

Explorando a videomodelação vocal para ensino de comportamento ecoico a crianças diagnosticadas com Transtorno do Espectro Autista (TEA)

Exploring vocal videomodeling to teach echoic behavior to children diagnosed with Autism Spectrum Disorder (ASD)

Explorando modelos vocales en video para enseñar comportamiento ecoico a niños diagnosticados con Trastorno del Espectro Autista (TEA)

Recebido: 10/05/2023 | Revisado: 19/05/2023 | Aceitado: 20/05/2023 | Publicado: 25/05/2023

Lady Anny Araújo do Espírito Santo

ORCID: <https://orcid.org/0000-0002-7808-342X>

Universidade Federal do Pará, Brasil

E-mail: lady_fono@yahoo.com.br

Romariz da Silva Barros

ORCID: <https://orcid.org/0000-0002-1306-384X>

Universidade Federal do Pará, Brasil

E-mail: romarizsb@gmail.com

Resumo

Indivíduos diagnosticados com Transtorno do Espectro Autista (TEA) podem apresentar atraso no desenvolvimento do repertório verbal, como nas respostas ecoicas. É comum na prática fonoaudiológica o uso de pistas multissensoriais para estabelecer e refinar o repertório ecoico. Devido ao contexto de Pandemia da COVID-19 em 2021, o Conselho Federal de Fonoaudiologia (CFFa) determinou o uso de máscara cirúrgica ou N95 durante o atendimento ao paciente, sendo desencorajado o uso de máscaras transparentes. Nesse contexto, para mostrar ao paciente um modelo vocal visual a ser reproduzido, o CFFa orientou ao fonoaudiólogo utilizar gravações em vídeo ilustrativas de modelos vocais. Visto que existe um déficit de conhecimento empírico relacionado a essa recomendação, o presente estudo buscou verificar se pistas visuais multissensoriais apresentadas por meio de vídeos seriam eficazes para atingir acurácia de fala de respostas ecoicas em indivíduos com TEA. Participaram três crianças com diagnóstico de TEA e Apraxia de Fala na Infância. Após a obtenção dos dados de controle com cada participante (em um desenho experimental de sujeito único com sondas múltiplas), o tratamento foi implementado com a introdução da videomodelação vocal. Os dados relatados mostram que apenas um participante conseguiu atingir a precisão de desempenho programada. O presente estudo, portanto, não confirmou a eficácia do procedimento e apontou a necessidade de pesquisas empíricas adicionais para aprimorar o uso de técnicas de videomodelação para ensinar comportamento ecoico para essa população. Um ponto de partida para estudos de tais variáveis de procedimento no ensino de repertório ecoico via videomodelação é fornecido aqui.

Palavras-chave: Comportamento verbal; Fonoaudiologia; Autismo; Transtorno da produção dos sons da fala.

Abstract

Individuals diagnosed with Autistic Spectrum Disorder (ASD) may have a delay in the development of verbal repertoire, such as echoic responses. It is common in speech therapist practice to use multisensory cues to establish and refine echoic repertoire. Due to the context of COVID-19 Pandemic in 2021, the Federal Council of Speech Therapy (CFFa) determined the use of surgical mask or N95 while delivering patient care, and the use of transparent masks was discouraged. In this context, in order to show the patient a visual vocal model that he or she needs to reproduce, the CFFa advised speech therapists to use illustrative video recordings of vocal models. Since there is a deficit of empirical knowledge related to this recommendation, the current study sought to verify whether visual multisensory cues presented through videos would be effective in speech accuracy of echoic responses in individuals with ASD. Three children diagnosed with ASD and Childhood Apraxia of Speech participated. After obtaining control data with each participant (in a single subject, multiple probe experimental design), the treatment was implemented with the introduction of the vocal video modelling. The data reported show that only one participant was able to reach the programmed performance accuracy. This study did not confirm the effectiveness of the procedure and point out the need for additional empirical research to improve the use of video modelling techniques to teach echoic behavior with this population. A starting point for studies of procedural variables in teaching echoic repertoire via videomodelling is provided here.

Keywords: Verbal behavior; Speech, language and hearing sciences; Autism; Speech sound disorders.

Resumen

Las personas diagnosticadas con Trastorno del Espectro Autista (TEA) pueden presentar retraso en el desarrollo del repertorio verbal, como en las respuestas ecoicas. Es común en la práctica de la terapia del habla usar señales multisensoriales para establecer y refinar el repertorio ecoico. Debido al contexto de la Pandemia del COVID-19 en 2021, el Consejo Federal de Fonoaudiología (CFFa) determinó el uso de mascarilla quirúrgica o N95 durante la atención al paciente, desaconsejándose el uso de mascarillas transparentes. En este contexto, para mostrarle al paciente un modelo vocal visual para ser reproducido, la CFFa instruyó al logopeda para utilizar grabaciones de video que ilustran modelos vocales. Dado que existe un déficit de conocimiento empírico relacionado con esta recomendación, el presente estudio buscó verificar si las señales visuales multisensoriales presentadas a través de videos serían efectivas para lograr la precisión del habla de las respuestas ecoicas en personas con TEA. Participaron tres niños diagnosticados con TEA y Apraxia del Habla en la Infancia. Después de obtener datos de control con cada participante (en un diseño experimental de un solo sujeto con múltiples sondas), el tratamiento se implementó con la introducción del modelado de video vocal. Los datos informados muestran que solo un participante pudo lograr la precisión de rendimiento programada. El presente estudio, por lo tanto, no confirmó la efectividad del procedimiento y señaló la necesidad de investigación empírica adicional para mejorar el uso de técnicas de modelado de video para enseñar el comportamiento ecoico para esta población. Aquí se proporciona un punto de partida para los estudios de tales variables de procedimiento en la enseñanza del repertorio ecoico a través del modelado de video.

Palabras clave: Comportamiento verbal; Fonoaudiología; Autismo; Trastorno de producción de sonido del habla.

1. Introdução

O transtorno do espectro autista (TEA) se caracteriza por alterações persistentes na comunicação e interação social em contextos diversos. Dentre as características relatadas, está o atraso no desenvolvimento da comunicação e da linguagem. Zanon et al. (2014) realizaram uma investigação a fim de conhecer quais dificuldades relacionadas ao quadro de TEA são percebidas pelos pais nos primeiros anos de vida dos seus filhos e ainda em que idade isso ocorre. No que se refere à linguagem, dentre as subcategorias criadas (a saber: “atraso na fala”, “atraso na comunicação gestual”, “atraso no balbucio” e “parou de falar”), o sintoma de “atraso na fala” foi o mais relatado pelos pais (61%, n = 11).

Perissinoto (2003) confirma que tais alterações são comumente descritas por familiares e profissionais que interagem com crianças no quadro do TEA, destacando como principais alterações no comportamento inicial da linguagem as dificuldades para reagir a sons, vocalizar, balbuciar e reproduzir sons. Além disso, indivíduos com TEA podem apresentar como comorbidade um transtorno motor de fala (TMF), que pode dificultar a produção vocal. Um TMF que vem sendo estudado por fonoaudiólogos (Rodrigues et al., 2019; Braz et al., 2020, Oliveira et al., 2021) é a *apraxia de fala na infância* (AFI), que pode ser uma comorbidade nos casos de TEA (Dawson, 2010; Shriberg et al., 2011; Tierney et al., 2015)

De acordo com o Comitê *Ad Hoc* sobre AFI, formado através de resolução aprovada pela American Speech-Language-Hearing Association (ASHA):

A apraxia de fala na infância (AFI) é um distúrbio neurológico da fala infantil (pediátrico) no qual a precisão e a consistência dos movimentos subjacentes à fala são prejudicadas na ausência de déficits neuromusculares (por exemplo, reflexos anormais, tônus anormal). A AFI pode ocorrer como resultado de comprometimento neurológico conhecido, em associação com distúrbios neurocomportamentais complexos de origem conhecida ou desconhecida, ou como um distúrbio neurogênico idiopático dos sons da fala. A deficiência central no planejamento e/ou programação dos parâmetros espaço-temporais das sequências de movimento resulta em erros na produção dos sons da fala e na prosódia. (ASHA, 2007, seção *Childhood Apraxia of Speech*)

Indivíduos com AFI podem apresentar imprecisão nas emissões vocais, o que dificulta a compreensão pelo interlocutor (ouvinte). Pode-se dizer que ocorre uma falha na emissão de respostas vocais com correspondência ponto-a-ponto ao estímulo ouvido, ou seja, uma falha na resposta ecoica. É sabido que, dada a importância da imitação da fala (ou seja, do comportamento ecoico) para aquisição de outros operantes verbais, ensinar a criança a reproduzir a verbalização do outro pode garantir respostas adequadas e reforçamento da comunidade verbal, o que não aconteceria com respostas vocais (ecoicas) imprecisas ou pouco inteligíveis. Desse modo, indivíduos no TEA e quadro concomitante de AFI podem vivenciar mais

dificuldades quanto ao desenvolvimento da fala e da linguagem, já que o repertório verbal é, em partes, moldado a partir da capacidade do indivíduo de produzir respostas ecoicas com correspondência ao estímulo antecedente. Essa precisão é importante para que ocorra o reforçamento e, por conseguinte, haja possibilidade de aumento da frequência de comportamento verbal vocal e ampliar seu repertório de linguagem (Abreu & Hubner, 2012).

É fato indiscutível que o repertório frequentemente referido como “linguagem”, por muitos teóricos e profissionais, constitui um conjunto de comportamentos complexos e importantes para a vida do indivíduo. A linguagem, para a análise do comportamento, é tratada com pouca ou nenhuma ênfase ao seu aspecto estrutural (tal como é comum na visão da linguística) e maior ênfase nas interações entre falante e ouvinte. Assim o interesse é no comportamento operante verbal, que é modelado e mantido pelas consequências ambientais mediadas por um ouvinte (Souza & Miguel, 2018). Pierce e Cheney (2013) descreveram operante como um comportamento (qualquer comportamento) que opere no ambiente e produza um efeito. A consequência (ou efeito) altera a probabilidade de ocorrência desse comportamento (operante), que pode vir a acontecer novamente em contexto semelhante.

Desse modo, as interações linguísticas, enquanto comportamento, estão sujeitas às mesmas influências das variáveis ambientais que afetam quaisquer outros comportamentos operantes (Skinner, 1957). A fim de direcionar a atenção às contingências operacionais envolvidas no fenômeno linguagem e diferenciar como essa seria estudada a partir dos conceitos da análise do comportamento, Skinner utilizou o termo *comportamento verbal*. Por essa razão, no presente trabalho, a partir desse ponto, “linguagem” será discutida e referida como comportamento verbal (CV), ou seja, comportamento em que as contingências operantes e condições ambientais são responsáveis pelo estabelecer e manter a performance do falante (Pierce & Cheney, 2013).

O comportamento ecoico, dentro da taxonomia de operantes verbais proposta por Skinner (1957) —que inclui também mando, textual, cópia, intraverbal, tato e autoclítico— é aquele operante verbal que está sob controle de um estímulo verbal e a resposta emitida tem um padrão sonoro que reproduz (total ou parcialmente) o padrão do estímulo. Existe, portanto, correspondência ponto-a-ponto e similaridade formal entre o estímulo e a resposta. Cada parte do antecedente controla cada mínima unidade da resposta e características dinâmicas não são relevantes no operante verbal ecoico, como a intensidade do som. Falar “ecoico” após ouvir alguém falar “ecoico” é um exemplo de comportamento ecoico (Peterson, 1978).

Greer e Ross (2007) elencaram o ecoico, assim como o tato e mando, como requisitos para aquisição de outros operantes por considerarem que eles estão presentes no curso do desenvolvimento típico das crianças, no período de aquisição das primeiras palavras. Quando está aprendendo a falar, a criança tem acesso a reforçadores sociais por atender à solicitação de um adulto para que repita a sua fala (Skinner, 1957). Adicionalmente, através da imitação da fala ou ecoico, os pais ensinam seus filhos os demais operantes verbais, como tatos e mandos. Essas respostas ecoicas são modeladas pela comunidade verbal, até que se tornem precisas e se igualem à fala do adulto (topograficamente), o que ocorre por volta dos 4 anos de idade nas crianças neurotípicas, fase na qual se espera que elas consigam produzir todos os fonemas da língua materna (Ferrante, 2011).

É através desse processo, de repetir a fala do adulto, enquanto ouvinte, que provavelmente a criança se torna um falante, adquirindo através da repetição da fala, outros operantes verbais (Barros, 2003). O comportamento ecoico continua sendo reforçado após essa fase “educacional” da criança (Skinner, 1957), só que de forma indireta, como por exemplo, numa conversa, quando uma das partes repete o que o outro falou, ao perceber que dessa forma a comunicação será mais efetiva. Percebe-se, portanto que, no caso do operante verbal ecoico, a análise vai além da sua função é envolve também sua forma, ou seja, a topografia da resposta. Desse modo, já que o ecoico influencia a modelagem de outros operantes verbais, é importante que a comunidade verbal consequencie respostas ecoicas precisas, já que é esse padrão que poderá ser generalizado para os demais operantes (Abreu & Hubner, 2012).

Ainda que o operante ecoico seja considerado fundamental para aquisição de comportamento verbal vocal, poucas pesquisas são realizadas com o intuito de desenvolver tecnologias de ensino voltadas para a instalação ou refinamento desse operante. Em uma revisão bibliográfica realizada por Esteves et al. (2014), sobre o ensino de tato, mando vocal e ecoico, na base de dados do periódico JABA, verificou-se que nos anos de 1968 a 2012 houve uma tendência decrescente dos estudos envolvendo ensino de ecoico, principalmente a partir da década de 1990 e uma maioria deles (40 artigos), citou o ecoico como um *prompt* para ensino de outros operantes verbais (tato e mando). De fato, várias pesquisas utilizam o ecoico apenas como *prompt*, ou seja, como parte da variável independente nos estudos que objetivam, principalmente, o ensino de mando, tato e intraverbal (Swerdan & Rosales, 2015; Kisamore et al., 2016; Pérez-González & García-Asenjo, 2016; Santos & Souza, 2016; Dass et al., 2018; Shillingsburg et al., 2018; Lee et al., 2020).

Uma revisão sistemática realizada por Guerra et al. (2019) encontrou apenas três trabalhos em que o ensino de ecoico foi estudado diretamente como variável dependente em pessoas com TEA: 1) estudo realizado por Drash et al. (1999), em que um dos procedimentos utilizados foi o uso de mando para produção de um repertório ecoico; 2) pesquisa conduzida por Esch et al. (2005), em que três experimentos foram realizados, porém apenas no Experimento 01 ocorreu ensino de ecoico; c) o terceiro estudo, realizado por Carroll e Klatt (2008), investigou o efeito do pareamento estímulo-estímulo para aumento de vocalizações de baixa frequência e produção de novas vocalizações, para em seguida colocar sob controle operante ecoico.

Um estudo mais recente, de Wang et al. (2023) utilizou o *Pivotal Response Treatment* (PRT), que é uma intervenção comportamental naturalística, para ensinar repertório verbal de tato, mando, ouvinte e ecoico a crianças no TEA. Participaram desse estudo 30 indivíduos, que tinham entre 04 e 08 anos de idade, que foram divididas em dois grupos, um que recebeu a intervenção e outro que foi o grupo controle. Com relação ao ecoico, as tarefas consistiam em a) a criança deveria repetir substantivos com concordância de 90%; b) repetir frases (substantivo + verbo) com 90% de concordância e c) repetir frases curtas. Todas essas tarefas foram realizadas em momentos de brincadeira, como brincar de cortar frutas de brinquedos. O resultado final demonstrou que houve melhoria do repertório ecoico dos participantes que participaram da intervenção, em comparação com o grupo controle, nas fases de intervenção, generalização e manutenção.

Observa-se que os autores desses estudos realizaram vários procedimentos na tentativa de estabelecer repertório ecoico em crianças no TEA, com grande variabilidade interindividual nos resultados. Uma lacuna importante observada nesses estudos foi a falta da descrição ou análise da precisão de fala. Em resposta a essa questão empírica, Guerra (2020) conduziu um experimento, a fim de responder quais os efeitos do ensino do ecoico na precisão de fala de dois indivíduos com TEA e apraxia de fala na infância. Os dois participantes eram do sexo masculino, um participante tinha 04 anos de idade e diagnóstico de TEA e o outro participante tinha 11 anos de idade e diagnóstico de TEA e paralisia cerebral. As palavras alvo escolhidas para o ensino de cada participante dependiam de um critério de desempenho de 50% a 75% de precisão da fala, que era analisada a partir da transcrição da vocalização emitida pelo participante, com posterior verificação da correspondência com o estímulo antecedente, seguindo critério de acurácia da fala. Como variável dependente, a autora determinou uma porcentagem de respostas independentes para ecoico, que era o objeto principal do estudo, além dos comportamentos de ouvinte-seleção e tato para os estímulos apresentados (estímulos 3D). Como variáveis independentes foram elencados os procedimentos de ensino por *multiple exemplar instruction* (MEI), ecoico para mando e ensino por *single exemplar instruction* (SEI). O delineamento utilizado foi de tratamento alternado adaptado. Os principais resultados apontaram que o participante com diagnóstico de TEA apresentou mais respostas corretas diante do ensino por SEI, enquanto o participante com TEA e paralisia cerebral respondeu melhor ao ensino de ecoico para mando.

A condução e os resultados do estudo de Guerra (2020) representam um marco importante na produção de dados acerca do ensino de ecoico e precisão de fala a indivíduos no TEA e quadro concomitante de transtornos dos sons da fala, mais especificamente, da AFI. Na prática profissional do fonoaudiólogo, é comum o uso de pistas multissensoriais no

estabelecimento e facilitação da produção precisa de repertório vocal, incluindo o repertório ecoico, de indivíduos com transtornos do neurodesenvolvimento e quadro concomitante de AFI (Rogers et al., 2006; Fish, 2019).

Segundo Fish (2019), pistas multissensoriais são o uso conjunto e generalizado de *inputs*, que podem ser: a) **táteis, cinestésicas e proprioceptivas** (e.g. a fonoaudióloga toca os lábios do indivíduo, para favorecer a produção de um som alvo); b) **metacognitivas** (e.g. a fonoaudióloga fornece instruções sobre como posicionar um órgão fonoarticulatório para produzir um fonema); c) **visuais** (e.g. imagem da boca da fonoaudióloga durante a produção do som alvo em vídeos ou cartões, uso do espelho); e d) **auditivas** (e.g. modelo auditivo vocal do som alvo, falar de forma lentificada, silabar o fonema alvo; p. 134).

O contexto da pandemia da COVID-19 dificultou a atuação clínica do fonoaudiólogo(a) que atua com ensino de fala a indivíduos com TEA, já que o uso de determinados equipamentos de proteção individual (EPIs), como as máscaras, impede a visualização, pelo indivíduo que está em intervenção, dos modelos de fala emitidos pela profissional e do uso de algumas pistas multissensoriais (e.g. pistas visuais). No dia 04 de março de 2021, o Conselho Federal de Fonoaudiologia (CFFa), através da sua rede social Instagram, recomendou que era imprescindível o uso de máscara cirúrgica ou N95 no atendimento dos pacientes, visando a proteção do profissional e do paciente. Ainda nessa recomendação, o uso de máscaras transparentes é colocado como uma prática não segura, uma vez que esse tipo de máscara não filtra o ar expirado ou inspirado e não tem uma boa adesão ao rosto. Finalmente, como uma alternativa ao atendimento fonoaudiológico em que haja necessidade de se mostrar ao paciente um modelo que ele precise reproduzir (e.g. modelo visual da posição da boca e língua para a produção vocal), na impossibilidade do uso da máscara transparente e na obrigatoriedade de uso da máscara cirúrgica ou N95, o CFFa orientou que os fonoaudiólogos (as) utilizassem gravações de vídeos ilustrativos desses modelos com os pacientes.

Esse tipo de procedimento já é utilizado nas intervenções para ensino de outros comportamentos dentro do contexto da análise do comportamento aplicada (ABA) ao TEA, sendo conhecido pelos termos: a) *video modelling*, em inglês (Corbett & Abdullah, 2005); b) videomodelação (Lobato et al., 2018) ou c) modelagem em vídeo (Rodrigues & Almeida, 2017), em português. A videomodelação ou modelagem em vídeo (MV) foi apontada pelo *The National Clearinghouse on Autism Evidence & Practice* (NCAEP), como uma prática baseada em evidência, no ano de 2020.

A modelagem em vídeo (ou videomodelação) é um tipo de intervenção comportamental que consiste no sujeito realizar a observação e posterior reprodução (prática e imitação), de um determinado comportamento, que é exibido por um modelo em um vídeo (Corbett & Abdullah, 2005). Em revisão de literatura realizada por Rodrigues e Almeida (2017), foram incluídos 11 artigos que envolviam ensino à indivíduos com TEA a partir de videomodelação, nos anos de 2010 a 2016. Alguns dos experimentos realizados envolviam o ensino de repertório verbal, como por exemplo o ensino de comunicação suplementar e alternativa com dispositivo gerador de voz (Banda et al., 2010), PECS (Smith et al., 2014) e ensino de mandos (Plavnick & Vitale, 2016).

Além desses, outros repertórios importantes também podem ser ensinados através de videomodelação, como por exemplo responder apropriadamente a chamados de estranhos e pessoas conhecidas (Abadir et al., 2021) e habilidades de vida diária (Wertalik, 2021). Esses estudos, além dos resultados positivos, foram considerados socialmente válidos, sinalizando que a videomodelação pode ser um procedimento eletivo para ensino repertórios variados, incluindo o ensino de operantes verbais.

Considerando, portanto, a importância fundamental de um repertório ecoico preciso para desenvolvimento de comportamento verbal vocal e o déficit de conhecimento empírico relacionado à recomendação do CFFa sobre o uso de gravações de vídeos ilustrativos como modelos ao paciente, o objetivo do presente estudo foi verificar se modelos vocais pré-gravados em vídeos ilustrativos tem efetividade no treino da precisão de respostas ecoicas, ou seja, na inteligibilidade de fala de indivíduos com TEA e quadro concomitante de Apraxia da Fala na Infância.

2. Metodologia

Participantes

Participaram dessa pesquisa três crianças com diagnóstico clínico de transtorno do espectro autista, transtorno do déficit de atenção e hiperatividade (TDAH) associado ao TEA e diagnóstico fonoaudiológico de apraxia de fala na infância, devidamente confirmados através dos laudos e relatórios apresentados pelos seus responsáveis. Além disso, para serem incluídos nesse estudo, os participantes deveriam ter: 1) idade igual ou superior a 5 anos; 2) repertório mínimo ecoico, de ouvinte e imitação do nível I do *Verbal Behavior Milestones Assessment and Placement Program* (VB-MAPP); e 3) apresentar comportamentos de colaboração, tais como permanecer sentado por pelo menos 10 minutos sem emitir comportamento disruptivo, esperar por 5 segundos, atender a instruções simples e entregar o item reforçador quando solicitado. A Tabela 1 contém a descrição resumida do repertório dos participantes.

Tabela 1 - Descrição resumida do repertório verbal dos participantes.

	Participantes		
	André	Otto	Gael
Idade	11	11	5
Diagnóstico	TEA, TDAH, AFI	TEA, TDAH, AFI	TEA, TDAH, AFI
APCE	Ecoico de sílabas simples e duplicadas (e.g. não, mama, piu, um, eu, ai, pão, dada, au au)/ ecoico de combinação de duas sílabas (e.g. copo, vaca, pote, minha, pato, meu pé, nada, tatu)/ ecoico de combinação de 03 sílabas apenas do estímulo “tudo bem”.	Ecoico de sílabas simples e duplicadas (e.g. au au, papa, tutu, pão)/ ecoico de combinação de duas sílabas (e.g. copo, gato, vaca, pato, nada, bota, tatu, boca, pote)/ ecoico de combinação de 03 sílabas apenas dos estímulos “pipoca” e “tomate”.	Ecoico de sílabas simples e duplicadas (e.g. au au, oi, piu, tutu, babá, ão, um, eu, boi)/ ecoico de combinação de duas sílabas (e.g. boca, nunca, pato, bota, copo, tatu, maçã)/ ecoico de combinação de 03 sílabas apenas do estímulo “batata”.
VB MAPP	<i>Nível 1</i> – pontuação máxima em todos os marcos/ <i>Nível 2</i> – pontuação máxima em todos os marcos, exceto no mando (4,5 pontos)/ <i>Nível 3</i> – pontuação máxima nos marcos de mando, ouvinte, pareamento, brincar independente, brincar social, grupo e matemática. Nos demais marcos: tato (4,5 pontos), RFCC (4 pontos), intraverbal (3,5 pontos), estrutura linguística (4 pontos), leitura (4 pontos), escrita (4 pontos).	<i>Nível 1</i> – pontuação máxima em todos os marcos, com exceção do comportamento social (2 pontos), imitação (4 pontos) e vocal (2,5 pontos)/ <i>Nível 2</i> – pontuação no tato (2,5 pontos), ouvinte (2,5 pontos), pareamento (1 ponto), brincar (2,5 pontos), imitação (3,5 pontos), RFCC (0,5 pontos), intraverbal (2,5 pontos) e linguística (1,5 pontos)/ <i>Nível 3</i> – pontuação em leitura (2,5 pontos), escrita (3,5 pontos), matemática (2 pontos).	<i>Nível 1</i> – pontuação máxima em todos os marcos, com exceção do brincar (3 pontos), social (4 pontos), ecoico (sem avaliação) e vocal (4 pontos)/ <i>Nível 2</i> – pontuação no tato (2 pontos), ouvinte (5 pontos), pareamento (5 pontos), brincar (2 pontos), imitação (4,5 pontos), RFCC (3 pontos), intraverbal (5 pontos), grupo (5 pontos) e linguística (2 pontos)/ <i>Nível 3</i> – ouvinte (2 pontos), pareamento (5 pontos), leitura (3 pontos), escrita (5 pontos), grupo (1 ponto), matemática (4 pontos).

Fonte: Elaborado pelos autores (2023).

Pode-se observar pelos dados dispostos na Tabela 1, que os três participantes possuíam repertório verbal de entrada semelhante, principalmente com relação ao repertório ecoico, pois conseguiam emitir respostas ecoicas de sílabas simples e duplicadas e de combinações de duas sílabas, com maior dificuldade para ecoar estímulos com três sílabas.

Após a seleção inicial das crianças e autorização da realização da pesquisa pelo Comitê de Ética em Pesquisa do Hospital Universitário João de Barros Barreto da Universidade Federal do Pará (HUIBB – UFPA – nº 4.857.364), a responsável por essa pesquisa entrou em contato com os responsáveis, realizando um convite para que comparecessem ao ambiente onde a coleta ocorreu, para que pudessem conversar pessoalmente sobre os objetivos da pesquisa. Após esse contato inicial e aceite dos responsáveis, os responsáveis assinaram o Termo de Consentimento Livre e Esclarecido, o Termo de

Autorização de uso da imagem e os participantes assinaram o Termo de Assentimento, concordando com os termos da pesquisa.

Ambiente e medidas de segurança

Por conta da Pandemia do COVID-19, o experimento foi realizado no domicílio das crianças ou em uma sala especificamente preparada para fins de pesquisa e atendimento. O ambiente da sala de pesquisa e atendimento era equipado com mesas e cadeiras grandes e pequenas, além de materiais lúdicos variados. A iluminação ocorria por luz artificial e a temperatura era controlada pelo uso do ar-condicionado. Todo o ambiente foi higienizado de acordo com as regras estabelecidas pela Organização Mundial de Saúde (OMS), considerando que ainda estávamos em Pandemia da COVID-19 no segundo semestre de 2021. Além disso, a pesquisadora estava devidamente vacinada e em todas as sessões utilizava os equipamentos de proteção individual (*e.g.* máscara, capote, faceshield etc.).

Instrumentos

Os seguintes instrumentos foram utilizados: a) o instrumento de Avaliação Precoce de Competências Ecoicas (APCE), que é um subteste do VB-MAPP e pode ser encontrado na tese de doutorado de Martone (2017); b) um questionário de tradução livre adaptado do instrumento proposto por Fisher et al. (1996), o *Reinforcement Assessment for Individuals with Severe Disability* (RAISD), que foi respondido pelos pais; e c) a avaliação de preferências de operante livre (ambiente livre), de Roane et al. (1998). O APCE foi utilizado para auxiliar na seleção dos fonemas-alvo do estudo, em que os participantes passaram apenas pela avaliação dos Grupos 1, 2 e 3 de sílabas e combinação de sílabas. Os demais instrumentos serviram para que o pesquisador pudesse selecionar itens ou atividades de preferência dos participantes, que pudessem ser usados como potenciais reforçadores.

Delineamento

O delineamento experimental utilizado neste estudo foi de sujeito único, de sondas múltiplas (Horner & Baer, 1978).

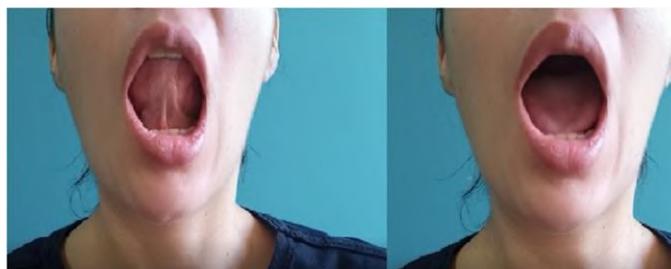
Variável dependente

As variáveis dependentes analisadas foram a porcentagem de ocorrência independente de respostas ecoicas, que deveriam ter correspondência ponto a ponto com os estímulos discriminativos auditivos, conforme descrito na Tabela 2. Esses estímulos eram compostos de fonemas-alvo (sílabas), que faziam parte ou da primeira ou da segunda sílaba de palavras que foram escolhidas a partir da avaliação do repertório ecoico de cada participante. A criança poderia responder simultaneamente ou após assistir o vídeo, em até 3 segundos.

Variável independente

A variável independente manipulada foi o tratamento implementado: videmodelação com pistas auditivo-visuais, em que os fonemas-alvos foram apresentados através de vídeos ilustrativos pré-gravados, conforme exemplificado na Figura 1.

Figura 1 - Demonstração visual da produção vocal do fonema /LA/.



Fonte: Elaborado pelos autores (2023).

Na Figura 1, é possível observar um exemplo de uma pista visual que pode auxiliar indivíduos com dificuldades na produção da fala a, por exemplo, produzir o fonema /LA/, através de um exemplo de como esse fonema ocorre, com a ponta da língua tocando nos alvéolos dentais, na direção localizada posteriormente aos dentes incisivos superiores e movimentando-se para baixo, seguida da produção quase concomitante da vogal A, momento em que a língua permanece no assoalho bucal (Silva, 2007).

A seleção dos estímulos utilizados nos vídeos seguiu alguns princípios norteadores: estrutura silábica, repertório fonêmico da criança, fatores motivacionais e funcionalidade (Fish, 2019). Dessa forma, foram selecionadas palavras que contivessem os fonemas-alvo, com estrutura silábica simples de consoante – vogal – consoante – vocal (*e.g.* palavra BOLA), considerando o repertório vocal de fonemas apresentado pelas crianças pós-avaliação de ecoico com o APCE, conforme exemplificado na Tabela 2.

Tabela 2 - Estímulos discriminativos.

Fonemas/alvo	Participantes e fonemas-alvo correspondentes					
	André		Otto		Gael	
	“za”	“rro”	“rro”	“la”	“rro”	“ga”
Estímulo	CASA	CARRO	CARRO	LAVO	CARRO	GABÍ
Resposta	CAZA	CARRO	CARRO	LAVO	CARRO	GABÍ

Fonte: Elaborado pelos autores (2023).

Na Tabela 2 estão demonstrados os estímulos discriminativos utilizados nesse estudo, que eram compostos de fonemas-alvo (sílabas), que faziam parte ou da primeira ou da segunda sílaba de palavras escolhidas a partir da avaliação do repertório ecoico de cada participante. Além disso, também foram selecionados alvos que pudessem ser utilizados pelos participantes com outras funções, como mando (solicitar item preferido, como carro) e que parte dessa palavra contivesse fonemas produzidos pelos participantes sem nenhuma dificuldade. Por exemplo, se o alvo fosse /LE/ de palavra “LEGO”, o participante deveria conseguir falar sem dificuldade a segunda sílaba da palavra, no caso “GO”.

Com relação às pistas que foram utilizadas, essas foram baseadas nos tipos de pistas multissensoriais utilizadas por fonoaudiólogos(as) nas intervenções aos transtornos de fala (Fish, 2019). Para essa pesquisa, as pistas, que foram acrescentadas ao estímulo auditivo original que deve controlar a resposta ecoica, foram a auditiva, que consistiu na variação de velocidade (*e.g.* falar lentamente) e a visual, que correspondia a mostrar a sobrearticulação (*e.g.* articulação exagerada de abertura da boca).

A pesquisadora principal, que é Fonoaudióloga, serviu como modelo para gravação dos vídeos, para os quais utilizou-se a câmera de um celular Samsung Galaxy A32, com câmera de 65 megapixels e resolução de 1920 x 1080 pixels para gravação de vídeos em alta definição (*Full HD*), posicionado em tripé, em ambiente silencioso, com luz natural e/ou artificial. A pesquisadora utilizou uma blusa básica azul marinho sem detalhes, cabelo preso, sem brincos e focando a câmera na direção

da boca, sem mostrar os olhos. Todas essas medidas foram tomadas para evitar distrações com outros estímulos que não fossem os alvos da pesquisa. Finalmente, os vídeos eram apresentados aos participantes em um tablet Samsung Galaxy Tab A7 Lite de 64 GB, cujo *display* (tela) é de 8,7 polegadas, velocidade de processador *octacore* de 2.3GHz, 1.8GHz, resolução alta de 1340 x 800 *pixels* e alto-falantes estéreos duplos e efeito *Dolby Atmos*, que otimizam a qualidade do som.

Medição

Foi utilizada a câmera de um celular para registrar todas as sessões em vídeos, os quais foram utilizados para verificação da integridade da implementação do procedimento e da qualidade do registro através de acordo entre observadores independentes. Cada sessão teve a duração máxima de 45 minutos. Os dados coletados foram quantificados em termos de porcentagem, considerando o número total de acertos por total de tentativas implementadas na sessão para cada fonema-alvo.

Acordo entre observadores

Foi realizada a avaliação de acordo entre observadores para aproximadamente 30% do total de sessões de cada participante, selecionados aleatoriamente através de sorteio ao longo do experimento. Os registros de dados coletados durante as sessões pelo experimentador foram comparados aos registros feitos por observador independente treinado (modelos de registros nos Apêndices 5, 6 e 7). Foi considerado acordo quando os registros produzidos pelo observador e pelo pesquisador para uma determinada resposta foram coincidentes. Foi considerado aceitável o nível mínimo de acordo entre observadores de 80%, sendo este obtido através da divisão do número total de acordos pelo somatório de acordos e desacordos multiplicado por 100 (Cooper et al., 2007). Para o participante André, o acordo foi de 100%, enquanto para o participante Otto foi de 82,5% e para o participante Gael foi de 85,84%. A Tabela 3 exemplifica a forma de registrar as repostas dos participantes.

Tabela 3 - Forma de registrar as respostas emitidas pelos participantes.

Respostas esperadas	Marcação
Resposta independente	+
Resposta incorreta ou ausência de resposta	-

Fonte: Elaborado pelos autores (2023).

A Tabela 3 demonstra como deveriam ser realizados os registros das respostas, que poderiam ser “resposta independente”, em caso de respostas corretas ou “resposta incorreta ou ausência de resposta”, considerando respostas erradas, além da devida marcação que deveria ser sinalizada na folha de registro.

Integridade do Tratamento

As sessões de intervenção do participante André (sessões 01, 02, 04 e 05 realizadas na sala de pesquisa e sessões 03 e 06 em seu domicílio), foram realizadas pela pesquisadora principal e uma auxiliar treinada pela pesquisadora, para evitar possíveis divergências quanto à aplicação dos procedimentos. Já as sessões de intervenção dos participantes Otto (realizadas na sala de pesquisa) e Gael (realizadas na sala de pesquisa, com exceção da sessão 02, realizada em domicílio) foram conduzidas apenas pela pesquisadora principal. Os dados para realização da integridade do procedimento foram coletados em aproximadamente 30% das sessões, a partir do total das sessões de intervenção, para cada participante, selecionadas aleatoriamente através de sorteio.

Foram considerados alguns passos para realização da avaliação da integridade: ao realizar a tentativa, o experimentador deveria: 1) obter a atenção da criança, 2) apresentar de forma adequada o estímulo antecedente 3) reforçar a resposta correta emitida pela criança (que deveria ser emitida em até cinco segundos) 4) reforçar a resposta, se correta, e em

seguida registrá-la e 5) respeitar o intervalo entre as tentativas (IET) de três segundos. A integridade do tratamento foi avaliada por um observador treinado que avaliou através dos vídeos a implementação do procedimento, dividindo-se o número total de componentes implementados adequadamente pelo número total de componentes aplicados adequadamente e multiplicando por 100 (Cooper et al., 2007). Para o participante André, a integridade foi de 96%, já para os participantes Otto e Gael foi de 98%.

Procedimento de coleta de dados

Sondagem e Linha de base

Estas fases têm por objetivo avaliar a frequência das respostas ecoicas dos participantes para os fonemas-alvo, diante do modelo emitido pelo pesquisador. Todos os participantes passaram pela sondagem ao longo do experimento, sempre antes da intervenção e ao final dela. O procedimento de sondagem, assim como o de linha de base, foi realizado utilizando-se tentativa discreta, com a diferença de que ainda que o participante respondesse corretamente, ele não receberia reforçamento positivo nessa fase. A Tabela 4 exemplifica os dados obtidos na sondagem e linha de base.

Tabela 4 - Fonemas-alvo e respostas dos participantes na sondagem inicial e linha de base.

Fonemas/alvo	André		Otto		Gael	
	“za”	“rro”	“rro”	“la”	“rro”	“ga”
Estímulos	CASA	CARRO	CARRO	LAVO	CARRO	GABÍ
Respostas	“CAJA” “CARRAS”	“CAO”	“CACO”	“UADU” “UAVO” “UAOTT” “UAPOTT”	“CAO”	“ABÍ”

Fonte: Elaborado pelos autores (2023).

Na Tabela 4 encontram-se os fonemas-alvo, identificados em vermelho para cada estímulo e as respostas que foram emitidas pelos participantes nas sessões de sondagem iniciais (pré-intervenção) e linha de base.

Intervenção

A intervenção toda ocorreu utilizando-se o Treino por Tentativas Discretas, em formato de single exemplar instruction (SEI). Inicialmente, todos os participantes deveriam ser expostos a 30 tentativas de ensino para cada fonema-alvo, porém, após o primeiro participante não ter atingido o critério de precisão de desempenho para nenhum fonema-alvo, decidiu-se que os demais participantes passariam por 15 tentativas de treino apenas, para cada fonema-alvo. A quantidade total de tentativas de cada fonema-alvo foi baseada na descrição realizada por Fish (2019), de que “Os alvos menos estáveis podem receber mais exercícios de prática (por exemplo aproximadamente 50 repetições por sessão) e alvos mais estáveis podem receber menos exercícios de prática (por exemplo, aproximadamente 10 repetições por sessão), na mesma sessão” (p. 64). A tentativa inicial ocorria com o pesquisador garantindo a atenção da criança, em seguida apresentando o estímulo discriminativo auditivo-visual em vídeo e emitindo a instrução: “repita”, “faz igual” ou “sua vez”. Era esperado um período de 3 segundos para que a criança emitisse a resposta e havia um intervalo entre tentativas (IET) de 3-5 segundos. Foram considerados acertos a emissão de uma resposta ecoica ponto-a-ponto e erros respostas com topografias diferentes do modelo emitido pela pesquisadora ou ausência de resposta. Após cada tentativa, era realizado o registro da resposta emitida por cada participante. Para garantir o engajamento dos participantes nas tarefas, foi utilizada uma combinação de esquema de reforçamento contínuo (FR1) para o reforço social com um esquema de reforçamento intermitente (VR2) para a apresentação de fichas (procedimento de economia de fichas), além de um esquema de razão fixa (FR6) para acesso a tangível e/ou intervalo (ao conseguir seis fichas, o participante poderia acessar um reforçador tangível e/ou intervalo). Caso não houvesse emissão de respostas corretas, respostas alternativas de baixo custo de resposta eram solicitadas, para liberação de reforço positivo de menor magnitude.

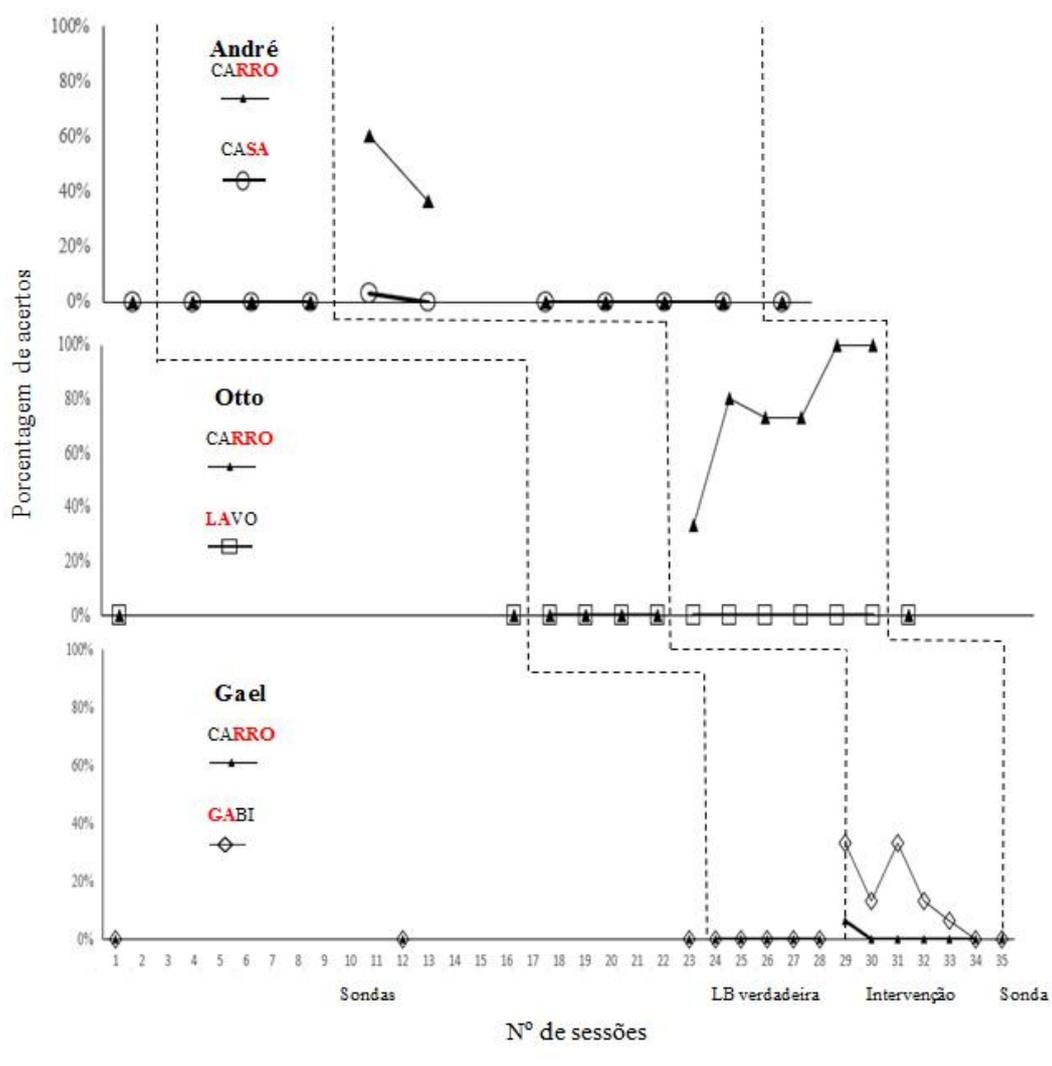
Análise de dados

Considerando que esse é um delineamento experimental de sujeito único, o desempenho de cada participante na linha de base foi comparado ao desempenho do mesmo participante sob a Intervenção. A cada sessão, foram contabilizados o número de acertos (respostas independentes sem ajuda), dividindo-se pelo total de respostas (acertos e erros) para cada um dos fonemas-alvo. Posteriormente, esse valor foi multiplicado por 100, para se obter o percentual de precisão de desempenho dos participantes em cada um dos fonemas-alvo separadamente.

3. Resultados

Os resultados obtidos a partir da intervenção são apresentados para cada participante, considerando os fonemas-alvos selecionados e o percentual de precisão de desempenho das respostas. Foram consideradas corretas respostas ecoicas com correspondência ponto-a-ponto para os fonemas-alvo (estímulos). Por exemplo, se o fonema-alvo era a segunda sílaba da palavra “CARRO”, no caso “RO” (lê-se ‘rro’), então seria considerado 100% de acerto se o participante ecoasse os fonemas que ele já conseguia vocalizar corretamente, mais o fonema-alvo, no caso “CA” e “RRO”. Na Figura 2 encontram-se os dados em porcentagem de acertos para cada fonema-alvo dos participantes.

Figura 2 - Percentagem de acertos para cada fonema-alvo para os participantes André, Otto e Gael.



Fonte: Elaborado pelos autores (2023).

Na Figura 2, estão os resultados dos três participantes, para cada fonema-alvo. Nas sessões de sondagem e linha de base verdadeira não houve respostas corretas emitidas por nenhum dos participantes. No caso do primeiro participante, André, logo na primeira sessão de intervenção ele apresentou 60% de acertos para o fonema-alvo “RO” (lê se ‘rro), de “CARRO”, porém após essa sessão apresentou tendência decrescente dos dados, com uma porcentagem de acertos de 36,66% na segunda sessão e a partir dessa sessão até o final da intervenção, não emitiu mais respostas corretas. Com relação ao fonema-alvo “SA”, de “CASA”, o participante emitiu uma única vez uma resposta correta durante a primeira sessão de intervenção (tentativa 29 de 30).

O participante Otto, que iniciou a intervenção logo após o primeiro participante, apresentou uma porcentagem de acertos de 33,33% logo na primeira sessão para o fonema-alvo “RO” (lê se ‘rro), de “CARRO”, e seguiu emitindo respostas corretas até atingir o percentual de 100% de respostas corretas para esse fonema nas duas últimas sessões de intervenção. Na sondagem pós-intervenção, porém, o participante não apresentou emissão de respostas corretas. Quanto ao fonema-alvo “LA”, de “LAVO”, não foram observados efeitos da intervenção sobre a emissão desse alvo.

O último participante, Gael, apesar de ter apresentado respostas corretas em cinco das seis sessões de intervenção, para o fonema-alvo “GA” de “GABI”, não atingiu o percentual de precisão de desempenho estipulado para esse estudo e, além disso, apresentou diminuição dessas respostas, finalizando a intervenção sem emissão de respostas corretas para esse fonema-alvo. Para o outro fonema-alvo, “RO” (lê se ‘rro), de “CARRO”, Gael apresentou uma única vez uma resposta correta, logo na primeira sessão de intervenção (tentativa 06 de 15).

4. Discussão

Esse estudo buscou verificar se o uso de pistas multissensoriais auditivo-visuais, apresentadas através de videmodulação, facilitariam a emissão de respostas ecoicas precisas (com correspondência ponto-a-ponto) de indivíduos no TEA e quadro concomitante de AFI. Os resultados mostraram que apenas um dos participantes (Otto) atingiu a porcentagem de acertos estipulada, com 100% de acertos nas duas últimas sessões de intervenção, e exclusivamente para o fonema-alvo “CARRO”. Desse modo, os resultados dessa pesquisa apontam para o fato de que as pistas auditivo-visuais utilizadas na apresentação dos estímulos, em formato de vídeo podem não ser suficientes para provocar mudanças significativas na fala de indivíduos diagnosticados no TEA e com AFI. Apesar das circunstâncias emergenciais relacionadas ao enfrentamento da pandemia de Covid-19 ter levado ao uso desse tipo de procedimento, a presente pesquisa chama a atenção para a necessidade de aperfeiçoamento e desenvolvimento de procedimentos alternativos para enfrentamento de situações semelhantes no futuro.

Com relação ao primeiro participante a passar pela intervenção (André), algumas variáveis foram identificadas e podem ter interferido nos resultados: a) na terceira sessão, logo após iniciar, o participante pediu para interromper, pois estava sentido dor de garganta e a sessão foi imediatamente suspensa; b) depois dessa sessão, por ter ficado doente, o participante retornou após 11 dias afastado (no geral as sessões eram realizadas com um intervalo máximo de 03 dias entre elas); c) nesse momento, ele compareceu sem óculos, o que pode ter interferido na sua acuidade visual (considerando o tipo de estímulo e dica que foram apresentados nesse estudo); d) o participante retornou para escola nesse período e passou por mudança de rotina, além de ter realizado retirada e mudança de medicação; e) o número de tentativas de intervenção eram 30, para cada fonema-alvo.

Não é possível afirmar a interferência dessas variáveis nos resultados, mas quando se fala no número total de tentativas, esse pode ter sido um fator importante, principalmente quando ocorreram mais erros que acertos, pois nesse caso havia diminuição ou ausência de consequências reforçadoras e, conseqüentemente, diminuição ou extinção das respostas consideradas corretas. Recomenda-se em pesquisas futuras a introdução imediata de procedimentos antecedentes que evitem a

ocorrência de erros, como repetições do estímulo e inserção de tentativas de imitação motora intercaladas com tentativas de repertório ecoico.

Quanto ao participante Gael (terceiro participante), apesar de ele não ter atingido critério de precisão de desempenho para nenhuma dos fonemas-alvos “CARRO” e “GABI”, observou-se que em algumas tentativas ele emitiu aproximações das respostas corretas, como por exemplo “CALHO” e “CABI”. Apesar de essas respostas não terem sido quantificadas, por não fazerem parte do objetivo do estudo, essas emissões mostram que houve mudança no padrão da fala, quando comparadas às respostas da sondagem e da linha de base, porém seriam necessários ajustes nos procedimentos utilizados na intervenção para resultados mais precisos.

Uma possibilidade seria o reforçamento diferencial de aproximações sucessivas da resposta final esperada, como ocorre na modelagem. Nesse tipo de intervenção, o comportamento ensinado é aprendido gradativamente e pode-se partir de uma linha de base sem respostas ou respostas mínimas emitidas pelo indivíduo (Martin & Pear, 2018). O ensino do repertório ecoico a partir de modelagem consistiria em o instrutor reforçar vocalizações aproximadas ao modelo, emitidas pelo aprendiz, inserindo níveis de exigências de modo gradativo, até que o aprendiz conseguisse emitir a resposta correta (Sundberg & Partington, 1998; Souza & Miguel, 2018). Esse procedimento também foi descrito por Fish (2019), que é fonoaudióloga, a qual descreveu a importância de reforçar produções vocais, ainda que simples, pois caso contrário o indivíduo pode ficar desmotivado e reduzir as tentativas de comunicação vocal.

O segundo participante, Otto, foi o único a apresentar respostas corretas para o fonema-alvo “CARRO” ao final da intervenção, com 100% de acerto nas duas últimas sessões (quinta e sexta sessões). Nesse caso, o procedimento utilizado de apresentação de modelos auditivo-vocais com pistas auditivo-visuais (lentificação e sobrearticulação da fala), por videmodulação, foi eficaz. Esse dado pode ser corroborado por um estudo realizado por Aravamudhan e Awasthi (2019), no qual duas crianças diagnosticadas no TEA e com transtornos dos sons da fala, com 11 e 15 anos, descritos como menino (TA) e menina (KS), foram submetidas à intervenção comportamental com o objetivo de melhorar a articulação da fala. Como apenas apresentar um modelo vocal não era suficiente para que os participantes emitissem respostas ecoicas corretas, prompts adicionais foram integrados ao modelo vocal (*e.g. within-stimulus*). Uma das pistas adicionais utilizadas, que foi aplicada aos dois participantes do estudo, foi o “exagero” na produção do fonema-alvo, por exemplo na palavra “stop”, o terapeuta poderia falar da seguinte forma “sssstop” e depois esvanecer a dica. Ao final da intervenção, ambos os participantes demonstraram melhorias na articulação da fala, sendo que para o participante TA apenas o prompt do tipo *within-stimulus* (exagero na emissão do fonema-alvo) foi suficiente para produzir mudanças, tal qual ocorreu com o participante Otto nesse estudo, porém apenas para um fonema-alvo (o alvo “RO” de “CARRO”).

Observa-se que os resultados dos três participantes à intervenção proposta nessa pesquisa foram significativamente variados: enquanto o primeiro participante iniciou emitindo respostas corretas e em seguida parou de emití-las para o fonema-alvo “CARRO” e apresentou apenas uma emissão correta do fonema-alvo “CASA”; o segundo participante atingiu precisão de desempenho para o fonema-alvo “CARRO” e não emitiu nenhuma resposta correta para o alvo “LAVO”; e o terceiro participante emitiu uma única vez uma resposta correta para o fonema-alvo “CARRO”, enquanto que para o alvo “GABI”, apresentou algumas respostas corretas em cinco das seis sessões de intervenção, porém com uma porcentagem menor que 50% em todas as sessões, finalizando a intervenção sem emitir nenhuma resposta correta.

Outros estudos em que o objetivo foi o ensino do ecoico também apresentaram essa variação, no que se refere aos resultados, o que pode ter relação com o repertório de aprendizagem de cada indivíduo, história de reforçamento, tipos de prompts utilizados, fonemas-alvo selecionados, quantidade de sessões de intervenção, intervalo entre sessões, idade dos participantes, magnitude das barreiras de aprendizagem determinadas pelos transtornos e comorbidades (Drash et al., 1999; Tarbox et al., 2009; Cividini-Motta et al., 2016; Shane, 2016; Aravamudhan e Awasthi, 2019; Guerra, 2020).

Esses dados mostram que cada indivíduo pode responder a diferentes tipos de intervenção no aprendizado de repertório ecoico e uma variável importante pode ser o tipo de fonema que se pretende ensinar. Provavelmente fonemas que podem ser visualmente exemplificados, como /p/, /b/, /f/ e /v/ podem ser ensinados a partir de encadeamento e prompts auditivos e/ou visuais, diferentemente de fonemas que não são “visíveis”, como /k/ e /g/. Neste último caso, o fonoaudiólogo(a) pode utilizar técnicas adicionais (e.g. *Prompts for Restructuring Oral Muscular Phonetic Targets* [PROMPT©]), aprendidas na sua formação, combinadas à procedimentos analítico comportamentais, para favorecer o ensino de ecoico. Aravamudhan e Awasthi (2019) também precisaram utilizar prompts variados, dependendo dos fonemas-alvos que pretendiam ensinar. Por exemplo, enquanto um participante (TA) respondeu à intervenção com apenas uma dica adicional (e.g. exagero na produção do fonema-alvo), a outra participante (KS), precisou de outras pistas para emitir respostas corretas (e.g. exagero ou prolongamento de sons alvos, modelo de posicionamento labial e dental mostrado pelo terapeuta, e uso de uma colher para ensinar o som “ck” isoladamente)

O uso de diversos tipos de pistas, principalmente as multissensoriais, é amplamente recomendando por estudiosos da área da Fonoaudiologia, para o tratamento da apraxia de fala na infância (Dale & Hayden, 2013; Square et al., 2014; Namasivayam et al., 2015; Fish, 2019). Essas pistas devem ser selecionadas a partir de uma avaliação minuciosa do indivíduo, o que permite selecionar quais delas podem ser efetivas no ensino de fala. Fish (2019) descreveu que “Por exemplo, algumas crianças têm mais sucesso quando recebem pistas táteis, enquanto outras preferem *input* visual e auditivo simultâneos” (p. 136). Cividini-Mota (2016) também propôs que uma avaliação de preferências sobre procedimentos de ensino de ecoico pode ser realizada, considerando a possibilidade de vários métodos serem eficazes nesse tipo de intervenção.

Apesar das limitações desse estudo, os dados aqui apresentados podem contribuir para o avanço da compreensão sobre ensino de ecoico em indivíduos no TEA e quadro concomitante de AFI, considerando o número limitado de pesquisas cujo objetivo é ensinar ecoico, especificamente para essa população (Rogers et al., 2006; Aravamudhan e Awasthi, 2019; Guerra, 2019). Além disso, considerando o sugerido por Wolf (1978) como elementos de análise da *validade social* de estudos em ABA —significância social dos objetivos, adequação social dos procedimentos e a importâncias social dos efeitos— pode-se dizer que esse estudo possui relevância social, já que é o desenvolvimento de tecnologias comportamentais para o aprimoramento de habilidades de fala de indivíduos no TEA é de grande interesse da comunidade acadêmica e também foi sinalizado como um objetivo de estudo importante pelos responsáveis dos participantes.

5. Considerações Finais

Finalmente, uma sistematização dos problemas e contribuições derivados do presente trabalho pode ser assim listada:

É possível que a aprendizagem fosse favorecida se algum procedimento ágil de remoção das pistas fosse implementado de forma a dar oportunidades constantes de respostas independentes. Isso, contudo, dependeria de trabalho intenso de edição de vídeos ou de elaboração de numerosos exemplares de vídeos com diferentes níveis de ajuda. É possível que esse custo de resposta para a implementação do procedimento o tornasse pouco prático para sua adição na prática clínica. Essa flexibilidade na apresentação e remoção das pistas é muito adequada para procedimento ao vivo, em oposição ao uso de vídeos como estímulos-modelo.

O treino extensivo, com elevado número de tentativas e dentro do formato de DTT pode ter diminuído o engajamento dos participantes, apesar das combinações de esquemas de reforçamento implementadas. Em pesquisa futuras, recomenda-se adaptar o procedimento a uma estratégia de ensino naturalístico com maior grau de incidentalidade.

A magnitude da barreira de aprendizagem determinada pela apraxia de fala na infância pode ser alta e isso pode significar que um número muito grande de sessões seja necessário para que se possa observar algum efeito de aprendizagem. Recomenda-se, em estudos futuros, que, além da adoção de estratégias naturalísticas de ensino e uso de modelagem, que a

implementação do procedimento seja duradoura até atingir um volume de tentativas elevado, mesmo que com baixa densidade dentro das sessões. Além disso, a inserção de procedimentos remediativos, como a inclusão de demandas por imitação motora intercaladas com a demanda por repertório ecoico deve ser também considerada em estudos posteriores. Por fim, porém não menos importante, que os estudiosos que venham a estudar o ensino de ecoico a indivíduos no TEA lembrem-se do quão importante é a participação de um fonoaudiólogo(a) no delineamento e execução de pesquisas que envolvam o ensino de comunicação, linguagem e fala.

Referências

- Abadir, C. M., DeBar, R. M., Vladescu, J. C., Reeve, S.A. & Kupferman, D. M. (2021). Effects of video modeling on abduction-prevention skills by individuals with autism spectrum disorder. *Journal of Applied Behavior Analysis*, 54(3), 1139-1156. <https://doi.org/10.1002/jaba.822>
- Abreu, P. R., & Hübner, M. M. C. (2012). O comportamento verbal para B. F. Skinner e para S. C. Hayes: Uma síntese com base na mediação social arbitrária do reforçamento. *Acta Comportamental*, 20(3), 367-381. <http://pepsic.bvsalud.org/pdf/actac/v20n3/a08.pdf>
- American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Technical Report]. www.asha.org/policy
- Aravamudhan, S., & Awasthi, S. (2020). Behavioral interventions to treat speech sound disorders in children with autism. *Behavior Analysis in Practice*, 13(1), 174-185. <https://doi.org/10.1007/s40617-019-00362-5>
- Baros, R. S. (2003). Uma introdução ao comportamento verbal. *Revista Brasileira de Terapia Comportamental e Cognitiva*, 5(1), 73-82. <https://doi.org/10.31505/rbtcc.v5i1.92>
- Banda, D. R., Cople, K. S., Koul, R. K., Sancibrian, S. L., & Bogschutz, R. J. (2010). Video modelling interventions to teach spontaneous requesting using AAC devices to individuals with autism: A preliminary investigation. *Disability and Rehabilitation*, 32(16), 1364-1372. <https://doi.org/10.3109/09638280903551525>
- Braz, C. H., Gonçalves, L. F., Oliveira, A. M. de, & Haas, P. (2020a). Sinais de risco para apraxia de fala infantil: revisão sistemática / Risk signs for children's speech apraxia: systematic review. *Brazilian Journal of Development*, 6(9), 71593-71608. <https://doi.org/10.34117/bjdv6n9-561>
- Carroll, R. A., & Klatt, K. P. (2008). Using stimulus-stimulus pairing and direct reinforcement to teach vocal verbal behavior to young children with autism. *The Analysis of Verbal Behavior*, 24(1):135-126. <https://dx.doi.org/10.1007%2F03393062>
- Cividini-Motta, C., Scharrer, N., & Ahearn, W. H. (2016). An assessment of three procedures to teach echoic responding. *The Analysis of Verbal Behavior*, 33(1), 41-63. <https://doi.org/10.1007/s40616-016-0069-z>
- Conselho Federal de Fonoaudiologia [@conselhofederalfonoaudiologia]. (2021, Março 04). *Proteção contra coronavírus: CFFa não recomenda o uso de máscaras transparentes de vinil*. Instagram. <https://www.instagram.com/p/CMA0ivDljcE/>
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Improving and assessing the quality of behavioral measurement. In Cooper, J. O., Heron, T. E. & Heward, W. L. (Eds.), *Applied Behavior Analysis* (2ª ed. Pp. 113-116). New Jersey: Pearson Merrill Prentice Hal.
- Corbett, B. A., & Abdullah, M. (2005). Video modeling: Why does it work for children with autism? *Journal of Early and Intensive Behavior Intervention*, 2(1), 2-8. <https://doi.org/10.1037/h0100294>
- Dale, P. S., & Hayden, D. A. (2013). Treating speech subsystems in childhood apraxia of speech with tactual input: The PROMPT Approach. *American Journal of Speech-Language Pathology*, 22(4), 644-661. [https://doi.org/10.1044/1058-0360\(2013\)12-0055](https://doi.org/10.1044/1058-0360(2013)12-0055)
- Dass, T. K., Kisamore, A. N., Vladescu, J. C., Reeve, K. F., Reeve, S. A., & Taylor Santa, C. (2018). Teaching children with autism spectrum disorder to tact olfactory stimuli. *Journal of Applied Behavior Analysis* 51(3), 538-552. <http://doi.wiley.com/10.1002/jaba.470>
- Dawson, E. (2010). *Current assessment and treatment practices for children with autism and suspected childhood apraxia of speech: A survey of speech-language pathologists*. [Thesis, Portland State University]. <https://doi.org/10.15760/etd.29>
- Drash, Philip W., High, R. L., & Tudor, R. M. (1999). Using mand training to establish an echoic repertoire in young children with autism. *The Analysis of Verbal Behavior*, 16(1),29-44. <https://dx.doi.org/10.1007%2F03392945>
- Esteves, R. C., Lucchesi, F. D. M., & Almeida-Verdu, A. C. M. (2014). Ensino de ecoico, tato e mando: Uma revisão bibliográfica dos artigos do Journal of Applied Behavior Analysis (JABA). *Revista Brasileira de Terapia Comportamental e Cognitiva*, 16(2). <https://doi.org/10.31505/rbtcc.v16i2.694>
- Esch, B. E., Carr, J. E., & Michael, J. (2005). Evaluating stimulus-stimulus pairing and direct reinforcement in the establishment of an echoic repertoire of children diagnosed with autism. *The Analysis of Verbal Behavior* 21(1), 43-58. <https://dx.doi.org/10.1007%2F03393009>
- Fish, M. (2019) *Como tratar apraxia de fala na infância*. (E. Giusti, Trad.). Pró-Fono. (Original work published 2016).
- Fisher, W. W., Piazza, C. C., Bowman, L. G., & Amari, A. (1996). Reinforcer assessment for individuals with severe disability. <https://www.dshs.wa.gov/sites/default/files/DDA/dda/documents/F.%20Module%205%20Handout%202.pdf>
- Guerra, B. T., Espírito Santo, L. A. A., Barros, R. S., & Almeida-Verdu, A. C. M. (2019). Ensino de ecoico em pessoas com transtorno do espectro autista: revisão sistemática de literatura. *Revista Brasileira de Educação Especial*, 25(4),691-708. <https://doi.org/10.1590/s1413-65382519000400010>

- Guerra, B.T. (2020). *Ensino de ecoico em crianças com transtorno do espectro autista*. [Tese de Doutorado, Universidade Estadual Paulista Júlio de Mesquita Filho]. <https://repositorio.unesp.br/handle/11449/193248>
- Kisamore, A.N., Karsten, A. M., & Mann, C. C. (2016). Teaching multiply controlled intraverbals to children and adolescents with autism spectrum disorders: multiply controlled intraverbals. *Journal of Applied Behavior Analysis*, 49(4), 826–47. <https://doi.org/10.1002/jaba.344>
- Lee, G. T., Hu, X., Liu, Y., Zou, C., Cheng, X., Zhao, Q., & Huang, J. (2020). Increasing response diversity to intraverbals in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 50(1), 292–307. <https://doi.org/10.1007/s10803-019-04250-3>
- Lobato, A.F.F., Nogueira, C.B., Santos, E.A.L. (2018). Modelação e videomodelação. In: C.P. Duarte, L. Coltri e Silva, R.L. Velloso (Orgs.), *Estratégias da análise do comportamento aplicada para pessoas com transtornos do espectro do autismo* (p. 164). Memnon.
- Martone, M. C. C. (2017). Tradução e adaptação do Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP) para a língua portuguesa e a efetividade do treino de habilidades comportamentais para qualificar profissionais [Tese de Doutorado, Universidade Federal de São Carlos]. <https://repositorio.ufscar.br/handle/ufscar/9315>
- Namasivayam, A. K., Pukonen, M., Goshulak, D., Hard, J., Rudzicz, F., Rietveld, T., Maassen, B., Kroll, R., & van Lieshout, P. (2015). Treatment intensity and childhood apraxia of speech: Treatment intensity and childhood apraxia of speech. *International Journal of Language & Communication Disorders*, 50(4), 529–546. <https://doi.org/10.1111/1460-6984.12154>
- Perissinoto, J. (2003). Linguagem da criança com Autismo. In Perissonoto, J., Marchesan, I.Q. & Zorzi, J. L. (Eds.), *Conhecimentos essenciais para atender bem as crianças com autismo* (pp. 39-44). São José dos Campos: Pulso.
- Peterson, N. (1998). *An introduction to Verbal Behavior*. Ottawa: Behavior Associates, Inc.
- Pierce, C. D & Cheney, C. D. (2013). Verbal Behavior. In Pierce, C. D & Cheney, C.D (Eds.), *Behavior analysis and learning* (5th ed. 357-358) Psychology Press.
- Plavnick, J. B., & Vitale, F. A. (2016). A comparison of vocal mand training strategies for children with autism spectrum disorders. *Journal of Positive Behavior Interventions*, 18(1), 52–62. <https://doi.org/10.1177/1098300714548800>
- Roane, H. S., Vollmer, T. R., Ringdahl, J. E. & Marcus, B. A. (1998). Evaluation of a brief stimulus preference assessment. *Journal of Applied Behavior Analysis*, 31(4):605–20. <https://doi.org/10.1901/jaba.1998.31-605>
- Rodrigues, V., & Almeida, M. A. (2017). Modelagem em vídeo para o ensino de habilidades de comunicação a indivíduos com autismo: Revisão de estudos. *Revista Brasileira de Educação Especial*, 23(4), 595–606. <https://doi.org/10.1590/s1413-65382317000400009>
- Rodrigues, G.S., Maciel, R. C. S., Correia, D.V., Rêgo, F.L.C., Coutinho, H.M.T.A., Delgado, I.C., Soares, J.F.R., & Vogeley, A.C.E. (2019). Apraxia de fala da infância e marcadores históricos, desenvolvimentais e comportamentais. *Revista Brasileira de Ciências da Saúde*, 23(2), 53-64. <https://periodicos.ufpb.br/ojs2/index.php/rbcs/article/view/48418>
- Rogers, S. J., Hayden, D., Hepburn, S., Charlifue-Smith, R., Hall, T., & Hayes, A. (2006). Teaching young nonverbal children with autism useful speech: A pilot study of the Denver Model and Prompt interventions. *Journal of Autism and Developmental Disorders*, 36(8), 1007–1024. <https://doi.org/10.1007/s10803-006-0142-x>
- Silva, T.C. (1998). *Fonética e fonologia do português*. Contexto.
- Silva, M. K., Ferrante, C., Borsel, J. V. & Pereira, M. M. B. (2012). Aquisição fonológica do português brasileiro em crianças do Rio de Janeiro. *Jornal da Sociedade Brasileira de Fonoaudiologia*, 24(3), 248–54. <https://doi.org/10.1590/S2179-64912012000300010>
- Shane, J. (2016). *Increasing vocal behavior and establishing echoic stimulus control in children with autism*. [Dissertation, Western Michigan University]. <https://scholarworks.wmich.edu/dissertations/1400/>
- Shillingsburg, M. A., Frampton, S. E., Wymer, S. C., & Bartlett, B. (2018). A preliminary procedure for teaching children with autism to mand for social information. *Behavior Analysis in Practice* 11(1), 34–38. <https://doi.org/10.1007/s40617-016-0163-7>
- Shriberg, L. D., Paul, R., Black, L. M., & van Santen, J. P. (2011). The hypothesis of apraxia of speech in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 41(4), 405–426. <https://doi.org/10.1007/s10803-010-1117-5>
- Skinner, B. F. (1957) *Verbal Behavior*. New York: Appleton-Century-Crofts.
- Smith, J., Hand, L., & Dowrick, P. W. (2014). Video feedforward for rapid learning of a picture-based communication system. *Journal of Autism and Developmental Disorders*, 44(4), 926–936. <https://doi.org/10.1007/s10803-013-1946-0>
- Square, P. A., Namasivayam, A. K., Bose, A., Goshulak, D., & Hayden, D. (2014). Multi-sensory treatment for children with developmental motor speech disorders: Multi-sensory treatment for motor speech disorders. *International Journal of Language & Communication Disorders*, 49(5), 527–542. <https://doi.org/10.1111/1460-6984.12083>
- Steinbrenner, J. R., Hume, K., Odom, S. L., Morin, K. L., Nowell, S. W., Tomaszewski, B., Szendrey, S., McIntyre, N. S., Yücesoy-Özkan, Ş., & Savage, M. N. ([s.d.]). *Children, Youth, and Young Adults with Autism*. <https://ncaep.fpg.unc.edu/sites/ncaep.fpg.unc.edu/files/imce/documents/EBP%20Report%202020.pdf>
- Swerdan, M. G., & Rosales, R. (2015). Comparison of prompting techniques to teach children with autism to ask questions in the context of a conversation. *Focus on Autism and Other Developmental Disabilities*, 32(2), 93–101. <https://doi.org/10.1177/1088357615610111>
- Tarbox, J., Madrid, W., Aguilar, B., Jacobo, W. & Schiff, A. (2009). Use of chaining to increase complexity of echoics in children with autism. *Journal of Applied Behavior Analysis*, 42(4), 901–906. <https://doi.org/10.1901/jaba.2009.42-901>

Tierney, C., Mayes, S., Lohs, S. R., Black, A., Gisin, E., & Veglia, M. (2015). How valid is the checklist for autism spectrum disorder when a child has apraxia of speech? *Journal of Developmental & Behavioral Pediatrics*, 36(8), 569–574. <https://doi.org/10.1097/DBP.0000000000000189>

Wang, L., Shuting, L. & Chongying, W. (2023). Using pivotal response treatment to improve language functions of autistic children in special schools: A randomized controlled trial. *Journal of Autism and Developmental Disorders*, 1-13. <https://doi.org/10.1007/s10803-023-05988-7>

Wertalik, J. L., & Kubina, R. M. (2023). Comparison of two video prompting interventions to teach daily living skills to adolescents with autism. *Behavioral Interventions*, 38(1), 39–61. <https://doi.org/10.1002/bin.1914>

Wolf, M. (1978). Social validity: The case for subjective measurement or How applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11(2), 203 – 214. <https://doi.org/10.1901%2Fjaba.1978.11-203>

Zanon, R. B., Backes, B. & Bosa, C. A. (2014). Identificação dos primeiros sintomas do autismo pelos pais. *Psicologia: Teoria e Pesquisa* 30(1), 25–33. <https://doi.org/10.1590/S0102-37722014000100004>