (BRESSMANN et al., 2010; CAMPBELL et al., 2010; DAWSON; CARL; WHALEN, 2015; DERRICK; SCHULTZ, 2013; KOCHETOV; POUPLIER; TROUNG, 2013, LAWSON; STUART-SMITH; SCOBIE, 2008; MAYER; GICK, 2012; MIELKE, 2011; e PROCTOR, 2009).

Imagens obtidas pela USG também podem ser utilizadas como um importante recurso no tratamento de crianças com distúrbios dos sons da fala (DSF). A USG dos movimentos da língua pode fornecer pistas visuais, táteis e sinestésicas, como objetivo de modificar o comportamento articulatório anormal. As imagens fornecidas pelo

ultrassom têm por objetivo tornar visíveis os mecanismos de produção que a criança utiliza. Esses estudos sugerem que a generalização pode ser ampliada quando a intervenção utiliza um mecanismo de biofeedback visual e proprioceptivo para a aprendizagem dos sons.

No processo terapêutico, a USG tem sido utilizada especialmente como um método de biofeedback visual direto na intervenção com o som 'r' (ADLER-BOCK et al., 2007; BACSFALVI, 2010; BERNHARDT et al., 2008; CLELAND; SCOOBIE; WRENCH, 2015; GICK et al., 2008; KLEIN et al., 2013; e MODHA et al. 2008). A USG para os fonemas líquidos, quando comparada a outros instrumentos que podem ser utilizados como recurso terapêutico, tem sido referida como aquela com maiores ganhos quando utilizada como recurso durante a intervenção, pois as imagens forneceram maior movimentação da língua (BERNHARDT et al., 2003).

Estudos recentes verificaram que a USG da fala é um meio de alcançar a generalização e a manutenção de ganhos para o tratamento (CLELAND; SCOOBIE; WRENCH, 2015; HITCHCOCK; BYUN, 2015; e KLEIN et al., 2013). Esses estudos geralmente apresentaram palavras com 'r' associadas a vários contextos de fala, sendo o tempo de tratamento entre 12 e 17 semanas. As imagens de USG foram utilizadas como apoio durante as sessões terapêuticas, sendo variável o tempo de estimulação, entre 10 e 40 minutos por sessão.

Adler-Bock et al. (2007); Bernhardt et al. (2008); Modha et al. (2008) também investigaram os efeitos de uma terapia utilizando a USG como biofeedback visual da produção da fala de crianças com distúrbios na fala, envolvendo sons de 'r'. Os participantes estavam em terapia para correção das alterações de fala, com distúrbios persistentes. A prática terapêutica incluía frequentemente a produção imitativa e espontânea do som, de forma isolada, em sílabas, palavras e frases, sendo que a USG contribuiu para avanços significativos em um menor período de tempo de tratamento, assim como o aumento na motivação e envolvimento dos participantes no processo terapêutico.

Klein et al. (2013) destacaram que as análises ultrassonográficas evidenciaram que a constrição na região da faringe é essencial para a produção correta de qualquer variante consonantal de 'r'. Hitchcock; Byun (2015) sugeriram que a generalização pode ser melhorada quando a intervenção utiliza um mecanismo de biofeedback visual para a aprendizagem dos sons. Cleland; Scobbie; Wrench (2015) verificaram que, após a intervenção, as crianças apresentaram

progressos significativos nos segmentos alvos do tratamento, evidenciados por mudanças na forma da língua.

Esta pesquisa buscou analisar os gestos articulatórios do som [r] tap em uma proposta de intervenção fonoaudiológica que utiliza as imagens de USG como estratégia terapêutica, sendo que também procurou analisar a evolução do tratamento em duas crianças com gravidade leve-moderado dos distúrbios dos sons da fala.

Métodos de estudo

Para este estudo, foi elaborada uma proposta de intervenção para crianças com distúrbios dos sons da fala (DSF), do tipo desvio fonológico, na qual o uso da ultrassonografia (USG) foi o recurso de apoio para a intervenção.

Trata-se de estudo do tipo caso-controle, com o objetivo de explicar variáveis causais em situações complexas, de caráter quantitativo e qualitativo.

As crianças foram selecionadas por conveniência, sendo do convívio/conhecimento dos pesquisadores, encaminhadas por escolas do Estado e/ou Município e/ou aquelas que aguardavam atendimento no Serviço de Atendimento Fonoaudiológico (SAF) de uma instituição de nível superior do sul do Brasil. Para participação no processo de intervenção foram selecionadas duas crianças com DSF, com idades cinco anos e dois meses e cinco anos e sete meses.

Crítérios de seleção

Os critérios de inclusão para a seleção das crianças participantes contemplaram a assinatura do Termo de Consentimento Livre e Esclarecido (TCLE) pelos responsáveis, bem como o assentimento da criança para participar da pesquisa (Resolução 466/2012). As crianças deveriam apresentar alterações no sistema de sons que comprometiam a produção do [r] tap, idades entre quatro e oito anos, limiares auditivos dentro do padrão de normalidade, não apresentar queixas auditivas no momento do exame, ausência de histórico de otites médias de repetição, além de estruturas e funções motoras orais adequadas.

As crianças com DSF também deveriam: ser membros de uma família de falantes monolíngues do Português Brasileiro (PB) da região central do Sul do Brasil; apresentar ausência de tratamento fonoaudiológico prévio; apresentar capacidades intelectuais adequadas para o desenvolvimento da linguagem falada; compreensão da linguagem falada apropriada à idade mental; não apresentar queixas socioambientais e emocionais relevantes ao desenvolvimento da fala/linguagem; apresentar capacidades de linguagem expressiva desenvolvidas em termos de abrangência de vocabulário e comprimento dos enunciados; frequentar o jardim de infância, pré-escola ou as primeiras séries do Ensino Fundamental; apresentar, no mínimo, omissão ou substituição do [r] *tap* em posição inicial de sílaba dentro da palavra (ISDP), [r] *tap* não-adquirido ou parcialmente-adquirido no sistema fonológico, de acordo com Bernhardt (1992). Além disso, foram consideradas crianças que apresentaram outras substituições ou omissões além de [r] *tap*.

Foram considerados critérios de exclusão para os participantes que apresentaram: alterações neurológicas e psicológicas consideradas pelo histórico de saúde e desenvolvimento neuropsicomotor; alterações auditivas (incluindo infecções temporárias repetitivas na orelha média); alterações nas estruturas e funções motoras orais; capacidades intelectuais inadequadas para a linguagem falada; falantes de outro dialeto que não o [r] *tap* falado na cidade na qual foi realizada a pesquisa; alterações na linguagem que comprometessem a abrangência de vocabulário e comprimento dos enunciados; além de crianças com o [r] *tap* adquirido.

Avaliações Fonoaudiológicas

Inicialmente os responsáveis pelas crianças foram entrevistados sobre dados de identificação, idades, históricos pré-natais e perinatal, desenvolvimento psicomotor, comportamento, desempenho escolar, desenvolvimento de fala, linguagem, voz, audição e condições gerais de saúde, sociais e afetivas.

Na avaliação da linguagem foi apresentada uma sequência lógica de quatro fatos fora de ordem (ilustração de um aniversário), utilizada no estudo de Vidor-Souza, Mota e Santos (2011). A criança deveria organizar e contar oralmente uma história, além de responder

questões a respeito das figuras. A história foi gravada e transcrita. Na avaliação do sistema estomatognático foi utilizado o Protocolo MBGR de Genaro et al. (2009), onde foram examinadas as estruturas e funções dos órgãos fonoarticulatórios.

A avaliação audiológica foi realizada com audiômetro portátil *Interacoustic Screening Audiometer*, modelo AS208, considerados limiares de via aérea a 20 dB em ambiente silencioso, nas frequências de 500 a 6000Hz. A partir dos resultados das avaliações iniciais, foram realizadas as avaliações fonológica e ultrassonográfica dos movimentos da língua, detalhadas a seguir.

Avaliação fonológica

Na avaliação da fala foi aplicado o instrumento INFONO - Instrumento de Avaliação Fonológica Computadorizada (CERON, 2015). Esse instrumento consta de 84 figuras que contém palavras-alvo que devem ser nomeadas pela criança. O instrumento avaliou todos os fonemas do PB, transcritos foneticamente.

O INFONO permite as análises dos inventários fonético e fonológico, bem como a análise das estruturas silábicas e encontros consonantais, mediante. Cada som foi considerado presente no inventário fonético quando ocorreu duas ou mais vezes, independentemente da posição na palavra. Os resultados foram conferidos por dois avaliadores experientes na área. A concordância entre observadores foi realizada por meio do Coeficiente de Tau-b de Kendall. A maioria dos itens analisados apresentou concordância boa (entre 60% e 80%) e excelente ($\geq 80\%$). De forma geral, a concordância entre os avaliadores foi boa, sendo a média de 76%.

A partir dessa análise, o sistema fonológico foi classificado de acordo Bernhardt (1992), sendo o som considerado: adquirido (com ocorrência adequada em 80% a 100% das possibilidades), parcialmente-adquirido (com 40% a 79%) e não-adquirido (com 0% a 39%).

Na avaliação da gravidade das alterações de fala foi utilizado o Percentual de Consoantes Corretas-Revisado (PCC-R), proposto por SHRIBERG et al. (1997), no qual foram considerados erros somente as substituições e as omissões dos sons, sem considerar as distorções de fala. Para a classificação da gravidade dos distúrbios dos sons da fala, foi considerado o índice de gravidade proposto por Shriberg e

Kwiatkowski (1982), sendo: leve (86 a 100%), leve-moderado (66% a 85%), moderado-grave (51% a 65%) e grave (<50%).

Análise ultrassonográfica dos movimentos da língua

As figuras que representaram as palavras-alvo foram previamente apresentadas aos participantes na tela do computador. A autora da pesquisa nomeou cada figura isoladamente e instruiu que a palavra correspondente à mesma deveria ser incluída na frase veículo: fala _____ "palavra" _____ de novo. Nas gravações, as crianças foram orientadas a sentar confortavelmente em uma cadeira no interior de uma cabine acústica, com a coluna e a cabeça eretas e os pés apoiados no chão. Os movimentos da língua obtidos pelo ultrassom foram registrados no computador, capturados e sincronizados no *software Articulate Assistant Advanced (AAA)* - (*Articulate Instruments, Ltd.*). Após, as crianças foram orientadas a falar as palavras que representavam as figuras. As figuras foram apresentadas na tela do computador, sendo possível visualizá-las através de um vidro transparente interno da cabine acústica.

Foram utilizados os seguintes equipamentos: microfone unidirecional (*Shure – SM48*) e ultrassom portátil modelo DP 6600 com transdutor endocavitário modelo 65EC10EA. Para melhor delineamento nas imagens que envolveram a produção de vogais altas, optou-se por ajustar o valor da frequência para 5.0 MHz. Durante a avaliação ultrassonográfica, a pesquisadora permaneceu dentro da cabine com o objetivo de auxiliar na memorização e a nomeação das figuras que deveriam ser inseridas na frase-veículo, enquanto outro examinador auxiliava na gravação dos dados no computador. Todos os participantes utilizaram o estabilizador de cabeça (*Articulate Instruments Ltd.*), o que permitiu a fixação da sonda na região submandibular durante a coleta.

Para a análise das imagens, optou-se pelo plano sagital/Modo B, pois este possibilitou a visualização da superfície do contorno da língua, desde a raiz até a ponta. Antes do início das gravações das listas de palavras, já com o estabilizador de cabeça, sonda e microfone posicionados, foi coletada e gravada uma imagem da cavidade oral com um gole d'água na boca, o que permitiu visualizar o palato de cada participante. Foram gravadas duas listas de palavras, considerando os

contextos vocálicos /a/, /i/ e /u/, analisados os sons [r] *tap* e seus pares [l] e [j] em palavras, totalizando 15 palavras, representadas em figuras.

Para fins de tabulação dos dados das avaliações ultrassonográficas para as análises estatísticas e apresentação em gráficos, as palavras foram classificadas em siglas, representadas entre parênteses (Tabela 1). A primeira lista de figuras consistiu na representação dos pares de palavras dissílabas em sílaba átona, nos contextos seguintes de /a/, /i/ e /u/, e a segunda lista de figuras consistiu na representação dos pares de palavras trissílabas em sílaba tônica no contexto seguinte de /a/, /i/ e /u/.

Tabela 1 – Palavras com [r], [l], [j] considerando os contextos vocálicos seguintes às vogais /a/, /i/, /u/.

Palavra	Som	Contexto
['kara] (CRA)	[r]	/a/
['kala] (CLA)	[l]	/a/
['kaja] (CIA)	[j]	/a/
['kari] (CRI)	[r]	/i/
['kali] (CLI)	[l]	/i/
['kaj] (CAI)	[j]	/i/
['karu] (CRO)	[r]	/u/
['kalu] (CLO)	[l]	/u/
['kaju] (CAO)	[j]	/u/
[ba' rata] (BRA)	[r]	/a/
[sa' lada] (SLA)	[l]	/a/
[fi' rida] (FRI)	[r]	/i/
[fa' lida] (FLI)	[l]	/i/
[ja' rutu] (CRU)	[r]	/u/
[kabe' ludu] (CLU)	[l]	/u/

Para cada palavra foram coletadas e analisadas seis repetições. Após as gravações dos movimentos da língua obtidos pela USG, foram realizadas análises dos contornos da superfície de língua no programa

que capturou e sincronizou áudio e vídeo (*Software AAA*), na visualização das imagens a partir do frame escolhido.

Inicialmente, a análise da espectrografia da fala permitiu demarcar o segmento a ser analisado em cada palavra, [r], [l] ou [j]. Na análise do segmento, foi observada simultaneamente à imagem dos movimentos da língua e selecionado como frame (imagem) o ponto de elevação máxima da ponta da língua em direção ao palato. A partir da escolha do frame, o contorno da língua entre as duas sombras (mandíbula e hioide), para cada repetição, foi então traçado manualmente por uma das pesquisadoras. Os recursos automáticos disponíveis no *software Best-fit* e *Snap-to-fit* também foram acionados na demarcação dos contornos da superfície da língua.

Após a seleção de todos os contornos da superfície da língua em cada palavra, para cada participante, os mesmos foram exportados para o recurso contido no próprio programa, denominado *Workspace*. Para o *Workspace*, foram exportados os 42 pontos que marcaram o contorno da superfície da língua em cada repetição de uma determinada palavra, denominados raios (*radius*). Os raios correspondem às distâncias desde o início do leque até os pontos que interceptam a língua (contorno da superfície da língua marcado). O leque é selecionado previamente às análises a partir do ângulo escolhido. Nesta pesquisa foi selecionado o ângulo de 120 graus, sendo esse o mais adequado para o ultrassom DP6600, disponível no próprio *Software* para seleção.

Os 42 pontos (0 até 41) aparecem no visor contido na área de trabalho (*Workspace*), os pontos válidos de uma única *spline* corresponderam, cada um, ao critério de 100% de confiança. Dessa forma, os valores válidos dos pontos entre 0 e 41, correspondentes ao contorno da superfície de língua, foram transferidos para uma planilha, através do recurso “copiar e colar”, para posterior análise estatística. A “marca” correspondente ao traçado do palato de cada participante também foi extraída da mesma forma.

As distâncias (0 a 41 pontos) correspondentes aos raios do ângulo escolhido, quanto os intervalos de confiança, em cada contorno da superfície de língua, foram exportados para a área de trabalho (*Workspace*) através dos recursos contidos no próprio *Software*, na opção *export – data*, ou solicitando-se a média, na área de trabalho, de cada curva de língua. Assim, para a planilha, foram exportados os

pontos de cada contorno de língua (sendo considerados válidos somente aqueles pontos que interceptaram a *spline*), em cada uma das seis repetições de cada palavra, e não a média deles. Após a delimitação das imagens de língua realizadas no *software* AAA, outro examinador, com experiência na visualização das imagens USG da língua, também analisou as imagens e indicou concordância ou discordância quanto ao frame de elevação máxima da ponta da língua, com resultado de 98,73% de concordância entre os avaliadores.

Análise quantitativa dos achados ultrassonográficos

Foram tabulados os pontos que interceptaram cada imagem dos contornos da língua na repetição de cada palavra. Os valores de cada um dos pontos foram organizados em planilhas. Após, todos os valores foram exportados para análise no ambiente R (*R Core Team*, 2015).

Em todas as figuras que apresentaram os contornos de língua, as linhas contínuas representaram os valores médios em milímetros da superfície da língua em cada ponto das *splines*. As linhas pontilhadas ao redor das linhas médias, também chamadas de bandas de confiança, foram obtidas através do cálculo de intervalos pontuais com 95% de confiança em torno de cada um dos pontos médios das *splines*.

Após essa etapa, foram também verificadas as regiões com diferenças significativas (p -valor $\leq 0,05$) para cada criança, através da análise do intervalo de 95% de confiança para a diferença média entre os contornos de língua. As regiões que apresentaram diferenças significativas foram aquelas que não contiveram a linha horizontal referente ao valor zero (representando a hipótese de diferença nula) contido dentro dos limites inferiores e superiores dos respectivos intervalos.

Apresentação da proposta de intervenção

Após as avaliações iniciais e duas crianças iniciaram o processo de intervenção, de acordo com o modelo de terapia proposto, detalhado a seguir. A terapia com o uso do ultrassom foi realizada utilizando-se figuras que representaram as palavras-alvo com o som [r] *tap* em ISDP, dividida em períodos. Cada período de intervenção se constituiu de duas sessões iniciais de avaliação, cinco etapas de intervenção e duas

sessões finais de avaliação. Para este estudo, cada etapa correspondeu a uma sessão de intervenção.

Desse modo, cinco sessões de intervenção, em cada período, foram distribuídas em duas sessões semanais, cada uma com duração de 45 minutos. Uma tela de computador com as imagens ultrassonográficas produzidas em tempo real pelo terapeuta e pela criança, de acordo com o objetivo da sessão, foi posicionada sobre uma mesa onde as atividades foram realizadas. As imagens ultrassonográficas da língua foram utilizadas com o mesmo aparelho detalhado nas sessões de avaliação inicial, porém sem o uso do estabilizador de cabeça para fixação da sonda. Nos cinco minutos iniciais e finais de cada sessão de intervenção, totalizando 10 minutos, foram apresentadas e reavaliadas, respectivamente, as produções das palavras-alvo e as atividades envolvendo os gestos articulatórios, dependendo dos objetivos em cada sessão.

Para esse registro, foram utilizadas duas fichas de evolução (Anexos 1,2 e 3), sendo: Ficha de Evolução 1 - na qual foi realizado o registro de produção correta/incorreta das palavras-alvo, mediante nomeação espontânea, sendo calculado o valor total referente ao percentual de produções corretas do som em cada sessão; e Ficha de Evolução 2 - na qual foram registrados os desempenhos nos objetivos propostos em cada sessão.

No preenchimento da Ficha de Evolução 2, após explicar a tarefa referente aos objetivos a serem alcançados na sessão, o terapeuta proporcionou três tentativas para que a criança demonstrasse sua evolução. Essa avaliação foi realizada no início e no final de cada sessão. Cada tentativa da criança foi pontuada, sendo: 0%, caso a criança não tivesse entendido e/ou cumprido adequadamente a atividade, 50%, caso a criança tivesse entendido e/ou cumprido a atividade parcialmente e 100%, caso a criança tivesse entendido e/ou cumprido a atividade totalmente. A média das três tentativas realizadas no início e no final para cada sessão indicou o desempenho no objetivo proposto para aquela sessão.

Cada sessão de intervenção correspondeu a uma etapa, sendo que a criança poderia progredir para a próxima etapa mediante a obtenção do desempenho nos objetivos propostos de, no mínimo, 50% em cada sessão, de acordo com a Ficha de Evolução 2. Durante 35 minutos entre os instantes iniciais e finais, foram abordados diferentes

objetivos e estratégias. As atividades com a USG foram demonstradas durante, aproximadamente, 10 minutos, quando utilizada na sessão. Todas as sessões envolveram as mesmas atividades para todas as crianças, de forma lúdica.

O som [r] foi considerado produzido corretamente quando alcançou a porcentagem mínima de 80% das produções nas sessões de avaliação final (AVF) de um determinado período. Para as avaliações fonológicas e USG foram utilizadas as palavras constantes na Tabela 1, diferentes das palavras utilizadas nas sessões terapêuticas. As palavras utilizadas nas sessões terapêuticas foram dissílabas ou trissílabas envolvendo o som [r] em ISDP, em diferentes contextos vocálico. Os participantes que persistiram com quaisquer alterações de fala após 20 sessões foram encaminhados para continuidade do tratamento no setor de Fonoaudiologia da própria instituição.

Princípios do modelo de intervenção

O modelo de terapia foi elaborado considerando os princípios da Fonologia Gestual (FREITAS, 2012). Esses princípios englobaram: (a) preocupação com o contraste, (b) produção e percepção, (c) conhecimento do funcionamento fonológico da criança, (d) trabalhar o contraste em diferentes contextos de uso de fala.

Objetivos das sessões de intervenção

As sessões de intervenção tiveram como objetivos: (1) proporcionar às crianças a conscientização da produção dos gestos articulatórios envolvidos na produção do som [r] *tap*, sendo: gesto de constrição na ponta da língua nos alvéolos e retração da raiz da língua, (2) auxiliar a criança a perceber a função contrastiva dos sons a partir de sua própria fala, no reconhecimento de pares mínimos com os sons [l] ou [j], (3) enfatizar o fato de que as mudanças na forma com que os sons são produzidos nos pares mínimos podem resultar em diferenças de sentido/significado, (4) demonstrar às crianças a percepção do som a partir da produção da fala do terapeuta, (5) proporcionar que a criança modifique os gestos articulatórios a partir de sua própria percepção e produção do som, observando também as produções do

terapeuta, (6) proporcionar atividades que utilizem a percepção e a produção dos sons em diferentes contextos de fala.

Estrutura das sessões de tratamento

Todos os períodos de intervenção constaram de nove sessões de 45 minutos cada, sendo duas sessões iniciais de avaliação, cinco sessões de intervenção direta do som [r] e duas sessões finais de avaliação. Nas sessões de avaliação, tanto iniciais quanto finais de cada período, foram realizadas avaliação fonológica e avaliação das imagens USG da língua.

Cada período envolveu, entre as sessões de avaliações iniciais e finais, cinco etapas de intervenção direta do som [r]. No início e no final de cada etapa foram utilizadas as fichas de evolução da fala da criança e de compreensão dos objetivos propostos.

Desse modo, os resultados das avaliações foram obtidos mediante o preenchimento da Ficha de Evolução 1, que demonstra a produção das palavras-alvo, e da Ficha de Evolução 2, que demonstra o desempenho nos objetivos propostos, ambas aplicadas no início e no final de cada sessão (Anexos 1, 2 e 3).

Primeiro período de intervenção

Primeira etapa de intervenção direta

A etapa de intervenção (sessão 3) direta do som-alvo [r] teve como objetivo a conscientização e significação dos gestos envolvidos na produção do som [r]. Nesse período foram utilizadas quatro figuras contendo o som-alvo [r] em diferentes contextos vocálicos. O terapeuta nomeou cada uma das figuras e solicitou que a criança também as nomeasse (Ficha de Evolução 1). O terapeuta demonstrou os gestos de língua para a criança, representados em massa de modelar e em um boneco (fantoche), detalhando a constrição da ponta da língua e retração da raiz da língua no som [r]. Foi solicitado à criança que também realizasse a tarefa de “modelar” e apresentasse no boneco os gestos de língua no som [r] (Ficha de Evolução 2).

Durante a sessão foi utilizada uma atividade lúdica como estratégia de produção das palavras-alvo contendo o som [r],

representadas por meio de figuras. O terapeuta lembrou a criança dos gestos envolvidos na produção do som [r] e solicitou que a criança também os produzisse. Nos minutos finais novamente foi avaliado o desempenho da criança na produção das palavras alvo e nos objetivos propostos (Fichas de Evolução 1 e 2).

Quando a criança apresentou, no mínimo, 50% de compreensão do objetivo proposto, iniciou-se a segunda etapa. Caso contrário, a criança repetiu a primeira etapa, utilizando-se outra estratégia.

Segunda etapa de intervenção direta

Essa etapa de intervenção (sessão 4) também envolveu a intervenção direta do som-alvo [r] contendo as mesmas quatro palavras-alvo da etapa anterior. O objetivo da sessão foi favorecer percepção dos gestos do som [r] a partir da produção da fala do “outro” (terapeuta) nas imagens USG.

Nos minutos iniciais e finais foram apresentadas novamente as figuras contendo as palavras-alvo e registradas as produções da criança. Nesses primeiros minutos, o terapeuta também mostrou a imagem de sua própria língua na USG para que a criança percebesse os gestos de língua no terapeuta na produção do som [r]. O terapeuta avaliou, nesses primeiros minutos, o reconhecimento da produção dos gestos de língua em si pela criança, mostrando, por vezes, o movimento da língua nas palavras com o [r] e em outro som, por exemplo, o [l], perguntando à criança qual gesto correspondia ao [r].

Durante essa etapa, em alguns momentos, a criança deveria observar a imagem da língua no terapeuta na produção das palavras contendo o som [r], obtido pela USG. A atividade lúdica consistiu em jogos infantis (associações figuras e números). Nos minutos finais novamente foi avaliado o desempenho da criança na sessão (Fichas de Evolução 1 e 2). Novamente, quando a criança apresentou, no mínimo, 50% de compreensão do objetivo proposto, iniciou-se a terceira etapa. Caso contrário, a criança repetiu a etapa utilizando-se outra estratégia.

Terceira etapa de intervenção direta

Essa etapa (sessão 5) com a intervenção direta do som-alvo [r] teve como objetivo auxiliar a criança a perceber a função contrastiva

dos sons a partir de sua própria fala e enfatizar o fato de que as mudanças na forma com que os sons são produzidos podem resultar em diferenças de sentido/significado. Nos minutos iniciais foram apresentadas novamente as figuras contendo as palavras-alvo com o som [r]. O terapeuta também mostrou os pares mínimos [l] e [j], formados com as mesmas quatro palavras-alvo que já estavam sendo trabalhadas nas etapas anteriores com som [r] e solicitou que a criança reconhecesse os sons nos pares de palavras, representadas em figuras. O desempenho na produção das palavras e no reconhecimento de pares mínimos foi novamente registrado nas fichas de evolução.

Durante essa etapa foram apresentados pares mínimos representados em figuras, associados a uma atividade lúdica. O terapeuta deveria destacar como a substituição de um som por outro pela criança poderia acarretar diferenças de significado. A atividade lúdica consistiu em: jogos de associação (animais e ambientes). Nos cinco minutos finais novamente foram registrados os desempenhos na produção das palavras-alvo e na atividade de reconhecimento dos pares mínimos. Os mesmos critérios das etapas anteriores foram considerados para progressão para as etapas seguintes.

Quarta etapa de intervenção direta

Essa etapa (sessão 6) com a intervenção direta do som-alvo [r] teve como objetivo proporcionar que a criança modifique os gestos articulatórios a partir de sua própria percepção e produção do som com apoio da USG, observando também as produções do terapeuta nas imagens de USG. Nos minutos iniciais foram avaliadas as produções das palavras-alvo, a percepção e a produção do gesto de língua pelo terapeuta e pela própria criança. Nesta sessão, a criança deveria identificar os gestos de língua nas imagens USG quando produzidas pelo terapeuta e nela mesma, detalhando a constrição da ponta e a retração da raiz da língua.

Durante essa etapa o terapeuta deveria auxiliar sobre as diferenças na produção dos diferentes jeitos de falar e a criança tentaria aproximar sua produção daquela considerada convencional. Terapeuta e criança deveriam produzir as palavras-alvo com e sem o auxílio das imagens dos movimentos da língua, obtidos pela USG. A

atividade lúdica consistiu em: desenhar e colorir as figuras contendo as palavras-alvo já impressas utilizando também materiais como folhas em branco, lápis de cor e canetas hidrográficas. Nos minutos finais, novamente, foram preenchidas as Fichas de Evolução 1 e 2. A criança deveria mostrar nas imagens USG os gestos de ponta de língua e retração de raiz de língua no terapeuta e também mostrar esses gestos tentando produzi-los na palavra.

Quinta etapa de intervenção direta

Essa etapa (sessão 7) com a intervenção direta do som-alvo [r] teve como objetivo proporcionar que a criança modifique os gestos articulatórios a partir de sua própria percepção e produção do som com apoio das imagens USG. Nos minutos iniciais foram preenchidas as fichas de avaliação das produções das palavras-alvo e o desempenho das tentativas de produção dos gestos de língua pela criança nas imagens ultrassonográficas, a criança deveria demonstrar os movimentos para a produção, detalhando a constrição da ponta da língua e retração da raiz da língua na produção do som [r]. Durante essa etapa a criança deveria modificar suas próprias produções, com auxílio das imagens dos movimentos da língua.

A atividade lúdica consistiu em: terapeuta contou uma história infantil com auxílio de um livro e a criança deveria incluir as figuras-alvo na estória, produzindo-as. A criança também poderia, em alguns momentos, desenhar os personagens da estória, ou fatos da estória. O desempenho da criança foi novamente registrado nas Fichas de Evolução 1 e 2.

Figura 1 – Primeiro período – Percepção e produção de sons com o apoio da USG

Sessão 1 (AVI)	Análise da fala	Nomeação espontânea	Avaliação fonológica
Sessão 2 (AVI)	Análise da fala	Avaliação USG	15 palavras
Sessão 3 Etapa 1	Intervenção direta 4 palavras-alvo	Conscientização Significação	Palavras
Sessão 4 Etapa 2	Intervenção direta 4 palavras-alvo	Percepção terapeuta USG	Palavras
Sessão 5 Etapa 3	Intervenção direta 4 palavras-alvo	Reconhecimento dos contrastes	Palavras
Sessão 6 Etapa 4	Intervenção direta 4 palavras-alvo	Percepção/produção USG terapeuta Percepção/produção USG criança	Palavras
Sessão 7 Etapa 5	Intervenção direta 4 palavras-alvo	Percepção/produção USG criança	Palavras
Sessão 8 (AVF)	Análise da fala	Nomeação espontânea	Avaliação fonológica
Sessão 9 (AVF)	Análise da fala	Avaliação USG	15 palavras

Legenda: AVI: avaliação inicial. AVF: avaliação final.

Segundo período de intervenção

Após a reavaliação da fala da criança (INFONO) no final do período anterior, caso o som [r] se apresentasse não-adquirido ou parcialmente-adquirido, a criança deveria iniciar o segundo período de intervenção.

Nesse período de intervenção foram utilizadas outras quatro palavras-alvo com o som-alvo [r] representadas em figuras, diferentes daquelas utilizadas no primeiro anterior. As estruturas das sessões e os objetivos foram os mesmos propostos no primeiro período, porém modificando as estratégias e enfatizando ainda mais os gestos de [r] nas produções realizadas pela criança nas palavras trabalhadas durante as sessões. Enquanto que no primeiro período, a criança poderia ser mais “passiva” e “observadora”, no segundo período a criança deveria “lembrar” dos gestos articulatórios realizados na própria língua na produção das novas palavras-alvo.

Em todas as sessões do segundo período de intervenção as crianças foram estimuladas a elevar a ponta da língua em direção ao rebordo alveolar e retraindo o dorso da língua nas suas produções. Os mesmos critérios mencionados no período anterior para progressão, atingindo no mínimo 50% de desempenho nos objetivos em cada etapa foram utilizados nesta etapa. Novamente, nos minutos iniciais e finais foram aplicadas as fichas de desempenho. Dependendo do desempenho no nível da palavra (mínimo 50%), na próxima sessão estimulavam-se as palavras em pequenas sentenças.

Primeira etapa de intervenção direta

Na primeira etapa (sessão 3) de intervenção direta do som-alvo [r] com o objetivo: conscientizar e estimular a produção dos gestos articulatórios envolvidos na produção do som foi utilizada como estratégia uma boca e uma língua em material de borracha para mostrar os gestos de língua. As figuras contendo o som-alvo foram produzidas pela criança, percebendo os movimentos na sua própria língua, em alguns momentos, a criança também mostrou os movimentos na língua de borracha. O terapeuta também utilizou uma folha com uma língua pontilhada desenhada, demonstrando, no desenho, os gestos de ponta e dorso da língua. A criança preencheu os pontos e produziu as palavras-alvo com o objetivo de reproduzir os mesmos gestos.

Segunda etapa de intervenção direta

Na segunda etapa (sessão 4) com intervenção direta do som-alvo [r] o objetivo foi favorecer a percepção dos gestos de língua do som [r] a partir da produção da fala do terapeuta via USG e propiciar tentativas de produção dos gestos articulatórios na própria criança. Nos minutos iniciais e finais foram aplicadas as fichas de desempenho (Fichas de Evolução 1 e 2).

Na estratégia o terapeuta fala as palavras com o apoio das figuras e, em alguns momentos, das imagens de USG, demonstrando os gestos da língua. A criança também tentou produzir os gestos em sua própria língua a partir da nomeação das figuras contendo as palavras-alvo. O terapeuta

auxiliou na produção correta do som, mostrando os gestos em si. Para esta sessão também foi utilizada uma atividade lúdica.

Terceira etapa de intervenção direta

Na terceira etapa (sessão 5) com intervenção direta do som-alvo [r] o objetivo da sessão foi auxiliar a criança a perceber a função contrastiva dos sons a partir de sua própria fala, e enfatizar o fato de que as mudanças na forma com que os sons são produzidos podem resultar em diferenças de sentido/significado, lembrando-se de tentar modificar os gestos articulatórios na própria língua durante a produção das palavras-alvo.

A estratégia consistiu em apresentar pares mínimos, sendo que a criança deveria visualizar as figuras e identificar as palavras correspondentes às mesmas. Novamente, o terapeuta deveria destacar como a substituição de um som por outro pela criança poderia acarretar mudanças de significado, também foram enfatizadas as produções do som-alvo pela criança.

Quarta etapa de intervenção direta

Na quarta etapa (sessão 6) com intervenção direta do som-alvo [r] o objetivo da sessão foi proporcionar que a criança modificasse os gestos articulatórios a partir de sua própria percepção e produção do som, observando também as produções do terapeuta, ambos com apoio da USG.

Na estratégia o terapeuta e criança deveriam produzir as palavras-alvo representadas nas figuras, com e sem o auxílio das imagens dos movimentos da língua obtidos pela USG. O terapeuta mostrou as figuras contendo as palavras-alvo e a criança produziu a palavra-alvo no contexto da atividade lúdica. Nessa sessão, como atividade lúdica, foi utilizado um jogo de encaixe e associação entre animais e ambientes.

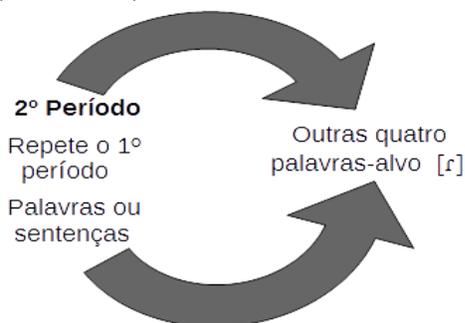
Quinta etapa de intervenção direta

Nessa etapa (sessão 6) com intervenção direta do som-alvo [r] o objetivo da sessão foi proporcionar que a criança modifique os gestos

articulatórios a partir de sua própria percepção e produção do som com e sem o apoio das imagens dos movimentos da língua obtidos pela USG.

Na estratégia o terapeuta apresenta uma estória e pede que a criança tente encaixar uma figura que contém a palavra-alvo no contexto da estória. Nesse momento, o terapeuta enfatiza a produção da palavra-alvo e a criança pode utilizar as imagens USG para produzir a palavra adequadamente, porém também deve lembrar-se de produzir os gestos de língua sem o apoio da USG. O terapeuta deve, novamente, relembrar os gestos articulatórios envolvidos para a produção do [r] na palavra.

Figura 2 – Segundo período –
Percepção e produção de sons com o apoio da USG



Terceiro período de intervenção

Após a reavaliação da fala, no período anterior, caso o som [r] se apresentasse não-adquirido ou parcialmente-adquirido, porém a criança apresentasse 50% de desempenho adequado nos objetivos propostos, iniciava-se o terceiro período. Caso a criança apresentasse 50% ou mais de produção do som [r] no inventário fonológico nas sessões finais de avaliação do período anterior, a mesma deveria associar duas ou mais palavras em vários momentos do processo de intervenção, procurando contextualizá-las a partir de frases simples, sendo as estratégias sempre associadas a uma atividade lúdica.

Caso a criança não apresentasse, no mínimo, 50% de aquisição do [r], o terapeuta utilizou sentenças e proporcionou contextualizações na sua própria fala, incluindo também aquelas que continham o som-

alvo [r], porém a criança foi estimulada modificar suas produções ainda no nível da palavra. Para isso, dez figuras que representavam o som-alvo [r], diferentes daquelas do período anterior, foram utilizadas para a intervenção nas sessões.

Nos primeiros cinco minutos iniciais e finais, em todas as sessões, novamente, foram aplicadas as fichas de evolução das produções do som-alvo em palavras (Ficha de Evolução 1). Nesse período a ficha de desempenho nos objetivos propostos (Ficha de Evolução 2) foi utilizada somente para aquelas crianças que não obtiveram, no mínimo, 80% de desempenho nos períodos anteriores. Nesse período, o objetivo da sessão foi proporcionar que a criança modificasse os gestos articulatórios a partir de sua própria percepção e produção do som, observando também as produções do terapeuta, quando necessário, ambos com apoio da USG. As imagens USG foram utilizadas em alguns instantes, não ultrapassando 10 minutos, podendo ser utilizadas no terapeuta ou na criança, dependendo do contexto de fala.

Para todas as etapas foram utilizadas diferentes estratégias: (1) apresentar cartelas com as figuras que representavam as palavras-alvo (a criança poderia pintar as cartelas e nomeá-las. A criança poderia criar desenhos para associar uma ou mais palavras-alvo. Formar diálogos com figuras, associar figuras aleatórias e figuras contendo o som-alvo, utilizar no diálogo preferencialmente frases simples, cada vez que o terapeuta ou a criança falarem uma palavra contendo o som-alvo, cada um deve assinalar na sua cartela um sinal. Nesse momento, o terapeuta mostra a produção adequada do som-alvo com ou sem o auxílio da USG e solicita que a criança também tente produzir). (2) utilizar trechos de histórias infantis, mostrar os personagens e estimular o diálogo, inserir no contexto de fala as figuras que representam as palavras-alvo, sendo utilizados jogos com sequência de fatos para a criança montar de acordo com a ordenação temporal. Estimular a inserção de figuras que contém as palavras-alvo na estória. O terapeuta inseria as figuras nas histórias e estimulava a formação de frases ou a produção isolada da palavra-alvo. Foram utilizados jogos de sequência de fatos e figuras que representavam as palavras-alvo, além de livros de desenhos com figuras para colorir e pintar e as figuras que representavam as palavras-alvo. O terapeuta associava diferentes figuras e estimulava a formação de frases simples

com as palavras-alvo, ou então estimulava a produção de uma palavra-alvo a partir do contexto.

Figura 3 – Terceiro período – Percepção e produção de sons com o apoio da USG

Sessão 1 (AVI)	Análise da fala	Nomeação espontânea	Avaliação Fonológica
Sessão 2 (AVI)	Análise da fala	Avaliação USG	15 palavras
Sessão 3 Etapa 1	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Sentenças
Sessão 4 Etapa 2	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Sentenças
Sessão 5 Etapa 3	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Sentenças
Sessão 6 Etapa 4	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Sentenças
Sessão 7 Etapa 5	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Sentenças
Sessão 8 (AVF)	Análise da fala	Nomeação espontânea	Avaliação Fonológica
Sessão 9 (AVF)	Análise da fala	Avaliação USG	15 palavras

Legenda: AVI: avaliação inicial. AVF: avaliação final

Quarto período de intervenção

Nesse período, para todas as etapas, o objetivo também foi constante: proporcionar que a criança modificasse os gestos articulatórios a partir de sua própria percepção e produção do som, observando também as produções do terapeuta, quando necessário, ambos com apoio da USG. Para isso, foram ampliadas atividades que utilizaram a percepção e a produção dos sons em diferentes contextos de fala, utilizando histórias. Os mesmos critérios para progressão para o terceiro período também foram utilizados para o quarto período de intervenção. No quarto período de intervenção, propiciou-se que as crianças realizassem associações entre as palavras-alvo e outras figuras em histórias infantis. Para isso, outras dez figuras que

representavam o som-alvo [r], diferentes daquelas do período anterior, foram utilizadas para a intervenção nas sessões.

A criança, com o auxílio do terapeuta, deveria associar palavras, procurando contextualizá-las em histórias. O terapeuta criou diálogos com associações da palavra no contexto da estória que estava sendo apresentada, incluindo aquelas figuras que continham o som-alvo [r]. As imagens USG foram utilizadas em alguns instantes, não ultrapassando 10 minutos, podendo ser utilizadas no terapeuta ou na criança, dependendo do contexto de fala. Nos primeiros cinco minutos iniciais e finais, em todas as sessões, novamente foram aplicadas as fichas de evolução das produções do som-alvo.

Nesse período, todas as crianças que participaram da pesquisa apresentaram desempenho de 80% ou mais nos objetivos no período anterior, por esse motivo a ficha de desempenho nas atividades propostas não foi aplicada. Foram utilizadas diferentes estratégias: (1) estimular a formação de histórias a partir de diferentes desenhos em livros infantis, inserir as figuras contendo as palavras-alvo no diálogo, contextualizando-as. Histórias infantis e figuras, inclusive aquelas contendo as palavras-alvo, (2) utilizar jogos com diferentes gravuras, associar as gravuras às figuras que contém as palavras-alvo, criando histórias e diálogos. Jogos e figuras contendo as palavras-alvo, (3) contar histórias enfatizando as palavras do livro, lembrando a produção dos gestos nas figuras que continham o som [r], incluído também as palavras-alvo, contextualizando-as, (4) utilizar jogo de associação com figuras, o terapeuta deve estimular a criação de uma estória a partir das associações, (5) associar as gravuras de jogos às figuras que continham as palavras-alvo.

Figura 4 – Quarto período – Percepção e produção de sons com o apoio da USG

Sessão 1 (AVI)	Análise da fala	Nomeação espontânea	Avaliação Fonológica
Sessão 2 (AVI)	Análise da fala	Avaliação USG	15 palavras
Sessão 3 Etapa 1	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Histórias
Sessão 4 Etapa 2	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Histórias
Sessão 5 Etapa 3	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Histórias
Sessão 6 Etapa 4	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Histórias
Sessão 7 Etapa 5	Intervenção direta 10 palavras-alvo	Percepção/produção USG terapeuta e criança	Histórias
Sessão 8 (AVF)	Análise da fala	Nomeação espontânea	Avaliação Fonológica
Sessão 9 (AVF)	Análise da fala	Avaliação USG	15 palavras

AVI: avaliação inicial. AVF: avaliação final.

Apresentação dos casos e resultados de intervenção

A Tabela 2 apresenta as duas crianças atípicas participantes do processo de intervenção, com idades entre cinco anos e dois meses e cinco anos e sete meses, ambas do sexo feminino. Quanto ao grau de gravidade das alterações na fala (PCC-R), as duas pertenciam ao grau leve-moderado. As crianças realizaram quatro períodos de tratamento (T), sendo: quatro períodos de intervenção em T1 (Terapia 1) e T2 (Terapia 2).

Tabela 2 – Idades, sexo, grau de gravidade e estimulabilidade de fala para o [r] das duas crianças participantes da intervenção

C	Idades	Sexo	PCC-R	Grau	Estimulável [r]
T1	05:02	F	79,06	Leve-Moderado	Não
T2	05:07	F	68,09	Leve-Moderado	Não

Legenda: C: criança, T: terapia, Idades: em anos e meses, PCC-R: Percentual de Consoantes Corretas Revisado, Grau: grau de gravidade das alterações na fala.

T1, do sexo feminino, cinco anos e dois meses, com grau de gravidade leve-moderado, apresentava a substituição de [r] por [l] e apresentava o som [r] não-adquirido no inventário fonológico, em coocorrência com outros sons não adquiridos sendo: [r] em final de sílaba dentro da palavra (FSDP) e em encontros consonantais. A Figura 5 demonstra o desempenho nos objetivos propostos e na produção das palavras-alvo em cada período e cada sessão de intervenção ao longo do processo terapêutico de T1.

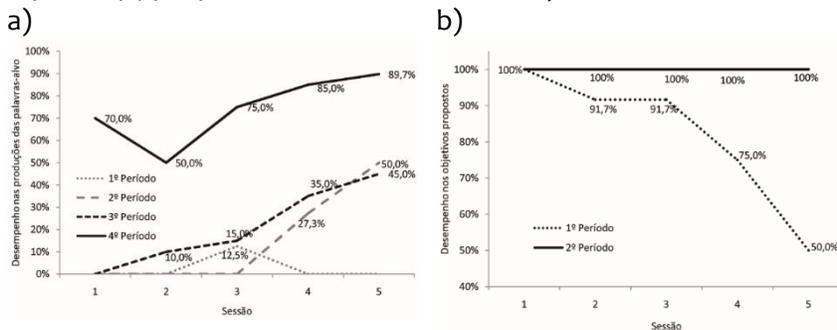
Foram realizados quatro períodos de intervenção no T1, sendo cinco sessões em cada período, totalizando 20 sessões. Em todas as sessões foi utilizado o som-alvo [r], iniciando no nível da palavra, com quatro palavras-alvo no primeiro e segundo períodos. No terceiro e quarto períodos foram utilizadas dez palavras, progredindo para sentenças e histórias infantis. Em cada período, diferentes palavras foram utilizadas.

Quanto às produções corretas das palavras-alvo em cada sessão, a partir do tratamento com o som-alvo [r], T1 alcançou 50 % de produções corretas do [r] na última sessão do segundo período. No terceiro período foram utilizadas outras dez figuras contendo o som-alvo [r], sendo as palavras contextualizadas em pequenas frases e a criança estimulada a falar o som-alvo em palavras simples e, caso acertasse os alvos durante a sessão, também foi estimulada a iniciar a produção com pequenas frases.

No quarto período foram utilizadas outras dez figuras contendo o som-alvo [r], sendo que T1 apresentou desempenho satisfatório na primeira sessão, com 70% de produção correta das palavras-alvo. O [r] foi adquirido pelo T1 no final desse período.

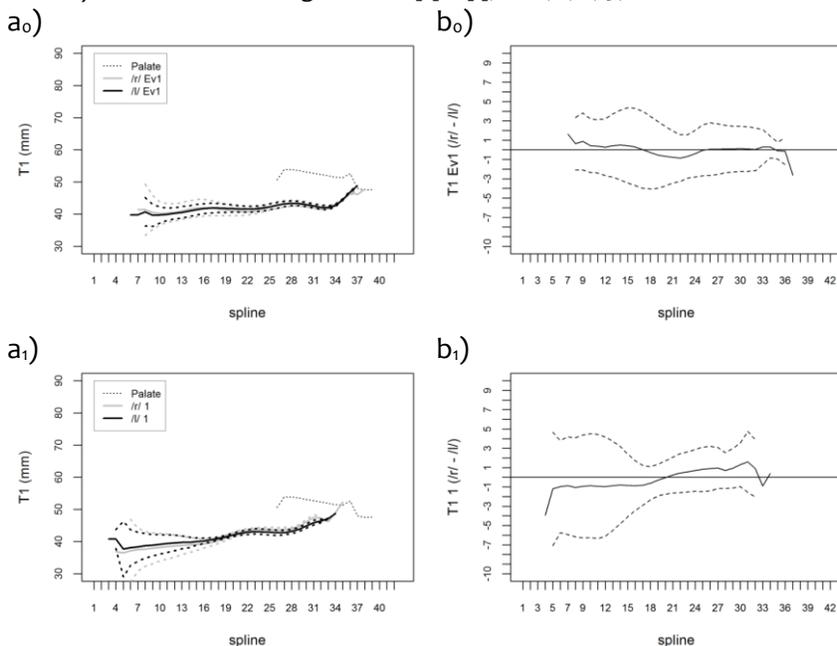
Quanto aos objetivos propostos em cada sessão, no primeiro período de intervenção verificou-se compreensão dos objetivos entre 50% e 100%, com 50% apenas na sessão que tinha como objetivo a produção dos gestos de língua pela criança com auxílio da USG. No segundo foram repetidos os mesmos objetivos do período anterior em cada sessão, porém utilizando-se outras estratégias, sendo que T1 atingiu 100% de desempenho na compreensão dos objetivos em todas as sessões. A partir do terceiro período o objetivo foi constante, o de produção dos gestos articulatórios no [r] pela criança e terapeuta, com apoio da USG na percepção e produção do gesto articulatório, com 100% de desempenho em todas as sessões.

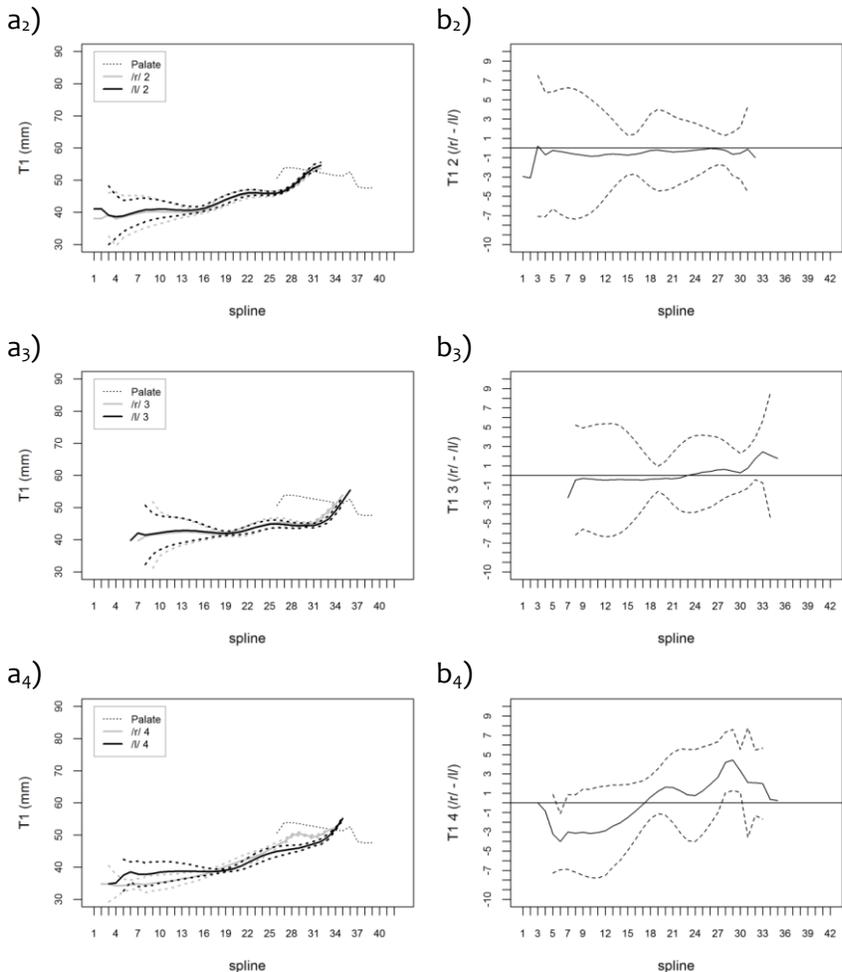
Figura 5 - Desempenho nas produções das palavras-alvo (a) e nos objetivos propostos (b) por períodos e sessões de intervenção do T1.



Os contornos médios de língua obtidos pela USG comparando-se [r] e [l] na avaliação inicial (0 = Ev1) e no final de cada período de intervenção (1, 2, 3 e 4) são apresentados na Figura 6.

Figura 6 – a_i) Contornos médios de língua para [r] em comparação a [l] na avaliação inicial (Ev1) e nos finais dos períodos de intervenção para T1; b_i) Diferenças das curvas de língua entre [r] e [l], i = 0, 1, 2, 3, 4.





No final de cada período de intervenção, observou-se progressiva coordenação entre o gesto de ponta em relação ao gesto de raiz de língua para a produção do [r], quando comparado ao [l] em cada momento das avaliações. As modificações nos gestos de língua, ao final de cada período no [r], quando comparado ao [l] demonstraram as tentativas de acertos gradientes na fala de T1.

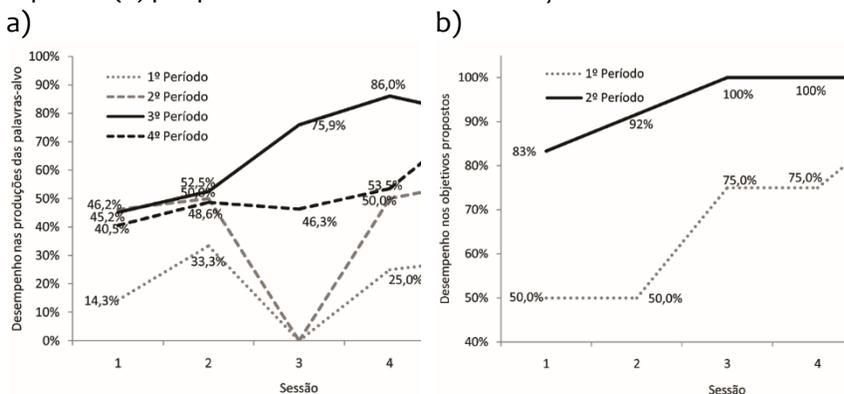
T₂, do sexo feminino, cinco anos e sete meses, com grau de gravidade leve-moderado, apresentava a substituição de [r] por [l] e apresentava o som [r] não-adquirido no inventário fonológico, em coocorrência com outros sons não-adquiridos, sendo: [z] e [ʒ]

(processo de dessonorização de fricativas), [r] em FSDP e encontros consonantais. No T3 também foram realizados quatro períodos de intervenção, utilizado o som-alvo [r], iniciando no nível da palavra, progredindo para sentenças e histórias.

Quanto às produções corretas do som [r] em cada sessão, T2 alcançou 50 % a partir da segunda sessão do segundo período de tratamento. No terceiro período houve melhora, alcançando 86 % e 79,1% de produções corretas nas duas últimas sessões desse período. No quarto período [r] foi produzido com 79,5% de acertos na última sessão e foi adquirido na avaliação final com 80%.

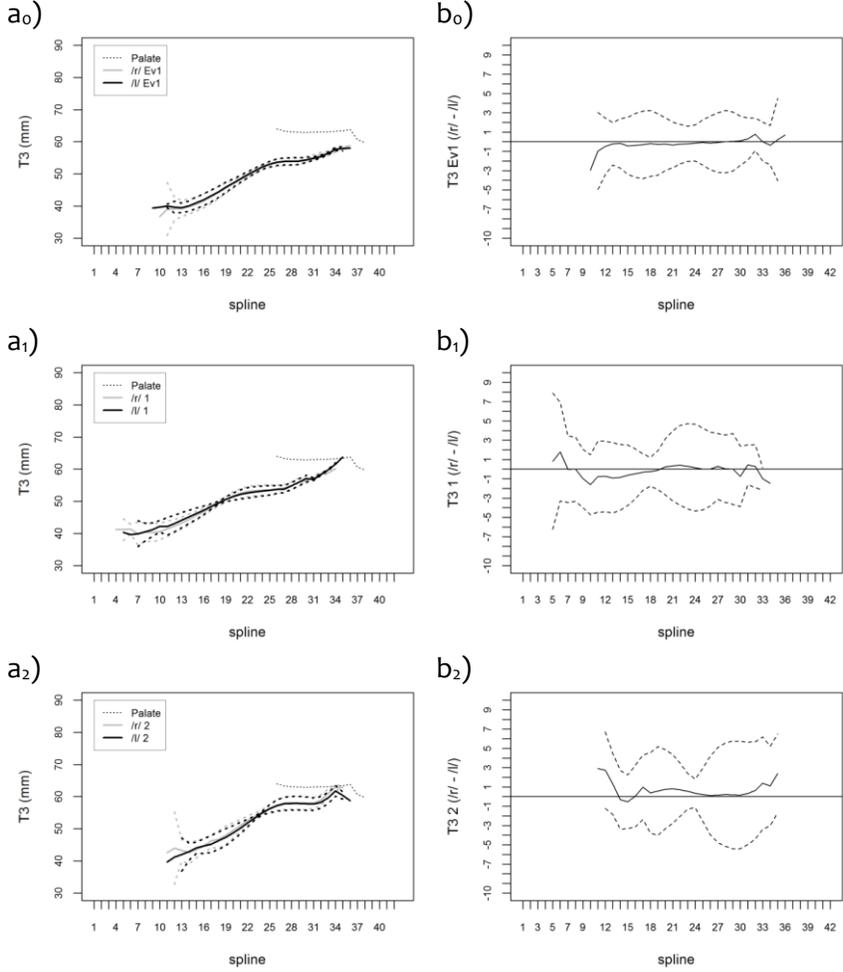
Quanto aos objetivos propostos, no segundo período, todas as tarefas foram cumpridas adequadamente, alcançando valores acima de 80%. No terceiro e quarto períodos as atividades de reconhecimento dos gestos articulatórios no terapeuta e na produção dos gestos pelo terapeuta e pela criança (USG) atingiram 100%.

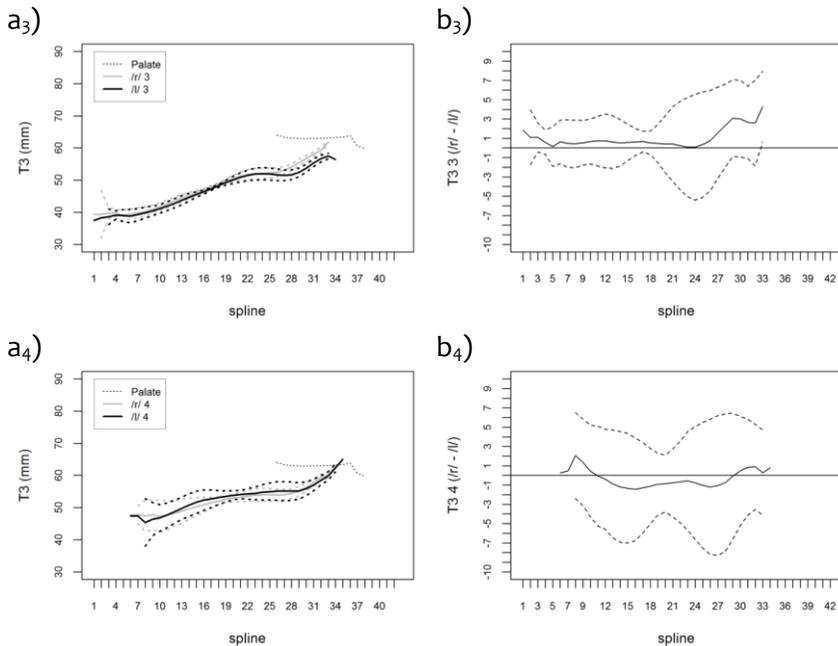
Figura 9 – Desempenho nas produções das palavras-alvo (a) e nos objetivos propostos (b) por períodos e sessões de intervenção do T2.



Quanto aos contornos médios de língua, no final de cada período de intervenção, observou-se também progressiva coordenação entre o gesto de ponta em relação ao gesto de raiz de língua para a produção do [r] quando comparado ao [l]. As modificações nos gestos de ponta de língua, ao final de cada período no [r] demonstraram as tentativas de acertos gradientes na fala de T2, embora, na análise perceptivo-auditiva, [r] tenha se tornado adquirido somente no final do terceiro período de intervenção.

Figura 10 – a_i) Contornos médios de língua para [r] em comparação a [l] na avaliação inicial (Ev1) e nos finais dos períodos de intervenção para T₂; b_i) Intervalos de confiança para diferenças das curvas de língua entre [r] e [l] i = 0, 1, 2, 3, 4.





Importante considerar que T1 e T2 apresentaram evoluções muito semelhantes, com substituições de [r] por [r]. T1 e T2 exigiram o tratamento mais prolongado a fim de que alcançassem refinamentos dos gestos de ponta e raiz de língua, necessários à produção do [r], menor diferenciação dos gestos articulatórios de [r] em comparação a [l] foi observada na avaliação inicial, sugerindo que esses sujeitos apresentavam maiores defasagens na habilidade em refinar e conseqüentemente diferenciar os duplos gestos de língua no [r] ao longo do tratamento.

As crianças não eram estimuláveis à produção do [r], o que poderia justificar maior dificuldade em diferenciar os gestos de língua, com um maior número de sessões para desenvolver a habilidade em refinar os duplos gestos de língua necessários à aquisição desse som. De acordo com Castro; Wertzner (2009), a estimulabilidade de fala reflete a habilidade em imitar corretamente um determinado som quando fornecida a orientação para olhar e ouvir o examinador, conseqüentemente detecta se a criança é capaz de produzir os sons ausentes de seu inventário fonético e está relacionada à execução dos diferentes gestos necessários à produção dos sons da língua.

A progressão na realização dos gestos articulatórios de [r] pela USG revelaram as possibilidades de acertos gradientes nas produções de fala. Desse modo, as substituições observadas como categóricas pela análise perceptivo-auditiva, revelaram-se gradientes na análise dos gestos articulatórios captados pela USG. Dentre esses processos gradientes, destacam-se os contrastes encobertos referidos em pesquisas (GIBBON, 1999; SCOBIE et al., 2000; HEWLETT; WATERS, 2004; BERTI, 2006; FREITAS, 2012; RODRIGUES, 2007). Os contrastes encobertos da fala revelados pela USG permitem considerar a construção do sistema de sons pela criança, demonstrando o conhecimento linguístico e as possíveis estratégias utilizadas na comunicação.

Estudos com a USG da fala têm utilizado como recurso terapêutico o uso de imagens da língua no tratamento dos distúrbios dos sons da fala e proporcionaram rápido aumento na produção adequada dos sons, especificamente quando envolveram os sons de /r/ (BERNHARDT et al., 2008; GICK et al., 2008; MODHA et al. 2008; ADLER-BOCK et al., 2007; BACSFALVI, 2010; KLEIN et al., 2013; CLELAND; SCOOBIE; e WRENCH, 2015).

O uso da USG nas análises de fala e como recurso terapêutico, considerando o gesto articulatório como unidade de ação e representação necessária para que a tarefa de realização de [r] se concretizasse, possibilitou desvendar detalhes existentes na fala da criança e que estavam em desenvolvimento, tanto na avaliação inicial, como durante o processo terapêutico. Para o PB, os achados de alguns estudos (ALBANO, 2001; BERTI, 2006; FREITAS; ALBANO, 2012; MELO et al., 2016; e RODRIGUES et al., 2008) permitiram supor implicações para o tratamento, indo além do diagnóstico e metodologia de avaliação.

Neste estudo, o biofeedback visual proporcionado pela USG durante o tratamento proporcionou a apreensão dos gestos articulatórios envolvidos na produção do [r] para as crianças participantes do processo de intervenção, facilitando a aquisição desse segmento entre em 20 sessões de tratamento.

Conclusão

Esta proposta de tratamento que utiliza a ultrassonografia dos movimentos da língua como facilitadora da percepção e produção dos gestos articulatórios proporciona a aquisição do som [r] foi eficaz para

as crianças que participaram do processo de intervenção. As crianças com distúrbios dos sons da fala desta pesquisa apresentam acertos gradientes tanto nas substituições de [r] por [l] ou [r] por [j], bem como nas omissões de [r], revelados pela ultrassonografia dos movimentos da língua.

A ultrassonografia utilizada no processo terapêutico como um método de biofeedback visual e proprioceptivo dos gestos articulatórios contribui para a aquisição do som [r], bem como incentiva e envolve os participantes durante a intervenção.

O uso das imagens captadas pela ultrassonografia dos movimentos da língua propicia a apreensão dos gestos articulatórios do som [r] e permite avaliar progressos durante o processo de intervenção. A aplicação de métodos quantitativos para explorar as curvas de língua dos indivíduos possibilita a identificação de regiões que demonstram diferenças entre os gestos articulatórios.

O modelo proposto neste estudo pretende trazer contribuições que podem ser adotadas como recursos na pesquisa e clínica fonoaudiológicas, sob a inspiração de uma teoria dinâmica de abordagem da fala e a ultrassonografia dos movimentos da língua como instrumento promissor utilizado no processo de intervenção.

Agradecimentos

Esta pesquisa teve incentivo financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Ministério da Educação, sob o número 483722/2013-8.

Os autores agradecem à CAPES e às crianças e suas famílias, que generosamente aceitaram participar deste estudo.

Anexo 1 - Ficha de evolução do desempenho na produção das palavras-alvo (para 4 palavras)

Nome da criança: _____

Terapeuta: _____

Etapa: _____

Data: _____

PALAVRA-ALVO	Produção (n)		Número total de produção nas palavras-alvo	Percentual de Produção Correta por palavra-alvo	
	C	I		C	%
	C			C	_____ %
	I				
	C			C	_____ %
	I				
	C			C	_____ %
	I				
	C			C	_____ %
	I				
Produções do alvo			n	%	
Total de produções do alvo					
Total de produções corretas (C) do alvo					
Total de produções incorretas (I) do alvo					

Anexo 2 - Ficha de evolução do desempenho na produção das palavras-alvo (para 10 palavras)

Nome da criança: _____

Terapeuta: _____

Etapa: _____

Data: _____

PALAVRA -ALVO	Produção (n)		Número total de produção nas palavras-alvo	Percentual de Produção Correta por palavra-alvo	
	C	I		C	%
	C			C	%
	I				
	C			C	%
	I				
	C			C	%
	I				
	C			C	%
	I				
	C			C	%
	I				
	C			C	%
	I				
	C			C	%
	I				
	C			C	%
	I				
	C			C	%
	I				
Produções do alvo			n	%	
Total de produções do alvo					
Total de produções corretas (C) do alvo					
Total de produções incorretas (I) do alvo					

Anexo 3 – Ficha de evolução do desempenho nos objetivos propostos

Nome da criança: _____

Sessão:Sessão:Sessão:Sessão:Sessão:

Data:Data:Data:Data:Data:

Objetivo:Objetivo:Objetivo:Objetivo:Objetivo:

ATIVIDADE	AVI	AVF	%
Tentativa1			
Tentativa2			
Tentativa3			
Escala	Resultado	%	
1	Não consegue nada	0	
2	Consegue parcial	50	
3	Consegue totalmente	100	
TOTAL/ MÉDIA			

ATIVIDADE	AVI	AVF	%
Tentativa1			
Tentativa2			
Tentativa3			
Escala	Resultado	%	
1	Não consegue nada	0	
2	Consegue parcial	50	
3	Consegue totalmente	100	
TOTAL/ MÉDIA			

ATIVIDADE	AVI	AVF	%
Tentativa 1			
Tentativa 2			
Tentativa 3			
Escala	Resultado	%	
1	Não consegue nada	0	
2	Consegue parcial	50	
3	Consegue totalmente	100	
TOTAL/ MÉDIA			

ATIVIDADE	AVI	AVF	%
Tentativa1			
Tentativa2			
Tentativa3			
Escala	Resultado	%	
1	Não consegue nada	0	
2	Consegue parcial	50	
3	Consegue totalmente	100	
TOTAL/ MÉDIA			

ATIVIDADE	AVI	AVF	%
Tentativa1			
Tentativa2			
Tentativa3			
Escala	Resultado	%	
1	Não consegue nada	0	
2	Consegue parcial	50	
3	Consegue totalmente	100	
TOTAL/ MÉDIA			

AVI: avaliação no início da sessão. AVF: avaliação no final da sessão

ULTRASSONOGRRAFIA DE LÍNGUA NO DESVIO FONOLÓGICO: DA AVALIAÇÃO À TERAPIA

Aline Mara de Oliveira
Rafaela Duarte Silva
Izabel Christine Seara

Introdução

No presente capítulo, apresentamos a ultrassonografia de língua, a partir de um programa de tratamento fonológico, associando o uso do *biofeedback* visual por meio da ultrassonografia de língua à terapia fonológica tradicional. Vamos explicitar as vantagens e desvantagens desse instrumento na análise de fala para aplicação na área de aquisição típica e atípica de fala. Para descrição de exemplos de suas aplicações, apresentaremos um estudo de caso, desenvolvido por Silva (2020), de caráter longitudinal, realizado em uma criança falante do português brasileiro com diagnóstico de desvio fonológico. Nesse estudo de caso, será observada a efetividade desse tipo de programa. O objetivo deste relato de caso é descrever os resultados da terapia fonoaudiológica, associada ao *biofeedback* ultrassonográfico no tratamento de uma criança com desvio fonológico (DF), comparando a produção das fricativas /s/ e /ʃ/ pré e pós terapia.

A ultrassonografia de língua (USG) pode ser utilizada para fornecer a imagem em tempo real da configuração e do movimento de língua para crianças com transtornos dos sons da fala (TSF) (LIMA et al., 2018). Além disso, tem proporcionado uma compreensão refinada e fidedigna de dados típicos e atípicos de fala (MELO et al. 2016). Há evidências quanto à eficácia no tratamento fonoaudiológico associado ao uso do USG quando comparado à terapia de abordagem tradicional, ou seja, aquela que não faz uso de um aparato instrumental (SUGDEN et al., 2019). A USG surgiu com o objetivo de oferecer melhores condições de avaliação, permitindo a realização de *biofeedbacks* na terapia fonoaudiológica (MELO et al. 2016), possibilitando realizar a detecção e o tratamento de erros dos sons da fala, tornando-se uma ferramenta com a qual a criança pode reconhecer e ajustar suas

produções de fala (PRESTON, BRICK e LANDI, 2013; PRESTON, LEECE e MAAS, 2017). Desse modo, as imagens captadas pelo ultrassom acabam por fornecer à terapia dos desvios fonológicos uma melhor compreensão e eficiência (MELO *et al.*, 2016).

Estudos (MELO *et al.*, 2016; e PRESTON, LEECE e MAAS, 2016), dentre outros) têm indicado que a terapia associada ao uso da ultrassonografia de língua como *biofeedback* visual pode fornecer ao terapeuta e ao paciente dados do movimento de língua, resultando na aquisição de padrões motores mais precisos e na melhoria da articulação, com níveis variados de generalização (correta produção de sons não tratados em terapia) e retenção (permanência da produção correta ao longo do tempo), beneficiando assim a produção da fala esperada. Além disso, há evidências de que a terapia fornecida com sessões frequentes, ou seja, de forma intensiva, pode revelar resultados superiores se comparada à abordagem tradicional (PRESTON, LEECE e MAAS, 2017).

As pesquisas pautadas na USG com produções de fala típica e atípica são realizadas, primeiramente, a partir de uma análise qualitativa, elaborada por meio de inspeção visual das imagens de ultrassom com o objetivo de averiguar a qualidade da imagem e identificar as partes da língua. Em seguida, parte-se para uma análise quantitativa, realizada por meio de *softwares* que permitem um exame minucioso dos gestos articulatórios, não apenas daqueles produzidos adequadamente, mas, em especial, dos erros de produção. As análises quantitativas geram valores numéricos oriundos do contorno de língua ou de parâmetros articulatórios como, por exemplo, distância e área de regiões da língua. Os valores extraídos, a depender do interesse clínico ou acadêmico, são analisados por meio de um *design* estatístico específico, tornando possível a comparação entre crianças do mesmo grupo clínico, de uma mesma criança no momento pré e pós terapia, dentre outras possibilidades.

Ao final deste capítulo, indicaremos também perspectivas futuras de estudos e terapias envolvendo a análise ultrassonográfica de fala típica e com desvios fonológicos. Apresentada a ultrassonografia de língua, passemos então à descrição das vantagens e das desvantagens de seu uso.

Ultrassonografia de língua: vantagens e desvantagens para a aplicação na área de aquisição típica e atípica de fala

O uso das imagens de ultrassom de língua no contexto clínico e no contexto de pesquisa vem sendo ampliado e tem evidenciado vantagens, tanto no que se refere à avaliação, quanto na intervenção de indivíduos com transtornos dos sons da fala (TSF) (PRESTON *et al.*, 2017). Isso porque, dentre as metodologias articulatórias, a ultrassonografia de língua (USG) tem apresentado melhor custo/benefício, uma vez que apresenta baixo custo, maior portabilidade e facilidade para manusear se comparado às técnicas de Eletropalatografia ou de Articulografia Eletromagnética.

A USG é um procedimento não invasivo, que leva pouco tempo para a preparação das imagens de ultrassom, sendo possível visualizar o contorno da língua da ponta até a raiz (PRESTON *et al.*, 2017). É possível observar os movimentos intra-orais de modo estático e dinâmico durante a produção motora da fala (WIETHAN *et al.*, 2015; MELO *et al.*, 2016; CLELAND, SCOBIE, WRENCH, 2015). A técnica consiste em colocar a sonda do aparelho de ultrassom na parte submandibular, permitindo capturar em tempo real o contorno da língua tanto no plano médio-sagital, quanto no coronal (CLELAND, SCOBIE e WRENCH, 2015). No que diz respeito ao tratamento, estudos com crianças com TSF persistentes ou erros residuais de fala (BACSFALVI e BERNHARDT, 2011; PRESTON, BRICK e LANDI, 2013; PRESTON *et al.*, 2014, 2017; PRESTON; LEECE e MAAS, 2016) têm mostrado que o *biofeedback* visual, por meio da USG, durante a terapia de fala tem alcançado melhores prognósticos terapêuticos (CLELAND, SCOBIE e WRENCH, 2015).

Uma pesquisa (SUDGEN *et al.*, 2019) que investigava evidências sobre a ultrassonografia no processo de intervenção para desenvolvimento de crianças com TSF mostrou que o *biofeedback* por meio da ultrassonografia pode ser eficaz para indivíduos quando usado nos estágios iniciais de aprendizagem motora, mas pode ser menos eficaz para promover a generalização de sons ou palavras não tratadas. Isso significa dizer que essa intervenção mostrou resultados positivos nos estágios iniciais da intervenção quando comparados com a abordagem tradicional, podendo facilitar a aquisição dos sons tratados, porém nem sempre a generalização é alcançada com sucesso. Um estudo recente

com crianças com TSF com implante coclear (GIBSON e LEE, 2020) indicou que o *biofeedback* visual por USG pode ser um meio de intervenção eficaz de tratamento desses transtornos.

O estudo de Preston, Brick e Landi (2013) com *biofeedback* por ultrassom, realizado com seis crianças com TSF, especificamente a apraxia de fala na infância (AFI), demonstrou maior precisão das sequências sonoras no nível da palavra e alguns participantes apresentaram generalização para sons-alvo não tratados. Além disso, dados de acompanhamento longitudinal de dois meses indicaram que geralmente os ganhos obtidos com este tratamento eram mantidos. Todos os participantes alcançaram o critério pré-estabelecido de 80% no nível da palavra em duas sessões consecutivas. O critério de desempenho foi atendido em 23 das 31 sequências-alvo incluídas no tratamento. O tempo médio para atingir esse critério foi de cinco sessões.

No entanto, a USG pode apresentar desvantagens previstas na técnica das imagens de ultrassom por conta de suas limitações. Uma delas é que apenas a língua pode ser visualizada, o que significa dizer que as estruturas como o palato duro, o palato mole e as paredes da faringe não são visíveis. Outro cuidado importante é a posição da sonda no queixo do participante da pesquisa, que não pode estar desalinhada no eixo médio, pois provocaria erros de interpretação. Além disso, nem todos os participantes suportam o uso de gel que deve ser aplicado no queixo para a captura das imagens de ultrassom (PRESTON *et al.*, 2017).

Apresentamos, na sequência, um estudo de caso que exemplifica o uso da USG como terapia de fala.

Programa de tratamento fonoaudiológico com apoio da USG

A ultrassonografia de língua pode ser utilizada para avaliar a produção de fala de qualquer criança com desenvolvimento de fala típico e ou com TSF, considerando alguns requisitos. O primeiro deles é que as crianças consigam cooperar com o procedimento de gravação, pois, em decorrência do uso do capacete estabilizador da sonda do ultrassom, a experiência tem mostrado que, para a avaliação, é recomendado crianças acima de quatro anos de idade. Já, para a intervenção terapêutica, a idade mínima seria de cinco anos, uma vez que essa terapia exige atenção e motivação da criança. Vale ressaltar

que é possível realizar avaliação e intervenção em crianças de outras faixas etárias e com comorbidades clínicas, como transtornos do espectro autista e com dificuldades de aprendizagem de graus mais leves. No entanto, para crianças com distúrbios de linguagem associadas às comorbidades de graus moderados a graves, o progresso terapêutico pode ser mais lento, portanto, o terapeuta deverá se preocupar em oferecer mais sessões (CLELAND *et al.* 2019).

Considerando estes critérios, o caso clínico aqui relatado é de um menino, com 7 anos e 3 meses, diagnosticado com desvio fonológico, e monolíngue do português brasileiro (SILVA, 2020). Para ser considerado como participante da pesquisa, a criança foi submetida à anamnese e à avaliação fonoaudiológica (linguagem compreensiva e expressiva oral, sistema estomatognático, exame articulatório, avaliação da fala por meio do estabelecimento dos inventários fonético e fonológico) e exames complementares (avaliação otorrinolaringológica, audiológica, neurológica e processamento auditivo central), para que se verificasse outras alterações que pudessem excluir o diagnóstico de desvio fonológico.

Feitas as considerações gerais para a aplicação desta técnica com crianças em aquisição da linguagem, vamos apresentar as consoantes que foram alvo do estudo de caso.

As consoantes focalizadas no tratamento

As imagens de ultrassom permitem visualizar todos os fonemas linguais. Assim, no relato de caso, propomo-nos a uma intervenção com as fricativas /ʃ/ e /s/. Análises de ultrassonografia de língua mostram que a consoante fricativa /ʃ/ é realizada por adultos com a elevação da língua acompanhada de uma curvatura do dorso, gerando assim uma turbulência palatina. Já a consoante fricativa /s/ é caracterizada com quatro padrões distintos: (1) ausência de curvatura acentuada do dorso ou raiz da língua; (2) presença de curvatura acentuada ascendente da raiz e do dorso da língua; (3) presença de curvatura acentuada descendente da raiz e do dorso da língua; (4) presença de curvatura acentuada e anteriorização do dorso da língua (FRANCISCO e WERTZNER, 2017 *apud* LIMA *et al.*, 2018). Acusticamente, as fricativas recebem esse nome pelo fato de o fluxo de ar percorrer um espaço de grande estreitamento na

cavidade oral, sendo suficiente para produzir uma turbulência aérea (BRASIL *et al.*, 2012).

A consoante fricativa /s/ é considerada forte e aguda, alcançando uma frequência de 4.500Hz, atingindo a faixa de 8.000Hz, enquanto a fricativa surda /ʃ/ é percebida como mais grave, variando entre a faixa de frequência de 2.500 e 6.000Hz (RUSSO e BEHLAU, 1993). No que se refere aos traços distintivos, o fonema /s/ é caracterizado como [coronal, +anterior], enquanto /ʃ/ como [coronal, -anterior] (BRASIL, *et al.*, 2012).

As crianças brasileiras com desenvolvimento típico tendem a adquirir as fricativas com 2 anos e 10 meses. Mais especificamente, /s/ e /ʃ/ são adquiridos, em média, aos 2 anos e 6 meses e aos 2 anos e 10 meses, respectivamente (BRASIL, *et al.*, 2012). Entretanto, se a criança não alcançar o contraste fonológico envolvendo as fricativas como os seus pares, será considerada uma criança com desvio fonológico, sendo, portanto, uma candidata ao uso da intervenção fonoaudiológica (WIETHAN e MOTA, 2012).

O desvio fonológico é caracterizado pela dificuldade na organização dos sons da fala, envolvendo o sistema fonológico. Crianças acometidas de desvio fonológico costumam lançar mão de estratégias de reparo por um período de tempo maior quando comparadas àquelas sem alterações fonológicas (MELO *et al.*, 2018; VASSOLER e BERTI, 2015). As estratégias de reparo ou processos fonológicos são caracterizadas pela substituição de uma classe de sons de difícil produção por outra idêntica, porém destituída de propriedades difíceis. Dentre os processos de substituição, tem-se o processo de anteriorização de fricativa, quando ocorre a substituição de um fonema produzido na região posterior por um som produzido na região anterior da cavidade oral, como, por exemplo, /ʃ/ → [s]. Já, no processo de posteriorização, ocorre o oposto, a substituição de um fonema produzido na região anterior por um som produzido na região posterior do trato oral, como, por exemplo, /s/ → [ʃ]. Do ponto de vista articulatório, a anteriorização ou posteriorização de fricativa pode ser uma consequência da dificuldade de elevação da língua, somada ao fechamento insuficiente para constrição do fluxo aéreo, resultando, assim, em dificuldades articulatórias e/ou fonológicas (CIELO e CASARIN, 2008; e BRASIL, *et al.*, 2012).

Por meio de uma análise de cluster (agrupamento), Francisco e Wertzner (2017) investigaram os contornos de língua para /ʃ/ e /s/ de

adultos, crianças com desenvolvimento fonológico típico e crianças com desenvolvimento fonológico atípico, falantes do português brasileiro. Essa análise mostrou semelhanças entre os contornos de adultos e de crianças com desenvolvimento típico, ou seja, foi possível identificar o comportamento do contorno de língua para /s/ como mais elevado na região da raiz da língua com abaixamento na região anterior da língua (ponta/lâmina da língua). Na produção do alvo /ʃ/, a região posterior da língua (dorso) apresentou uma curvatura na região do palato. No entanto, o contorno de língua das crianças com desvio fonológico mostrou-se semelhante para /ʃ/ e /s/, ou seja, a análise estatística não diferenciou os contornos da língua para /s/ e /ʃ/. Isso significa dizer que, nesse caso, houve indiferenciação gestual entre /ʃ/ e /s/.

Caracterizados os sons-alvo do programa terapêutico e as limitações que podem ocorrer em suas produções, passamos às estratégias metodológicas empregadas neste estudo de caso.

Procedimentos metodológicos

A terapia fonoaudiológica aqui proposta como exemplo de aplicação da USG foi realizada por Silva (2020). Esse programa de tratamento foi dividido em oito sessões com *feedback* visual por ultrassom da língua (com duração de 20 minutos) associado à abordagem tradicional (com duração de 40 minutos) durante quatro semanas. O tratamento consistiu em promover a produção e percepção das fricativas /ʃ/ e /s/, buscando acurácia da produção, bem como generalização e retenção. Na pré-terapia, a gravidade do desvio, apresentado pelo participante da pesquisa, foi classificada como grau leve a moderado.

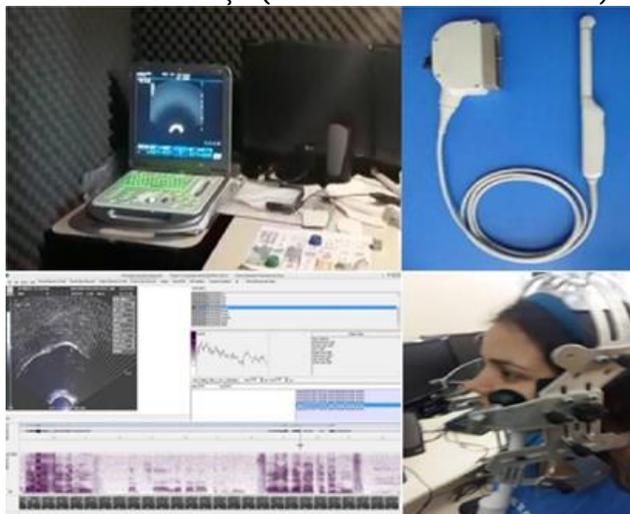
O estudo de caso, exemplificado neste estudo, trata de um relato de caso clínico, longitudinal, qualitativo realizado no Laboratório de Fonética Aplicada (FONAPLI). Este estudo de caso provém de uma pesquisa mais ampla, aprovada pelo Comitê de Ética em Pesquisa da Universidade Federal de Santa Catarina, conforme Resolução CNS nº 466/2012, na Plataforma Brasil sob nº 3.254.573 9. A responsável pela criança assinou o Termo de Consentimento Livre e Esclarecido, concordando com os termos e objetivos do estudo.

Para a avaliação dos sistemas fonético e fonológico, utilizou-se a Avaliação Fonológica da Criança (AFC) (YAVAS, HERNANDORENA,

LAMPRECHT, 2002). Assim, a partir da coleta dos dados, foi realizada a análise contrastiva e estabelecido o sistema fonético e fonológico, por meio de análise perceptivo-auditiva.

Para compor um *setting* de gravação, faz-se necessário os seguintes equipamentos: aparelho de ultrassom inserido em uma cabine tratada acusticamente, microfone unidirecional, transdutor micro-convexo acoplado ao aparelho de ultrassom, estabilizador de cabeça (SCOBBIE *et al.*, 2008), sincronizador de áudio e vídeo e câmera para a captura de imagens de boca. Veja Figura 1.

Figura 1 – Instrumentos de pesquisa: cabine acústica, microfone unidirecional, transdutor micro-convexo acoplado a um aparelho de ultrassom portátil (*Mindray M5*), computador, sincronizador, caixa de som, estabilizador de cabeça¹ (*Articulate Instruments Ltda*).



Para o diagnóstico clínico, a sonda de ultrassom pode variar minimamente entre uma frequência de aproximadamente 3 a 8 kHz, com uma taxa de *frames* de 30 *frames* por segundo (PRESTON *et al.* 2017). Entretanto, equipamentos de ultrassom com taxa de *frames* por volta de 120 *frames* por segundo (CLELAND *et al.* 2015; e 2019) já são utilizados em contextos clínicos, vislumbrando mais detalhamento dos gestos

¹ A imagem da posição de uso do capacete estabilizador de cabeça é de uma das autoras e teve sua aprovação para publicação.

articulatórios e o seu refinamento durante a produção de fala. A captura e análise dos dados de ultrassom é feita por meio de *software* específico para esse fim, que pode ser o *Articulate Assistant Advanced* ou o *Sonospeech software* (*Articulate Instrument Ltda.*), por exemplo.

Assim, a gravação e o *biofeedback* ultrassonográfico foram realizados utilizando os equipamentos acima descritos e exibidos na Figura 1. Para o procedimento de gravação, o participante da pesquisa foi orientado a se manter sentado, com postura ereta, no interior de uma cabine acústica. O transdutor do ultrassom foi alocado na região submandibular de modo fixo, preso ao estabilizador de cabeça (*Articulate Instruments Ltda.*).

O *corpus* utilizado para a avaliação perceptivo-auditiva foi: chave /'ʃave/, chique /'ʃike/, chuva /'ʃuva/, sapo /'sapo/, sica /'sika/ e suco /'suko/. Todas as palavras selecionadas são paroxítonas. Referente ao contexto vocálico, foi escolhido a vogal /i/ por corresponder à vogal mais anterior e mais alta, a vogal /a/ por ser mais baixa e central, e por último, a vogal /u/ por ser à vogal mais alta e mais posterior do Português Brasileiro (BERTI; PAGLIUSO e LACAVA, 2009).

O participante foi instruído, então, a incluir a palavra-alvo na frase-veículo “*Digo ____ bem bonito*”. Cada palavra-alvo foi repetida três vezes, totalizando 18 tentativas. A frase-veículo foi utilizada para tentar preservar a ocorrência da palavra-alvo no mesmo contexto fonético e prosódico, tentando dessa forma, evitar possíveis modificações nos parâmetros acústicos (BERTI; PAGLIUSO e LACAVA, 2009). O tempo de coleta dos sinais acústicos na avaliação pré e pós terapia variou entre 15 e 20 minutos.

Avaliação perceptivo-auditiva: pré e pós-terapia fonoaudiológica

O julgamento perceptivo-auditivo da produção de /s/ e /ʃ/ das palavras-alvo foi realizado por três juízes que as classificaram quanto à gravidade do desvio fonológico e ao grau de inteligibilidade. A gravidade do desvio fonológico foi baseada no cálculo do Percentual de Consoantes Corretas Revisado (PCC-R). A partir dos resultados obtidos do PCC-R, é possível classificar a gravidade do desvio como: severo (PCC-R < 50%), moderadamente-severo (50% < PCC-R < 65%), levemente-moderado (65% < PCC-R < 85%) ou leve (85% < PCC-R < 100%), conforme Shriberg et al. (1997). Para o grau de inteligibilidade

de fala foram utilizadas, para cada palavra-alvo, as seguintes qualificações: insuficiente e/ou incompreensível, regular e/ou pouco compreensível, e boa e/ou compreensível (BARLOW e HERSEN, 1982). Ao final, foi calculada uma média para classificação dos resultados percentuais.

Cada juiz recebeu uma tabela contendo o som-alvo, renomeados de forma que não fossem identificados como produzidos no pré ou pós-tratamento. Os juízes foram orientados para que realizassem a análise perceptivo-auditiva e, posteriormente, fizessem a transcrição fonética dos estímulos ouvidos. Como resultado, obtém-se uma tabela com as análises perceptivo-auditivas pré e pós-terapia de fala realizadas pelos três juízes. Esses resultados permitiram a comparação entre o pré e pós-tratamento. Foram consideradas corretas as produções que alcançam $\frac{3}{5}$, ou seja, três produções corretas em cinco repetições.

Tratamento fonoaudiológico: biofeedback por meio da ultrassonografia de língua

Na estimulação com *biofeedback* visual por ultrassom de língua, é fornecida a descrição do gesto articulatório, por meio da descrição verbal do terapeuta, explicando os sons-alvo e orientando a criança para que realize a produção dos gestos articulatórios da palavra contendo o som-alvo, representada através de uma figura. A criança deveria então observar, no ultrassom, a imagem em tempo real da sua produção, enquanto realizava o som-alvo. Na Figura 2, são exibidos exemplos do contorno de língua na produção pré e pós-terapia, bem como no período com retenção do som-alvo trabalhado em terapia.

Figura 2 – Exemplo do contorno de língua para /s/ na produção na pré e pós-terapia e retenção do som-alvo, trabalhado em terapia.



Pela Figura 2, percebemos, na produção referente ao fonema /s/, mudanças articulatórias no contorno de língua no período pós-terapia e de retenção se comparado ao período pré-terapia, em que parecia haver uma indiferenciação dos gestos entre as produções referentes aos fonemas /s/ e /ʃ/. Entretanto, vale ressaltar que, apesar de as análises perceptivo-auditiva (ver adiante na Tabela 1) e qualitativa indicarem que a terapia proporcionou refinamento dos gestos articulatórios, no futuro, seria importante implementar medidas quantitativas articulatórias para detalhar qual(is) partes da língua e em que medida ocorreu a diferenciação gestual entre o período pré e pós terapia.

O participante da pesquisa já havia recebido tratamento tradicional, anteriormente a este programa de tratamento intensivo com *biofeedback* visual por meio de ultrassom, porém sem sucesso terapêutico. O estudo de Preston, Brick e Landi (2013) também fez uso do *biofeedback* associado à terapia tradicional de intervenção, porém com crianças com diagnóstico de AFI. Esse estudo mostrou eficácia em todos os participantes que já haviam recebido tratamento tradicional sem sucesso, e a associação do *biofeedback* também contribuiu no processo de melhoria da produção da fala.

Na abordagem tradicional, foram realizadas atividades com jogos, usando palavras que continham os fonemas-alvo /s/ e /ʃ/, como: jogo de cartas, jogo de memória, jogo de dominó, jogo do bingo, jogo do labirinto, jogo do tabuleiro e diversos outros jogos. Todas essas atividades utilizaram técnicas de formação dos sons-alvo, imitação de palavras com os sons-alvo, pistas visuais, gestos articulatórios e treinamento auditivo.

Assim, a abordagem de intervenção fonológica foi pautada em estudo anterior (CLELAND, SCOBIE e WRENCH, 2015), abordando os aspectos de produção, percepção e organização dos gestos articulatórios, conforme detalhamento apresentado no Quadro 1.

Quadro 1 – Organização da terapia associada ao *biofeedback* ultrassonográfico de acordo com o objetivo, instrumentos e resultados alcançados.

Dados Terapêuticos						
Semana	Dias	Sessões	Objetivos	Instrumentos ou metodologia		Resultados alcançados
1ªsem.	1º	1ª	Coleta de amostra de fala	Gravação de dados de imagem (ultrassom) e áudio por meio da interface de áudio.		Avaliação pré-terapia.
2ª sem.	1º	2ª	Terapia tradicional com <i>biofeedback</i> visual	Percepção do som dos fonos [s] e [ʃ]; introdução do ponto articulatorio com o ultrassom.	Jogo da memória: a cada vez que um jogador virar um par de cartas, deverá repetir três vezes o nome da figura para que o outro jogador tenha percepção do som e possa observar o ponto articulatorio das fricativas [s] e [ʃ]. Jogo do Labirinto: para chegar até o final do labirinto, passe apenas pelas figuras que tenham o som de [s] e [ʃ].	Percepção e produção dos fonos [s] e [ʃ].
	2º	3ª	Terapia tradicional com <i>biofeedback</i> visual	Percepção e estimulação articulatória dos fonos [s] e [ʃ].	Jogo de Dominó: sempre que um jogador comprar uma pedra, deverá automaticamente "comprar" uma carta, e repetir em voz alta o nome da figura contida na carta, ressaltando o som das fricativas [s] e [ʃ]. O número de vezes da extremidade aberta da última pedra comprada é colocada em jogo.	Pronúncia sistemática das palavras com os fonos [s] e [ʃ].
3ªsem.	1º	4ª	Terapia tradicional com <i>biofeedback</i> visual	Terapia do contraste fonológico dos fonos [s] e [ʃ].	Jogo do bingo: cada jogador deverá escolher uma cartela. A cada rodada, um número será sorteado e quem tiver o número na cartela deverá marcar com uma peça. Cada vez que o jogador marcar um número na cartela, deverá comprar uma carta e, conseqüentemente, repetir a palavra-alvo pelo tempo marcado atrás da carta, que varia de 0 segundo a 1 minuto e 39 segundos.	Pronúncia sistemática das palavras com os fonos [s] e [ʃ].

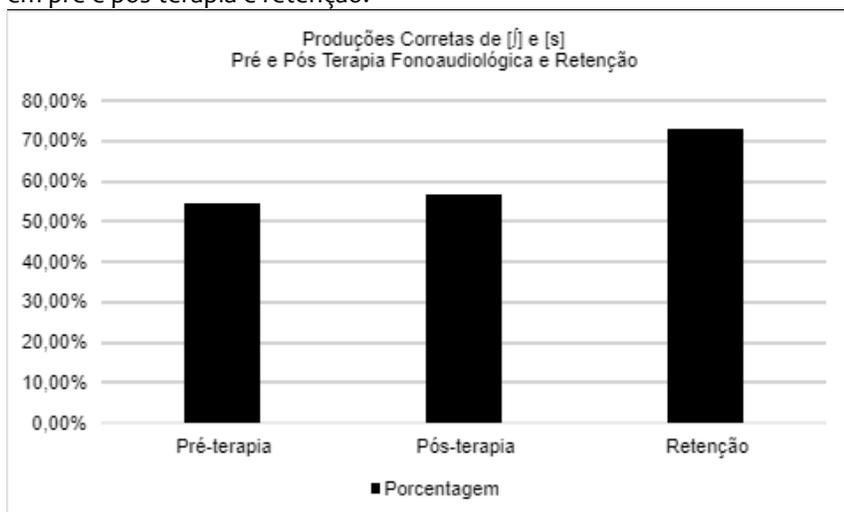
Avaliações fonoaudiológicas usadas em pré e pós-terapia

A avaliação do sistema fonético e fonológico pré e pós-terapia foi realizada por meio do protocolo de Avaliação Fonológica da Criança - AFC (YAVAS, HERNANDORENA e LAMPRECHT, 2002). Foram coletados dados do sistema fonológico com o apoio das figuras do AFC, nas quais o participante tinha de nomear as figuras a ele apresentadas.

Na avaliação pré-terapia, foi possível observar a substituição //j/ → [s], e a presença dos processos fonológicos de ensurdecimento de fricativa e plosiva, frontalização de palatal, omissão de líquida, simplificação de consoante final e de encontro consonantal. Já na reavaliação pós-terapia, observou-se a substituição //j/ → [s], no entanto, em onset inicial, foi considerada adquirida e, em onset medial, parcialmente adquirida. Verificou-se ainda a presença dos processos fonológicos de redução de encontro consonantal, apagamento de líquida intervocálica, anteriorização, apagamento de líquida final, substituição de líquida, assimilação e semivocalização de líquida.

A gravidade do desvio fonológico foi classificada a partir do Percentual de Consoantes Corretas Revisado (PCC-R) como grau leve a moderado na pré-terapia e grau leve na pós-terapia. Veja Figura 3.

Figura 3 –. Histograma dos resultados obtidos na análise perceptivo-auditiva em pré e pós-terapia e retenção.



Com a análise perceptivo-auditiva dos juízes, foi possível classificar a gravidade da distorção como moderadamente-severa na fase pré-terapia (média de 54,44%) e pós-terapia (média de 56,66%). Já na avaliação da retenção, a gravidade da distorção foi classificada como levemente-moderada (média de 72,83%). Veja, na Tabela 1, mais detalhes.

Tabela 1 – Resultados obtidos na análise perceptivo-auditiva dos três juízes na pré e pós-terapia e na retenção.

	Pré terapia	Pós terapia	Retenção
Juiz 1	53,33%	46,66%	62,96%
Juiz 2	60,00%	63,33%	55,55%
Juiz 3	50,00%	60,00%	100%
Média	54,44%	56,66%	72,83%

Os resultados obtidos na análise perceptivo-auditiva foram relevantes (Figura 3 e Tabela 1), observando-se melhora do PCC-R pré e pós-terapia e na retenção, na qual houve progresso da média de 54,44% (pré-terapia), 56,66% (pós-terapia), para 72,83%, classificado, então, como levemente-moderado.

No que se refere à generalização, foi possível verificar, na reavaliação pós-terapia, a correta produção de sons não tratados em terapia, ou seja, resultando na generalização para outros sons. Na avaliação pré-terapia dos aspectos fonológicos, além das substituições de /ʃ/→[s], notou-se a ocorrência de outras substituições. Já na reavaliação pós-terapia, nota-se ainda a ocorrência de substituições, mas também é verificada a superação. Observe as Tabelas 2 e 3.

Tabela 2 – Resultados da avaliação fonológica da criança no momento pré-terapia.

Onset Inicial		
Esperado	Realizado	Adquirido/Não adquirido
/b/	[p]	Parcialmente adquirido
/g/	[k]	Parcialmente adquirido
/v/	[f]	Não adquirido
/ʃ/	[s]	Não adquirido
/ʒ/	[ʃ] e [z]	Não adquirido
/l/	∅	Parcialmente adquirido
/dʒ/	[tʃ]	Parcialmente adquirido
Onset Medial		
Esperado	Realizado	Adquirido/Não adquirido
/g/	[k]	Parcialmente adquirido
/v/	[f]	Parcialmente adquirido
/ʃ/	[s]	Não adquirido
/z/	[s]	Não adquirido
/ʒ/	[ʃ] e [z]	Não adquirido
/λ/	[y]	Parcialmente adquirido
/r/	∅ e [l]	Parcialmente adquirido
Coda Medial		
Esperado	Realizado	Adquirido / Não adquirido
/R/	∅	Não adquirido
Encontros consonantais		
Tratando-se dos encontros consonantais /pr/, /br/, /tr/, /dr/, /kr/, /gr/, /fr/ e /vr/, observou-se apagamento da líquida /r/ em todas as posições das palavras, e apagamento da líquida /l/ nos encontros /pl/, /kl/, /gl/ e /fl/.		

Tabela 3 – Resultados da avaliação fonológica da criança no momento pós-terapia.

Onset Inicial		
Esperado	Realizado	Adquirido/Não adquirido
/ʃ/	[ʃ]	Adquirido
Onset Medial		
Esperado	Realizado	Adquirido/Não adquirido
/ʃ/	[s]	Parcialmente adquirido
/ʒ/	[z]	Adquirido
/ʎ/	[i]	Adquirido
/r/	[r]	Adquirido
Coda Medial		
Esperado	Realizado	Adquirido/Não adquirido
/R/	∅	Não adquirido
Encontros consonantais		
Tratando-se dos encontros consonantais /pr/, /br/, /tr/, /dr/ e /gr/, houve apagamento da líquida /r/ em todas as posições das palavras e nos encontros /kr/ e /fr/ na posição onset inicial. Nos encontros /pl/, /kl/ e /fl/, houve apagamento da líquida /l/.		

Quanto ao grau de inteligibilidade de fala (GIF) (BARLOW e HERSEN, 1982), analisado pelos três juízes, verificou-se que, em pré-tratamento, o paciente apresentou GIF insuficiente e, em pós-tratamento, GIF suficiente. Houve retenção, em pós-tratamento, retratando GIF suficiente.

Na avaliação fonológica, durante o período de pré-terapia, foi identificada a presença de substituição de /ʃ/ → [s] (Tabela 2). No entanto, na avaliação fonológica pós-terapia, depois de 8 sessões, foi constatada a reorganização dos gestos articulatórios envolvidos na produção de fala, sendo o som-alvo avaliado auditivamente como [ʃ] (Tabela 3).

Acredita-se que há a necessidade de realizar novas pesquisas que analisam a inserção de equipamentos para *biofeedback* visual dos articuladores envolvidos na produção dos sons-alvos, tanto em

crianças com desvio fonológico, em diferentes graus de severidade, como também para outras crianças com transtornos dos sons da fala.

Os resultados deste estudo certamente fornecem evidências de que esse método instrumental pode levar ao aprimoramento na precisão da produção dos sons da fala, à generalização de sons não tratados (Tabela 2 e 3), com resultados positivos de retenção, corroborando outros estudos (PRESTON, LEECE e MAAS, 2016; CLELAND, SCOBIE e WRENCH, 2015) que foram realizados com crianças em idade escolar, utilizando programa de tratamento semelhante. Nesses estudos, também houve generalização para outros sons, e retenção para algumas crianças, e ainda todos os participantes mostraram capacidade aumentada de realizar movimentos de fala, tornando o método uma opção para aumentar a precisão de fala.

Em resumo, podemos afirmar que estudos envolvendo o uso de tecnologias na área da fala vêm ganhando espaço no Brasil. A utilização desse aparato tecnológico reafirmou a efetividade do modelo de terapia com base fonológica, porém é necessário que os estudos ampliem a amostra de participantes a fim de verificar a robustez e aprimorar técnicas de *biofeedback*, vislumbrando aperfeiçoar o delineamento da intervenção.

Perspectivas de estudos com a ultrassonografia de língua

Este estudo explorou um programa de tratamento intensivo para crianças com desvio fonológico, associando o uso do *biofeedback* visual por meio da ultrassonografia de língua à terapia fonológica tradicional, objetivando descrever e analisar a sua efetividade na terapia fonológica dos fonemas /s/ e /ʃ/. Com o avanço dos recursos tecnológicos, a ultrassonografia de língua tem se mostrado de grande importância para o sucesso no processo terapêutico fonoaudiológico, na medida em que as avaliações são realizadas de forma minuciosa e aprofundada, podendo assim se intervir de maneira mais precisa nas dificuldades de cada paciente (WIETHAN e MOTA, 2012). Aliado às análises perceptivo-auditivas e às análises visuais (qualitativas), a literatura tem apontado para medidas quantitativas que são capazes de mensurar com acurácia as mudanças gestuais que ocorrem após o tratamento por meio do *biofeedback* com a USG. Dessa forma, em

trabalhos futuros, pretende-se incorporar parâmetros articulatórios quantitativos a fim de comparar a produção pré e pós-terapia dos sujeitos com TSF.

São inúmeras as possibilidades de uso desse aparato instrumental que é a ultrassonografia de língua, como exemplo na área da linguagem, em que se pode contribuir em pesquisas relacionadas à fonética, à organização de sequências silábicas, à coarticulação e às organizações estruturais dos sons. Além disso, pode vir a enriquecer o trabalho de outras áreas da Fonoaudiologia por suas possibilidades de interface entre áreas.

Na área da voz, por exemplo, pode auxiliar como parâmetro para diminuição do risco de complicações pós-cirúrgicas e determinar, com melhor exatidão, segmentos laríngeos e variações de comprimento muscular. Pode ser útil também na área de motricidade orofacial em distúrbios relacionados a alterações musculares decorrentes de distúrbios oclusais, possibilitando o acompanhamento e a definição de critérios de alta para os pacientes. E por fim, pode auxiliar na área de disfagia na avaliação do deslocamento do osso hióide na investigação da deglutição, e na correlação entre a espessura da língua e o estado nutricional de pacientes com disfagia pós Acidente Vascular Encefálico (AVE) (BARBERENA *et al.*, 2014).

ULTRASSONOGRAFIA E ERROS DE FALA

Letícia Bitencourt Uberti
Caroline Rodrigues Portalete
Denis Altieri de Oliveira Moraes
Márcia Keske-Soares

Introdução

Este capítulo apresenta um estudo que utiliza a ultrassonografia (USG) para caracterizar os contornos médios de língua de fricativas e líquidas (/s/, /z/, /l/ e /r/), do Português Brasileiro, em adolescentes e adultos com a produção de fala típica e com Erros Residuais de Fala.

Os distúrbios dos sons da fala (DSF) referem-se a qualquer combinação de dificuldades com a percepção, produção motora e/ou a representação fonológica dos sons e segmentos da fala que afeta a inteligibilidade da comunicação oral (ASHA, 2014). A base articulatória/fonética é a base de acometimento que mais prevalece nos quadros de DSF. Nesses casos, os erros de produção dos sons (omissões, substituições e distorções de sons) permanecem na fala criança/adolescente após a faixa etária de 8 a 9 anos, idade em que a criança já deveria ter desenvolvido a fala do padrão adulto, sendo assim chamado de Erros Residuais de Fala (ERF) (SHRIBERG; GRUBER e KWIATKOWSKI, 1994).

Nesse público, alterações em fricativas, róticos e/ou laterais são os erros de fala mais frequentes (BOYCE, 2015; CAMPBELL; MCALLISTER BYUN, 2018; MCALLISTER BYUN; CAMPBELL, 2016; PRESTON *et al.*, 2018; PRESTON; KOENIG, 2011; SHRIBERG *et al.*, 2010; VERÍSSIMO; BORSEL; e BRITTO PEREIRA, 2012). Sendo a distorção do /r/ o caso que mais acontece e o mais estudado (ADLER-BOCK *et al.*, 2007; BOYCE, 2015; CAMPBELL; MCALLISTER BYUN, 2018; MCALLISTER BYUN; CAMPBELL, 2016; PRESTON *et al.*, 2018; e SHRIBERG *et al.*, 2010).

Estudo traz que os ERF envolvem posicionamento ou movimento anormal da língua. Devido à localização visualmente inacessível da língua no trato vocal, a velocidade e a complexidade de seus movimentos, a visualização da língua na produção desses sons é um desafio para os

clínicos e pesquisadores. Sendo assim, o diagnóstico e o tratamento dessa patologia tornam-se difíceis (GIBBON, F.; LEE, 2015).

Quanto à percepção de fala, indivíduos com ERF mostram comprometimento variável nesse aspecto (BROEN *et al.*, 1983; RVACHEW e JAMIESON, 1989) e mostram diferenças neurobiológicas na ativação cerebral durante a percepção auditiva (CABBAGE; HOGAN e CARRELL, 2016; PRESTON *et al.*, 2012). Além disso, estudo com o objetivo de identificar diferenças cerebrais estruturais em crianças com ERF utilizando ressonância magnética, trouxe evidências que pode-se haver massa cinzenta aumentada no giro supratemporal bilateral e no giro supramarginal no cérebro desses pacientes, o que pode contribuir para a dificuldade na percepção fonética e na detecção de erros (PRESTON *et al.*, 2014).

Em relação aos róticos, a dificuldade de aquisição dessa classe de sons, devido a sua complexidade articulatória, é a principal causa da incidência de casos nos ERF (BOYCE, 2015). Além disso, as configurações complexas específicas da língua usadas para obter a qualidade aceitável acusticamente podem ser variáveis entre os falantes, dificultando a caracterização das características comuns das formas linguísticas (MIELKE; BAKER e ARCHANGELI, 2016).

A avaliação fonoaudiológica, incluindo avaliação da fala, é essencial para o diagnóstico dos ERF. Com o propósito de auxiliar as avaliações da base motora articulatória, a área da saúde dispõe de diversos instrumentos tecnológicos que tornam a avaliação dos ERF mais objetiva e melhor de ser mensurada, como a Ultrassonografia dos movimentos de língua (USG).

A literatura de estudos com ERF apresenta mais trabalhos utilizando a ultrassonografia principalmente para terapia como *biofeedback* visual. Em 1985, estudo descreveu a utilização do ultrassom como *biofeedback* e a experiência com o uso na terapia de fala em um paciente de 9 anos de idade (SHAWKER; SONIES, 1985). Em 2005 e 2007, autores examinaram a eficácia do uso do ultrassom como *biofeedback* visual na remediação de erros persistentes do /r/, tendo efeito benéfico na remediação do erro em dois sujeitos adolescentes (ADLER-BOCK *et al.*, 2007; e ADLER-BOCK *et al.*, 2005).

Em 2008, pesquisa comparou o tratamento para o /ɹ/ sem e com o uso da ultrassonografia. Neste estudo, a USG pareceu para facilitar a aquisição do /ɹ/ para o participante (MODHA *et al.*, 2008). Em 2014, um

grupo de pesquisadores testou a eficácia de um tratamento baseado no aprendizado motor que incluía biofeedback visual ultrassonográfico para indivíduos com ERF, e foram observados aumentos na precisão dos contextos sonoros alvo no nível de palavras com o programa de tratamento para a maioria dos participantes (PRESTON *et al.*, 2014)

Também no mesmo ano, estudiosos realizaram um estudo com objetivo de documentar a eficácia do tratamento de biofeedback por ultrassonografia para a má articulação dos róticos do inglês norte-americano em crianças com ERF, novamente com sucesso na utilização do instrumento (BYUN, T. M. A.; HITCHCOCK, E. R.; e SWARTZ, 2014).

Pode-se perceber que, após os primeiros dados relatando efeitos positivos do uso da USG nos ERF, a metodologia dos estudos começou a ficar mais complexa. Estudo, em 2015, verificou se a generalização poderia ser aprimorada quando a intervenção de *biofeedback* fosse estruturada de acordo com uma estrutura de "ponto de desafio" para a aprendizagem fonoaudiológica, apresentando eficácia nos resultados preliminares (HITCHCOCK; HAREL; BYUN, 2015).

Pesquisadores, em 2015, verificaram o potencial de uma técnica de *biofeedback* visual diferente, utilizando a ultrassonografia, para o tratamento de DSF persistentes, e após a intervenção, todas as crianças realizaram progressos significativos, evidenciados por medidas perceptuais e mudanças na forma da língua (CLELAND; SCOBIE; WRENCH, 2015). Em 2016, trabalho descreveu mudanças no deslocamento perceptivo, duracional e de língua ao longo de 10 sessões de terapia articulatória com o uso do ultrassom para o /ɹ/ em seis crianças (BRESSMANN *et al.*, 2016).

A partir de 2017, Preston e colegas realizaram diferentes estudos com indivíduos com ERF. O primeiro explorou as mudanças na precisão dos sons após programa intensivo de terapia de 1 semana, utilizando o ultrassom como *biofeedback* visual e teve como resultado melhora mensurável e estatisticamente significativa na precisão do som da fala (PRESTON; LEECE, 2017). O segundo comparou a generalização dos róticos após terapia com e sem ultrassonografia, entretanto ambos os tratamentos resultaram em aumento da generalização após a terapia (PRESTON; LEECE; MAAS, 2017).

O terceiro estudo, explorou como a frequência com que o *biofeedback* visual de ultrassom é fornecida durante a fonoterapia afeta a aprendizagem do som da fala (PRESTON *et al.*, 2018). E o último,

explorou a validade de uma nova medida ultrassonográfica, o Número de INFLections, para caracterizar os róticos produzidos por sujeitos sem e com ERF. Essa medida foi capaz de captar mudanças na forma da língua ao longo do tempo, bem como a variação específica do contexto em falantes típicos, e os sujeitos com fala típica produziram os róticos com maior precisão e com formas de língua mais complexas do que os com ERF (PRESTON *et al.*, 2019).

Pode-se observar que todos os estudos sobre ERF estão relatando casos de sujeitos com dificuldades em sons da fala, principalmente róticos, e são falantes do inglês, além de todos apresentarem objetivo de verificar a eficiência da USG no uso como biofeedback visual em terapias de fala. Apenas um estudo, de Preston e colaboradores (2018), comparou o grupo com ERF com indivíduos de fala típica, porém na língua inglesa.

Sendo assim, foi realizado um estudo com o objetivo de caracterizar os contornos médios de língua de fricativas e líquidas (/s/, /z/, /l/ e /r/), do Português Brasileiro, em adolescentes e adultos com a produção de fala típica e com Erros Residuais de Fala.

Métodos de estudo e participantes

Trata-se de uma pesquisa descritiva prospectiva, transversal, observacional, quase experimental, em que os participantes analisados se submeteram a avaliações fonoaudiológicas instrumentais quantitativas e qualitativas.

Adolescentes e adultos, de ambos os sexos, falantes do PB compuseram a amostra da pesquisa. Todos foram submetidos às avaliações fonoaudiológicas, para averiguar os aspectos de fala, motricidade orofacial e audição; e a avaliação instrumental ultrassonográfica.

Para participarem da avaliação instrumental, os adolescentes e adultos participantes deveriam estar de acordo com os seguintes critérios: possuir idade entre 12:00 e 44:11; ter realizado todas as avaliações fonoaudiológicas; apresentar linguagem compreensiva e expressiva oral adequadas; apresentar ou não alteração de fala; e apresentar limiares auditivos dentro dos padrões de normalidade.

Foram considerados critérios de exclusão adolescentes e adultos que apresentem DSF de base estrutural, especificamente, fissura

labiopalatina, utilizassem prótese dentária, fossem usuários de medicamentos controlados, álcool, tabaco e/ou drogas ilícitas, ou fatores neurológicos e/ou psiquiátricos, e estivessem realizando terapia fonoaudiológica.

Após a realização das avaliações fonoaudiológicas, os participantes de pesquisa foram separados em dois grupos conforme a presença ou não de alteração de fala: Grupo Sem Erros Residuais de Fala (GSE) formado por pessoas que apresentaram a produção de fala com padrão típico e o Grupo Com Erros Residuais de Fala (GCE) formado por pessoas que apresentaram distorções e/ou substituições na avaliação de produção da fala. Os participantes deste grupo apresentavam alterações nos sons fricativos (/s/ e /z/) e alterações nos sons líquidos (/l/ e /r/).

Avaliação fonoaudiológica

Dentre as avaliações fonoaudiológicas, foram realizadas avaliação linguagem expressiva e compreensiva oral, avaliação da produção dos sons da fala, por meio de um teste de imitação de palavra única, avaliação miofuncional orofacial e avaliação audiológica. Após essa bateria de avaliações, os participantes foram submetidos à avaliação ultrassonográfica, devendo reproduzir uma lista de palavras.

O Protocolo de Avaliação Instrumental de Fala – PRAIN (PORTALETE; FERNANDES e PAGLIARIN, 2018), que apresenta um corpus de 19 palavras, cada uma contendo uma consoante-alvo do PB, foi utilizado para a avaliação. As palavras deste instrumento seguiram os seguintes critérios linguísticos: segmento em posição inicial da sílaba dentro da palavra, tendo o som alvo com a vogal /a/ como precedente e seguinte, em palavras trissílabas e paroxítonas, como descrito em estudo prévio (BERTI; PAGLIUSO e LACAVA, 2009).

Foram apresentadas figuras na tela do computador, referentes à lista de palavras do PRAIN, para que o participante falasse espontaneamente a palavra-alvo durante a avaliação, considerando que essa é forma mais fidedigna de obtenção das amostras de fala. Solicitou-se, de cada sujeito, cinco repetições de cada palavra-alvo inserida em frase veículo “fale [palavra-alvo] de novo”. Para a classe de sons das fricativas foram analisados os sons /s/ e /z/, presentes nas palavras [ma.'sa.3ẽj] (massagem) e [ka.'za.dus] (casados). Para a

classe de sons das líquidas foram analisados os sons /l/ e /r/, presentes nas palavras [sa.'la.dɐ] (salada) e [ba.'ra.tɐ] (barata).

O ultrassom portátil, modelo DP 6600, e o microfone unidirecional foram utilizados para captação de imagem e de áudio. Para capturar, sincronizar os sinais de áudio e vídeo, e analisar as imagens obtidas pela USG foi utilizado o software *Articulate Assistant Advanced (AAA)*. O participante utilizou o estabilizador de cabeça para o posicionamento da sonda da USG, desenvolvido pela *Articulate Instruments Ltd*. As imagens foram selecionadas considerando o plano sagital, Modo B.

A cabine tratada acusticamente foi utilizada para que não houvesse interferência do ruído externo na captação e gravação dos sinais acústicos da fala. Após o examinador verificar o ajuste do capacete e a posição da sonda, o paciente ingeriu um gole de água (para mensurar a localização da área anterior do palato) e, na sequência, reproduziu a lista de palavras do PRAINP com os sons-alvo.

O exame foi realizado a fim de analisar os gestos articulatórios de língua presentes na produção dos fonemas do PB. Após a coleta, foi realizada a análise dos movimentos de língua, em que foram observadas as imagens dos movimentos de língua e selecionado o ponto de maior constrição para cada som analisado. O contorno da superfície da língua, em corte sagital, foi traçado e exportado para o recurso *Workspace*, onde foram obtidas 42 *splines* representando a continuidade do traçado em cada segmento, considerando os pontos válidos da curva da língua aqueles correspondentes ao intervalo de 100% de confiança.

Esses pontos foram separados por parâmetros e tabulados em programa Excel, para análise estatística na ferramenta estatística R (R Core Team). Os valores médios dos contornos de língua nas repetições das palavras para cada grupo e os respectivos intervalos de confiança em torno dos valores médios (para cada um dos pontos das curvas de língua) foram calculados. Os intervalos foram obtidos com 95% de confiança em torno de cada um dos pontos médios das curvas de língua (*splines*). Assim, as linhas médias e os intervalos de confiança dos contornos de superfície da língua para cada som avaliado foram obtidos.

Resultados e Discussão

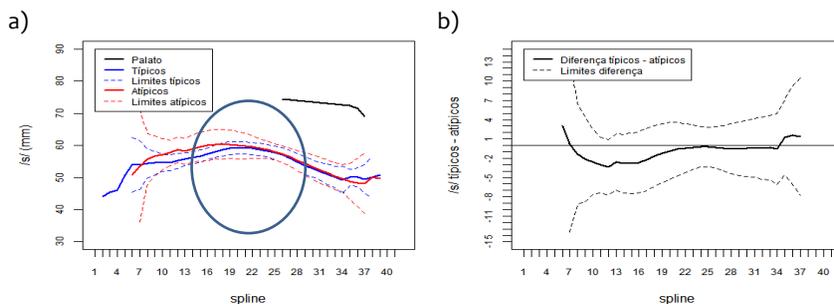
Dos 26 participantes do estudo, 54% (n=14) apresentaram alterações de fala após a realização das avaliações fonoaudiológicas, 50% (n=7) apresentou alteração na classe de sons das fricativas e 36% (n=5) apresentaram alteração na classe de sons das líquidas e 14% (n=2) apresentaram alteração em ambas classes de sons.

Na avaliação miofuncional orofacial, o GCE e o GSE diferiram nos aspectos de aparência e postura de lábios e mandíbula, mobilidade de lábios, língua, mandíbula e bochechas e nas funções avaliadas, porém, de acordo com o Teste Exato de Fischer as diferenças não foram significativas estatisticamente. A dificuldade (tremor e falta de precisão) no movimento de lábios foi o que mais se aproximou da diferenciação entre os grupos. Nenhum participante apresentou alteração em relação à fixação do frênulo lingual e, em relação à força da língua, os sujeitos do GSE apresentaram melhor desempenho na tarefa.

Os gráficos permitem visualizar os contornos médios de língua dos fonemas fricativos (/s/ e /z/) e líquidos (/l/ e /r/), em corte sagital, comparando com GCE e GSE. Nas figuras (a) foram expostas as linhas médias dos fonemas avaliados com intervalo de 95% de confiança para os grupos e nas figuras (b) o intervalo de 95% de confiança para a diferença entre os grupos. À esquerda do gráfico pode-se ver a base da língua e à direita do gráfico a ponta da língua.

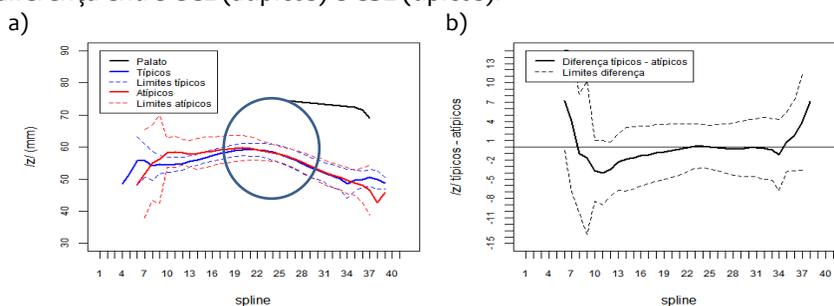
Todos os gráficos apresentam o contorno médio de língua (linha contínua). As linhas pontilhadas em torno das médias representam os intervalos de confiança em cada ponto da curva de língua, conforme pode ser visualizado nas figuras (a). Os intervalos em torno das bandas de confiança foram calculados com 95% de confiança. Quando as bandas, entre dois segmentos, se afastam, há tendência à diferenciação naquele(s) pontos (s), porém, quando as bandas de confiança entre as médias de duas curvas separam-se, é possível afirmar que houve diferença significativa, conforme pode ser visualizado nas figuras (b).

Gráfico 1 – (a) Linha média do /s/ com intervalo de 95% de confiança para o GCE (atípicos) e GSE (típicos) e (b) intervalo de 95% de confiança para a diferença entre GCE (atípicos) e GSE (típicos).



No Gráfico 1, a representação dos contornos médios de língua para a produção do /s/, através da linha média e dos intervalos de confiança para diferença. Em (a), o contorno médio de /s/, apresentou maior elevação à esquerda, em direção à raiz da língua, no GCE, que ocorre mais na região central da língua no GSE. Em (b), na região à direita, na ponta de língua, ao comparar a produção do GCE e GSE, as bandas dos limites superiores e inferiores separaram-se da linha média em direções contrárias, porém não há diferença significativa.

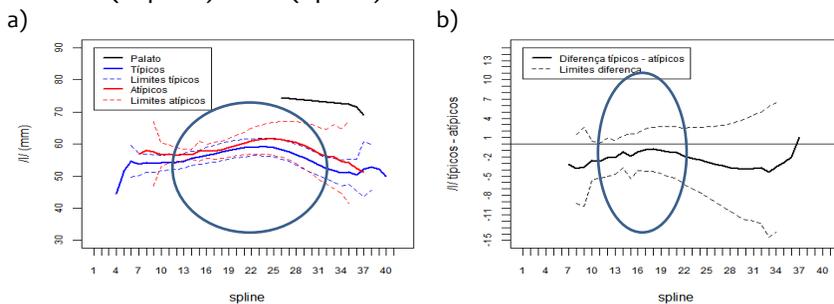
Gráfico 2 – (a) Linha média do /z/ com intervalo de 95% de confiança para o GCE (atípicos) e GSE (típicos) e (b) intervalo de 95% de confiança para a diferença entre GCE (atípicos) e GSE (típicos).



No Gráfico 2, a representação dos contornos médios de língua para a produção do /z/, através da linha média e dos intervalos de confiança para diferença. Em (a), no contorno de /z/, ambos os grupos

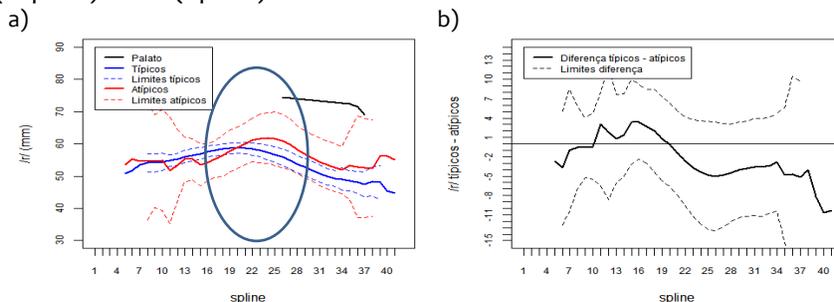
realizaram elevação da região central da língua, entretanto, o GCE apresentou elevação maior em direção à raiz. À direita, na região de ponta de língua, no GSE há elevação de ponta na produção do som e no GCE essa elevação não ocorre. Em (b), a banda superior e a banda inferior convergem para a mesma direção, significando que não há diferença significativa entre a produção de ambos os grupos.

Gráfico 3 – (a) Linha média do /l/ com intervalo de 95% de confiança para o GCE (atípicos) e GSE (típicos) e (b) intervalo de 95% de confiança para a diferença entre GCE (atípicos) e GSE (típicos).



No Gráfico 3, a representação dos contornos médios de língua para a produção do /l/, através da linha média e dos intervalos de confiança para diferença. Na figura (a), à esquerda, as produções de ambos os grupos tenderam a significância em relação à raiz de língua. À direita, as bandas estão muito afastadas da linha média e a diferença entre os limites superiores e inferiores é muito alta, a partir do spline 33, demonstrando que houve maior variabilidade da produção do /l/ nos GCE. Além disso, no GSE houve maior elevação de ponta de língua do que no grupo atípico. Em (b), as bandas de confiança entre as médias das curvas se separaram, no intervalo do spline 26 a 33, porém, sem diferença significativa.

Gráfico 4 – (a) Linha média do /r/ com intervalo de 95% de para o GCE (atípicos) e GSE (típicos) e (b) intervalo de 95% de confiança para a diferença entre GCE (atípicos) e GSE (típicos).



No gráfico 4, a representação dos contornos médios de língua para a produção do /r/, através da linha média e dos intervalos de confiança para diferença. Em (a), tanto a banda superior quanto a inferior estão muito afastadas da linha média na produção do /r/ no GCE, com valores maiores de diferença nos splines de ponta da língua, demonstrando a grande variabilidade de produção deste som, comparado ao GSE em que as bandas estão bem próximas da linha média. Em (b), à direita, a banda superior e inferior estão se separando da língua média em direção opostas, porém não há valores próximos à diferenciação na região de ponta de língua.

A dificuldade de articulação das consoantes fricativas e líquidas são as mais comuns na vida adulta (PRESTON *et al.*, 2018; PRESTON e KOENIG, 2011; VERÍSSIMO; BORSEL e BRITTO PEREIRA, 2012; e WIETHAN *et al.*, 2015). Esse dado foi confirmado no presente estudo visto que, dos 26 sujeitos que participaram da pesquisa, 14 apresentaram alterações de fala nessas classes de sons.

Estudo sugere que sujeitos com ERF podem ter dificuldade em estabelecer um plano motor adequado para os sons ou podem ter dificuldades em conseguir o tempo adequado dos movimentos (como coordenar os lábios e a língua apropriadamente com movimentos que acontecem antes ou depois do som) (PRESTON *et al.*, 2014). Outro trabalho refere que nos casos de ERF há o posicionamento ou movimento anormal da língua e que o “ambiente articulatório” (posturas orais habituais de um indivíduo) pode causar um impacto geral na articulação dos segmentos (GIBBON, F. e LEE, 2015). Além disso, muitos estudos demonstram que o desenvolvimento motor da

fala ainda não está completo na adolescência, sendo caracterizado por ser um processo vagaroso (KOENIG; LUCERO; PERLMAN, 2008; WALSH; SMITH, 2002).

Na produção articulatória da fricativa /s/, em ambos os grupos, houve elevação na parte posterior da língua em ambos os grupos, concordando com estudos anteriores, pois para a produção deste som há um estreitamento da passagem de ar entre a parede faríngea e a raiz de língua, conduzindo o ar até o ponto de turbulência na constrição da ponta da língua (BRESSMANN *et al.*, 2005; e FRANCISCO, 2015). Ainda no presente estudo, houve diferenciação entre o GCE e o GSE no contorno médio de língua em relação à parte anterior da língua, podendo ser justificado pela falta de independência do controle da ponta de língua. Isto é característico de crianças com DSF com um grau de imaturidade de fala (GIBBON, F. E., 1999; GICK *et al.*, 2007; ZHARKOVA; HEWLETT e HARDCASTLE, 2012).

Os contornos médios de língua na produção do /z/ são semelhantes aos da produção do /s/. Porém, na realização do /z/, no GSE, houve elevação de ponta de língua, que não ocorreu no GCE. Em falantes típicos, a produção das fricativas [s] e [z] é caracterizada pelo contato lateral da língua ao longo do palato, além de um contato incompleto da língua na parte anterior do rebordo alveolar, resultando em um sulco nesta região (BAUMANN-WAENGLER, 2008). Porém, estudo evidenciou após realização de medidas eletropalatográficas, um maior contato linguo/palatal para [z] do que para [s], justificando esta diferença como uma necessidade de um maior volume de ar para a produção de [s], o que empurraria a língua lateralmente, numa tentativa de criar uma passagem mais ampla para o fluxo aéreo (MCLEOD; ROBERTS; SITA, 2006). Como esse maior contato foi verificado na produção do GSE na elevação da ponta da língua, pode-se sugerir que a produção do /z/ no GCE foi mais próxima da produção do /s/.

Em relação às líquidas, no GSE ocorreram dois gestos articulatórios simultâneos, tanto na produção do /l/ quanto na do /r/: elevação de ponta e de corpo de língua produzidos concomitantemente. O que pode ser justificado pela semelhança fonética entre esses segmentos e nos processos que eles partilham (SILVA, 2010).

A presença de duplos gestos na produção das líquidas refere-se a maior dificuldade de coordenação dos movimentos, pois requer

habilidades coordenadas e refinadas de movimentos dissociados e simultâneos da língua (BERTI, 2010). O presente estudo corrobora com trabalhos de análise ultrassonográfica das línguas no Português Brasileiro e no Inglês (BARBERENA; KESKE-SOARES e BERTI, 2014; MAYER e GICK, 2012; SPROAT e FUJIMURA, 1993).

Em ambos os contornos médios, para o GCE, as bandas superiores e inferiores estão bem afastadas, evidenciando grande variabilidade articulatória para a produção do /l/ e do /r/, o que pode ser justificado pela imaturidade do desenvolvimento do controle motor (SMITH e ZELAZNIK, 2004; ZHARKOVA; HEWLETT; HARDCASTLE e 2011, 2012), inconsistências em timing, velocidade e amplitude dos movimentos articulatórios (WALSH; SMITH e WEBER-FOX, 2006).

A análise articulatória, realizada neste estudo, possibilitou a visualização direta dos movimentos dos articuladores, proporcionando informações mais precisas a respeito dos processos envolvidos na produção da fala de adolescentes e adultos sem e com ERF. A análise da fala na população de adolescentes e adultos deve ser alvo de maiores estudos, principalmente com estudos nacionais e com a utilização de outros métodos de análise quantitativa, visto que os instrumentos para avaliação de fala mostram-se cada vez mais sensíveis para a exploração das bases motoras envolvidas na fala (BARBERENA; KESKE-SOARES; BERTI, 2014). Além disso, a investigação da fala de adolescentes e adultos é muito importante, uma vez que a presença de alterações de fala pode influenciar de maneira negativa a vida desses sujeitos no meio social, educacional e profissional.

ANÁLISE ULTRASSONGRÁFICA DE FRICATIVAS POSTERIORES EM ONSET SILÁBICO EM DIFERENTES CONTEXTOS VOCÁLICOS

Carine Haupt
Izabel Christine Seara

Introdução

Os estudos de análise de imagens ultrassonográficas voltados para a descrição de variedades em fala normal de adultos no Brasil são ainda bastante restritos. O foco da descrição fonética nesse grupo de sujeitos ainda recai sobre as análises acústicas e, muito lentamente, os estudos articulatórios estão se desenvolvendo, a fim de discutir correlatos acústico-articulatórios e de debater questões de representação fonológica. Com o intuito de expandir os conhecimentos acerca da articulação dos segmentos do português brasileiro, apresentamos este estudo articulatório a partir de dados de ultrassonografia de língua.

O objeto de análise escolhido são as fricativas posteriores, encontradas predominantemente no português brasileiro na produção do /r/ em início de sílaba. Este estudo visa aprofundar achados de uma investigação que discutiu correlatos acústico-articulatórios desses segmentos (HAUPT e SEARA, 2020). Nessa investigação, encontramos a possibilidade de constrictões na laringe, além da constrictão no trato oral de fricativas velares e uvulares, e fortes influências do contexto vocálico na produção dessas fricativas. Desse modo, expandimos os contextos vocálicos, restritos às vogais [a, i, u] no estudo de Haupt e Seara (2020), para todas as sete vogais orais do português. Os objetivos do presente estudo, portanto, são verificar o comportamento das fricativas posteriores diante de diferentes contextos vocálicos e contribuir para a descrição articulatória da categoria dos róticos.

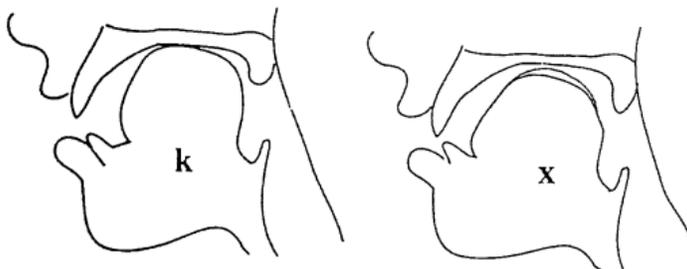
O texto que segue é dividido em três seções, além da introdução e considerações finais. A seção “Articulação de fricativas posteriores” traz informações articulatórias de fricativas velares, uvulares, faringais e glotais a partir de textos clássicos, como de Maddieson e Ladefoged (1996) e de um estudo atual de Spreafico e Vietti (2013), que traz uma

análise da pronúncia de róticos de Tirol do Sul por falantes bilíngues, a partir de dados ultrassonográficos. Há outros estudos sobre róticos com dados ultrassonográficos; no entanto, o único que encontramos que inclui uma fricativa posterior foi o dos falantes bilíngues tiroleses. Por isso, vamos descrever somente esse estudo. Na seção seguinte, “Metodologia”, descrevemos todo o protocolo de coleta de dados, além de trazermos informações sobre a análise. Na sequência, apresentamos a seção da análise com discussão. Nosso foco de análise será qualitativo e visa comparar as configurações articulatórias nos diferentes contextos fonéticos.

Articulação de fricativas posteriores

Ladefoged e Maddieson (1996) classificam como fricativas posteriores aquelas produzidas a partir da região do palato duro, compreendendo as fricativas palatais até as epiglóticas (as glóticas são tratadas a parte, como vogais). As fricativas posteriores são não sibilantes, uma vez que não encontram nenhuma obstrução além do estreitamento que gera o próprio ruído. Interessam-nos as fricativas posteriores produzidas a partir do palato mole (as velares) e as subsequentes, até a região da glote, pois não temos indícios de fricativa palatal para a produção de róticos. Os autores comparam a produção das fricativas palatais e velares com suas correspondentes oclusivas e observam que a forma do trato vocal é muito próxima. Observe, pela Figura 1, a elevação do corpo da língua e abaixamento da parte anterior da língua e da raiz.

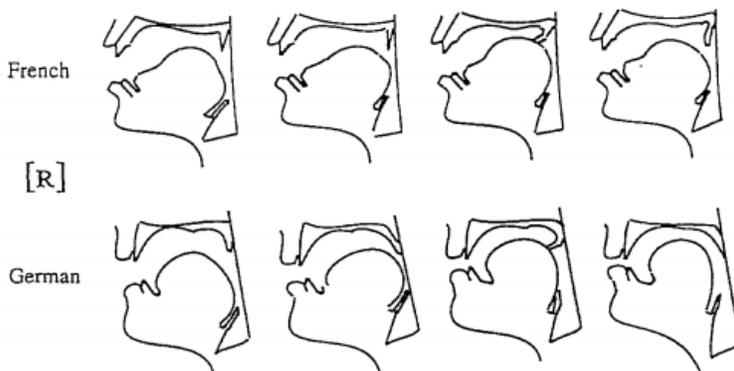
Figura 1 – Comparação entre a oclusiva velar e a fricativa velar do chinês padrão (ZHOU e WU, 1963)



Fonte: LADEFOGED, P.; MADDIESON, I. **The sounds of the words's language**. Blackwell Publishers: Oxford, 1996, p. 166.

As fricativas uvulares também são citadas como fricativas, mas com a ressalva de que “uvular fricatives in that the shape of the vocal tract may be such that the uvula vibrates” (LADEFOGED e MADDIESON, 1996, p. 167). Assim, elas são tratadas como vibrantes. A Figura 2 mostra uma sequência de frames da produção de uma uvular no francês e no alemão, a partir de dados de raio-X. Os autores salientam a combinação entre vibração e fricção, no entanto, afirmam também ser possível a produção apenas da fricção sem o componente de vibração.

Figura 2 – Sequência de frames da produção de vibrantes uvulares no francês e no alemão, a partir de dados de raio-X.

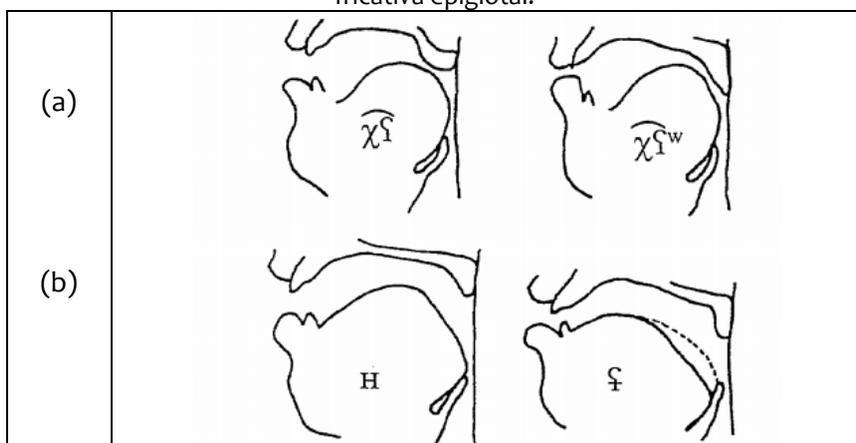


Fonte: LADEFOGED, P.; MADDIESON, I. **The sounds of the words's language**. Blackwell Publishers: Oxford, 1996, p. 229.

Na Figura 2, observamos, no segundo frame de cada sequência, um movimento de recuo (retração) da língua. Em seguida, no terceiro frame de cada sequência, verificamos um movimento de retorno da raiz da língua, acompanhado da elevação do corpo da língua e da úvula. Às fricativas faringais e epiglóticas, Ladefoged e Maddieson (1996) dedicam uma longa exposição. Os autores apresentam uma língua (Burkikham, dialeto de Angul) que contrasta as duas fricativas e outras línguas que apresentam em seu inventário apenas a faringal ou a epiglotal. A diferença articulatória entre as duas está no fato de que a epiglotal é realizada na região da epiglote, enquanto a faringal é produzida na parte anterior da faringe, logo após a úvula. Nas epiglóticas, a epiglote tem papel ativo (LAUFER e CONDAX, 1981). A partir de dados do árabe, os autores explicam que, muitas vezes, a língua não é envolvida na produção

da fricativa epiglotal, já que ela é produzida entre a epiglote e a parede posterior da faringe, compreendendo o contato da epiglote com a artenoide. A Figura 3 exibe imagens da produção da fricativa faringal e da epiglotal. Percebemos a diferença articulatória supracitada entre os dois tipos de fricativas: em (a), observamos a constrição faringal se estendendo da região uvular (baseado em BGAZHBA, 1964 *apud* LADEFOGED e MADDIESON, 1996) e, em (b), a constrição da região epiglotal (baseado em Gaprindashvili, 1966 *apud* LADEFOGED e MADDIESON, 1996).

Figura 3 – Imagens da produção de: (a) uma fricativa faringal e (b) uma fricativa epiglotal.



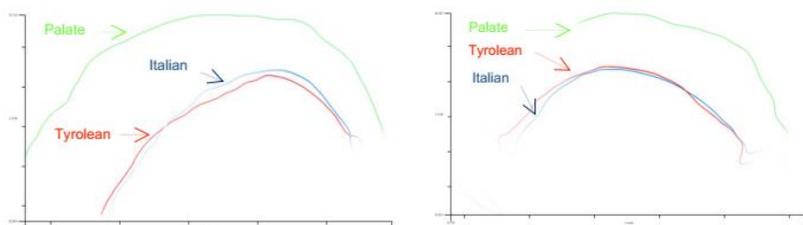
Fonte: LADEFOGED, P.; MADDIESON, I. **The sounds of the world's languages**. Blackwell Publishers: Oxford, 1996, p. 170.

Através das descrições apresentadas anteriormente, é possível identificar diferentes configurações da língua para a produção da fricativa velar, uvular, faringal e epiglotal. Obviamente temos de considerar as diferenças entre línguas e o fato de essas fricativas não terem valor contrastivo no PB, o que poderá resultar em articulações um pouco diferentes das apresentadas aqui. No entanto, as descrições existentes serão o nosso ponto de partida. Para a fricativa glotal, não temos uma configuração específica do corpo da língua, uma vez que a fricção acontece na glote e a língua assume a configuração da vogal seguinte (KENT e READ, 2015). Dessa forma, temos de observar se há uma constrição na realização da fricativa que difere da configuração da vogal.

Spreafico e Vietti (2013) realizaram um estudo ultrassonográfico da produção de róticos por falantes bilíngues italianos e tiroleses da

região de Tirol. A comparação sociolinguística foi feita entre bilíngues tardios e bilíngues simultâneos. Os pesquisadores encontraram apenas produções alveolares para ambas as línguas para o bilíngue tardio italiano e apenas dorsais para a falante bilíngue tardia tirolesa. Para os bilíngues simultâneos, houve predominância de produções uvulares para o dialeto tirolês e algumas produções alveolares para a língua italiana. A Figura 4 contrasta essas produções.

Figura 4. Contorno de língua de róticos alveolares do tirolês (linha vermelha) e do italiano (linha azul) produzido por um bilíngue simultâneo (à esquerda) e de róticos uvulares do tirolês (linha vermelha) e do italiano (linha azul) produzidos por outro falante bilíngue simultâneo.



Fonte: SPREAFICO, L. VIETTI, A. On rothic in bilingual community. A preliminar UTI research. In: SPREAFICO, L. VIETTI, A. **Rothics. New data and perspectives.** Bozen-Bolzano University Press: Bolzano, 2013: p. 67.

Comparando as imagens à esquerda e à direita da Figura 4, percebemos a elevação da raiz da língua e o abaixamento da lâmina na produção das uvulares. Em relação às fricativas posteriores do PB, os estudos focam em descrições sociolinguísticas. E, de forma geral, ainda são poucos os estudos existentes e estes exploram ainda poucas variedades brasileiras. Enumeramos aqui algumas localidades em que já se fez algum estudo sobre fricativas posteriores: Tocantins (HAUPT, 2018), Belo Horizonte (RENNICKE, 2015), Florianópolis (REINECKE, 2006), São Paulo, Rio de Janeiro, Paraná e Rio Grande do Sul (SILVA, 2002). Alguns desses estudos apontam para uma fricativa velar ou uvular, mas nem sempre são indicados correlatos acústicos para esses pontos de articulação. Apesar de esses estudos indicarem diferentes variantes, Abaurre e Sandalo (2003) afirmam que a variante mais frequente, no PB, é a fricativa glotal para o r forte. Nesse estudo, conduzido no Rio de Janeiro (ABAURRE e SANDALO, 2003), os autores defendem que a glotal é predominante.

Clements (1985) afirma que os sons laríngeos assimilam traços de vogais adjacentes; assim, em inglês, o /h/ de *heat* assume qualidades palatais, enquanto o /h/ de *hot* assume qualidade dorsais. Este é exatamente o quadro que encontramos no português. Vários autores afirmam que a fricativa velar é “típica do dialeto carioca” (cf. Sila, 2002 [1998], p. 38), mas mesmo neste dialeto somente encontramos em ataque silábico, nos nossos dados, uma fricativa glotal que assume qualidades de vogais adjacentes (ABAURRE e SANDALO, 2003, p. 165)

Diante deste quadro, o esperado, então, seria que não houvesse nenhuma constrição para a realização no trato oral, uma vez que a constrição se dá na glote. No entanto, o estudo de Haupt e Seara (2020) atestou a realização de fricativas velares, faringais e glotais em seus dados com falantes florianopolitanos, condicionados aos contextos vocálicos. A vogal [a] condicionou a constrição no trato vocal, enquanto as vogais [i, u] condicionaram a produção de glotais. Por se tratar de um estudo que se restringiu apenas a três vogais, não foi possível fazer generalizações. Desse modo, precisamos investigar mais a fundo como se dá a produção das fricativas posteriores nos diferentes contextos vocálicos. Na seção a seguir, descrevemos a metodologia de coleta e análise de dados.

Metodologia

O presente estudo foi elaborado com produções de duas informantes com a mesma faixa etária (em torno de 20 anos), estudantes universitárias e naturais de Florianópolis (SC), nomeadas de Informante A e Informante B.

A coleta dos vídeos ultrassonográficos foi feita com os seguintes equipamentos:

(a) microfone oral (*headworn Shure*);

(b) ultrassom *Mindray M5* com sonda endocavitária (6CV1S), operando em uma taxa de frequência de 6 MHz;

(c) capacete de estabilização de movimentos e sincronizador de imagem e áudio (SBU 1.0).

A taxa de frames foi de 12 Hz, taxa máxima permitida pelo equipamento. Antes de iniciarmos a gravação, foi feita a demarcação do palato das informantes, com auxílio de água. Dados acústicos também foram gravados em outro computador, com *hardware* MOTU

UltraLite mk3, numa taxa de amostragem de 44 kHz. As informantes usaram outro microfone, marca e modelo Audio-technica PRO 8HEX, fixo no capacete de estabilização. Toda coleta foi feita em cabine com isolamento acústico.

Os itens lexicais consistem em palavras reais, de modo a garantir maior naturalidade nas produções. Optamos por montar um *corpus* apenas com o rótico em sílaba tônica para facilitar a visualização da articulação, considerando que a taxa de frames é relativamente baixa. Todos os róticos (em posição inicial e medial das palavras) são seguidos pelas sete vogais orais do PB ([a], [ɛ], [e], [i], [ɔ], [o], [u]) e, no meio de palavra, são antecidos das vogais [a], [i] e [u]. As palavras foram pronunciadas isoladamente, sem frase veículo, para evitar efeitos de coarticulação da palavra anterior. Seguindo esses critérios, temos a seguinte lista de palavras: *rato, reta, rego, Rita, rota, rolo, ruga, arrasta, arreta, carreta, arrisca, arrota, arrote, arruda, pirraça, irrita, burrada e corrupto*.

Para análise dos dados, executamos paralelamente uma análise auditiva e articulatória com base nas gravações de áudio e nas imagens sincronizadas do ultrassom. A partir disso, traçamos automaticamente a superfície da língua nos frames relativos aos róticos e às vogais seguintes das imagens ultrassonográficas no plano sagital, usando o programa AAA (*Articulate Assistant Advanced*), versão 2.17.06. Quando necessário, corrigimos manualmente o traçado. Por fim, transferimos as curvas traçadas na imagem bruta do ultrassom em coordenadas cartesianas para uma planilha como base para a análise qualitativa.

A discussão das análises das imagens ultrassonográficas foi realizada em duas etapas:

(i) análise de um ponto fixo: nessa análise, utilizamos, para cada rótico, o frame que corresponde à parte medial da fricativa, considerando que não há um ponto de fechamento específico para ser coletado. Depois, comparamos a produção desse segmento nos diferentes contextos vocálicos, sobrepondo os traçados;

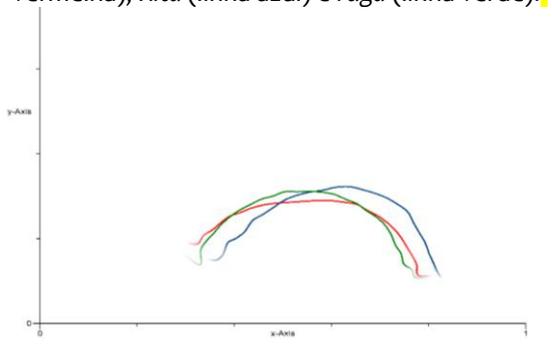
(ii) análise dinâmica: para essa análise, sobrepomos os frames do rótico e da vogal seguinte de cada palavra, de modo a visualizar o percurso articulatorio da fricativa para a vogal. Em seguida, comparamos os diferentes exemplares para verificar se é possível identificar um padrão.

A partir dessas análises, podemos apontar, com maior precisão, o tipo de fricativa produzido: se velar, uvular, faringal ou glotal.

Análise dos dados

A análise dos dados da Informante A evidenciou diferentes produções para o rótico em início de palavra e entre vogais. No início de palavra, a informante produziu uma glotal. A Figura 5 mostra a média das curvas da língua do ponto central dos róticos das palavras *rato*, *Rita* e *rua*.

Figura 5 – Linha média das repetições dos róticos das palavras *rato* (linha vermelha), *Rita* (linha azul) e *rua* (linha verde).¹

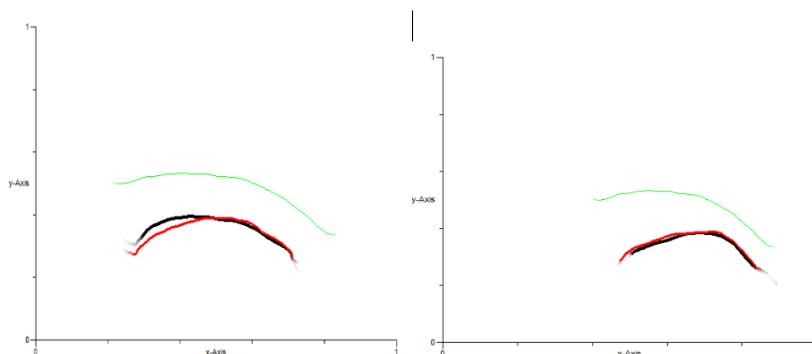


Pela Figura 5, que apresenta os contornos de língua nos róticos produzidos em posição inicial, podemos verificar diferentes contornos, evidenciando contornos de língua aracterísticos da vogal subsequente, o que não ocorreu com a realização da fricativa intervocálica, como discutiremos mais adiante.

Na Figura 6, então, comparamos o contorno de língua na produção do rótico com o contorno de língua na produção das vogais seguintes nos itens lexicais iniciados com r. Nessa figura, é possível observar uma sobreposição de contornos do ponto central do rótico e da vogal, confirmando tratar-se de uma fricativa glotal, conforme também já averiguado por Haupt e Seara (2020). Esse padrão, demonstrado na Figura 6, se repete nas outras palavras iniciadas com r.

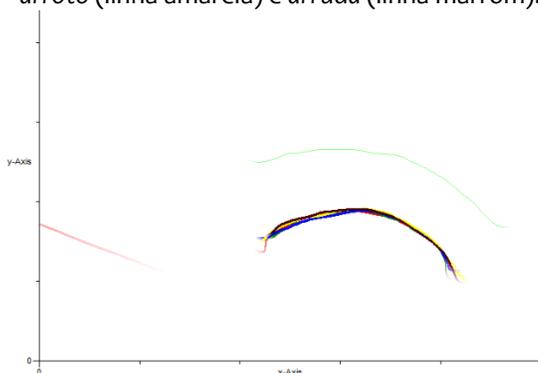
¹ Em todas as figuras de contorno de língua, a raiz da língua se encontra à esquerda e a ponta, à direita.

Figura 6 – Sobreposição da média do contorno da língua do rótico (linha preta) e da vogal seguinte (linha vermelha) da palavra *rota*, à esquerda; e sobreposição da média do contorno da língua do rótico (linha preta) e da vogal seguinte (linha vermelha) da palavra *reta*, à direita.



No entanto, ao analisarmos o rótico intervocálico, o padrão apresentado na Figura 6 não se manteve. E, nesse caso, obtivemos a mesma curva da língua para as palavras *arrasta*, *arreta*, *arrisca*, *arrota*, *arroto* e *arruda*, ou seja, para os seis contextos vocálicos distintos. Na Figura 7, apresentamos o contorno da língua de algumas dessas palavras.

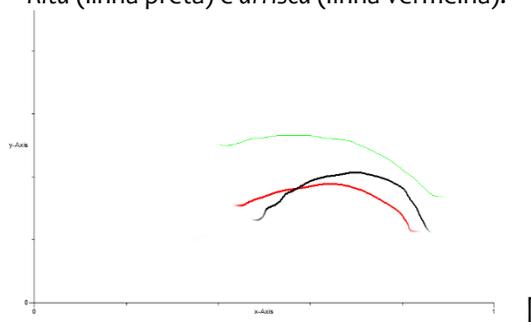
Figura 7 – Média do contorno da língua dos róticos intervocálicos nas palavras *arrasta* (linha vermelha), *arreta* (linha verde), *arrisca* (linha azul), *arroto* (linha amarela) e *arruda* (linha marrom).



O fato de não haver a antecipação articulatória da vogal seguinte é um indicativo de não se tratar de uma fricativa glotal. Considerando o desenho da curva da língua, podemos apontar para a produção de

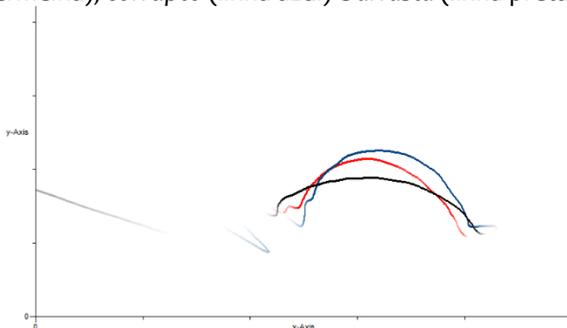
uma fricativa uvular (observe os dados das fricativas do tirolês, na Figura 4) ou até de uma fricativa faringal². Ao compararmos a produção do rótico na palavra *Rita* (glotal) e *arrisca* (uvular), observamos a diferença. Veja na Figura 8.

Figura 8 – Média do contorno da língua dos róticos nas palavras: *Rita* (linha preta) e *arrisca* (linha vermelha).



É interessante observarmos que todas as ocorrências de uvulares têm o mesmo contexto vocálico precedente, ou seja, a vogal [a]. Haupt e Seara (2020) encontraram em uma das informantes a produção de uma consoante faringal quando o rótico era precedido de [a] e outras realizações para outros contextos vocálicos. Para discutir a influência do contexto vocálico precedente, analisamos as palavras *arrasta*, *burrada*, *pirraça* e *corrupto*. Considere a Figura 9.

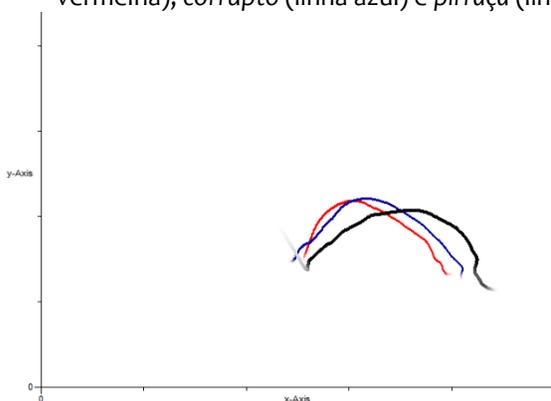
Figura 9 – Média dos contornos da língua das palavras *burrada* (linha vermelha), *corrupto* (linha azul) e *arrasta* (linha preta).



² A qualidade das imagens do dorso da língua não nos permite afirmar com precisão qual é a fricativa, uma vez que tanto a uvular quanto a faringal apresentam configurações do corpo da língua semelhantes, conforme Figuras 2 e 3.

A vogal precedente parece ter forte influência na produção da fricativa, como podemos ver na Figura 9, na qual apresentamos os contornos médios do rótico nas palavras *burrada* e *corrupto*, com realização velar, e *arrasta*, com uma fricativa mais recuada, possivelmente uvular. Ao analisarmos a palavra *pirraça*, não identificamos nenhuma fricção posterior, e sim uma elevação da parte anterior da língua, semelhante à configuração da vogal [i]³. Observe a Figura 10.

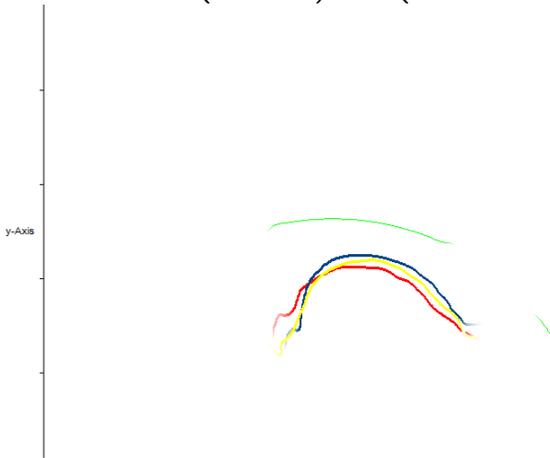
Figura 10 – Média dos contornos da língua das palavras *burrada* (linha vermelha), *corrupto* (linha azul) e *pirraça* (linha preta).



A informante B apresentou um comportamento distinto para a produção dos róticos, em início de palavra. Ao contrário da informante A, não produziu uma glotal. Podemos observar isso na Figura 11, na qual há um mesmo contorno para todos os róticos, independentemente da vogal seguinte. O mesmo ocorreu para os róticos intervocálicos. Pela articulação, julgamos tratar-se de uma fricativa velar.

³ Na palavra *irrita*, não encontramos qualidade de imagem suficiente para traçar o contorno da língua.

Figura 11 – Média dos contornos da língua das palavras *rato* (linha vermelha), *Rita* (linha azul) e *reta* (linha amarela).



As diferenças entre as informantes se devem provavelmente pelo fato de as fricativas posteriores não distinguirem significado no português. No caso da Informante B, não identificamos a influência vocálica na produção das fricativas posteriores, como ocorreu com a Informante A. Questões de ordem sociolinguística devem estar envolvidas na produção dos róticos, mas não foram alvo do presente estudo.

Considerações finais

A análise articulatória das fricativas posteriores tem se mostrado um caminho auxiliar para classificar essas fricativas. Como já apontaram vários estudos sociolinguísticos, no português brasileiro, há a produção de fricativas glotais, uvulares e velares, dependente da variável região. As imagens podem nos indicar com mais precisão qual fricativa foi produzida e, inclusive, mostrar outras realizações, como uma fricativa faríngea. Em particular, este estudo aponta também para a influência do contexto vocálico na produção das fricativas posteriores (Informante A). O mesmo já havia sido constatado em pesquisa anterior (HAUPT e SEARA, 2020).

Obviamente, os dados devem ser considerados ainda com cautela, uma vez que são em número reduzido. Um número maior de dados é necessário para confirmar se os padrões são recorrentes ou se são apenas características idiossincráticas. Além disso, um estudo

envolvendo informantes de diferentes regiões poderia acrescentar muito à discussão, no sentido de verificar o peso do papel de variáveis extralinguísticas na escolha da fricativa que ocupa a posição de rótico em contraposição com fatores linguísticos, como o contexto vocálico. Consideramos, no entanto, que esta investigação articulatória trouxe contribuições para os estudos sobre os róticos que são produzidos como fricativas posteriores.

CONFIGURAÇÃO GESTUAL DO TAP NA FALA DE PELOTENSES

Patrícia Pereira Melcheque
Giovana Ferreira-Gonçalves
Mirian Rose Brum-de-Paula

Introdução

No presente capítulo, busca-se, por meio de dados acústicos e articulatórios, caracterizar o tap [r] produzido por locutoras do português brasileiro, monolíngues, falantes do dialeto pelotense. Pretende-se investigar a configuração gestual na produção do segmento rótico em sílabas CV, CCV e CVC, bem como a natureza do elemento vocálico adjacente presente nas sílabas do tipo CCV e CVC.

Descrever acústica e articulatoriamente os segmentos róticos é tarefa sempre relevante para os estudos do português brasileiro, tendo em vista as formas variáveis com que são produzidos, bem como a discussão recorrente acerca de sua configuração gestual (SILVA, 2002; RECASENS, 2016; MELCHEQUE, FERREIRA-GONÇALVES e BRUM-DE-PAULA, 2018; ALBANO, 2020) e do elemento vocálico que o acompanha (SILVA, CLEMENTE e NISHIDA, 2006; MELCHEQUE, FERREIRA-GONÇALVES e BRUM-DE-PAULA, 2018; FERREIRA-GONÇALVES e BRUM-DE-PAULA, 2019; BILHARVA-DA-SILVA, 2019).

Em português, a posição do segmento rótico na sílaba é capaz não apenas de sofrer diferentes processos fonético-fonológicos, com a realização de formas variáveis, como também determina a emergência de diferentes unidades fonológicas, ou seja, de gestos articulatórios. Assim, em início de sílaba e início de palavra, temos um “r” forte, como em *raiz*; já em início de sílaba, mas em posição medial de palavra, um “r” fraco, como em *faro* (CÂMARA JR., 1977). Essa pronúncia mais branda do rótico também ocorre em outros contextos silábicos, como em encontros consonantais e posição final de sílaba, em coda medial e final.

Dependendo do dialeto, o rótico no português brasileiro pode ser articulado como: tepe alveolar, vibrante múltipla, retroflexa alveolar, fricativa velar vozeada ou desvozeada, fricativa glotal vozeada ou

desvozeada e aproximante alveolar. (SILVA *et al*, 2019). Por conta desses vários modos de articulação, é importante que esses segmentos sejam descritos de forma mais específica.

Descrições acústicas dos róticos em português podem ser encontradas em vários trabalhos, como Silva (1996, 2002), Barbosa e Madureira (2015), Silva, Clemente e Nishida (2006) e Silva *et al* (2019), dentre outros. Embora a análise acústica revele, por meio de inferências, aspectos ligados à articulação, a análise articulatória por intermédio do ultrassom lhe é complementar, possibilitando a visualização dos dados em tempo real. Além disso, em comparação com outras técnicas empregadas para a coleta de dados articulatórios, a ultrassonografia possui a vantagem de não inserir nenhum dispositivo no trato oral do informante, o que dá lugar a uma produção mais natural.

No Brasil, os estudos que utilizam a ultrassonografia como ferramenta metodológica para a descrição e análise de dados linguísticos ainda são poucos (FERREIRA-GONÇALVES e BRUM-DE-PAULA, 2013), mas podem ser encontrados em pesquisas voltadas para a aquisição de língua materna (VASSOLER, 2016; BARBIERI, 2017), sociofonética (ROSINSKI, 2017), descrição do português (SILVA, 2017) e terapias de fala (MELO *et al*, 2016; BARBERENA *et al*, 2020). No que concerne às pesquisas acerca da aquisição de línguas estrangeiras, o ultrassom também tem sido utilizado como ferramenta para: (i) descrição e análise do processo de aquisição de uma L2, como em Dias-Cavalheiro (2016) – que investigou a aquisição da vogal [a] do Espanhol, em contexto nasal, por falantes do português brasileiro – e Teixeira-Correa (2017) e Teixeira-Correa, Ferreira-Gonçalves e Brum-de-Paula (2017) – que investigaram a aquisição das vogais nasais do francês por aprendizes brasileiras; (ii) realização de atividades de instrução, como em Ferreira-Gonçalves, Pereira e Lemes (2019) – que investigaram a aquisição do segmento retroflexo do inglês /ɹ/ por falantes do português – e Silva-Garcia e Ferreira-Gonçalves (2019) – que estudaram a aplicabilidade do aparelho de ultrassom como recurso de instrução explícita no ensino de Espanhol como Língua Estrangeira (ELE).

Para o presente estudo, foi, assim, utilizada a técnica da ultrassonografia na captação de dados articulatórios relativos à produção dos róticos produzidos por falantes do dialeto pelotense.

Espera-se contribuir com a descrição articulatória do português brasileiro do extremo sul do Brasil, auxiliando nas pesquisas voltadas para terapias de fala, aquisição do português como língua materna e como língua estrangeira. As coletas realizadas passam também a contribuir para a constituição do banco de dados US LELO (Laboratório Emergência da Linguagem Oral).

Este capítulo está dividido em cinco partes: Introdução, Revisão teórica – apresentando uma breve descrição acústica e articulatória do *tap* no português brasileiro, incluindo discussões acerca da configuração gestual do segmento (simples ou complexo) e da natureza do elemento vocálico que acompanha a vogal nuclear –, Metodologia – com informações acerca dos informantes, do instrumento de coleta, dos equipamentos necessários para a coleta articulatória e dos parâmetros utilizados na análise dos dados –, Resultados e Conclusão.

Revisão teórica

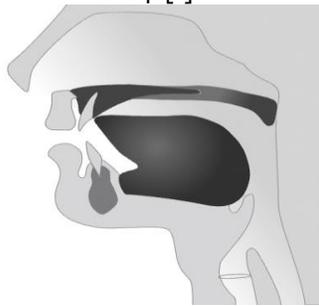
Com pontos e modos de articulação distintos, os sons de “r” – os chamados róticos – se apresentam com qualidades fonéticas diferentes. Esses sons podem apresentar de quatro a cinco pontos de articulação (alveolar/dental, pós-alveolar, retroflexo ou uvular) e vários modos de articulação (vibrantes, *taps*, fricativos, aproximantes). Nesse sentido, não há uma propriedade articulatória específica que consiga unificar todos os sons definidos como róticos (LINDAU, 1985). Há no alfabeto Fonético Internacional, para os sons representados com a letra R, sete símbolos fonéticos diferentes (WIESE, 2001). Ladefoged e Maddieson (1996) afirmam que o uso do termo rótico tem motivação na escrita, pois, de fato, as palavras grafadas com R representam as variações dos sons de “r”, comumente encontradas nos diferentes dialetos do português brasileiro. Os sons produzidos variam, assim, entre *taps*, vibrantes, aproximantes, fricativos e retroflexos.

No sul do Brasil, as variantes mais produzidas são o *tap* e a vibrante (MONARETTO, 1996). Ainda assim, o *tap* é o segmento mais encontrado quando em final de sílaba. De acordo com Botassini (2011), em Porto Alegre-RS, o *tap* apresenta frequência de 86% e a vibrante alveolar, 14%. Brescancini e Monaretto (2008), com base nos dados do

VARSUL, encontraram resultados semelhantes para a posição de coda silábica na capital gaúcha, com a predominância do *tap*. Na região sul do estado, em Pelotas/RS, os resultados de Melcheque, Brum-de-Paula e Ferreira-Gonçalves (2018) indicam que o rótico, presente em sílabas CV e CCV, também é predominantemente articulado como *tap*. Para a estrutura silábica CVC, no entanto, foram identificadas produções de *taps* e vibrantes.

Conforme a literatura da área, o *tap* se caracteriza como um som produzido a partir de um contato de curta duração entre a ponta da língua e o céu da boca. De modo mais preciso, a ponta/lâmina da língua toca os alvéolos durante a articulação desse segmento (LADEFOGED e MADDIESON, 1996). Na Figura 1, é possível visualizar o instante de máxima constrição desse gesto de ponta de língua. Durante sua produção, há uma breve oclusão do fluxo de ar na cavidade oral, enquanto o véu palatino levantado⁴ impede a passagem de ar pela cavidade nasal.

Figura 1 – Momento de máxima constrição do gesto de ponta de língua para o *tap* [r].



Fonte: Adaptado do site fonologia.org de CRISTÓFARO-SILVA; YEHA (2009)

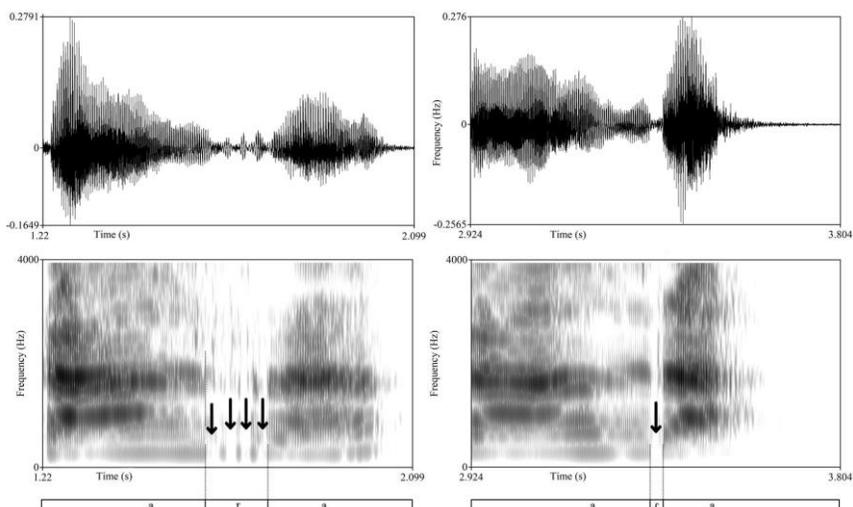
Dois momentos facultam a identificação do *tap* durante a análise acústica. O primeiro coincide com um espaçamento em branco no espectrograma, correspondente a uma ausência de energia promovida pelo bloqueio do ar no trato vocal. O segundo momento

⁴ No presente capítulo, assume-se, assim como Albano (2020), que o movimento de levantamento do véu palatino, na produção dos sons orais, caracteriza-se como um gesto articulatório. Assim, temos no véu palatino, para os sons orais, também um articulador.

coincide com o surgimento de um *burst* que representa uma rápida descarga de energia, causada pela soltura dos articuladores e consequente passagem do ar pelo trato (BARBOSA e MADUREIRA, 2015; e SILVA *et al.*, 2019).

Com duração em torno de 30 (ms), o *tap* é um dos segmentos com menor duração do português brasileiro; já uma vibrante múltipla, por exemplo, pode ter duração absoluta em torno de 200 (ms). A diferença é que o *tap* é realizado com apenas uma batida da língua na região alveolar, o chamado gesto balístico, pois é rápido e pontual. Há, assim, uma breve obstrução da passagem da corrente de ar, sendo que, na sequência, a língua já retoma sua posição articulatória para a produção do segmento vocálico. Para a vibrante, no entanto, há uma repetição dos gestos de fechamento e de abertura. Conforme Silva *et al* (2019), é justamente nos gestos de soltura da ponta da língua que surgem a produção de múltiplos elementos vocálicos. Na Figura 2, podem ser visualizadas as diferenças entre a vibrante e o *tap* aqui reportadas.

Figura 2 – Da esquerda para a direita, vibrante e *tap* na produção das sequências [ara] e [ara], respectivamente, da palavra “cará” por um falante pelotense

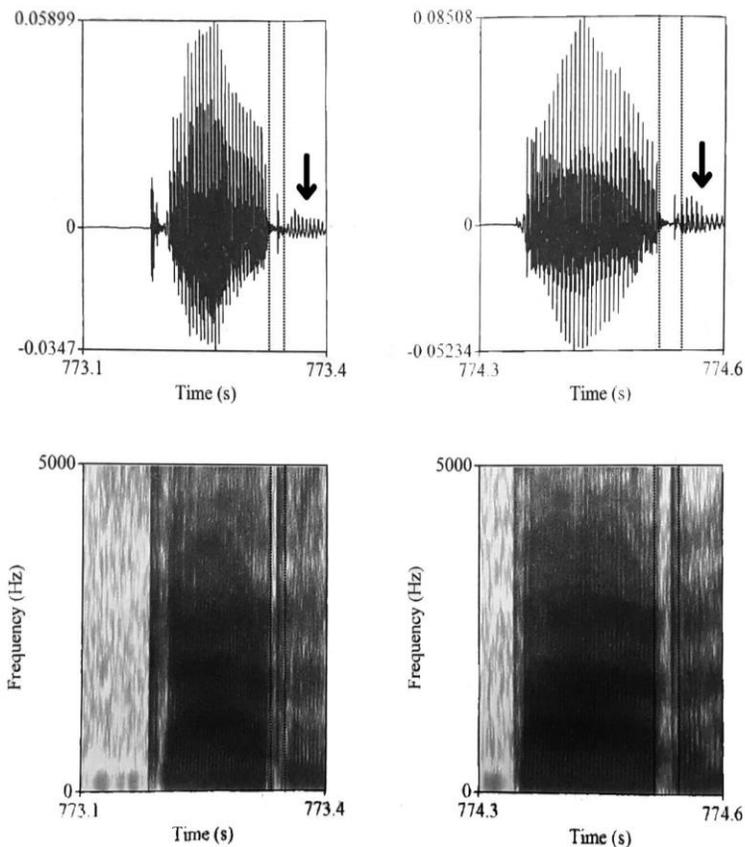


Fonte: as autoras

No oscilograma e no espectrograma da Figura 2, é possível observar as características acústicas apontadas por Silva *et al* (2019) na configuração dos dois segmentos róticos. A vibrante múltipla, com 143 (ms), apresenta repetidas ausências de energia – caracterizando o fechamento oral – seguidas de elementos vocálicos; no *tap*, com 32 (ms), pode-se visualizar igualmente a ausência de energia e uma barra de explosão. O elemento vocálico, ou a chamada vogal de apoio, não pode ser identificada devido ao posicionamento do *tap* em sílaba CV.

A Figura 3 apresenta a produção de um *tap*, em final de sílaba, seguido do elemento vocálico.

Figura 3 – Forma de onda e espectrograma da plosiva, vogal e tap das palavras *carta* e *par*.



Fonte: adaptado de Silva *et al.* (2019, p. 196)

É possível visualizar, na Figura 3, destacada pelas setas, a configuração acústica do elemento vocálico que ocorre logo após o *tap*. A descrição do elemento vocálico constatado na produção do *tap* no português brasileiro em sílabas CCV e CVC é essencial para a compreensão da coordenação dos gestos articulatórios em tais padrões silábicos. Conforme Silva, Clemente e Nishida (2006), a natureza do elemento vocálico é subordinada à posição silábica do *tap* na sílaba. Assim, na posição CCV, o elemento vocálico constitui parte da vogal nuclear; já em posição CVC, assume características de uma vogal neutra. Ao contrário, Silveira e Seara (2006) constataram que o elemento vocálico que antecede o *tap* em sílaba CCV não parece estar relacionado à vogal núcleo da sílaba, funcionando, também nessa posição, como uma vogal neutra. Em estudo mais recente, Bilharva-da-Silva (2019) propõe que, tanto em sílaba CCV como em CVC, o elemento vocálico que acompanha o *tap* é, na verdade, uma porção da vogal nuclear que emerge da sobreposição do *tap*.

Outro aspecto a ser destacado é a constituição do *tap* como um segmento simples ou complexo. Para Silva *et al* (2006) e Barberena, Soares e Berti (2014), o *tap* é caracterizado como um segmento de difícil aquisição por apresentar dois gestos articulatórios distintos, sendo assim, como um segmento complexo, composto por movimento de ponta de língua e anteriorização de dorso. Em estudo mais recente, no entanto, Recasens (2016) questiona essa configuração de gestos, propondo que o movimento de retração do corpo da língua e o abaixamento do pré-dorso – constatados na produção do *tap* por meio de inspeção ultrassonográfica – são apenas movimentos decorrentes da constrição alveolar e da configuração de língua necessária para a produção do *tap*, mas não gestos em si, intencionais. Desse modo, esse seria considerado como segmento simples, constituído de gestos complexos. Segundo o autor, o som é articulado com uma posição do corpo da língua próximo da neutralidade, subordinado à coarticulação com a vogal.

Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2018) apresentam dados da variedade pelotense que evidenciaram constrições concomitantes no dorso e na ponta da língua durante a produção do *tap*. Esses resultados vão ao encontro da tese de Recasens (2016), na qual o autor sustenta que esse segmento pode ser

articulado com um único gesto capaz de formar duas constrictões, sendo assim, um segmento simples.

Albano (2020), em acordo com Recasens (2016), destaca haver evidências articulatórias claras para identificar o quão complexo um gesto unitário pode ser, sendo a resistência à coarticulação um de seus argumentos sólidos. Ainda, a autora afirma que não é adequado reproduzir pautas gestuais abrangentes – como apresentado em Albano (2001) – para formalizar a articulação dos róticos, uma vez que essas estão defasadas em face das informações atuais sobre as variações das configurações articulatórias desses segmentos. Sendo assim, para formalizar a complexidade dos gestos unitários, que constituem segmentos simples, a autora apresenta um gráfico de acoplamento hipotético para a vibrante alveolar, conforme a Figura 4.

Figura 4 – Gráfico de acoplamento hipotético para a vibrante alveolar como um gesto unitário complexo (ALBANO, 2020, p. 131)

Ponta da língua aproximação, com voz soante e vibração autossustentada



Dorso da língua Vogal

Para a construção desse gráfico, a autora parte da tese proposta por Recasens (2016), uma vez que há apenas o gesto de ponta de língua associado à produção da vibrante. A vogal está em relação de antifase com a vibrante, seguindo Recasens e Espinosa (2009) e Recasens e Rodriguez (2016), evidenciando uma forte resistência à coarticulação entre a vibrante e a vogal. Importante destacar que a autora associa descritores sensoriais ao gesto de ponta de língua, como *com voz soante* – o que abarca a naturalidade do vozeamento espontâneo das soantes – e *vibração autossustentada* – o que expressa o *efeito de Bernoulli que sustenta a vibração por toda a duração da consoante* (ALBANO, 2020, p. 129).

Observe-se, pois, que as interpretações acerca dos gestos que constituem os róticos nem sempre convergem, o que aponta a necessidade do uso de técnicas diversas para que se possa melhor

apreender gestos tão peculiares: mutantes e, frequentemente, breves e velozes.

Metodologia

Na investigação aqui desenvolvida, foram realizadas coletas de dados orais e ultrassonográficos de seis informantes do sexo feminino falantes do português brasileiro, monolíngues, com idade entre 18 e 35 anos. Todas as informantes estavam cursando o ensino superior e eram naturais da cidade de Pelotas, localizada ao Sul do Brasil. O baixo índice de massa corporal também foi um critério considerado, pois, segundo Stone (2005), tal aspecto possibilita uma maior qualidade das imagens ultrassonográficas obtidas.

Os dados foram coletados em uma cabine acústica, localizada no Laboratório Emergência da Linguagem Oral (LELO), na Universidade Federal de Pelotas, utilizando um gravador digital, modelo Zoom H4n – para a obtenção de dados acústicos – e um aparelho de ultrassom, modelo Mindray DP-6600, e o software *Articulate Assistant Advanced (AAA)*⁵ – para a obtenção de dados articulatórios. Os sujeitos utilizaram um estabilizador de cabeça⁶ desenvolvido estritamente para esse tipo de análise.

Figura 5 – Equipamentos para coletas de dados de ultrassom e capacete para estabilização dos movimentos de cabeça



⁵ Software utilizado para coleta e análise de dados articulatórios desenvolvido pela *Articulate Instruments*: <http://www.articulateinstruments.com/aaa/>.

⁶ O equipamento foi desenvolvido e é comercializado pela *Articulate Instruments* (<http://www.articulateinstruments.com/ultrasound-imaging/>).

Os informantes produziram doze palavras – apresentadas por meio de imagens –, nas quais o *tap*, em sílaba tônica, ocupava a posição de CV, CVC e CCV em contexto de [a], [i] e [u]. As palavras foram produzidas na frase veículo “*Eu digo – palavra – pra você*”.

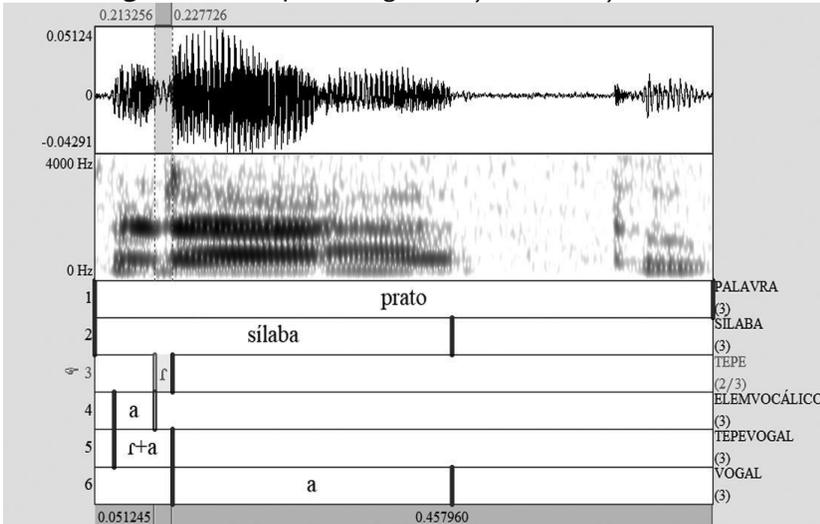
Quadro 1 – Palavras-alvo do instrumento de coleta

Contextos	[a]	[i]	[u]
onset medial	barata	siri	sururu
onset complexo	prato	frita	frutas
coda medial	carta	circo	curto
coda final	mar	vir	cur

Os itens lexicais foram repetidos aleatoriamente por cinco vezes. Também foram inseridas seis palavras distratoras, para que o informante não identificasse o som que estava sendo investigado, evitando, assim, uma possível fala monitorada.

Após a coleta, os dados foram segmentados e rotulados no *software* PRAAT, com a utilização de *Textgrids*. Para a segmentação do dado acústico, foram delimitadas (i) as bordas inicial e final das sílabas, do *tap* e do segmento vocálico que acompanha o *tap*, bem como (ii) a duração do *tap* mais o segmento vocálico. Para a delimitação do *tap*, foram considerados aspectos da forma de onda e do espectrograma, os quais caracterizam uma diminuição da energia do sinal acústico, devido à interrupção da passagem de ar quando do contato dos articuladores. Por fim, foram extraídos os valores de duração da palavra, do segmento rótico e do elemento vocálico. Na Figura 6, observam-se a segmentação e a rotulação do alvo *prato*.

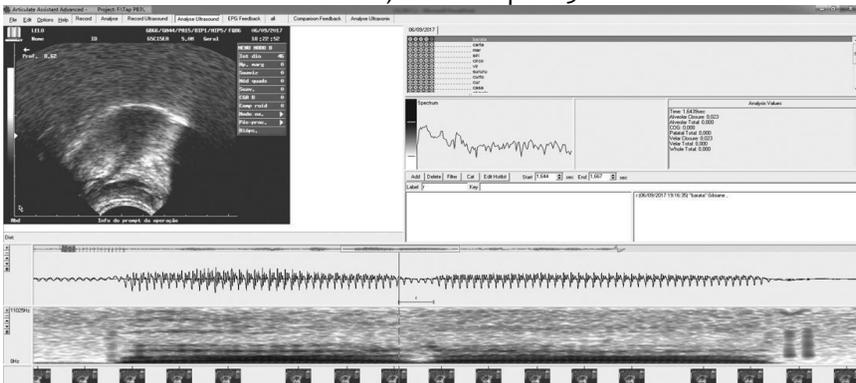
Figura 6 – Exemplo de segmentação e rotulação dos dados



Para palavras com o *tap* em sílaba CCV e CVC, foi controlada, ainda, a duração da vogal núcleo da sílaba. Com o controle da vogal nuclear adjacente ao *tap*, esperava-se encontrar resultados acerca da caracterização do elemento vocálico.

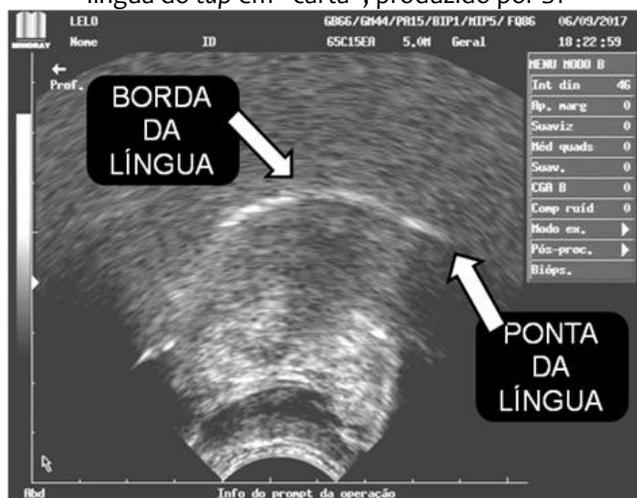
A análise dos dados articulatórios ocorreu por meio do software *Articulate Assistant Advanced* (AAA), conforme exemplo disposto na Figura 7.

Figura 7 – Software AAA apresentando imagem ultrassonográfica contendo o ápice do movimento de ponta de língua na produção do *tap* na palavra “barata”, realizado por S5



No ultrassom, a imagem do contorno da língua é formada por meio da reflexão de ondas sonoras de alta frequência produzidas por um transdutor constituído de materiais piezoelétricos. A imagem ultrassonográfica é gerada quando há diferença de impedância acústica entre os tecidos, assim, quanto maior é essa diferença, maior é a intensidade da reflexão. É o que ocorre quando o feixe encontra o ar presente entre a língua e o palato. Na Figura 8, pode ser observada a borda da língua (em branco) em um dos *frames* gerados durante a produção da palavra “carta”.

Figura 8 – Frame do gesto de máxima constrição do movimento de ponta de língua do *tap* em “carta”, produzido por S1



A utilização do ultrassom, na pesquisa voltada a dados linguísticos, requer, assim como outras técnicas, cuidados especiais em relação à coleta do dado articulatório. No caso do *tap*, por conter um gesto balístico envolvendo o ápice ou lâmina da língua como articulador ativo, há dificuldades na captura da imagem da ponta da língua pela sonda transdutora, em função da sombra da mandíbula. A baixa taxa de *frames* fornecida pelo aparelho também pode dificultar a sincronização do áudio e da imagem. Nesse sentido, ajustes relativos a alguns parâmetros devem ser realizados, como profundidade (controle da distância entre a região submandibular e o palato), frequência da sonda (quanto maior a frequência, menor a profundidade e mais nítida a imagem obtida), ganho (permite regular

a intensidade da imagem em diferentes regiões) e posição do foco (maior nitidez na região selecionada para o foco). Tais parâmetros foram ajustados antes do início de cada coleta realizada.

Para a análise dos dados articulatórios, foram realizados: (i) o traçado da borda da língua do *frame* contendo a máxima constrição do movimento de ponta de língua do *tap*; (ii) os traçados de todos os *frames* presentes nas sequências *V1+ tap* e *tap + V2* na palavra “barata”. Objetiva-se, nesse caso, verificar a presença de um ou mais gestos articulatórios na produção do segmento, corroborando, assim, os resultados constatados por Silva *et al* (2006) ou por Recasens (2016); (iii) o traçado do ponto médio do elemento vocálico presente em sílabas CCV e CVC; (iv) o traçado do ponto médio dos 25% iniciais da vogal nuclear em sílaba CV.

Resultados

Em relação à análise acústica dos dados, realizada com o *software Praat*, foram obtidos os resultados dispostos nos Gráficos 1, 2 e 3.

Gráfico 1 – Médias de duração absoluta do elemento vocálico, do rótico e da vogal núcleo em sílabas CCV

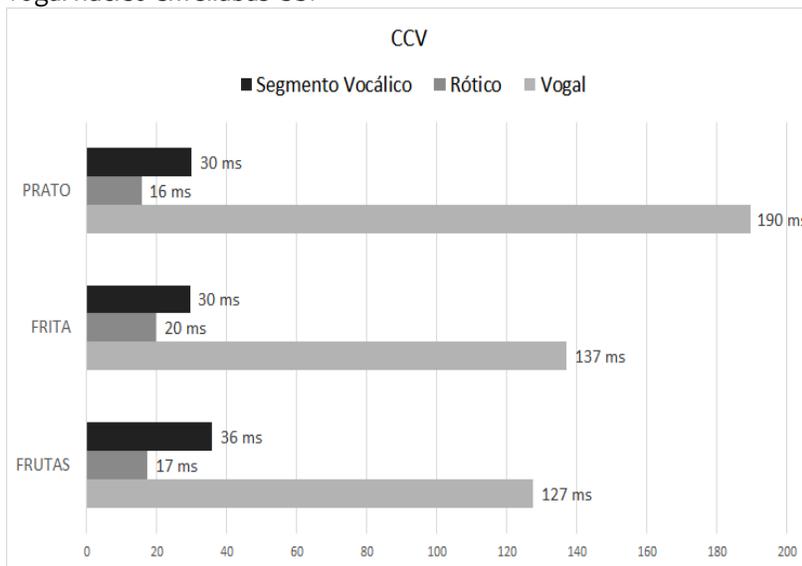


Gráfico 2 – Médias de duração absoluta do elemento vocálico, do rótico e da vogal núcleo em sílabas CVC medial

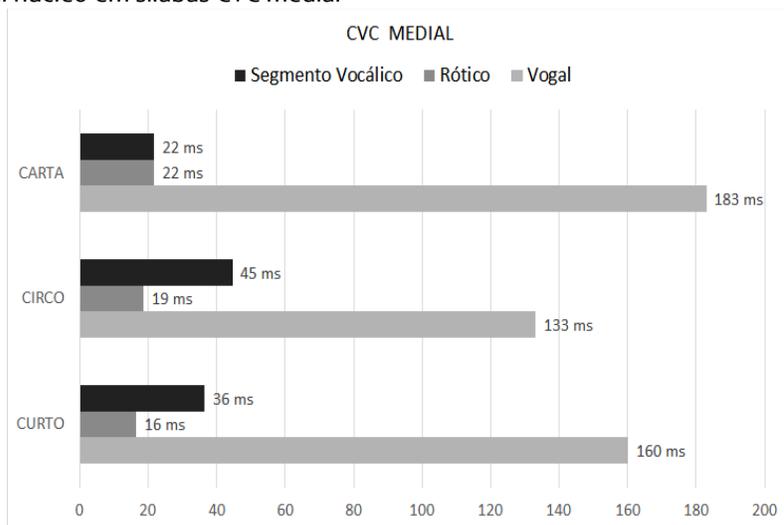
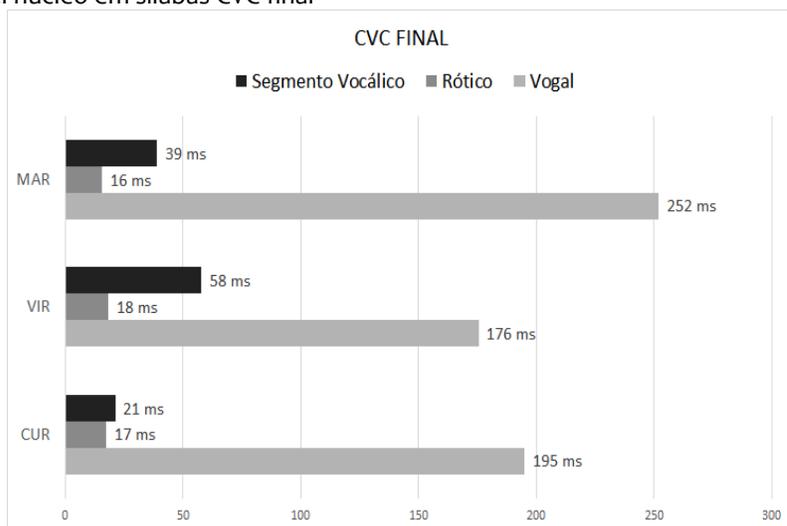


Gráfico 3 – Médias de duração absoluta do elemento vocálico, do rótico e da vogal núcleo em sílabas CVC final



Os resultados apontam o padrão de *tap* em sílabas CCV – médias entre 16 (ms) e 28 (ms) –, com uma possível ausência do papel da vogal

na duração do rótico. O elemento vocálico apresentou-se sensivelmente mais longo quando a vogal nuclear era alta e posterior.

Já no que concerne às posições CVC medial e CVC final, foram encontrados casos de vibrante múltipla⁷, principalmente reportados para as produções de S4 e S3. Em relação ao elemento vocálico, observaram-se diferenças mais expressivas na posição de coda, com uma maior duração quando a vogal nuclear também era alta, no entanto, anterior. Em posição medial, a segunda maior duração do elemento vocálico foi constatada com a vogal nuclear alta e posterior; já em posição de coda final, com a vogal baixa. A falta de um padrão mais homogêneo do elemento vocálico nos dois contextos de coda pode sinalizar para a ausência de relação entre a sua natureza e a da vogal nuclear.

No vocábulo “carta”, produzido pelas informantes S3 e S4, e no vocábulo “mar”, produzido pelas informantes S3, S4 e S6, todas as repetições foram caracterizadas como vibrantes pela análise acústica. A inspeção do espectrograma de banda larga no PRAAT leva à identificação de múltiplas batidas durante a produção do rótico.

Quadro 2 – Média de duração absoluta da palavra e do rótico nas palavras “carta” e “mar” das informantes S3, S4, e S6

Sujeitos	carta		mar	
	Duração da palavra	Duração do rótico	Duração da palavra	Duração do rótico
S3	543 (ms)	91 (ms)*	413 (ms)	62 (ms)*
S4	532 (ms)	68 (ms)*	448 (ms)	52 (ms)*
S6	573 (ms)	26 (ms)	512 (ms)	125 (ms)*

*Valores de duração relativos às produções de vibrantes

Como pode ser observado, o rótico do vocábulo “mar” foi produzido com uma média de duração entre 52 (ms) e 125 (ms). O rótico da palavra “carta” foi produzido com um valor médio de 68 (ms) a 91 (ms) pelas informantes S3 e S4.

A aplicação do teste estatístico Wilcoxon (SPSS v. 17.0) identificou diferenças marginalmente significativas entre a média de duração do rótico, em contexto da vogal alta anterior, quando comparadas as

⁷ Importante reportar que, para as médias dispostas nos Gráficos 1, 2 e 3, apenas os valores de duração do tap foram considerados.

seguintes estruturas: CV e CCV ($z = -1,753$, $p = 0,080$); CV e CVC medial ($z = -2,032$, $p = 0,042$) e CV e CVC final ($z = -1,753$, $p = 0,080$). Esses resultados indiciam um papel da estrutura silábica na duração do rótico: CV > CCV, CVC medial e final.

Quando comparadas as durações absolutas e relativas do rótico, em uma mesma estrutura silábica, nos diferentes contextos vocálicos – [a], [i] e [u] –, por meio da aplicação do teste *Freedman*, não foram constatadas diferenças significativas. Infere-se, portanto, uma ausência do papel da vogal nessas produções.

Em relação ao papel do contexto vocálico na duração absoluta e na duração relativa do elemento vocálico, a aplicação do teste estatístico *Wilcoxon* identificou diferença significativa entre as produções de CC[i] e CC[u] ($z = -2,023$, $p = -0,043$). O resultado, nesse caso, pode sinalizar para a natureza do elemento vocálico, quando em sílaba CCV, em assonância com a vogal nuclear da sílaba.

Os resultados acústicos, concernentes às medidas de duração do rótico, apontam, portanto, a produção de *taps* em sílabas CV e CCV no dialeto pelotense, no entanto, é recorrente a presença de vibrantes em sílaba CVC. Achados similares foram encontrados por Bilharva-da-Silva (2019). Com a produção predominante do *tap*, passa-se ao segundo ponto de análise aqui delineado, ou seja, a constituição do segmento, se formado por um ou dois gestos articulatórios.

Para essa discussão, serão apresentados os resultados articulatórios por meio das imagens ultrassonográficas. Nas Figuras 9 a 11, podem ser observados os traçados do contorno da borda da língua para a sequência V1+*tap*; ápice do *tap* e sequência V2+*tap*.

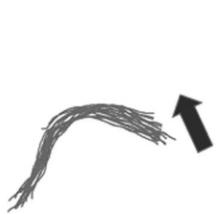


FIGURA 9: Splines da sequência V1+*tap* em “barata_2 e 3” de S1.

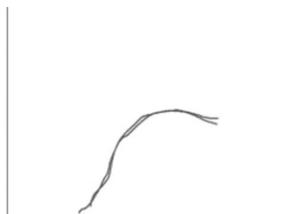


FIGURA 10: Splines do ápice do *tap* em “barata_2 e 3” da S1.



FIGURA 11: Splines da sequência *tap*+V2 em “barata_2 e 3” da S1.

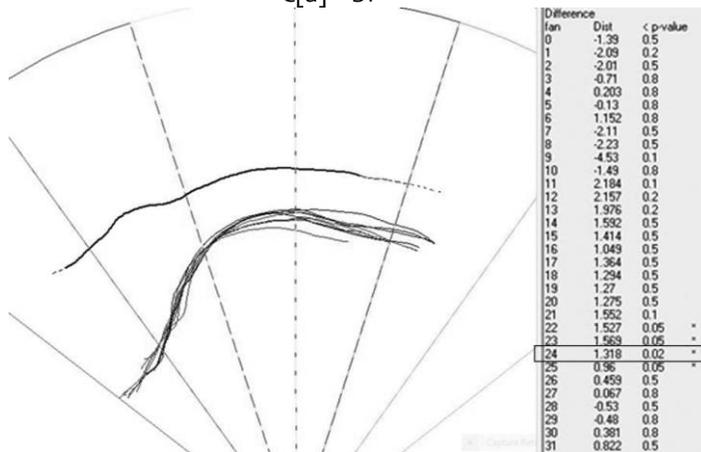
Na Figura 9, a análise articulatória evidencia a presença de gesto da ponta da língua da produção do *tap*, conforme previsto pela literatura. A Figura 10 evidencia, ainda, que os movimentos de retração do dorso e elevação do corpo ocorrem de forma concomitante à elevação da ponta na sequência “ar”; já na Figura 11, ao contrário, pode ser constatado o abaixamento do corpo concomitante ao abaixamento da ponta na sequência “ra”, como evidenciado no estudo de Recasens (2016). Assim, ainda que a inspeção articulatória tenha sido preliminar – já que reporta apenas os dados produzidos por uma informante, e em contexto vocálico específico – os resultados articulatórios parecem indiciar a presença de apenas um gesto na produção do *tap*, o de ponta de língua.

Em relação ao elemento vocálico, a diferença estatística relativa a sua duração em contexto CCV com as vogais nucleares [i] e [u] pode indiciar a relação entre a natureza da vogal núcleo da sílaba e do elemento vocálico. Os resultados são, obviamente, pouco conclusivos. Procedeu-se, então, à inspeção dos dados articulatórios.

Primeiramente, considerando a hipótese de que, em sílaba CCV, a vogal nuclear é entrecortada pelo *tap*, foi realizada uma comparação entre o traçado da borda da língua do ponto médio do elemento vocálico com o traçado da borda da língua do ponto médio da porção inicial – 25% iniciais – da mesma vogal em sílaba CV. A hipótese tecida é que, ao entrecortar a vogal núcleo da sílaba CCV, o elemento vocálico à esquerda do *tap* corresponderia aproximadamente à porção inicial de uma vogal em sílaba CV.

Os resultados articulatórios indiciam que o elemento vocálico se comporta como parte da vogal eclipsada em sílaba CCV, corroborando os resultados acústicos obtidos por Silva, Clemente e Nishida (2006) e Bilharva-da-Silva (2019). A análise estatística dos dados articulatórios foi realizada por meio de um teste-T rodado no *software* AAA.

Figura 12 – Teste-T das médias dos *frames* do elemento vocálico em sílaba CC[a] e das médias dos *frames* dos 25% iniciais da vogal nuclear em sílaba C[a] – S1



Na Figura 12, é possível visualizar a comparação, por meio da realização de um Teste-T, entre as médias dos *frames* do elemento vocálico – presente nas produções da palavra “prato” realizadas por S1 – e as médias dos *frames* da porção inicial (25%) da vogal nuclear da palavra “barata”, igualmente produzida por S1. Como pode ser observado, foi constatado apenas um ponto de diferença significativa entre os traçados ($p = 0.002$), o que indicia a mesma natureza entre o elemento vocálico e a vogal núcleo da sílaba.

Na sequência, procedeu-se, então, à análise dos traçados da borda da língua do elemento vocálico e da porção final da vogal nuclear em sílabas CVC. Considerando a hipótese, em acordo com Bilharva-da-Silva (2019), de que, em sílaba CVC, a vogal nuclear também é entrecortada pelo *tap*, foi realizada uma comparação entre o traçado da borda da língua do ponto médio do elemento vocálico à direita do rótico com o traçado da borda da língua do ponto médio da porção final – 25% finais – da mesma vogal em sílaba CV.

Em sílabas CVC, no entanto, o elemento vocálico parece se comportar de maneira distinta, conforme pode ser constatado nas Figuras 13 e 14.

Figura 13 – Teste-T das médias dos *frames* do elemento vocálico em sílaba C[a]C medial (apontado pela seta) e das médias dos *frames* dos 25% finais da vogal nuclear em sílaba C[a] – S1

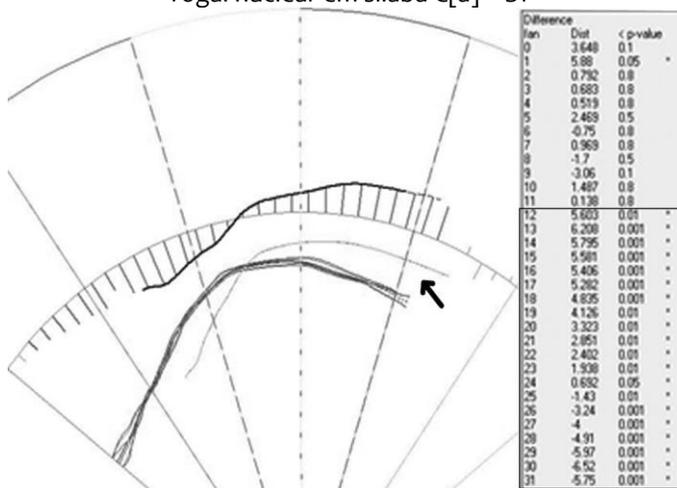
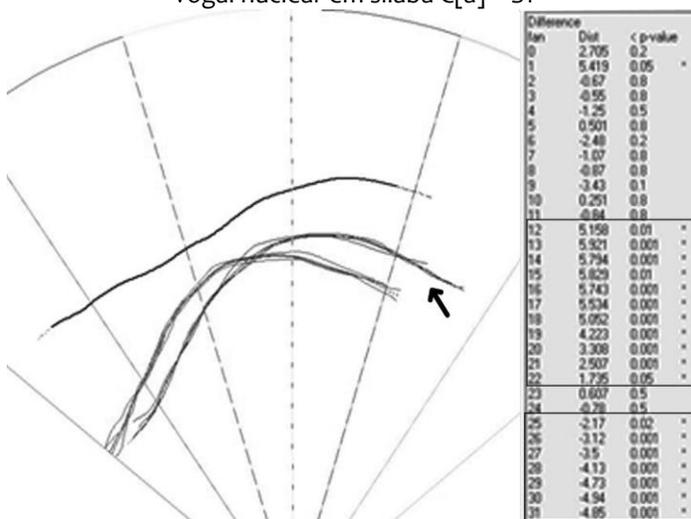


Figura 14 – Teste-T das médias dos *frames* do elemento vocálico em sílaba C[a]C final (apontado pela seta) e das médias dos *frames* dos 25% finais da vogal nuclear em sílaba C[a] – S1



As médias dos *frames* do elemento vocálico da palavra “carta” apresentam vários pontos de diferenças significativas quando

comparadas às médias dos *frames* da porção final (25%) da vogal nuclear da palavra “barata”, como pode ser constatado nos traçados da borda da língua expostos na Figura 13 e nos valores de *p*, sinalizados com asteriscos, no quadro que acompanha a referida Figura. Em posição final, conforme pode ser visualizado na Figura 14, o resultado é semelhante, ou seja, há vários pontos de diferenças significativas quando a mesma comparação é feita com as médias dos *frames* do elemento vocálico da palavra “mar”. Os resultados, nesse caso, indiciam que a natureza do elemento vocálico em sílabas CVC medial e final difere da natureza da vogal nuclear da sílaba.

A análise articulatória dos dados ultrassonográficos evidenciou, assim, similaridades entre o elemento vocálico e a vogal nuclear em sílaba CCV, sinalizando para a ocorrência de uma vogal eclipsada, ou seja, entrecortada pelo *tap*. Já em sílabas CVC, o elemento vocálico apresentou diferenças em relação à vogal nuclear. Os resultados aqui obtidos corroboram os constatados, por meio de análise acústica, por Silva, Clemente e Nishida (2006).

Conclusão

Com esta pesquisa, foi possível discorrer sobre padrões acústicos e articulatórios dos segmentos róticos produzidos por falantes pelotenses, os quais, em sua maior parte, seguem aspectos previstos pela literatura da área. Constatações acerca da produção de vibrantes múltiplas, por parte das informantes, em posição de coda final e medial devem ser destacadas. A análise acústica apontou que o rótico se apresenta como *tap* em sílabas CV e CCV, e como *tap* e vibrante em sílabas CVC.

A inspeção articulatória indicou que o *tap* é constituído por apenas um gesto articulatório, de ponta de língua, sendo as demais configurações gestuais decorrentes do movimento de ponta, conforme proposta de Recasens (2016), ou seja, temos um único gesto composto de movimentos complexos. Os dados articulatórios também parecem corroborar a proposta de Silva, Clemente e Nishida (2006) em relação à natureza do elemento vocálico que acompanha o *tap*, ou seja, apresenta a mesma natureza da vogal nuclear em sílabas CCV, mas comporta-se como uma vogal neutra em sílabas CVC medial e final.

Referências

- ABAURRE, M. B. M.; SANDALO, M. F. S. Os róticos revisitados. In: HORA, D. da; COLLISCHONN, G. *Teoria Linguística: fonologia e outros temas*. João Pessoa: Editora Universitária, 2003, p. 144-180.
- ADLER-BOCK, M.; BERNHARDT, B. M.; BACSFALVI, P.; GICK, B. Perceptual, acoustic, and tongue shape measures during /r/ production pre- and post-treatment using visual feedback from ultrasound: case studies of two adolescents. *The Journal of the Acoustical Society of America*, v. 117, n. 4, p. 2605–2605, 2005.
- ADLER-BOCK, M.; BERNHARDT, B.; GICK, B.; BACSFALVI, P. The use of ultrasound in remediation of North American English /r/ in two adolescents. *American Journal of Speech-Language Pathology*, v. 16, n. 2, p. 128–139, 2007.
- AHN, S. Tongue root contributions to voicing in utterance-initial stops in American English. *Proceedings of Meetings on Acoustics*, 2015, v. 25, n. 1, 2017.
- ALBANO, E. C. *O gesto e suas bordas: esboço de Fonologia Acústico-Articulatória do Português*. Campinas: Mercado de Letras, 2001.
- _____. Uma introdução à dinâmica em fonologia, com foco nos trabalhos desta coletânea. *Revista da ABRALIN*, v. 11, n. 1, P. 1–30, 2012.
- _____. *O gesto audível: fonologia como pragmática*. São Paulo: Cortez, 2020.
- ASHA. Speech Sound Disorders-Articulation and Phonology. *Clinical Topics*, 2014.
- BACSFALVI, P. Attaining the lingual components of /r/ with ultrasound for three adolescents with cochlear implants. *Revue canadienne d'orthophonie et d'audiologie*, v. 34, n. 3, p. 206-207, 2010.
- BACSFALVI, P.; BERNHARDT, B. M. Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: ultrasound and electropalatography. *Clinical Linguistics & Phonetics*, v. 25, n. 11-12, p. 1034-1043, 2011.

BAHGMAN, K.; ZARIFIAN, T.; ADIBI, A.; SHATI, M.; DERAKHSHANDEH, F. The quantitative ultrasound study of tongue shape and movement in normal Persian speaking children. *International Journal of Pediatric Otorhinolaryngology*, 134, 2020.

BAKER, E.; WILLIAMS, A. L.; MCLEOD, S.; MCCAULEY, R. Elements of Phonological Interventions for Children With Speech Sound Disorders: The Development of a Taxonomy. *American Journal of Speech-language Pathology*, v. 27 n. 3, p. 906-935, 2018.

BALL, M. J.; GRACCO, V.; STONE, M. A Comparison of imaging techniques for the investigation of normal and disordered speech production. *Advances in Speech Language Pathology*, v. 3, n. 1, p. 13-24, 2001.

BALL, M.J.; KENT, R.D. *The new phonologies. Developments in clinical Linguistics*. Wisconsin: Singular publishing Ltd, 1997.

BALLARD, K. J.; ROBIN, D. A.; MCCABE, P.; MCDONALD, J. A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, v. 53, n. 5, p. 1227–1245, 2010.

BARANOWSKI, M. Sociophonetics. In.: BAYLEY R.; CAMERON R.; LUCAS C. (Orgs.) *The Oxford Handbook of Sociolinguistics*, Oxford: Oxford University Press, 2013, p. 403-424.

BARBERENA, L. S.; BRASIL, B. de C.; MELO, R. M.; MEZZOMO, C. L.; MOTA, H. B.; KESKE-SOARES, M. Ultrasound applicability in Speech Language Pathology and Audiology. *Revista Cogas*, v. 26, n. 6, p. 520-530, 2014.

BARBERENA, L. S.; ROSADO, I. M.; MORAES, D. A. de O.; BERTI, L. C.; KESKE-SOARES, M. Ultrasonography of tongue movements for the /l/ sound and characterization by gender in children with typical and atypical speech development. *Revista CEFAC*, v. 22, p. 1-10, 2020.

BARBARENA, L. S.; SOARES, M. K.; BERTI, L.C. Descrição dos gestos articulatórios envolvidos na produção dos sons /r/ e /l/. *Audiology Communication Research*, v. 19, n. 4, p. 338-344, 2014.

BARBIERI, T. T. *Aquisição de encontros consonantais com tap no português brasileiro: análises acústica e articulatória*. Dissertação de Mestrado. Programa de Pós-Graduação em Letras. Centro de Letras e Comunicação, Universidade Federal de Pelotas, Pelotas, 2019.

- BARBOSA, P. A.; MADUREIRA, S. *Manual de fonética acústica experimental: aplicações a dados do português*. São Paulo: Cortez, 2015.
- BARLOW, D. H.; HERSEN, M. *Single case experimental designs: strategies for studying behavior change*. New York: Pergamon Press, 1982.
- BAUMAN-WAENGLER, J. *Articulatory and phonological impairments: A clinical focus*. Boston: Allyn and Bacon, 2004.
- BAYLEY, R.; CAMERON, R.; LUCAS, C. (eds.). *The Oxford Handbook of Sociolinguistics*. Oxford: Oxford University Press, 2013.
- BERNHARDT, B. Developmental implications of nonlinear phonological theory. *Clinical Linguistics & Phonetics*, v. 6, n. 4, p. 259-281, 1992.
- BERNHARDT, B.; BACSFALVI, P.; ADLER-BOCK, M.; SHIMIZU, R.; CHENEY, A.; GIESBRECHT, N.; O'CONNELL, M.; SIRIANNI, J.; RADANOV, B. Ultrasound as visual feedback in speech habilitation: Exploring consultative use in rural British Columbia, Canada. *Clinical Linguistics & Phonetics*, v.22, p.149-162, 2008.
- BERNHARDT, B.; GICK, B.; BACSFALVI, P.; ADLER-BOCK, M. Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics & Phonetics*, v. 19, n. 6/7, p. 605-617, 2005.
- BERNHARDT, B.; GICK, B.; BACSFALVI, P.; ASHDOWN, J. Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clinical Linguistics & Phonetics*, v. 17, n. 3, p. 199-217, 2003.
- BERTI, L. C. *Aquisição incompleta entre /s/ e /'int'/ em criança falantes do português brasileiro*. Tese de Doutorado. Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, 2006.
- _____. Investigaç o da produç o de fala a partir da ultrassonografia do movimento de l ngua. In: *18^o congresso brasileiro de fonoaudiologia: exerc cio profissional: bases te ricas, avanço e realidade nacional*, sociedade brasileira de fonoaudiologia, v.1, 2010, p. 661-1.
- BERTI, L. C.; BOER, G. D.; BRESSMANN, T. Tongue displacement and durational characteristics of normal and disordered Brazilian Portuguese liquids. *Clinical linguistics & phonetics*, v. 30, n. 2, p. 131-149, 2016.

BERTI, L. C.; MARINO, V. C. C. Marcas linguísticas constitutivas do processo de aquisição do contraste fônico. *Revista do GEL*, v. 5, n. 2, p. 103-21, 2008.

BERTI, L. C.; PAGLIUSO, A.; LACAVA, F. Instrumento de avaliação de fala para análise acústica (IAFAC) baseado em critérios linguísticos. *Revista da Sociedade Brasileira de Fonoaudiologia*, v. 14, n. 3, p. 305-314, 2009.

BERTHOZ, A. *Le sens du mouvement*. Paris: Odile Jacob, 1997.

BILHARVA-DA-SILVA, F. *O contato português-pomerano na produção dos grupos [cr] e [rc]: o caso das vogais suarabáticas*. Tese de Doutorado. Programa de Pós-Graduação em Letras. Escola de Humanidades. Pontifícia Universidade Católica do Rio Grande do Sul, 2019.

BIRD, S.; GICK, B. Ultrasound biofeedback in pronunciation teaching and learning. In: *ISAPh 2018 – International Symposium on Applied Phonetics*, 2018, p. 5-11.

BOTASSINI, J. O. M. A variação no uso dos róticos em Porto Alegre. *Estudos Linguísticos*, v. 40, n. 2, p. 1060-1072, 2011.

BOYCE, S. E. The Articulatory Phonetics of /r/ for Residual Speech Errors. *Seminars in Speech and Language*, v. 36, n. 4, p. 257-270, 2015.

BRASIL, B. C.; MEZZOMO, C. L.; MOTA, H. B.; MELO, R. M.; LOVATTO, L.; ARZENO, L. Características acústicas dos fones [s] e [ʃ] de adultos e crianças com desenvolvimento fonológico típico. *Revista da Sociedade Brasileira de Fonoaudiologia*, v. 17, n. 2, p. 182-188, 2012.

BRAUN, M. The Photographic work of W. J. Marey. *Studies in Visual Communication*, v. 9, n. 4, 1983.

BRESCANCINI, C. R.; FOULKES, P. (Orgs.) *Letras de Hoje*, v. 52, n. 1, 2017.

BRESCANCINI, C. R.; MONARETTO, V. Os róticos no Sul do Brasil: panorama e generalizações. *SIGNUM: Estudos Linguísticos*, v. 11, n. 2, p. 51-66, 2008.

BRESSMANN, T. An ultrasonographic study of lingual contortion speech. *Journal of Phonetics*, v. 25, p. 1028-1033, 2011.

BRESSMANN, T.; FLOWERS, H.; WONG, W.; IRISH, J. C. Coronal view ultrasound imaging of movement in different segments of the tongue during paced recital: Findings from four normal speakers and a speaker

with partial glossectomy. *Clinical Linguistics & Phonetics*, v. 24, n. 8, p. 589-601, 2010.

BRESSMANN, T.; HARPER, S.; ZHYLICH, I.; KULKARNI, G. V. Perceptual, durational and tongue displacement measures following articulation therapy for rhotic sound errors. *Clinical linguistics & phonetics*, v. 30, n. 3-5, p. 345-362, 2016.

BRESSMANN, T.; THIND, P.; Uy, C.; BOLLIG, C.; GILBERT, R. W.; IRISH, J. Quantitative three-dimensional ultrasound analysis of tongue protrusion, grooving and symmetry: Data from 12 normal speakers and a partial glossectomee. *Clinical Linguistics and Phonetics*, v. 19, p. 573-588, 2005.

BROEN, P. A.; STRANGE, W.; DOYLE, S. S.; HELLER, J. H. Perception and Production of Approximant Consonants by Normal and Articulation-Delayed Preschool Children. *Journal of Speech, Language, and Hearing Research*, v. 26, n. 4, p. 601-608, 1983.

BROWMAN, C. P.; GOLDSTEIN, L. Articulatory Phonology: An Overview. *Phonetica*, v. 49, n. 3-4, p. 155-180, 1992.

BRUMBAUGH, K. M.; SMIT, A. B. Treating children ages 3-6 who have speech sound disorder: a survey. *Language, Speech, and Hearing Services in Schools*, v. 44, n. 3, p. 306-19, 2013.

BRUM-DE-PAULA, M. R.; FERREIRA-GONÇALVES, G. Aquisição da linguagem: metodologias voltadas para a produção da fala. In: ERNST, A.; LEFFA, V. J. (Orgs.). *Linguagens: metodologias de ensino e pesquisa*. Pelotas: EDUCAT, p. 150-169, 2012.

BYUN, T. M.; BUCHWALD, A.; MIZOGUCHI, A. Covert contrast in velar fronting: An acoustic and ultrasound stud. *Clinical Linguistics and Phonetics*, v. 30, n. 3-5, p. 249-276, 2016.

BYUN, T. M.; HITCHCOCK, E. R.; SWARTZ M. T. Retroflex versus bunched in treatment for rhotic misarticulation: evidence from ultrasound biofeedback intervention. *Journal of Speech, Language, and Hearing Research*, v. 57, n. 6, p. 2116-2130, 2014.

BYUN, T. M.; SWARTZ, M.T.; HALPIN, P.F.; SZEREDI, D.; MAAS, E. Direction of attentional focus in biofeedback treatment for /r/

misarticulation. *International Journal of Language & Communication Disorders*, v. 51, n. 4, p. 384–401, 2016.

CABBAGE, K. L.; HOGAN, T. P.; CARRELL, T. D. Speech perception differences in children with dyslexia and persistent speech delay. *Speech Communication*, v. 82, p. 14–25, 2016.

CAGLIARI, C.F.; JURKIEWICZ, A. L.; SANTOS, R. S.; MARQUES, J. M. Análise dos sons da deglutição pelo sonar Doppler em indivíduos normais na faixa etária pediátrica. *Brazilian Journal of Otorhinolaryngology*, v. 75, n. 5, p. 706-15, 2009.

CÂMARA JR, J. M. *Estrutura da língua portuguesa*. Rio de Janeiro: Vozes, 1977.

CAMPBELL, F.; GICK, B.; WILSON, I.; VATIKIOTIS-BATESON, E. Spatial and temporal properties of gestures in North American English /r/. *Language and Speech*, v.53, n.1, p.49-69, 2010.

CAMPBELL, H.; MCALLISTER BYUN, T. Deriving individualised /r/ targets from the acoustics of children’s non-rhotic vowels. *Clinical Linguistics and Phonetics*, v. 32, p. 70-87, 2018.

CANDEA, M.; TRIMAILLE, C. Introduction. Phonétique, sociolinguistique, sociophonétique: histoires parallèles et croisements. *Langage et société*, n. 151, p. 7-25, 2015.

CASTRO; M. M.; WERTZNER, H. F. Influence of Sensory Cues on the Stimulability for Liquid Sounds in Brazilian Portuguese-Speaking Children. *Folia Phoniatria et Logopaedica*, v. 61, p. 283-287, 2009.

CERON, M. I. *Instrumento de avaliação fonológica (INFONO): desenvolvimento e estudos psicométricos*. Tese de Doutorado, Programa de Pós-Graduação em Distúrbios da Comunicação Humana. Universidade Federal de Santa Maria, Santa Maria, 2015.

CHARLES. S.; LULICH S. M. Case study of Brazilian Portuguese laterals using a novel articulatory-acoustic methodology with 3D/4D ultrasound. *Speech Communication*, v. 103, p. 37-48, 2018.

CHEN, Y.; LIN, H. Analysing tongue shape and movement in vowel production using SS ANOVA in ultrasound imaging. *The 17th International Congress of Phonetics Sciences*, 2011, p. 124-127.

CIELO, C. A.; CASARIN, M. T. Sons fricativos surdos. *Revista CEFAC*, v. 10, n. 3, p. 352-358, 2008.

CLELAND, J.; SCOBIE, J. M.; ROXBURGH, Z.; HEYDE, C.; WRENCH, A. Enabling New Articulatory Gestures in Children with Persistent Speech Sound Disorders Using Ultrasound Visual Biofeedback. *Journal of Speech, Language, And Hearing Research*, v. 62, n. 2, p. 229-246, 2019.

CLELAND, J.; SCOBIE, J. M.; WRENCH, A. A using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics*, v. 29, n. 8-10, p. 575-597, 2015.

COLLISCHONN, G. QUEDNAU, L. R. As Laterais variáveis na região Sul. In: BISOL, L. COLLISCHON, G. *Português do Sul do Brasil: variação fonológica*. Porto Alegre: EDIPUCRS, 2009, p. 129-147.

CORREA, B. T.; FERREIRA-GONÇALVES, G.; BRUM-DE-PAULA, M. R. Aquisição das vogais nasais francesas *ê*, *ã*, e *õ* por Aprendizes Brasileiros: Aspectos Articulatorios. *Ilha do Desterro*, v. 70, n. 3, p. 131-149, 2017.

CRISTÓFARO-SILVA, T.; YEHIA, H. C. Sonoridade em Artes, Saúde e Tecnologia. *Revista Docência do Ensino Superior*, v. 1, 2011.

DAWSON, K.; CARL, M.; WHALEN, D. H. Shape Coarticulation in the Spatial Frequency Domain: An Example Using /ɹ/. *ICPhS*, 2015.

DE DECKER, P. M.; NYCZ, J. R. Are tense [æ]s really tense? The mapping between articulation and acoustics. *Lingua*, v. 122, p. 810-821, 2012.

DERRICK, D; SCHULTZ, B. Acoustic correlates of flaps in North American English. *Journal of the Acoustical Society of America*, v.133, n.5, p.35-64, 2013.

DERWING, T.; MUNRO, M. Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, v. 39, n. 3, p. 379-397, 2005.

DESHAIES-LAFONTAINE D. *A socio-phonetic study of a Quebec French community: Trois Rivières*. Tese de Doutorado. University of London, 1974.

DIAS-CAVALHEIRO, B. S. *Aquisição da vogal [a] espanhola por falantes de português brasileiro*. Dissertação de Mestrado, Programa de Pós-Graduação em Letras, Universidade Federal de Pelotas, Pelotas, 2016.

DI PAOLO, M.; YAEGER-DROR, M. *Sociophonetics. A Student's Guide*. Routledge, 2010.

DODD, B.; REILLY, S.; EECEN, K. T.; MORGAN A. T. Articulation or phonology? Evidence from longitudinal error data. *Clinical Linguistics & Phonetics*, v. 32, n.11, p. 1027-1041, 2018.

DUARTE, N; FERREIRA-GONÇALVES, G. Aquisição do rótico retroflexo do inglês: Instrução explícita por meio da ultrassonografia. In: XXVII CONGRESSO DE INICIAÇÃO CIENTÍFICA, 2018. Pelotas. Anais do XXVII Congresso de Iniciação Científica: 2018, p. 1-4.

EDDINGTON, D. *Statistics for Linguists: A Step-by-Step Guide for Novices*. Newcastle: Cambridge Scholars Publishing, 2015.

FERREIRA-GONÇALVES, G.; BRUM-DE-PAULA, M. R. (Orgs.) Dinâmica dos movimentos articulatórios: sons, gestos e imagens. Pelotas: Editora UFPel, 2013.

FERREIRA-GONÇALVES, G.; BRUM-DE-PAULA, M. R. A ultrassonografia em pesquisas linguísticas. In: FERREIRA-GONÇALVES, G.; BRUM-DE-PAULA, M. R. (Orgs.). *Dinâmica dos Movimentos Articulatórios: sons, gestos, imagens*. Pelotas: Editora UFPel, 2013, p. 87-109.

_____. A vogal eclipsada e a coordenação gestual em sílabas CCV e CVC. *Anais do IV Encontro intermediário do Grupo de Trabalho de Fonética e Fonologia da ANPOLL*. Fortaleza/CE. Araraquara: Letraria, 2019.

FERREIRA-GONÇALVES, G.; PEREIRA, O; LEMES, M. Aquisição do rótico retroflexo do inglês: instrução explícita por meio de ultrassonografia. *Caderno de Letras*, v. 1, n. 33, p. 127-145, 2019.

FLIPSEN, P. Emergence and Prevalence of Persistent and Residual Speech Errors. *Seminars in Speech and Language*, v. 36, n. 4, p. 217-23, 2015.

FRANCISCO, D. T. *Contorno de língua na produção do /s/ e /ʃ/ na fala de adulto e crianças com e sem transtorno fonológico*. Dissertação de Mestrado. Faculdade de Medicina da Universidade de São Paulo, 2015.

FRANCISCO, D. T.; WERTZNER, H. F. Differences between the production of [s] and [ʃ] in the speech of adults, typically developing children, and children with speech sound disorders: An ultrasound study. *Clinical linguistics & phonetics*, v. 31, n. 5, p. 375-390, 2017.

FREITAS, M. C. C.; ALBANO, E. C. O gesto fônico na aquisição “desviante”: movimentos entre a produção e a percepção. *Revista da ABRALIN*, v. 11, n. 1, 2012.

FOULKES, P.; SCOBIE, J. M.; WATT, D. Sociophonetics. In: HARDCASTLE, W. J., LAVER, J., GIBBON, F. *The Handbook of Phonetic Sciences*. Wiley Online Library, 2010, p. 703-754.

GENARO, K. F.; BERRETIN-FELIX, G.; REHDER, M. I. B. C.; MARCHESAN, I. Q. Avaliação miofuncional orofacial: protocolo MBGR. *Revista CEFAC*, v. 11, n. 2, p. 237-255, 2009.

GHISLENI, M.R.L.; KESKE-SOARES, M.; MEZZOMO, L. M. The use of repair strategies considering the severity of the evolutionary phonological disorder. *Revista CEFAC*, v. 12, n. 5, p. 766-771, 2010.

GIBBON, F. E. Undifferentiated Lingual Gestures in Children With Articulation/Phonological Disorders. *Journal of Speech, Language, and Hearing Research*, v. 42, n. 2, p. 382–397, 1999.

GIBBON, F.; LEE, A. Electropalatography for Older Children and Adults with Residual Speech Errors. *Seminars in Speech and Language*, v. 36, n. 4, p. 271-281, 2015.

GIBSON, T.; LEE, S. A. S. Use of ultrasound visual feedback in speech intervention for children with cochlear implants. *Clinical Linguistics & Phonetics*, p. 1-20, 2020.

GICK, B. The use of ultrasound for linguistic phonetic fieldwork. *Journal of the International Phonetic Association*. v. 32, n. 2, p. 113-121, 2002.

GICK, B.; BACSFALVI, P.; BERNHARDT, B. M.; OH, S.; STOLAR, S.; WILSON, I. A motor differentiation model for liquid substitutions in children’s speech. *Proceedings of Meetings on Acoustics*, v.1, p. 060003–060003, 2007.

_____. A Motor Differentiation Model for Liquid Substitutions: English /r/ Variants in Normal and Disordered Acquisition. *The Journal of the Acoustical Society of America*, v. 6, n. 1, p. 1-9, 2008.

GICK, B.; BERNHARDT, B. M.; BACSFALVI, P.; WILSON, I. Ultrasound imaging applications in second language acquisition. *Phonology and Second Language Acquisition*. The Chinese University of Hong Kong/Le Moyne College, 2008, p. 309–322.

GOLDSTEIN, L.; BYRD, D.; SALTZMAN, E. The role of vocal tract gestural action units in understanding the evolution of phonology. In: ARBIB, M. A. (ed.). *Action to Language via the Mirror Neuron System*. Cambridge University Press, Cambridge, 2006, p. 215-249.

GONÇALVES, G. F.; KESKE-SOARES, M.; CHECALIN, M. A. Estudo do papel do contexto linguístico no tratamento do desvio fonológico. *Revista da Sociedade Brasileira de Fonoaudiologia*, v. 15, n. 1, p. 96-102, 2010.

GREEN, J. R.; MOORE, C. A.; HIGASHIKAWA, M.; STEEVE, R. W. The Physiologic Development of Speech Motor Control Lip and Jaw Coordination. *Journal of Speech, Language, and Hearing Research*, v. 43, n. 1, p. 239-255, 2000.

GUENTHER, F. H.; ESPY-WILSON, C. Y.; BOYCE, S. E.; MATTHIES, M. L.; ZANDIPOUR, M.; PERKELL, J. S. Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *The Journal of the Acoustical Society of America*, v. 105, p. 2854-2865, 1999.

HALL, K. G.; MAGILL, R. A. Variability of practice and contextual interference in motor skill learning. *Journal of Motor Behavior*, v. 27, n. 4, p. 299-309, 1995.

HAUPT, C. Estudo acústico dos róticos no português tocantinense: contribuições a partir da teoria dos exemplares. *Diadorim*, Rio de Janeiro, v. 20, n. 2, p. 191-208, 2018.

HAUPT, C.; SEARA, I. C. As fricativas posteriores: caracterização articulatória e acústica do /r/ em onset silábico. *Entrepalavras*, v. 10, n. 1, p. 77-97, 2020.

HAY, J.; DRAGER, K. Sociophonetics. *Annual Review of Anthropology*, v. 36, p. 89-103, 2007.

HEGARTY, N.; TITTERINGTON, J.; MCLEOD, S.; TAGGART, L. Intervention for children with phonological impairment: Knowledge, practices and intervention intensity in the UK. *International Journal of Language & Communication Disorders*, v. 53, n. 5, p. 995-1006, 2018.

HENG, Q.; MCCABE, P.; CLARKE, J.; PRESTON, J. L. Using ultrasound visual feedback to remediate velar fronting in preschool children: A pilot study. *Clinical linguistics & phonetics*, v. 30, n. 3-5, p. 382-397, 2016.

HEWLETT, N.; WATERS, D. Gradient change in the acquisition of phonology. *Clinical Linguistics and Phonetics*, v. 18, n.6-8, p. 523–533, 2004.

HITCHCOCK, E.; HAREL, D.; BYUN, T. Social, Emotional, and Academic Impact of Residual Speech Errors in School-Aged Children: A Survey Study. *Seminars in Speech and Language*, v. 36, n. 4, p. 283–294, 2015.

HODGES, N. J.; FRANKS, I. M. Learning a coordination skill: Interactive effects of instruction and feedback. *Research Quarterly for Exercise and Sport*, v. 72, n. 2, p. 132–142, 2001.

HODSON, S. L.; JARDINE, B. R. Revisiting Jarrod: Applications of gestural phonology theory to the assessment and treatment of speech sound disorder. *International Journal of Speech-Language Pathology*, v. 11, n. 2, p. 122-134, 2009.

HOFFMANN, T. C.; GLASZIOU, P. P.; BOUTRON, I.; RUAIRIDH MILNE, R.; PERERA, R.; MOHER, D.; ALTMAN, D. G.; BARBOUR, V.; MACDONALD, H.; JOHNSTON, M.; LAMB, S. E.; DIXON-WOODS, M.; MCCULLOCH, P.; WYATT, J. C.; CHAN, A.-W.; MICHIE, S. *Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide*. BMJ, 2014.

HUANG, H.; WOLF, S. L.; HE, J. Recent developments in biofeedback for neuromotor rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, v. 3, n. 11, 2006.

HUEBER, T.; DENBY, B. Analyse du conduit vocal par imagerie ultrasonore. In.: MARCHAL, A.; CAVÉ, C. *L'imagerie médicale pour l'étude de la parole*. Paris: Lavoisier, 2009, p. 147-174.

HWANG, Y.; CHARLES, S.; LULICH, S. M. Articulatory characteristics and variation of Korean laterals. *Phonetics and Speech Sciences*, v. 11, n. 1, p. 19-27, 2019.

JESUS, L.M.T.; SHADLE, C.H. A parametric study of the spectral characteristics of European Portuguese fricative. *Journal of Phonetics*, v. 30, p. 437-464, 2002.

JONGMAN, A.; WAYLAND, R.; WONG, S. Acoustic characteristics of English fricatives. *Working Papers of the Cornell phonetics Laboratory*, v. 12, p. 195-205, 1998.

KEEGSTRA, A.L., KNIJFF, W.A., POST, W.J.; GOORHUIS-BROUWER, S.M. Children with language problems in a speech and hearing clinic: background variables and extent of language problems. *International Journal of Pediatric Otorhinolaryngology*, v. 71, p. 815-21, 2007.

KENT, R. D.; READ, C. *Análise acústica da fala*. São Paulo: Cortez, 2015.

KIER, W. M.; SMITH, K. K. Tongues, tentacles and trunks: the biomechanics of movement in muscular-hydrostats. *Zoological Journal of the Linnean Society*, v. 83, n. 4, 307-324, 1985.

KIRKHAM, S.; WORMALD, J. Acoustic and articulatory variation in British Asian English liquids. *Proceedings of the XVIII International Congress of Phonetic Sciences*. Glasgow, UK, 2015, p. 1-5.

KLEIN, H. B.; BYUN, T. M.; DAVIDSON, L.; GRIGOS, M. I. A multidimensional investigation of children's/r/productions: Perceptual, ultrasound, and acoustic measures. *American Journal of Speech-Language Pathology*, v. 22, n. 3, p. 540-553, 2013.

KOCHETOV, AI; POUPLIER, M.; TROUNG, S. A preliminary ultrasound study of Nepali lingual articulations. *Proceedings of Meetings on Acoustics*, v. 19, n. 1., 2013, p.1-9.

KOENIG, L. L.; LUCERO, J. C.; PERLMAN, E. Speech production variability in fricatives of children and adults: Results of functional data analysis. *The Journal of the Acoustical Society of America*, v. 124, v. 5, p. 3158-3170, 2008.

LADEFOGED, P. *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell Publishing, 2003.

LADEFOGED, P.; MADDIESON, I. Rhotics. In: *The sounds of the world's languages*. Oxford: Blackwell, 1996. p. 215-245.

LAMONTAGNE, J.; MIELKE, J. Perception of Canadian French rhotic vowels. *The Journal of the Acoustic Society of America*, v. 133, n. 5, 2013.

LAMPRECHT, R. R.; FERREIRA-GONÇALVES, G.; FREITAS, G. C. M. de; MATZENAUER, C. L. B.; MEZZOMO, C. L.; OLIVEIRA, C. C.; RIBAS, L. P. *Aquisição fonológica do Português: perfil de desenvolvimento e subsídio para terapia*. Porto Alegre: ARTMED, 2004.

LAUFER, A.; CONDAX, I. D. The function of the epiglottis in speech. *Language and Speech*, v. 24, p. 39-62, 1981.

LAW, J.; BOYLE, J.; HARRIS, F.; HARKNESS, A.; NYE, C. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *International Journal of Language & Communication Disorders*, v. 35, n. 2, p. 165–188, 2000.

LAWSON, E.; SCOBIE, J.M.; SMITH, J.S. The social stratification of tongue shape for postvocalic /r/ in Scottish English. *Journal of Sociolinguistics*, v. 15, n.2, p. 256-268, 2011.

LAWSON, E.; STUART-SMITH, J.; SCOBIE, J. Articulatory insights into language variation and change: preliminary findings from an ultrasound study of derhoticization in Scottish English. *University of Pennsylvania Working Papers in Linguistics*, v. 14, n. 2, p. 102-110, 2008.

LEMES, M.; FERREIRA-GONÇALVES, G. Efeitos da instrução explícita mediada pela ultrassonografia na aquisição das vogais altas anteriores do inglês/L2. In: XXIX CONGRESSO DE INICIAÇÃO CIENTÍFICA, 2020. Pelotas. Anais do XXIX Congresso de Iniciação Científica, 2020, p. 1-4.

LÉON, P.; CICHOCKI, W. Bilan et problématique des études sociophonétiques franco-ontariennes. *Le Français Canadien Parlé Hors Québec: Aperçu Sociolinguistique*, Presses de l'Université Laval, Québec, 1989, p. 37-51.

LIMA, F. L. de C. N.; SILVA, C. E. E. da; SILVA, L. M. da; VASSOLER, A. M. de O.; FABBRON, E. M. G.; BERTI, L. C. Ultrasonographic analysis of lateral liquids and coronal fricatives: judgment of experienced and non-experienced judges. *Revista CEFAC*, v. 20, n. 4, p. 422-431, 2018.

LINDAU, M. The story of /r/. In: FROMKIN, V. *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, p. 157–168, 1985.

MAAS, E.; BUTALLA, C. E.; FARINELLA, K. A. Feedback frequency in treatment for childhood apraxia of speech. *American Journal of Speech-Language Pathology*, v. 21, n. 3, p. 239–257, 2012.

MAAS, E.; GILDERSLEEVE-NEUMANN, C. E.; JAKIELSKI, K. J.; STOECKEL, R. Motor-based intervention protocols in treatment of childhood apraxia of speech (CAS). *Current Developmental Disorders Reports*, v. 1, n. 3, p. 197–206, 2014.

MAAS, E.; ROBIN, D. A.; HULA, S. N. A.; FREEDMAN, S. E.; WULF, G.; BALLARD, K. J.; SCHMIDT, R. A. Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, v. 17, n. 3, p. 277-98, 2008.

MANTIE-KOZLOWSKI, A.; PITT, K. Electropalatography as an adjunct to nonspeech orofacial myofunctional disorder assessments: a feasibility study. *International Journal of Orofacial Myology*, v. 39, p. 31-44, 2013.

MAREY, E.-J. *La méthode graphique dans les sciences expérimentales et principalement en physiologie et en médecine*. Paris: G. Masson, 1878.

MARINI, C.; BRANCALIONI, A. R.; GUBIANI, M. B.; DE FREITAS, G. P.; KESKE-SOARES, M.; CECHELLA, C. O fonema /r/ e as alterações do sistema estomatognático, idade, gênero e gravidade no desvio fonológico. *Revista da Sociedade Brasileira de Fonoaudiologia*, v. 16, n. 4, p. 422-9, 2011.

MATSUI, M.; KOCHETOV, A. Tongue Root Positioning for Voicing vs. Contrastive Palatalization: An Ultrasound Study of Russian Word-Initial Coronal Stops. *Journal of the Phonetic Society of Japan*, v. 22, N. 2, p. 81–94, 2018.

MAYER, C.; GICK, B. Talking while chewing: Speaker response to natural perturbation of speech. *Phonetica*, v. 69, p. 109-123, 2012.

MCALLISTER BYUN, T.; CAMPBELL, H. Differential Effects of Visual-Acoustic Biofeedback Intervention for Residual Speech Errors. *Frontiers in Human Neuroscience*, v. 10, p. 567, 2016.

MCGOWAN, R. S.; NITTROUER, S.; MANNING, C. J. Development of [ɹ] in young, Midwestern, American children. *The Journal of the Acoustical Society of America*, v. 115, n. 2, p. 871-884, 2004.

MCLEOD, S.; BAKER, E. Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics & Phonetics*, v. 28, n. 7-8, p. 508-531, 2014.

MCLEOD, S.; ROBERTS, A.; SITA, J. Tongue/palate contact for the production of / <sc>s</sc> / and /z/. *Clinical Linguistics & Phonetics*, v. 20, n. 1, p. 51–66, 2006.

MCKINNON, D.H.; MCLEOD, S.; REILLY, S. The prevalence of stuttering, voice, and speech-sound disorders in primary school students in Australia. *Language Speech and Hearing Services in Schools*, v. 38, n. 1, p. 5-15, 2007.

MEADOWS, B. Implications of ultrasound technology in the L2 classroom. *Arizona Working Papers in SLA & Teaching*, v. 14, p. 15-41, 2007.

MELCHEQUE, P. P.; BRUM-DE-PAULA, M. R.; FERREIRA-GONÇALVES, G. Os róticos produzidos por pelotenses: características acústicas e articulatórias do tap. In: *Anais do XXVII Congresso de Iniciação Científica*, Universidade Federal de Pelotas, Pelotas, 2018.

MELCHEQUE, P. P.; FERREIRA-GONÇALVES, G.; BRUM-DE-PAULA, M. R. The gestural configuration of the rhotic sounds in the pelotense dialect. *Trabalho apresentado no Workshop em Estudos Ultrassonográficos de Dados de Fala*. Pontifícia Universidade Católica de São Paulo, São Paulo: 10 dez, 2018.

MELO, R. M.; DIAS, R. F.; MOTA, H. B.; MEZZOMO, C. L. Imagens de ultrasonografia de língua pré e pós terapia de fala. *Revista CEFAC*, v. 18, n. 1, p. 286-297, 2016.

MÉNARD, L.; AUBIN, J.; THIBEAULT, M.; RICHARD, G. Measuring tongue shapes and positions with ultrasound imaging: A validation experiment using an articulatory model. *Folia Phoniatria et Logopaedica*, v. 64, n. 2, p. 64-72, 2012.

MIELKE, J. *An articulatory study of rhotic vowels in Canadian French*. University of Ottawa: Members of Ottawa-Carleton Phonology Discussion group, 2011.

MIELKE, J. Ultrasound and corpus study of a change from below: vowel rhoticity in Canadian French. *University of Pennsylvania Working Papers in Linguistics*. v. 19, n. 2, p. 141-150, 2013.

MIELKE, J.; BAKER, A.; ARCHANGELI, D. Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /j/. *Language*, v. 92, n. 1, p. 101-140, 2016.

MIELKE, J.; BAKER, A.; ARCHANGELI, D.; RACY, S. Palatron: a technique for aligning ultrasound images of the tongue and palate. *Coyote Papers*. v. 14, 2005.

MILESKI, I. *A elevação das vogais médias átonas finais no Português falado por descendentes de imigrantes poloneses em Vista Alegre do Prata – RS*. Dissertação de Mestrado. Programa de Pós-Graduação em Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2013.

_____. *Variação no Português de contato com o Polonês no Rio Grande do Sul: vogais médias tônicas e pretônicas*. Tese de Doutorado. Programa de Pós-Graduação em Letras, Pontifícia Universidade Católica do Rio Grande do Sul, 2017.

MODHA, G.; BERNHARDT, B. M.; CHURCH, R.; BACSFALVI, P. Case study using ultrasound to treat /ɹ/. *International Journal of Language & Communication Disorders*, v. 43, n. 3, p. 323–329, 2008.

MOHER, D.; LIBERATI, A.; TETZLAFF, J.; ALTMAN, D.G. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses. *The PRISMA Statement*, v. 6, n. 6, 2009.

MONARETTO, V. N. de O. *Um reestudo da vibrante: análise variacionista e fonológica*. 1996. Tese (Doutorado em Letras). Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Porto Alegre, 1996.

MOTA, H.B.; KESKE-SOARES, M.; BAGETTI, T.; CERON, M.I.; MELO FILHA, M.G.C. Análise comparativa da eficiência de três diferentes modelos de terapia fonológica. *Pró-Fono Revista de Atualização Científica*, v. 19, n. 1, p. 67-74, 2007.

NEWLIN-ŁUKOWICZ, L. Polish stress: Looking for phonetic evidence of a bidirectional system. *Phonology*, v. 29, n. 2, 2012.

NISHIDA, M.; ISHIKAWA, Y.; YAMAMOTO, S. Speaker Clustering in Multi-party Conversation. *Nealt proceedings series*, v. 15, 2011, p. 56.

NOIRAY, A.; ABAKAROVA, D.; RUBERTUS, E.; KRUGER, S.; TIEDE, M. How Do Children Organize Their Speech in the First Years of Life? Insight From Ultrasound Imaging. *Journal of Speech Language and Hearing Research*, v. 61, n. 6, p. 1355-1368, 2018.

PERDICES, M.; TATE, R. Single-subject designs as a tool for evidence-based clinical practice: Are they unrecognised and undervalued? *Neuropsychological Rehabilitation*, v. 19, n. 6, p. 904-27, 2009.

PEREIRA, O. T. A.; FERREIRA-GONÇALVES, G. A instrução explícita aliada à ultrassonografia: aquisição do rótico retroflexo. 2016. Pelotas. *Anais do VIII SENALE – Seminário Nacional sobre Linguagens e Ensino*, 2016.

PILLOT-LOISEAU, C.; ANTOLÍK, T. K.; KAMIYAMA, T. Contribution of ultrasound visualisation to improving the production of the French / y / - / u / contrast by four Japanese learners. *PPLC13: Phonetics, phonology, languages in contact: varieties, multilingualism, second language learning*, Paris, 2013.

PORTALETE, C. R.; FERNANDES, E. G.; PAGLIARIN, K. C. Elaboração de um Protocolo de Avaliação Instrumental da Fala (PRAIN) baseado em critérios linguísticos e psicométricos. *CoDAS*, v. 30, n. 1, 2018.

PRESTON, J. L.; BRICK, N.; LANDI, N. Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech-Language Pathology*, v. 22, n. 4, p. 627-43, 2013.

PRESTON, J. L.; BYUN, T. M.; BOYCE, S. E.; HAMILTON, S.; TIEDE, M.; PHILLIPS, E.; RIVERA-CAMPOS, A.; WHALEN, D. H. Ultrasound Images of the Tongue: a tutorial for assessment and remediation of speech sound errors. *Journal of Visualized Experiments*, n. 119, p. 1-15, 2017.

PRESTON, J. L.; EDWARDS, M. L. Phonological processing skills of adolescents with residual speech sound errors. *Language, Speech, and Hearing Services in Schools*, v. 38, n. 4, p. 297–308, 2007.

PRESTON, J. L.; FELSENFELD, S.; FROST, S. J.; MENCL, W. E.; FULBRIGHT, R. K.; GRIGORENKO, E. L.; LANDI, N.; SEKI, A.; PUGH, K. R. Functional Brain Activation Differences in School-Age Children With Speech Sound Errors: Speech and Print Processing. *Journal of Speech, Language, and Hearing Research*, v. 55, n. 4, p. 1068–1082, 2012.

PRESTON, J. L. ; KOENIG, L. L. Phonetic variability in residual speech sound disorders: Exploration of subtypes. *Topics in Language Disorders*, v. 31, n. 2, p. 168-184, 2011.

PRESTON, J.L.; LEAMAN, M. Ultrasound visual feedback for acquired apraxia of speech: A case report. *Aphasiology*, v. 28, n. 3, p. 278- 95, 2014.

PRESTON, J.L.; LEECE, M. C. Intensive Treatment for Persisting Rhotic Distortions: A Case Series. *American Journal of Speech-Language Pathology*, v. 26, n. 4, p. 1066–1079, 2017.

PRESTON, J. L.; LEECE, M. C.; MAAS, E. Intensive Treatment with Ultrasound Visual Feedback for Speech Sound Errors in Childhood Apraxia. *Frontiers in Human Neuroscience*, v. 10, p. 1-9, 2016.

_____. Motor-based treatment with and without ultrasound feedback for residual speech-sound errors. *International Journal of Language & Communication Disorders*, v. 52, n. 1, p. 80-94, 2017.

PRESTON, J.L.; LEECE, M.C.; MCNAMARA, K.; MAAS, E. Variable practice to enhance speech learning in ultrasound biofeedback treatment for childhood apraxia of speech: A single case experimental study. *American Journal of Speech-Language Pathology*, v. 26, n. 3, 2017.

PRESTON, J.L.; MAAS, E.; WHITTLE, J.; LEECE, M.C.; MCCABE, P. Limited acquisition and generalization of rhotics with ultrasound visual feedback in childhood apraxia. *Clinical Linguistics & Phonetics*, v. 30, n. 3-5, p. 363-381, 2016.

PRESTON, J. L.; MCALLISTER, T.; PHILLIPS, E.; BOYCE, S. TIEDE, M.; KIM, J. S.; WHALEN, D. H. Treatment for residual rhotic errors with high-and low-frequency ultrasound visual feedback: A single-case experimental design. *Journal of Speech, Language, and Hearing Research*, v. 61, n. 8, p. 1875-1892, 2018.

PRESTON, J. L.; MCCABE, P.; RIVERA-CAMPOS, A.; WHITTLE, J. L.; LANDRY, E.; MAAS, E. Ultrasound visual feedback treatment and practice variability for residual speech sound errors. *Journal of Speech, Language, Hearing Research*, v. 57, n. 6, p. 2102-2115, 2014.

PRESTON, J.L.; MCCABE, P.; TIEDE, M.; WHALEN, D.H. Tongue shapes for rhotics in school-age children with and without residual speech errors. *Clinical Linguistics & Phonetics*, v. 33, n. 4, p. 334-348, 2019.

PROCTOR, M.I. *Gestural characterization of a phonological class: The liquids*. Tese de Doutorado. Yale University, 2009.

PROCTOR, M. Towards a gestural characterization of liquids: Evidence from Spanish and Russian. *Laboratory Phonology*. v. 2, n. 2, p.451-485, 2011.

RECASENS, D. Darknesse in [l] as scalar phonetic property: implications for phonology and articulatory control. *Clinical Linguistics e phonetics*, v. 18, n. 6-8, p. 593-603, 2004.

_____. What is and what is not an articulatory gesture in speech production: The case of lateral, rhotic and (alveolo)palatal consonants. *Gradus*, v. 1, n. 1, p. 24-42, 2016.

RECASENS, D.; ESPINOSA, A. An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America*, v. 125, n. 4, p. 2288-98, 2009.

RECASENS, D.; FONTDEVILA, J.; PALLARÈS, M. D. Velarization degree and coarticulatory resistance for /l/ in Catalan and German. *Journal of Phonetics*, v. 23, p. 37-52, 1995.

RECASENS, D.; RODRIGUEZ, C. A study of coarticulatory resistance and aggressiveness for front lingual consonants and vowels using ultrasound. *Journal of Phonetics*, v. 59, p. 58-79, 2016.

REINECKE, K. Os róticos intervocálicos na gramática individual de falantes de Blumenau e Lages. Tese (Doutorado em Linguística), Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, 2006.

RENNICKE, I. E. *Variation and Change in the Rhotics of Brazilian Portuguese*. Tese (Doutorado em Estudos Linguísticos), Programa de Pós-Graduação em Estudos Linguísticos, Faculdade de Letras, UFMG, Belo Horizonte, 2015.

RODRIGUES, L.L. *Aquisição dos róticos em crianças com queixa fonoaudiológica*. Dissertação (Mestrado). Universidade de Campinas, Campinas, 2007.

RODRIGUES, L.L.; FREITAS, M.C.C.; ALBANO, E.A.; BERTI, L.C. Acertos gradientes nos chamados erros de pronúncia. Aquisição de Língua Materna: Heterogeneidade da Pesquisa. *Revista do Programa de Pós-Graduação em Letras*, n. 36, p.85-112, 2008.

ROSINSKI, A. *A produção da lateral pós-vocálica em uma comunidade bilíngue: aspectos do Português sob a influência do Polonês como língua de imigração*. 2019. Dissertação de Mestrado. Programa de Pós-Graduação em Letras, Universidade Federal de Pelotas, Pelotas, 2019.

FERREIRA-GONÇALVES, G.; ROSINSKI, A. O contato entre português e polonês como língua de imigração: uma descrição acústico-articulatória de // pós-vocálico. *Linguagem & Ensino*, v. 23, p. 918-936, 2020.

RUBEN, R.J. Redefining the survival of the fittest: communication disorders in the 21st century. *Laryngoscope*, v. 110, n. 2, p. 241, 2000.

RUSCELLO, D. M. Visual feedback in treatment of residual phonological disorders. *Journal of Communication Disorders*. v. 28, n. 4, p. 279-302, 1995.

RUSSO, I. C. P.; BEHLAU, M. As pistas acústicas das vogais e consoantes. In: RUSSO, I. C. P.; BEHLAU, M. *Percepção da fala: análise acústica do português brasileiro*. São Paulo: Lovise, 1993, p. 25-50.

RVACHEW, S.; JAMIESON, D. G. Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders*, v. 54, n. 2, p. 193-208, 1989.

SCHMIDT, R. A.; LEE, T. D. *Motor control and learning: A behavioral emphasis*. Champaign, IL: Human Kinetics, 2011.

SCOBIE, J. M. Interactions between the acquisition of phonetics and phonology, 34th Annual Regional Meeting of the Chicago Linguistic Society, Chicago Linguistics Society, 1998, p. 343-358.

SCOBIE, J. M.; GIBBON, F.; HARDCASTLE, W. J.; FLETCHER, P. Covert contrast as a stage in the acquisition of phonetics and phonology. *Papers in laboratory phonology V: Acquisition and the lexicon*, v. 5, 2000, p. 194-207.

SCOBIE, J.M; STUART-SMITH, J.; LAWSON, E. Looking variation and change in the mouth: developing the sociolinguistic potential of Ultrasound Tongue Imaging, 2008.

SCOBIE, J. M.; WRENCH, A. A.; LINDEN, M. L. van der. Head-Probe Stabilisation in Ultrasound Tongue Imaging Using a Headset to Permit Natural Head Movement. *The 8th International Seminar on Speech Production*, Strasbourg, Université Marc Bloch, 2008, p. 373-376.

SECORD, W. A.; BOYCE, S. E.; DONOHUE, J. S.; FOX, R. A.; SHINE, R. E. *Eliciting sounds: techniques and strategies for clinicians*. Clifton Park, NY: Thomson Delmar Learning, 2007.

SHAWKER, T. H.; SONIES, B. C. Ultrasound biofeedback for speech training: Instrumentation and preliminary results. *Investigative Radiology*, v. 20, n. 1, p. 900-93, 1985.

SHRIBERG, L. D. A response evocation programme for /ʒ/. *Journal of Speech & Hearing Disorders*, v. 40, p. 92-105, 1975.

_____. Research in Idiopathic and Symptomatic Childhood Apraxia of Speech. *5th International Conference on Speech Motor Control Nijmegen*, 2006.

SHRIBERG, L. D.; AUSTIN, D.; LEWIS, B. A.; MCSWEENEY, J. L.; WILSON, D. L. The Speech Disorders Classification System (SDCS): Extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research*, v. 40, n. 4, p. 723-740, 1997.

_____. The percentage of consonants correct (PCC) metric: extensions and reliability data. *Journal of Speech, Language, and Hearing Research*. Madison, p. 708-22, 1997.

SHRIBERG, L. D.; FOURAKIS, M.; HALL, S. D.; KARLSSON, H. B.; LOHMEIER, H. L.; MCSWEENEY, J. L.; POTTER, N. L.; SCHEER-COHEN, A. R.; STRAND, E. A.; TILKENS C. M.; WILSON D. L. Extensions to the Speech Disorders Classification System (SDCS). *Clinical Linguistics & Phonetics*, v. 24, n.10, p. 795-824, 2010.

SHRIBERG, L. D.; GRUBER, F. A.; KWIATKOWSKI, J. Developmental Phonological Disorders III. *Journal of Speech, Language, and Hearing Research*, v. 37, n. 5, p. 1151-1177, 1994.

SHRIBERG, L. D.; KWIATKOWSKI, J. Phonological disorders III: a procedure for assessing severity of involvement. *The Journal of Speech and Hearing Disorders*, v. 47, n. 3, p. 256-270, 1982.

SILVA, A. F. Contrastes na produção de vogais da língua inglesa: um estudo com o uso de técnicas de fonética acústica e ultrassonografia. In: Workshop em estudos ultrassonográficos de dados de fala, 2018. São Paulo. *Anais do Workshop em estudos ultrassonográficos de dados de fala*, 2018, p. 3.

SILVA, A. H. P. *Para a descrição fonético-acústica das líquidas no português brasileiro: dados de um informante paulistano*. 1996.

Dissertação de Mestrado. Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, 1996.

_____. As fronteiras entre Fonética e Fonologia e a alofonia dos róticos iniciais em PB: dados de dois informantes do sul do país. Campinas, UNICAMP, 2002. Tese (Doutorado em Linguística), Programa de Pós-graduação em Linguística, Instituto de Estudos da Linguagem, UNICAMP, Campinas, 2002.

_____. O estatuto da análise acústica nos estudos fônicos. *Caderno de Letras UFF - Dossiê: Letras e cognição*, v. 41, n. 1, p. 213–29, 2010.

SILVA, A. H. P.; CLEMENTE, F. C.; NISHIDA, G. Para a representação dinâmica do tap em grupos e codas: evidências acústicas. *Revista Virtual de Estudos da Linguagem. ReVEL*, v. 4, n. 7, p. 1–26, 2006.

SILVA-GARCIA, L.; FERREIRA-GONÇALVES, G. Instrução explícita por meio da ultrassonografia: uma nova ferramenta para a aquisição da lateral // do espanhol. In: SOUSA, A. M.; GARCIA, R.; SANTOS, T. C. *Perspectivas para o ensino de línguas*. Rio Branco: EDUFAC, 2019, p. 61-74.

SILVA, L. M. *Parâmetros ultrassonográficos do movimento da língua em 14 fonemas consonantais do português brasileiro*. 2017. Dissertação de Mestrado. Pós-Graduação em Fonoaudiologia, Universidade Estadual Paulista, 2017.

SILVA, L. M.; VASSOLER, A. M. de O.; MARINO, V. C. de C.; BERTI, L. C. Quantitative analysis of tongue movement in 14 phonemes of Brazilian Portuguese. *CoDAS*, São Paulo, v. 29, n. 4, 2017.

SILVA, R. D. *Terapia fonoaudiológica por meio da ultrassonografia no distúrbio do som de fala: relato de caso*. TCC (Fonoaudiologia), Universidade Federal de Santa Catarina, Florianópolis, 2020.

SILVEIRA, F.; SEARA, I. C. Vogal de apoio em grupos consonantais CCV no Português Brasileiro. *Revista da ABRALIN*, v. 7, n. 1, p. 27–48, 2008.

_____. A vogal epentética em encontros consonantais heterossilábicos no português brasileiro: um estudo experimental. *Revista do GEL*, v. 6, p. 09-35, 2009.

SJOLIE, G. M.; LEECE, M. C.; PRESTON, J. L. Acquisition, retention, and generalization of rhotics with and without ultrasound visual feedback. *Journal of Communication Disorders*, v. 64, p. 62-77, 2016.

- SMIT, A. B.; HAND, L.; FREILINGER, J. J.; BERNTHAL, J. E.; BIRD, A. The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, v. 55, n. 4, p. 779–798, 1990.
- SMITH, A.; ZELAZNIK, H. N. Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology*, v. 45, n. 1, p. 22-33, 2004.
- SOMEFUN, O.A.; LESI, F.E.A.; DANFULANI, M.A.; OLUSANYA, B.O. Communication disorders in Nigerian children. *International Journal of Pediatric Otorhinolaryngology*, v. 70, n. 4, p. 697-702, 2006.
- SOUZA, T. N. U.; AVILA, C. R. B. D. Severity of phonological disorder, phonological awareness and articulatory praxis in preschoolers. *Revista da Sociedade Brasileira de Fonoaudiologia*, v. 16, n. 2, p. 182-188, 2011.
- SPREAFICO, L.; VIETTI, A. On rothic in bilingual community: A preliminar UTI research. In: SPREAFICO, L.; VIETTI, A. (eds.), *Rothics. New data and perspectives*. Bozen-Bolzano University Press, p. 57-77, 2013.
- SPROAT, R.; FUJIMURA, O. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, v. 21, n. 3, p. 291-311, 1993.
- STOLAR, S.; GICK, B. An index for quantifying tongue curvature. *Canadian Acoustics*, v. 41, n. 1, 2013.
- STONE, M. A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics*, v. 19, n. 6-7, p. 455-501, 2005.
- STRAND, E.A.; MCCAULEY, R.J.; WEIGAND, S.D.; STOECKEL, R.E.; BAAS, B.S. A motor speech assessment for children with severe speech disorders: reliability and validity evidence. *Journal of Speech, Language, and Hearing Research*, v. 56, n. 2, p. 505-20, 2013.
- STUDDERT-KENNEDY, M.; GOLDSTEIN, L. Launching language: The gestural origin of discrete infinity. *Studies in the Evolution of Language*, v. 3, p. 235-254, 2003.
- SUGDEN, E.; BAKER, E.; MUNRO, N.; WILLIAMS, A. L.; TRIVETTE, C. M. Service delivery and intervention intensity for phonology-based speech sound disorders. *International Journal of Language & Communication Disorders*. v. 53, n. 4, p. 718-734, 2018.

SUGDEN, E.; LLOYD, S.; LAM, J.; CLELAND, J. Systematic review of ultrasound visual biofeedback in intervention for speech sound disorders. *International Journal of Language & Communication Disorders*. v. 54, n. 5, p. 705-728, 2019.

SVICERO, M. A. Caracterização acústica e de imagens de ultrassonografia das vogais orais do Português Brasileiro. Dissertação de Mestrado. Pontifícia Universidade Católica de São Paulo. Programa de estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem. São Paulo, 2012.

SWAN, O. E. *A Grammar of Contemporary Polish*. Bloomington: Indiana University, Slavica Publisher, 2002.

SZREDER, M. The acquisition of consonant clusters in Polish: a case study. In: VIHMAN, M. M., KEREN-PORTNORY, T. (Orgs). *The emergence of phonology: Whole-word Approaches and Cross-linguistic Evidence*. Cambridge, Cambridge University Press, 2013, p. 343-361.

TATE, R.; ROSENKOETTER, U.; WAKIM, D.; SIGMUNDSDOTTIR, L.; DOUBLEDAY, J.; TOGHER, L.; MCDONALD, S.; PERDICES, M. *The risk-of-bias in N-of-1 trials (RoBiNT) scale: an expanded manual for the critical appraisal of single-case reports*. Sydney; St Leonards NSW, 2015.

TATEISHI, M. *Effects of the Use of Ultrasound in Production Training on the Perception of English /r/ and /l/ by Native Japanese Speakers*. Tese (Master of Arts) - Faculty of Graduate Studies (Department of Linguistics), University of Calgary, 2013.

TEIXEIRA-CORREA, B. *Aquisição das vogais nasais francesas [ɛ], [ã] e [ɔ] por aprendizes brasileiros: aspectos acústico-articulatórios*. 2017. Dissertação de Mestrado, Programa de Pós-graduação em Letras, Universidade Federal de Pelotas, Pelotas, 2017.

TEIXEIRA-CORREA, B.; FERREIRA-GONÇALVES, G.; BRUM-DE-PAULA, M. R. Aquisição das vogais nasais francesas [ɛ], [ã] e [ɔ] por aprendizes brasileiras: aspectos articulatórios. *Iha do Desterro*, v. 70, n. 3, p. 131-149, 2017.

TEIXEIRA, K.; BRUM-DE-PAULA, M. R.; FERREIRA-GONÇALVES, G. A Ultrassonografia aplicada ao ensino de L2: aquisição da lateral pós-vocálica do inglês. In: XXVIII CONGRESSO DE INICIAÇÃO CIENTÍFICA, 28., 2019. Pelotas. Anais do XXVIII Congresso de Iniciação Científica, p. 1-4, 2019.

TEIXEIRA, K.; FERREIRA-GONÇALVES, G. A instrução explícita e o ultrassom como ferramentas auxiliares na aquisição da lateral pós-vocálica do inglês. In: XXIX CONGRESSO DE INICIAÇÃO CIENTÍFICA, 29, 2020. Pelotas. Anais do XXIX Congresso de Iniciação Científica: 2020. p. 1-4.

THOMAS, E. *Sociophonetics: an introduction*. Basingstoke: Palgrave Macmillan, 2011.

TSUI, H. *Ultrasound Speech Training for Japanese Adults Learning English as a Second Language*. Tese (Master of Science) – The Faculty of Graduate Studies (Audiology and Speech Sciences), University of British Columbia, 2012.

TURTON, D. Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, v. 8, p. 1-31, 2017.

VASSOLER, A. M. de O. *Coordenação gestual na produção de encontros consonantais em crianças com desenvolvimento típico e atípico*. Dissertação de Mestrado. Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2016.

VASSOLER, A. M. O.; BERTI, L. C. Análise ultrassonográfica quantitativa da produção de encontros consonantais realizadas por crianças típicas e atípicas. In: XXIII CONGRESSO BRASILEIRO DE FONOAUDIOLOGIA E IX CONGRESSO INTERNACIONAL DE FONOAUDIOLOGIA, 2015, p. 1-7.

VERÍSSIMO, A.; BORSEL, J. VAN; BRITTO PEREIRA, M. DE. Residual /s/ and /r/ distortions: The perspective of the speaker. *International Journal of Speech-Language Pathology*, v. 14, n. 2, p. 183–186, 2012.

VIDOR-SOUZA, D.; MOTA, H. B.; SANTOS, R. M. A consciência fonoarticulatória em crianças com desvio fonológico. *Revista CEFAC*, v.13, n. 2, p. 196-204, 2011.

WALSH, B.; SMITH, A. Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research*, v. 45, n. 6, p. 1119-33, 2002.

WALSH, B.; SMITH, A.; WEBER-FOX, C. Short-term plasticity in children's speech motor systems. *Developmental Psychobiology*, v. 48, n. 8, p. 660-74, 2006.

WARING, R.; KNIGHT, R. How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems. *International Journal of Language & Communication Disorders*, v. 48, n. 1, p. 25-40, 2013.

WERTZNER, H.F.; FRANCISCO, D.T.; NEVES, L.O.P. Contorno de língua dos sons /s/ e /ʃ/ em crianças com transtorno fonológico. *Revista Cogas*, v. 26, n. 3, p. 248-251, 2014.

WESTERBERG, F. E. *An auditory, acoustic, articulatory and sociophonetic study of Swedish Viby-i*. Dissertação de Mestrado. College of Arts, University of Glasgow, Glasgow, Escócia, 2016.

WHITE, D.; GANANATHAN, R.; MOK, P. Teaching dark /l/ with ultrasound technology. In: O'BRIEN, M. ; LEVIS, J. (eds). *Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference*, Calgary, AB, August 2016. Ames, IA: Iowa State University, 2017, p. 155-175.

WIETHAN, F.; CERON, M. I.; MARCHETTI, P.; GIACCHINI, V.; MOTA, H. B. O uso da eletroglotografia, eletromiografia, espectrografia e ultrassom nos estudos de fala – revisão teórica. *Revista CEFAC*, v. 17, Supl. 1, p. 115-125, 2015.

WIETHAN, F. M.; MOTA, H. B. Emprego de estratégias de reparo para os fonemas fricativos no desvio fonológico. *Revista da Sociedade Brasileira de Fonoaudiologia*, v. 17, n. 1, p. 28-33, 2012.

WIESE, R. The phonology of /r/. In: HALL, T. A. (Org.). *Distinctive Feature Theory*. Berlim: Mouton de Gruyter, 2001.

WILSON, I. Using ultrasound for teaching and researching articulation. *Acoustical Science and Technology*, v. 35, n. 6, p. 285-289, 2014.

WILSON, I.; GICK, B. Ultrasound technology and second language acquisition research. *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA)*, 2006, p. 148-152.

WREN et al. Prevalence and Predictors of Persistent Speech Sound Disorder at Eight Years Old: Findings From a Population Cohort Study. *Journal of Speech, Language, and Hearing Research*, v. 59, p. 647–673, 2016.

WRENCH, A. A.; SCOBIE, J. M. Categorising vocalisation of English // using EPG, EMA and ultrasound. *Proceedings of the 6th International Seminar on Speech Production*, 2003, p. 314-319.

WULF, G.; SCHMIDT, R. A. Variability of practice and implicit motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, v. 23, n. 4, p. 987–1006, 1997.

YAVAS, M.; HERNANDORENA, C. L. M.; LAMPRECHT, R. R. *Avaliação Fonológica da Criança Reeducação e Terapia*. Porto Alegre: Artmed, 2002.

YI, M.; KAMBHAMETTU, C.; STONE, M. Automatic contour tracking in ultrasound images. *Clinical Linguistics and Phonetics*. v. 19, n. 6-7, p. 545-54, 2005.

ZHARKOVA, N. A normative-speaker validation study of two indices developed to quantify tongue dorsum activity from midsagittal tongue shapes. *Clinical linguistics & phonetics*, v. 27, n. 6-7, p. 484-496, 2013.

ZHARKOVA, N.; GIBBON, F. E.; HARDCASTLE, W. J. Quantifying lingual coarticulation using ultrasound imaging data collected with and without head stabilisation. *Clinical linguistics & phonetics*, v. 29, n. 4, p. 249-265, 2015.

ZHARKOVA N.; HEWLETT, N.; HARDCASTLE, W.J. Coarticulation as an indicator of speech motor control development in children: an ultrasound study. *Motor Control*, v. 15, n. 1, 118-40, 2011.

_____. An ultrasound study of lingual coarticulation in/s V/syllables produced by adults and typically developing children. *Journal of the International Phonetic Association*, v. 42, n. 2, p. 193-208, 2012.

ZHOU, D. e WU, Z. *Articulation Album of Putonghua*. Shangwu Yingshuguan, Beijing, 1963.

Sobre as autoras e os autores

Aline Mara de Oliveira

Fonoaudióloga, mestre em Linguística pela Universidade de São Paulo (USP), Doutora em Estudos Linguísticos pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP/São José do Rio Preto) e pós-doutora em Fonoaudiologia pela Universidade Estadual Paulista (UNESP/ Marília). Atualmente, coordena projeto de pesquisa que investiga os distúrbios dos sons da fala e transtornos de comunicação e/ou de linguagem em crianças brasileiras.

E-mail: fgaalinemaraoliveira@gmail.com

Aline Rosinski

Bacharel em Letras – Redação e Revisão de Textos pela Universidade Federal de Pelotas (UFPEL), mestra em Letras e Doutoranda do Programa de Pós-Graduação em Letras da mesma instituição. Desde 2015, desenvolve pesquisa envolvendo metodologias vinculadas à Sociofonética para a descrição do Português Brasileiro (PB) em contato com o Polonês como língua de imigração. É bolsista CAPES do Doutorado em Letras da UFPEL.

E-mail: rosinskivieira@gmail.com

Bruna Teixeira-Correa

Graduada em Letras Português/Francês pela Universidade Federal de Pelotas (UFPEL), especialista em Linguística Aplicada, mestre em Estudos da Linguagem pela mesma instituição. Atualmente, é bolsista CAPES do Doutorado em Letras da UFPEL. Tem interesse por detalhamento acústico, aerodinâmico e articulatório dos segmentos da fala, aquisição fonético-fonológica de L2 e fonologia do português, com foco especial em vogais nasais.

E-mail: bukacorrea@hotmail.com

Carine Haupt

Professora Associada da Universidade Federal do Tocantins (UFT), onde atua no curso de Letras e no Programa de Pós-graduação em Letras, do *campus* de Porto Nacional, na área de Teoria e Análise

Linguística. Pós-Doutora pelo Programa de Pós-graduação em Linguística, da Universidade Federal de Santa Catarina (UFSC), concentra sua pesquisa no campo da Fonética e Fonologia, com ênfase nas teorias cognitivas e de uso. Atualmente, estuda as fricativas posteriores, investigando sua articulação e correlatos acústicos. Interessa-se também por estudos e aplicação da Fonética e Fonologia ao ensino de línguas.

E-mail: carineh@uft.edu.br

Caroline Rodrigues Portalete

Fonoaudióloga. Doutoranda do Programa de Pós-Graduação em Distúrbios da Comunicação Humana da Universidade Federal de Santa Maria (UFSM). É pesquisadora no Laboratório de Fala (LabFala) da UFSM, onde desenvolve e participa de pesquisas que permitem conhecer, prevenir, identificar, diagnosticar e intervir nos Distúrbios dos Sons da Fala. Possui experiência e interesses na área de Fonoaudiologia, com ênfase em Linguagem e Voz, atuando principalmente nos seguintes temas: linguagem no adulto e no idoso; distúrbios motores de fala; distúrbios neurológicos adquiridos; fonética acústica e articulatória; Fonologia Gestual; avaliações instrumentais de fala; ultrassonografia dos movimentos de língua; nasometria e eletroglotografia.

E-mail: cportalete@gmail.com

Denis Altieri de Oliveira Moraes

Bacharel em Estatística, doutor em Probabilidade e Estatística pela Universidade Federal de Minas Gerais (UFMG), atua como Docente do Departamento de Estatística da Universidade Federal de Santa Maria (UFSM), atuando nas disciplinas de Estatística Aplicada, Inferência Bayesiana, Análise Multivariada, Modelos de Regressão, Análise de Dados Categóricos, Classificação de Padrões e Controle Estatístico de Processos.

E-mail: d_altieri@yahoo.com.br

Giovana Ferreira-Gonçalves

Doutora em Letras pela Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS), pós-doutora pela Universidade Federal de Santa Catarina (UFSC), atualmente é professora Associada da Universidade Federal de Pelotas (UFPel), onde coordena o Laboratório Emergência

da Linguagem Oral (LELO). Desenvolve pesquisas na área de Fonologia, com foco em teorias fonológicas e aquisição fonológica de língua materna e de língua estrangeira. Em seu percurso acadêmico, publicou trabalhos cujas análises foram realizadas sob diferentes perspectivas teóricas, com ênfase em Teoria da Otimidade. Nos últimos dez anos, tem voltado suas pesquisas para estudos em Fonologia Articulatória. É Bolsista de Produtividade do CNPq.
E-mail: giovanaferreiragoncalves@gmail.com

Helena Bolli Mota

Fonoaudióloga, Graduada em Letras, Especialista em Língua Inglesa, possui Mestrado e Doutorado em Letras pela Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS), e pós-doutorado pela *City University London*. Atualmente, é professora aposentada da Universidade Federal de Santa Maria. Tem experiência na área de Fonoaudiologia, com ênfase em Terapia dos Distúrbios de Fala e Linguagem. Atua, principalmente, nos seguintes temas de pesquisa: aquisição da linguagem e teorias linguísticas que envolvem a complexidade de traços.
E-mail: nebolli@hotmail.com

Isadora Mayer Rosado

Fonoaudióloga. Mestre em Distúrbios da Comunicação Humana pela Universidade Federal de Santa Maria (UFSM).
E-mail: isadora.mayer@hotmail.com

Izabel Christine Seara

Mestre e doutora em Linguística pela Universidade Federal de Santa Catarina (UFSC), é professora na Graduação em Letras – Português e no Programa de Pós Graduação em Linguística na área: Teoria e Análise Linguística. Realizou Pós-Doutorado pela *Université Paris 3 – Sorbonne Nouvelle*, no *Laboratoire de Phonétique et Phonologie* (LPP-Paris_França), e pela Universidade de Aveiro, no Laboratório de Fonética (Aveiro-Portugal). É bolsista de produtividade do CNPq. Seus temas de pesquisa são: entoação – prosódia e suas interfaces, detalhamento acústico-articulatório de segmentos de fala, síntese e reconhecimento de fala, francês como língua estrangeira e línguas indígenas.
E-mail: izabel.seara@gmail.com

Lais Silva Garcia

Graduada em Letras-Português e Espanhol pela Universidade Federal de Pelotas (UFPEL), atualmente, é mestranda pelo Programa de Pós-graduação em Letras, na área de Aquisição, Variação e Ensino, com bolsa CAPES, na mesma instituição. No ano de 2016, passou a integrar o Laboratório Emergência da Linguagem Oral (LELO), onde realiza pesquisas relacionadas ao ensino de línguas via ultrassom.

E-mail: laisg16@gmail.com

Larissa Cristina Berti

Fonoaudióloga, mestre em Educação pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) e Doutora pela Universidade Estadual de Campinas (UNICAMP). Realizou Pós-Doutorado em *Speech Language Pathology* na *University of Toronto*. Atualmente é professora do Departamento de Fonoaudiologia da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP–Campus de Marília), onde é docente do Programa de Pós Graduação em Fonoaudiologia. Atua também como docente no Programa de Pós Graduação em Estudos Linguísticos da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP–Campus de São José do Rio Preto). Tem experiência na área de Fonoaudiologia e Linguística, atuando principalmente nas seguintes áreas: aquisição da linguagem oral, fonética e fonologia, análise acústica e análise ultrassonográfica. Bolsista de Produtividade do CNPq.

E-mail: berti.larissa@gmail.com

Leticia Bitencourt Uberti

Fonoaudióloga, doutoranda pelo Programa de Pós-Graduação em Distúrbios da Comunicação Humana (PPGDCH) da Universidade Federal de Santa Maria (UFSM), é pesquisadora no Laboratório de Neuropsicologia da Linguagem (LabNeurolin) da UFSM, onde desenvolve e participa de pesquisas que permitem conhecer, prevenir, identificar, diagnosticar e intervir nos Distúrbios dos Sons da Fala. Possui experiência e interesses na área de Fonoaudiologia, com ênfase em Linguagem e Fala, atuando principalmente nos seguintes temas: fonoaudiologia; neuropsicologia da linguagem; distúrbios adquiridos e desenvolvimentais da fala, da linguagem e da comunicação; avaliação e reabilitação da linguagem e dos distúrbios dos sons da fala.

E-mail: leticiab_uberti@hotmail.com

Luciana da Silva Barberena

Fonoaudióloga, especialista em Distúrbios da Comunicação Humana pela Universidade Federal de São Paulo (Escola Paulista de Medicina), possui Mestrado e Doutorado em Distúrbios da Comunicação Humana pela Universidade Federal de Santa Maria (UFSM). Atualmente, é Fonoaudióloga da Empresa Brasileira de Serviços Hospitalares (EBSERH), desempenhando suas funções no Hospital Universitário de Santa Maria. Tem experiência na área de Fonoaudiologia, com ênfase em Linguagem, atuando principalmente nos seguintes temas: fala, fonoaudiologia e fonoterapia.

E-mail: lucianabarberena@hotmail.com

Márcia Keske-Soares

Fonoaudióloga, doutora em Linguística Aplicada pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), é professora do Departamento de Fonoaudiologia da Universidade Federal de Santa Maria (UFSM). Pesquisadora nos seguintes temas: avaliação, diagnóstico e terapia dos transtornos dos sons da fala.

E-mail: marcia-keske.soares@ufsm.br

Marieli Barichello Gubiani

Fonoaudióloga. Mestranda no Programa de Pós-Graduação em Distúrbios da Comunicação Humana da Universidade Federal de Santa Maria (UFSM).

E-mail: marielibarichellog@gmail.com

Marileda Barichello Gubiani

Fonoaudióloga. Doutora pelo Programa de Pós-Graduação em Distúrbios da Comunicação Humana da Universidade Federal de Santa Maria (UFSM), atualmente realiza Pós-Doutorado na mesma instituição.

E-mail: mari_gubiani@yahoo.com.br

Mirian Rose Brum-de-Paula

Realizou Maîtrise de Français Langue Étrangère pela Université de Franche-Comté, especialização em Formation de formateurs pelo Centre d'Études Pédagogiques de Sèvres, mestrado em Linguistique-Sémiotique-Communication pela Université de Franche-Comté e

doutorado em Sciences du langage pela Université de Paris X. Atualmente, é professora Associada da Universidade Federal de Pelotas (UFPEL) e coordenadora científica do Laboratório Emergência da Linguagem Oral (LELO). Tem experiência na área de Linguística, com ênfase em Aquisição da Linguagem (L1 e L2), atuando nos seguintes temas: aquisição fonético-fonológica, categorização e relações entre linguagem e cognição.

E-mail: brumdepaula@yahoo.fr

Misael Krueger Lemes

Graduado em Licenciatura em Letras – Português/Inglês pela Universidade Federal de Pelotas (UFPEL), atualmente, é mestrando do Programa de Pós-Graduação em Letras pela mesma instituição, com bolsa CAPES. Tem interesse na área de Linguística, com ênfase nos seguintes temas: aquisição fonético-fonológica do inglês como língua estrangeira, ensino de pronúncia e instrução explícita mediada pela ultrassonografia.

E-mail: misaelkrueger@gmail.com

Patrícia Pereira Melcheque

Mestranda em Linguagem, Texto e Imagem pelo Programa de Pós-Graduação em Letras da Universidade Federal de Pelotas (UFPEL), bolsista CAPES, é Graduada em Letras Português e Literatura pela mesma instituição. Vinculada ao Laboratório Emergência da Linguagem Oral (LELO), atualmente, desenvolve pesquisas voltadas à aquisição dos sons róticos em final de sílaba na fala infantil, com ênfase aos aspectos gestuais que compõem essa classe de sons.

E-mail: patriciamelcheque@hotmail.com

Rafaela Duarte Silva

Fonoaudióloga pela Universidade Federal de Santa Catarina (UFSC). Seu Trabalho de Conclusão de Curso, defendido em 2020, é intitulado: *Terapia Fonológica por meio da ultrassonografia no distúrbio do som da fala: relato de caso*. e-mail: rafaela.duartes@hotmail.com

Thalena Evangelista Santos

Mestranda em Letras (Aquisição, Variação e Ensino) no Programa de Pós-Graduação em Letras da Universidade Federal de Pelotas (UFPEL)

e bolsista CAPES, é Licenciada em Letras Português/Francês pela Universidade Federal de Rio Grande (FURG). Atualmente, investiga aspectos acústicos relacionados à aquisição das consoantes fricativas alveolares e alveopalatais do português brasileiro produzidas por imigrantes senegaleses.

E-mail: thalena_x3@hotmail.com

A tradição relativa aos estudos experimentais ligados à fala sempre promoveu alianças, possibilitando a elaboração de projetos comuns, o compartilhamento de materiais, a criação de novos artefatos. A exemplo dessa tradição, encontram-se no ambiente universitário – em que cada área defende o seu próprio território e suas especificidades –, frestas mal vedadas para a realização de trabalhos conjuntos, interdisciplinares e/ou de pesquisas cujos métodos e resultados transitam em *papers*, disciplinas e práticas docentes. Assim sendo, o leitor encontrará, nesta obra, contribuições advindas da linguística e da fonoaudiologia, e verificará que o ultrassom compõe a metodologia dos trabalhos apresentados. O ultrassom surgiu como uma boa opção de equipamento para o estudo dos movimentos da fala: é pequeno, pouco oneroso (comparado a outras tecnologias), sem riscos para os informantes e fornece imagens em tempo real dos movimentos efetuados pela língua durante a articulação de sons significativos. Permite, pois, que o mais veloz e flexível órgão do corpo humano e suas evoluções possam ser observados e analisados no seu tempo e espaço. Os autores dos capítulos desse livro atentam para o corpo que fala, especialmente. Muito do que se produziu sobre a fala excluiu o corpo que a produz. Não somente o corpo, o que o constitui, mas o que o associa à vida, ou seja, o movimento. A atração pelo movimento aproxima os autores desta coletânea.



ISBN 978-65-5869-256-0



9 786558 692560