

Marizete Ilha Ceron¹
Marileida Barichello Gubiani¹
Camila Rosa de Oliveira²
Márcia Keske-Soares¹

Descritores

Fala
Testes de Articulação da Fala
Transtornos da Articulação
Avaliação
Criança

Keywords

Speech
Speech Articulation Tests
Articulation Disorders
Assessment
Child

Endereço para correspondência:
Marizete Ilha Ceron
Av. Roraima, 1000, Prédio 26, CCS,
4º andar, Sala 1418, Bairro Camobi,
Santa Maria (RS), Brasil, CEP:
97105-900.
E-mail: marizeteceron@hotmail.com

Recebido em: Outubro 23, 2016

Aceito em: Outubro 23, 2017

Evidências de validade e fidedignidade de um instrumento de avaliação fonológica

Evidence of validity and reliability of a phonological assessment tool

RESUMO

Objetivo: Apresentar evidências de validade e fidedignidade de um instrumento de avaliação fonológica (INFONO) desenvolvido para avaliar os fonemas do Português Brasileiro. **Método:** Participaram do estudo 866 crianças com idades entre 3 e 8:11 anos, divididas em grupos: típico, controle e clínico. Os participantes foram avaliados pelo INFONO por nomeação espontânea. A produção da criança foi gravada e transcrita no momento da avaliação, no próprio *software*. Para análise de validade e fidedignidade, foram analisadas a consistência interna a partir da técnica *Alpha de Cronbach*. Para a validade de critério, comparou-se o desempenho entre os grupos através do teste *t* de Student para amostras independentes. A fidedignidade foi analisada pela concordância intra e interavaliadores por meio do Teste de Kendall. Considerou-se significância quando $p \leq 0,05$. **Resultados:** O INFONO apresentou evidências de validade e fidedignidade (consistência interna), indicando uma confiabilidade satisfatória dos itens, bem como excelente concordância entre os avaliadores em relação aos escores do teste (confiabilidade intra e interavaliador). Em relação à validade de critério, o desempenho do grupo clínico para todos os fonemas foi inferior ao grupo controle, mostrando que os escores são sensíveis para identificar crianças com desvio fonológico. **Conclusão:** O INFONO apresentou fortes evidências de validade e fidedignidade.

ABSTRACT

Purpose: To present evidence of the validity and reliability of a phonological assessment tool developed to assess the phonological inventory of Brazilian Portuguese. **Methods:** The study included 866 children aged between 3 and 8:11 years, divided into three groups: typical, control and clinical. Participants were evaluated using a phonological assessment software, which prompted the spontaneous naming of a series of images. The children's responses were audio recorded and transcribed at the time of the assessment, by the software itself. The Cronbach's alpha coefficient was used to evaluate the internal consistency of the instrument for reliability and validity purposes. Criterion validity was examined by comparing the performance of different groups using Student's t-test for independent samples. Intra- and inter-rater agreement were investigated using Kendall's tau. Results were considered significant at $p \leq 0.05$. **Results:** The present study provided evidence of validity and reliability (internal consistency) for this phonological assessment tool, confirming the reliability of its items and demonstrating excellent agreement rates between examiners regarding its scoring (intra- and inter-rater reliability). The criterion validity assessment demonstrated that the control group outperformed the clinical group across all phonemes, showing that test scores were successful in identifying children with speech sound disorders (phonological disorders). **Conclusion:** The present findings provide strong evidence of the validity and reliability of this phonological assessment tool.

Trabalho realizado no Programa de Pós-graduação em Distúrbios da Comunicação Humana, Universidade Federal de Santa Maria – UFSM - Santa Maria (RS), Brasil.

¹ Departamento de Fonoaudiologia, Universidade Federal de Santa Maria – UFSM - Santa Maria (RS), Brasil.

² Faculdade Meridional – IMED - Passo Fundo (RS), Brasil.

Fonte de financiamento: Programa Pesquisador Gaúcho (PqG) - Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), processo nº 11/2184-8.

Conflito de interesses: nada a declarar.

INTRODUÇÃO

Os distúrbios dos sons da fala podem ser decorrentes de diversas etiologias e resultam em prejuízo nos diferentes níveis de produção da fala, incluindo os de base linguístico-fonológica e/ou no aspecto motor da fala⁽¹⁻⁴⁾. Um desafio para o fonoaudiólogo é diagnosticar corretamente esses distúrbios diferenciando-os para posterior planejamento de uma intervenção eficaz.

O aspecto fonológico dos distúrbios dos sons da fala, denominado desvio fonológico, será abordado neste estudo, sendo esta uma das alterações de maior incidência na população infantil^(4,5) e alvo de muitas pesquisas^(4,6,7). Esse distúrbio é caracterizado por uma desorganização linguística do inventário de fonemas. Crianças com desvio fonológico apresentam omissões e substituições de fonemas, especialmente de consoantes e encontros consonantais em idade em que estas não deveriam ocorrer mais^(8,9).

Para a avaliação e o diagnóstico do desvio fonológico, destacam-se alguns instrumentos internacionais como o *Goldman Fristoe 2 – Test of Articulation – GFTA 2*⁽¹⁰⁾ e o *Clinical Assessment of Articulation and Phonology – CAAP*⁽¹¹⁾. No Brasil, os mais utilizados são a Avaliação Fonológica da Criança – AFC⁽¹²⁾ e o ABFW - Teste de Linguagem Infantil – Fonologia⁽¹³⁾. Estes últimos são bem elaborados e bastante completos, estando disponíveis para uso na clínica, também sendo utilizados na pesquisa fonoaudiológica. No entanto, esses instrumentos não apresentam estudos psicométricos de validade e fidedignidade para a avaliação de crianças brasileiras falantes do Português Brasileiro (PB). A literatura internacional^(14,15) aponta como sendo fundamental que instrumentos utilizados para fins de diagnóstico passem por estudos psicométricos de validade e fidedignidade. Somente após esses estudos, deveriam ser disponibilizados para uso.

O desenvolvimento de qualquer instrumento de avaliação deve garantir que estes mensurem o que se pretende, e que os resultados reflitam a habilidade em análise de forma não influenciada pela relação de outros aspectos que não os estabelecidos nos objetivos de uso^(2,10,14,15). Para garantir isso, o instrumento deve passar por estudos de validade e fidedignidade.

A validade é o processo que examina a precisão de uma determinada inferência realizada a partir de escores do instrumento⁽²⁾. As evidências de validade de um teste determinam o quão bem o teste mede o que ele se propõe a mensurar^(10,16). Este é um parâmetro importante de ser verificado ao utilizar um instrumento para o diagnóstico, pois, quanto mais preciso forem os seus escores, mais certeza teremos sobre as decisões diagnósticas.

Da mesma forma que as evidências de validade, a evidência de fidedignidade também é fundamental no desenvolvimento de um teste, porque essas medidas estimam em que grau um teste é vulnerável às várias fontes de erro (exemplos de fontes de erros são: a variação de avaliadores ou inconsistências do mesmo avaliador). Essas fontes de erro constituem ameaças à validade do teste^(2,17). Neste caso, é importante analisar as evidências de fidedignidade interavaliadores e intra-avaliadores dos escores do instrumento. Isso nos mostra o quanto os escores do teste são vulneráveis a erros.

Internacionalmente, há uma crescente busca por instrumentos que apresentem propriedades psicométricas para a avaliação dos distúrbios dos sons da fala^(2,14,15,17-19). Alguns desses estudos^(12,13,15) analisaram vários instrumentos de avaliação fonológica que apresentaram evidências de validade e fidedignidade. No entanto, no Brasil, esta prática é ainda recente, sendo desconhecida a existência de instrumentos de avaliação fonológica que tenham passado por estudos psicométricos^(14,20,21).

Essa limitação na área da avaliação fonológica do PB pode comprometer o processo diagnóstico e o planejamento terapêutico de crianças com alterações de fala, principalmente se o avaliador for ainda inexperiente. Dessa forma, o desenvolvimento da avaliação com critérios psicométricos é extremamente importante para a clínica e a pesquisa fonoaudiológica^(14,15,20). Ainda, aliado a estudos psicométricos, a utilização da tecnologia de informação facilita ao fonoaudiólogo a realização da avaliação fonológica. Portanto, viabilizar a avaliação fonológica em *software*, além de ser atrativo e interessante para as crianças por ser realizado no computador, permite a aplicação da avaliação a partir de um instrumento simples e de aplicação rápida.

Considerando o que foi exposto acima e a crescente preocupação em oferecer aos clínicos e pesquisadores instrumentos válidos e confiáveis, este estudo foi desenvolvido com o objetivo de apresentar evidências de validade e de fidedignidade dos escores do Instrumento de Avaliação Fonológica (INFONO) desenvolvido para avaliar os fonemas do Português Brasileiro (PB).

MÉTODO

Participantes

Foram contactadas 1448 crianças, entre 3 anos e 8 anos e 11 meses, em 12 escolas (8 públicas e 4 privadas) na região sul do Brasil, das quais 1076 (73%) foram autorizadas a participar do estudo. Inicialmente, foram excluídos da amostra participantes bilíngues e/ou que apresentavam queixas ou suspeitas de perda auditiva, problema neurológico e/ou psicológico, déficits intelectuais, diagnóstico de autismo, síndrome de Down e tratamento fonoaudiológico anterior. Essas informações foram obtidas em questionários respondidos pelos pais/responsáveis e professores que visavam detectar estes aspectos. Também, foram excluídos os participantes com indicativos de prováveis alterações fonoaudiológicas que pudessem prejudicar a produção da fala, como: mordida aberta anterior, ceceo, interposição lingual anterior, prováveis déficits de linguagem e/ou vocabulário, etc. Essas prováveis alterações foram detectadas durante uma breve conversa informal com as crianças, em que eram feitas perguntas sobre a escola, a sua idade, que dia era, do que gostavam de brincar, o que gostavam de comer, o que estavam fazendo na sala de aula, se gostavam de bichos, quais bichos, etc. e/ou durante a realização do INFONO. Dessa forma, considerando os critérios de exclusão, 210 crianças foram excluídas da amostra da pesquisa por não se enquadrarem nos critérios de inclusão da pesquisa.

Assim, a amostra total deste estudo foi constituída por $n = 866$ participantes, divididos em dois grupos denominados de grupo típico e grupo clínico. O grupo típico foi composto

Quadro 1. Caracterização dos grupos

Grupos	N	Idade em anos e meses		Gênero		Tipo de escola	
		M	DP	Feminino n (%)	Masculino n (%)	Privada n (%)	Pública n (%)
Típico	733	6,08	1,58	332 (45,3)	401 (54,7)	422 (57,6)	311 (42,4)
Atípico	133	5,87	1,23	54 (40,6)	79 (59,3)	58 (43,6)	75 (56,4)
Controle	Típico	228	5,89	100 (43,9)	128 (56,1)	104 (45,6)	124 (54,4)
	Clínico	114	5,88	1,32	50 (43,9)	64 (56,1)	52 (45,6)

Legenda: N = número total de indivíduos; n = número de indivíduos; M = Média; DP = Desvio Padrão

por crianças com desenvolvimento fonológico típico ($n = 733$), ou seja, não possuíam fonemas ausentes, conforme o esperado para sua faixa etária, por exemplo, uma criança que aos 3 anos não falou o /r/ foi considerada como pertencendo a este grupo. O grupo atípico ($n = 133$) foi formado por crianças com desenvolvimento fonológico atípico, ou seja, crianças que apresentavam omissões ou substituições de fonemas não mais esperado para sua idade. Por exemplo, uma criança que aos 3 anos tenha alteração na produção de /b/ ou que aos 5 anos não tenha o /r/. Esse julgamento foi realizado pela fonoaudióloga avaliadora conforme sua experiência clínica em avaliação e terapia, e conforme a literatura nacional^(22,23) em relação à aquisição fonológica.

Um terceiro grupo foi formado, a partir dos grupos típico e atípico, que foi denominado de grupo controle. Os participantes do grupo controle foram recrutados do grupo típico ($n = 228$) e emparelhados ao grupo clínico ($n = 114$) pela faixa etária, tipo de escola e gênero. Um total de 342 crianças participaram desta análise. Os grupos controle e clínicos não diferiram significativamente em relação à média de idade ($F = 0,000$; $p = 0,977$), distribuição de gênero ($X^2 = 0,000$; $p = 1,000$) e tipo de escola ($X^2 = 0,000$; $p = 1,000$). A amostra está caracterizada no Quadro 1.

Instrumento

O INFONO é um instrumento que foi desenvolvido em *software* para uso do fonoaudiólogo com auxílio de um computador. A elaboração deste instrumento foi realizada em quatro etapas: revisão de literatura/seleção das palavras⁽²⁴⁾; análise de critérios por juízes especialistas⁽²⁴⁾; análise de critério por juízes não especialistas⁽²⁴⁾; e estudo piloto⁽²⁵⁾. Estas etapas são fundamentais para a elaboração do *Software* INFONO. Este instrumento já está concluído e encontra-se em fase de testes finais (estudos psicométricos) necessários para, em seguida, ser disponibilizado e divulgado para o uso fonoaudiológico. A literatura internacional^(14,15) refere que a realização de estudos psicométricos são importantes para qualquer instrumento de avaliação para fins de diagnóstico e somente após esses estudos o instrumento pode ser disponibilizado.

O INFONO permite que a coleta de dados para a avaliação fonológica seja gravada no próprio *software* para posterior conferência das transcrições. A coleta foi realizada por nomeação espontânea dos estímulos. Durante a avaliação, o avaliador deve selecionar a transcrição da produção da criança (o instrumento oferece uma lista de transcrições para cada palavra-alvo), a qual pode ser conferida após a avaliação mediante a gravação.

Esta lista de transcrição foi elaborada com base nos processos fonológicos que ocorrem na fala da criança, tanto na aquisição típica quanto com desvio fonológico. Caso não tenha a opção registrada na lista de transcrições, esta pode ser digitada no teclado fornecido pelo *software*, assim como as distorções presentes na produção da criança.

A avaliação por nomeação espontânea é composta por 84 desenhos coloridos e “animados” com *gif's* (termo dado às animações formadas por várias imagens *GIF* compactadas numa só dando movimento à figura), representando a palavra-alvo e, ainda, fornece uma pergunta-chave (para o avaliador fazer para a criança no momento da avaliação) para facilitar a identificação e produção da palavra-alvo desejada, como: “*Ele usa o lápis para...?*” (escrever), “*Que bicho é este?*” (cachorro), etc. Também podem ser usadas outras perguntas para auxiliar na produção do alvo desejado. É importante que todas as palavras-alvo sejam produzidas, uma vez que elas avaliam todos os fonemas do PB nas diferentes posições na sílaba e na palavra.

Após finalizar a avaliação e transcrição, deve-se clicar em gerar resultados. O *software* disponibiliza os seguintes resultados para serem visualizados: análise dos dados da avaliação (lista das palavras-alvo e das transcrições fonéticas da produção da criança), análise contrastiva (quantidade de acertos, omissões e substituições na avaliação para cada fonema do PB) que permite determinar o inventário fonético (presença ou ausência de som) e o fonológico (percentuais de acerto, omissão e substituição para cada fonema e para o *Onset* Complexo (OC)); análises dos traços distintivos; análises de processos fonológicos; e gravidade do desvio fonológico.

Esse instrumento não analisa as vogais, pois, como são adquiridas cedo, elas são muito menos propensas a serem produzidas com alterações em casos de desvio fonológico no PB.

Procedimentos

Todos os procedimentos éticos para o desenvolvimento desta pesquisa foram respeitados como a aprovação do Comitê de Ética em Pesquisa (CEP) da Universidade Federal de Santa Maria (UFSM) sob o protocolo 23081.005433/2011-65. Os pais ou representantes legais das crianças assinaram o Termo de Consentimento Livre e Esclarecido. Só foram avaliadas as crianças que concordaram em participar da avaliação (assentimento oral da criança).

Primeiramente foi realizada uma breve conversa com a criança a fim de detectar possíveis alterações que justificassem a exclusão do participante do estudo. Essa breve conversa,

assim como a aplicação do INFONO, foi realizada por três fonoaudiólogas doutorandas e por uma fonoaudióloga com mais de 10 anos de experiência na área do estudo. Todas tiveram um treinamento prévio sobre os procedimentos da pesquisa e aplicação do INFONO.

Após a conversa, os participantes foram avaliados com o INFONO, os estímulos foram obtidos por nomeação espontânea em sessão individual em um ambiente disponibilizado pela escola. A aplicação do INFONO leva aproximadamente 20 minutos e seguiu os procedimentos de administração preconizados para a aplicação do teste. A produção da criança foi gravada e transcrita no momento da avaliação no próprio *software*. Após as avaliações, os resultados do inventário fonológico (gerados pelo *software*) foram utilizados para a elaboração de um banco de dados no programa estatístico SPSS com o objetivo de realizar os procedimentos psicométricos para a busca de evidências de validade e fidedignidade dos escores do INFONO, que serão detalhados a seguir:

(a) Validade

Para examinar a precisão dos seus escores do INFONO, as evidências de validade analisadas foram: consistência interna e a validade de critério (através da análise discriminante). A consistência interna mostra o quão bem os itens do teste foram projetados para atingir o objetivo pelo qual o instrumento foi elaborado⁽¹⁴⁾. A validade de critério verifica a eficácia do instrumento em identificar o desempenho de um grupo específico de indivíduos⁽¹⁹⁾. Dessa forma, a consistência interna foi calculada a partir do grupo típico ($n = 733$), enquanto que para a análise da validade de critério utilizou-se o grupo controle.

(b) Fidedignidade

A fidedignidade analisa a consistência dos escores do INFONO, isso garante que os resultados do teste não flutuem quando aplicado por uma mesma pessoa em momentos diferentes (confiabilidade intra-avaliador) ou quando aplicado por pessoas diferentes (confiabilidade interavaliadores). Isso nos mostra o quanto os resultados obtidos podem ser confiáveis. Assim, foram analisados três componentes da fidedignidade: consistência interna, confiabilidade intra-avaliador e confiabilidade interavaliadores.

A consistência interna foi calculada da mesma forma que para a validade, pois os escores para serem válidos devem também ser consistentes (medir sempre da mesma forma). Para a confiabilidade intra-avaliador, o mesmo avaliador analisou a gravação de áudio da administração original do instrumento, realizando manualmente novas transcrições (sem o uso do *software*) e as análises dos resultados foram realizadas de forma independente da primeira avaliação. Este procedimento foi realizado, para um subgrupo de 77 crianças (aproximadamente 10% da amostra) escolhidas aleatoriamente por sorteio, pelo mesmo avaliador no mínimo um mês após a avaliação da criança.

Da mesma forma, para o procedimento de confiabilidade interavaliadores, os áudios de outras 120 crianças (aproximadamente 15% da amostra) foram analisados por duas estudantes do último ano do curso de graduação com experiência em fonologia (bolsistas de iniciação científica do laboratório). As novas transcrições e análises dos resultados foram realizadas de forma independente da primeira avaliação.

Análise dos dados

Os dados foram analisados através do *software* SPSS versão 22 para Windows. A consistência interna auxilia tanto para a validade quanto para a fidedignidade e foi analisada a partir da técnica *Alpha de Cronbach*. Este método foi utilizado porque os participantes foram avaliados com um teste em uma única ocasião. A técnica *Alpha de Cronbach* consiste em calcular a correlação entre cada item do teste e o restante dos itens ou o escore total dos itens. Essa técnica mede o grau de covariância e os resultados variam de 0 a 1, sendo considerado que, quanto mais elevada a contagem, maior a confiabilidade. Valores acima de 0,7 refletem uma validade e fidedignidade adequada.

Para a validade de critério, realizou-se a comparação de desempenho entre os escores das crianças com desenvolvimento fonológico típico (grupo controle) e atípico (grupo clínico) e análise discriminante. Primeiramente comparou-se o desempenho entre as crianças que formaram o grupo clínico e grupo controle através do teste *t* de Student para amostras independentes. A seguir, realizou-se a análise discriminante *stepwise* com o método do Λ de Wilks a fim de verificar a capacidade dos escores do INFONO em estimar a qual grupo a criança pertencia (desenvolvimento fonológico típico ou atípico). Os pressupostos da normalidade e de homogeneidade das matrizes de variâncias-covariâncias de cada grupo foram testados, respectivamente, com o teste de Shapiro-Wilk e o teste *M* de Box.

Ainda, foi investigada a fidedignidade do INFONO pela concordância intra e interavaliadores por meio do Teste de Kendall, sendo que os valores acima de 0,6 indicam um bom nível de concordância e, dessa forma, uma fidedignidade adequada para os escores. Para analisar a concordância intra e interavaliadores, os fonemas foram divididos nas posições: *Onset Geral* (OG), *Onset Inicial* (OI), *Onset Medial* (OM), *Coda Geral* (CG), *Coda Medial* (CM) e *Coda Final* (CF). Os resultados foram considerados significativos quando $p \leq 0,05$.

RESULTADOS

Os escores obtidos pelo INFONO foram testados em relação às evidências de validade e fidedignidade. A Tabela 1 apresenta os resultados referentes à consistência interna, analisados pelos coeficientes *Alpha de Cronbach* para as diferentes faixas etárias em relação aos fonemas do PB avaliados pelos instrumentos.

Observou-se que todas as faixas etárias do estudo apresentaram alta consistência interna (acima de 0,7), variando entre 0,713 e 0,922 (mediana 0,816), o que sugere uma adequada validade e fidedignidade dos itens que avaliam os fonemas do PB por faixa etária.

Quanto às evidências de fidedignidade, a Tabela 2 apresenta a confiabilidade intra-avaliadores que foi medida pela concordância da produção correta dos fonemas e encontro consonantal do PB testado pelo INFONO.

Verificou-se que a concordância intra-avaliadores variou de 0,622 a 1,0. A concordância foi excelente para a maioria dos fonemas nas diferentes posições, com valores acima de 0,8. A concordância foi considerada boa apenas para o /s/ nas posições de CG, CM e CF; o /k/ em OG e OM; o /m/ em OG e OI; o /r/ em CM; a Plosiva + /r/ em OG e OI; e a fricativa + /r/ em OG, OI e OM. Os percentuais médios de concordância intra-avaliadores por posição foram considerados excelentes, variando entre 0,822 e 0,944.

Em relação à confiabilidade interavaliadores, a Tabela 3 mostra os valores de concordância para a produção correta dos fonemas e encontro consonantal do PB testado pelo INFONO.

Observou-se que a concordância interavaliadores variou de 0,663 a 1,0. Os menores percentuais de concordâncias ocorreram

para /d/ em OG, /k/ em OG e OM, /m/ em OG e OI e /l/ em OG, os quais foram considerados bons. Ressalta-se que as demais concordâncias interavaliadores foram excelentes, todas acima de 0,826. Os percentuais médios de concordância interavaliadores por posição foram excelentes, todos acima de 0,93.

No que se refere às evidências de validade, a Tabela 4 apresenta a comparação dos escores de desempenho (médias e desvios padrão) entre os grupos clínico e controle para a produção de todos os fonemas em OG, CG e OC.

Os resultados mostraram que a média de produção do grupo clínico para todos os fonemas foi inferior ao grupo controle. A análise mostrou diferenças significativas na produção de vários fonemas entre os grupos clínico e controle com exceção para /p/, /d/ e /j/ em OG e /n/ em CG; /p/ em OI; /b/, /d/ e /m/ em OM; /n/ em CM; e, /l/ e /s/ em CF. Isso mostra que os escores do INFONO discriminam crianças com alterações no desenvolvimento fonológico daquelas que não apresentam alterações.

Com o objetivo de analisar quais desses fonemas em *onset* simples, coda e/ou OC mais auxiliavam na discriminação entre os grupos clínico e controle, realizou-se a análise discriminante *stepwise* que extraiu uma função discriminante, retendo como estatisticamente significativa a produção de 18 fonemas e OC, explicando 100% da variabilidade entre os grupos ($\Lambda = 0,494$; $\chi^2(18) = 602,403$; $p \leq 0,001$). Na Tabela 5, foram apresentados os coeficientes estandardizados e os fonemas incluídos na função discriminante. Sendo esses os principais fonemas que auxiliaram na diferenciação entre os grupos com desenvolvimento fonológico típico e atípico.

Tabela 1. Análise da consistência interna do INFONO por faixa etária

Faixa etária	n	Alpha de Cronbach
3:0-3:3	25	0,922
3:4-3:7	38	0,916
3:8-3:11	34	0,922
4:0-4:3	31	0,917
4:4-4:7	34	0,903
4:8-4:11	35	0,862
5:0-5:3	44	0,820
5:4-5:7	40	0,815
5:8-5:11	51	0,809
6:0-6:3	53	0,746
6:4-6:7	59	0,861
6:8-6:11	59	0,743
7:0-7:6	72	0,713
7:7-7:11	58	0,714
8:0-8:6	61	0,732
8:7-8:11	39	0,718
Mediana		0,816

Legenda: n = número de indivíduos

Tabela 2. Confiabilidade intra-avaliadores medida pela concordância da produção correta dos fonemas

	Posições					
	Onset Geral	Coda Geral	Onset Inicial	Onset Medial	Coda Medial	Coda Final
/p/	0,811		1,000	1,000		
/b/	1,000		1,000	1,000		
/t/	0,973		1,000	0,973		
/d/	1,000		1,000	1,000		
/k/	0,702		1,000	0,702		
/g/	1,000		1,000	1,000		
/f/	1,000		1,000	1,000		
/v/	0,964		1,000	0,964		
/s/	1,000	0,690	0,935	1,000	0,631	0,678
/z/	1,000		1,000	1,000		
/ʃ/	1,000		1,000	1,000		
/ʒ/	1,000		1,000	1,000		
/x/	1,000		1,000	1,000		
/m/	0,653		0,653	1,000		
/n/	1,000	0,856	1,000	1,000	0,962	1,000
/ɲ/	1,000			0,974		
/l/	0,816	1,000	1,000	0,816	1,000	1,000
/ʎ/	0,945			0,945		
/r/	0,973	0,830		0,971	0,700	0,878
Plos.+/l/	0,872		0,920	0,891		
Plos.+/r/	0,765		0,622	0,862		
Fric.+/l/	0,801		0,801			
Fric.+/r/	0,676		0,795	0,661		
Média	0,911	0,844	0,936	0,944	0,823	0,889

Tabela 3. Confiabilidade interavaliadores medida pela concordância da produção correta dos fonemas

Fonemas	Posições					
	Onset Geral	Coda Geral	Onset Inicial	Onset Medial	Coda Medial	Coda Final
/p/	1,000		1,000	1,000		
/b/	1,000		1,000	1,000		
/t/	0,941		1,000	1,000		
/d/	0,663		1,000	1,000		
/k/	0,772		1,000	0,704		
/g/	1,000		1,000	1,000		
/f/	1,000		1,000	1,000		
/v/	0,920		1,000	0,899		
/s/	1,000	0,909	1,000	1,000	0,946	0,863
/z/	1,000		1,000	1,000		
/ʃ/	0,996		0,997	1,000		
/ʒ/	1,000		1,000	1,000		
/x/	1,000		1,000	1,000		
/m/	0,791		0,658	1,000		
/n/	1,000	0,952	1,000	1,000	1,000	1,000
/ɲ/	0,993			0,993		
/l/	0,748	1,000	1,000	0,866	1,000	1,000
/ʎ/	0,958			0,958		
/r/	0,960	0,873		0,960	0,996	0,870
Plos.+/l/	0,866		0,970	0,826		
Plos.+/r/	0,892		0,897	0,892		
Fric.+/l/	0,972		0,972			
Fric.+/r/	1,000		1,000	1,000		
Média	0,934	0,934	0,975	0,959	0,986	0,933

Tabela 4. Desempenho na produção dos fonemas entre os grupos clínico e controle

		Controle <i>n</i> = 228		Clínico <i>n</i> = 114		<i>F</i>	<i>p</i> -valor
		<i>M</i>	<i>DP</i>	<i>M</i>	<i>DP</i>		
Onset Geral	/p/	8,99	0,09	8,95	0,26	21,215	0,084
	/b/	9,90	0,34	9,54	1,33	50,876	0,004
	/t/	18,73	0,67	18,12	1,80	34,793	≤ 0,001
	/d/	7,02	0,42	6,77	1,35	44,188	0,060
	/k/	17,94	0,27	16,72	3,71	81,367	≤ 0,001
	/g/	6,97	0,20	6,18	1,86	193,940	≤ 0,001
	/f/	6,98	0,17	6,64	1,23	63,277	0,004
	/v/	7,68	0,51	6,97	1,52	34,942	≤ 0,001
	/s/	8,96	0,25	8,28	1,78	111,313	≤ 0,001
	/z/	5,99	0,28	5,17	1,54	225,076	≤ 0,001
	/j/	6,82	0,85	4,69	2,79	395,252	≤ 0,001
	/ʒ/	5,93	0,77	4,20	2,40	406,058	≤ 0,001
	/χ/	5,93	0,45	5,28	1,60	128,677	≤ 0,001
	/m/	5,96	0,21	5,86	0,40	36,231	0,016
	/n/	7,91	0,31	7,67	0,67	73,707	≤ 0,001
	/ɲ/	3,75	1,20	3,75	1,21	0,044	1,000
/l/	10,92	0,32	9,73	2,63	146,593	≤ 0,001	
/ʎ/	3,60	0,60	2,92	1,12	27,320	≤ 0,001	
/r/	7,56	1,68	3,72	3,47	241,044	≤ 0,001	
Coda Geral	/n/	7,89	0,74	7,70	1,05	27,162	0,087
	/l/	5,96	0,23	5,80	0,67	47,966	0,011
	/r/	6,37	1,69	3,72	2,67	74,978	≤ 0,001
Onset Inicial	/s/	11,41	1,93	10,64	2,97	15,361	0,013
	/p/	3,99	0,09	3,95	0,26	21,215	0,084
	/b/	6,92	0,29	6,61	1,04	56,702	0,003
	/t/	2,99	0,11	2,93	0,29	28,152	0,045
	/d/	2,98	0,13	2,81	0,64	71,780	0,004
	/k/	11,00	0,07	10,32	2,19	87,960	≤ 0,001
	/g/	2,98	0,17	2,59	0,96	182,586	≤ 0,001
	/f/	3,00	0,00	2,89	0,53	46,051	0,023
	/v/	2,99	0,11	2,75	0,62	136,383	≤ 0,001
	/s/	3,36	0,49	3,16	0,88	9,707	0,021
	/z/	2,00	0,07	1,70	0,61	297,829	≤ 0,001
	/j/	3,90	0,46	2,61	1,61	413,576	≤ 0,001
	/ʒ/	2,95	0,31	2,08	1,21	400,694	≤ 0,001
	/χ/	2,96	0,28	2,60	0,85	154,203	≤ 0,001
/m/	2,96	0,20	2,87	0,36	38,125	0,013	
/n/	2,93	0,27	2,77	0,50	54,132	0,002	
/l/	4,00	0,00	3,65	0,97	145,847	≤ 0,001	
Onset Medial	/p/	5,00	0,00	5,00	0,00	*	*
	/b/	2,99	0,11	2,92	0,38	23,865	0,073
	/t/	15,75	0,65	15,19	1,60	38,803	≤ 0,001
	/d/	4,01	0,37	3,92	0,82	27,029	0,256
	/k/	6,94	0,26	6,39	1,62	92,131	≤ 0,001
	/g/	3,99	0,09	3,60	1,02	157,564	≤ 0,001
	/f/	3,98	0,17	3,75	0,77	71,917	0,003
	/v/	4,69	0,48	4,22	1,05	37,185	≤ 0,001
	/s/	5,60	0,53	5,12	1,19	24,948	≤ 0,001
	/z/	4,00	0,24	3,46	1,11	199,564	≤ 0,001
	/j/	2,92	0,43	2,08	1,26	315,886	≤ 0,001
	/ʒ/	2,97	0,54	2,12	1,30	263,857	≤ 0,001
	/χ/	2,97	0,21	2,68	0,83	110,945	≤ 0,001

Legenda: M = Média; DP = Desvio Padrão; *n* = número de indivíduos; *F* = Razão; OI = Onset Inicial; OM = Onset Medial; * = Não foi possível rodar o teste t em função dos desvios padrão serem equivalentes a zero

Tabela 4. Continuação...

	Controle <i>n</i> = 228		Clínico <i>n</i> = 114		<i>F</i>	<i>p</i> -valor	
	<i>M</i>	<i>DP</i>	<i>M</i>	<i>DP</i>			
	/m/	3,00	0,07	2,99	0,09	1,001	0,617
	/n/	4,98	0,13	4,89	0,31	59,123	0,004
	/ɲ/	3,75	1,20	3,75	1,21	0,044	1,000
	/l/	6,92	0,32	6,08	1,77	154,948	≤ 0,001
	/ʎ/	3,60	0,60	2,92	1,12	27,320	≤ 0,001
	/r/	7,56	1,68	3,72	3,47	241,044	≤ 0,001
Coda Medial	/n/	5,12	0,47	5,11	0,73	13,032	0,862
	/l/	2,96	0,23	2,82	0,63	41,904	0,016
	/r/	2,74	0,75	1,54	1,35	175,508	≤ 0,001
Coda Final	/s/	5,79	0,92	4,90	1,80	67,513	≤ 0,001
	/n/	2,83	0,38	2,71	0,47	23,184	0,021
	/l/	3,00	0,00	2,98	0,13	16,785	0,158
Onset Complexo	/r/	3,63	1,19	2,18	1,55	12,968	≤ 0,001
	/s/	5,62	1,56	5,74	1,80	1,383	0,531
	Plosiva +/l/	3,82	1,69	1,83	1,83	5,643	≤ 0,001
	Plosiva +/r/	16,87	6,68	6,52	7,45	11,713	≤ 0,001
	Fricativa + /l/	1,40	0,79	0,57	0,74	1,399	≤ 0,001
	Fricativa + /r/	4,10	1,67	1,70	1,99	20,797	≤ 0,001
	Plosiva +/l/ - OI	1,59	0,74	0,78	0,86	14,035	≤ 0,001
	Plosiva +/l/ - OM	2,23	1,06	1,05	1,11	0,448	≤ 0,001
	Plosiva +/r/ - OI	8,16	3,15	3,54	3,80	25,418	≤ 0,001
	Plosiva +/r/ - OM	8,71	3,62	2,98	3,79	4,261	≤ 0,001
	Fricativa + /l/ - OI	1,40	0,79	0,57	0,74	1,399	≤ 0,001
	Fricativa + /r/ - OI	1,63	0,71	0,72	0,92	41,572	≤ 0,001
	Fricativa + /r/ -OM	2,46	1,02	0,98	1,20	10,794	≤ 0,001

Legenda: M = Média; DP = Desvio Padrão; *n* = número de indivíduos; *F* = Razão; OI = Onset Inicial; OM= Onset Medial;
 *= Não foi possível rodar o teste t em função dos desvios padrão serem equivalentes a zero

Tabela 5. Coeficientes da função discriminante

Fonemas	Coeficientes da função discriminante
/p/ em Onset Geral	-1,066
/s/ em Onset Geral	0,232
/ʎ/ em Onset Geral	0,417
/ɲ/ em Onset Geral	-0,368
/r/ em Onset Geral	0,179
/s/ em Coda Geral	-0,119
/k/ em OI	0,324
/v/ em OI	0,768
/j/ em OI	0,802
/l/ em OI	0,253
/p/ em OM	-1,784
/ʎ/ em OM	-0,330
/n/ em OM	0,610
/ɲ/ em OM	0,153
/r/ em CF	0,095
plosiva + /r/	-0,172
plosiva + /r/ em OM	0,484
fricativa + /r/ em OI	-0,494
Constante	3,081
Eigenvalue	1,023

Legenda: OI = Onset Inicial; OM= Onset Medial; CF = Coda Final

Quadro 2. Resultados da classificação e função discriminatória utilizada

Grupo original	Grupo predito		Total
	Típica <i>n</i> (%)	Atípico <i>n</i> (%)	
Típica <i>n</i> (%)	704 (96%)	29 (4%)	733
Atípico <i>n</i> (%)	43 (32%)	90 (68%)	133

Legenda: *n* = número de indivíduos

O Quadro 2 apresenta os resultados estatísticos da classificação e função discriminatória utilizada. Observa-se que o percentual de crianças classificadas corretamente com a classificação original foi de 91,7%, demonstrando que o INFONO apresentou um alto percentual de discriminação entre crianças com desenvolvimento fonológico típico e atípico.

DISCUSSÃO

Considerando os objetivos deste estudo e os resultados exibidos, verificou-se que os escores do INFONO apresentaram adequados indicadores de validade e fidedignidade, sendo um instrumento adequado para ser aplicado como instrumento de avaliação fonológica em crianças com suspeitas de alterações fonológicas.

No que se refere às evidências de validade e confiabilidade, a consistência interna analisada pelo coeficiente *Alpha de Cronbach* mostrou-se adequada para o INFONO (Mediana 0,816), indicando uma confiabilidade satisfatória dos itens que compõem o instrumento para avaliar os fonemas do PB. Outros instrumentos internacionais^(10,26) de avaliação fonológica também analisaram a consistência interna pelo coeficiente *Alpha de Cronbach*. O GFTA-2⁽¹⁰⁾ analisou a consistência interna por faixa etária e gênero e obteve resultados variando de 0,85 a 0,98, o que é considerado adequado. O *Test para Evaluar Procesos de Simplificación Fonológica - TEPROSIF-R*⁽²⁶⁾ e o CAAP⁽¹¹⁾, ambos instrumentos de avaliação fonológica, também obtiveram resultados que indicaram alta confiabilidade interna (0,90) para os escores obtidos.

Um teste tem alta confiabilidade dos itens quando todos os itens ou tarefas do instrumento apresentam o mesmo tipo de desempenho ou domínio do conteúdo⁽¹⁰⁾. Neste estudo, a análise de consistência interna do INFONO não foi realizada entre faixa etária e gênero ou faixa etária e tipo de escola, pois não foi encontrada variabilidade suficiente dos escores para a realização dessa análise dos dados.

O AFC⁽¹²⁾ e ABFW-Fonologia⁽¹³⁾ que avaliam a fonologia do PB são os instrumentos mais utilizados em pesquisas brasileiras^(7,27-29) atualmente, porém eles ainda não fornecem estudos de validade e fidedignidade. Portanto, este estudo torna-se relevante para o avanço da Fonoaudiologia, pois contribuiu para a elaboração de um instrumento em *software*, o qual passou por rigorosos estudos de validade e fidedignidade para poder ser utilizado para avaliar criança com suspeita de alterações no desenvolvimento fonológico. O uso de instrumentos normatizados que apresentem estudos de validade e fidedignidade pressupõe que mensurem corretamente a presença ou a ausência de um determinado distúrbio.

Existem alguns fatores que podem comprometer a exatidão das medidas de um instrumento e, dessa forma, comprometer o diagnóstico⁽¹⁸⁾ e, conseqüentemente, a terapia. Devido a isso, é importante que os instrumentos passem por estudos de validade e fidedignidade para garantir que mensurem o que se pretende e que os resultados reflitam a habilidade em análise^(2,10,14,15). Por exemplo, os resultados de um instrumento que foi elaborado para avaliar a fonologia deve permitir a análise de todos os fonemas da língua em todas as posições que ocorrem na sílaba e na palavra para se concluir se há alteração no desenvolvimento ou não. A validade de um instrumento pode ser afetada pela qualidade e confiabilidade de suas medidas e pela habilidade do aplicador no teste.

A fidedignidade pode ser definida como a consistência de uma medida, embora ocorram mudanças no tempo, na administração ou nos escores das medidas⁽¹⁹⁾. Uma fidedignidade aceitável permite a generalização dos resultados obtidos na situação de avaliação a um conjunto mais amplo de situações. Neste estudo, a fidedignidade foi testada pela análise da confiabilidade intra-avaliador e interavaliador. A confiabilidade intra-avaliador foi mensurada através da concordância nos julgamentos do mesmo avaliador das transcrições dos fonemas realizadas em momentos distintos. A média de concordância intra-avaliador obtida pelos julgamentos da produção dos fonemas por posição foi considerada excelente (percentuais variaram entre 0,822 e 0,944). Isso demonstrou nos dois julgamentos das avaliações que o mesmo avaliador em dois momentos distintos concordou em 0,822, demonstrando boa reprodutibilidade e confiabilidade diagnóstica dos escores do INFONO.

Na literatura, não foram encontrados instrumentos de avaliação fonológica que apresentem resultados referentes à confiabilidade intra-avaliador. No entanto, este tipo de confiabilidade é referido na literatura⁽¹⁵⁾ como importante para instrumentos que avaliam a fala, porque permite mensurar o grau de consistência da medida quando aplicado pelo mesmo clínico em momentos diferentes. É esperado que não haja variabilidade nessa medida, pois um instrumento deve medir sempre da mesma forma, independentemente do momento em que for realizado. A confiabilidade intra-avaliador foi relatada para um instrumento de avaliação de apraxia de fala, sendo que esta variou entre 0,81 e 0,95⁽²⁾, valores próximos aos encontrados para o INFONO.

A confiabilidade obtida pela concordância média interavaliadores das transcrições dos fonemas por posição foi excelente, sendo que todas as concordâncias foram acima de 0,93. Isso demonstra que os dois julgadores concordaram em mais de 93% das transcrições. Por se tratar de um método de avaliação subjetivo, em que sua precisão depende da habilidade/experiência do avaliador, a consistência dos resultados das transcrições interavaliadores para o INFONO refletiu uma boa reprodutibilidade e confiabilidade diagnóstica do instrumento entre os avaliadores no processo de normatização dos escores. Da mesma forma, no GFTA-2⁽¹⁰⁾, foram observados percentuais de concordância interavaliadores similares (variaram de 70% a 100%) com média de 93% ou mais, dependendo da posição do fonema. O coeficiente de confiabilidade interavaliadores no CAAP⁽¹¹⁾ foi 99%, indicando

que os examinadores foram altamente consistentes em suas pontuações.

Em relação às evidências de validade analisadas, a análise discriminante realizada entre os grupos clínico e controle mostrou fortes evidências de validade de critério para os escores do INFONO. Como era esperado, os escores médios de produção para todos os fonemas do grupo clínico foram inferiores aos do grupo controle, mostrando que os escores do INFONO são sensíveis para identificar crianças com desenvolvimento fonológico típico e atípico. O desempenho entre os grupos clínico e controle teve diferenças significativas na produção para a maioria dos fonemas e OC, com exceção para /p/, /d/ e /ŋ/ em OG e /n/ em CG; /p/ em OI; /b/ e /m/ em OM; /n/ em CM; e, /l/ e /s/ em CF. Tal fato pode estar relacionado à aquisição e estabilização desses fonemas antes dos 3 anos, faixa etária não avaliada neste estudo. Outras pesquisas^(22,23) confirmaram que a aquisição desses fonemas ocorre antes dos 3 anos. Em geral, a ordem de aquisição dos fonemas do PB é plosivas e nasais, seguidas por fricativas e, finalmente, pelas líquidas^(23,30) e, por último, a estrutura silábica de OC.

Obteve-se um percentual alto (91,7%) de discriminação entre as crianças com desenvolvimento fonológico típico e atípico. Sendo que os principais fonemas que auxiliaram nesta discriminação/classificação entre os grupos clínico e controle foram /p/, /s/, /z/, /ʒ/ e /r/ em OG; /s/ em CG; /k/, /v/, /ʃ/ e /l/ em OI; /p/, /ʒ/, /n/ e /ŋ/ em OM; /r/ em CF; plosiva + /r/; plosiva + /r/ em OM; fricativa + /r/ em OI. Dessa forma, foi possível verificar que o INFONO atende a um de seus objetivos principais, o de auxiliar no diagnóstico diferencial entre crianças com desenvolvimento fonológico típico e de crianças com desvio fonológico. Outros instrumentos que analisam a fonologia (CAAP e GFTA-2) também referiram apresentar alto percentual de identificação de crianças com desvio fonológico^(10,11).

O desenvolvimento do INFONO foi motivado pela necessidade de desenvolver uma avaliação fonológica formal normatizada para a população do sul do Brasil. Dessa forma, esse instrumento foi desenvolvido em *software* para se tornar mais atrativo para a criança e facilitar a aplicação e obtenção dos resultados pelo fonoaudiólogo. No entanto, existem algumas limitações deste estudo, por exemplo, não apresentar coeficientes de confiabilidade para cada subgrupo para o qual o teste fornece dados normativos (faixa etária, gênero, tipo de escola). Os coeficientes de confiabilidade podem variar de acordo com as diferenças de habilidade entre os grupos, por exemplo, grupos mais jovens tendem a ter menor confiabilidade⁽¹⁴⁾.

Por fim, o processo inicial de investigação das propriedades psicométricas do INFONO mostrou evidências de validade e fidedignidade importantes, mostrando ser confiável para aplicação e obtenção dos escores para o diagnóstico. A falta de um instrumento padrão-ouro de avaliação fonológica no Brasil que tenha estudos de validade e fidedignidade impossibilita a realização de outras análises comparativas, especialmente. Porém, outros estudos devem ser realizados em busca de outras evidências de validade e fidedignidade para o instrumento INFONO. Apesar dessas limitações, o desenvolvimento deste instrumento tem importantes implicações clínicas para a avaliação e diagnóstico do desvio fonológico no sul do país

por representar a primeira avaliação fonológica normatizada que apresenta evidências de validade e fidedignidade para a avaliação e diagnóstico do desvio fonológico.

CONCLUSÃO

O INFONO apresentou adequadas evidências de validade (consistência interna e validade de critério - análise discriminante) e evidências de fidedignidade (consistência interna e confiabilidade intra e interavaliador). Essas análises indicaram uma confiabilidade satisfatória dos itens utilizados para avaliação, bem como excelente concordância entre os escores do instrumento, quando aplicado pelo mesmo avaliador em momentos distintos ou por avaliadores diferentes. Esses achados podem ser consequências dos procedimentos de administração, gravação e pontuações do instrumento, além de ter sido administrado por avaliadores experientes, o que sugere que, antes de aplicar o instrumento, o avaliador deve conhecer e dominar o instrumento.

Esses resultados demonstraram que o INFONO teve boa confiabilidade diagnóstica no processo de normatização. Portanto, é um instrumento válido e confiável para a aplicação na clínica e em pesquisas fonoaudiológicas brasileiras para avaliação da aquisição fonológica típica e atípica.

AGRADECIMENTOS

Agradecemos aos bolsistas de Iniciação Científica: Jonatas Corteze, do Curso de Ciências da Computação, por desenvolver as várias versões do INFONO para as diferentes etapas do projeto; e, do Curso de Desenho Industrial, Matheus Tanuri, por desenvolver os desenhos do INFONO, e Matheus Moreira, por desenvolver a Logomarca. Bem como à Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) o apoio concedido no Edital PqG para a realização deste projeto (processo nº 11/2184-8).

REFERÊNCIAS

1. Namasivayam AK, Pukonen M, Goshulak D, Yu VY, Kadis DS, Kroll R, et al. Relationship between speech motor control and speech intelligibility in children with speech sound disorders. *J Commun Disord*. 2013;46(3):264-80. <http://dx.doi.org/10.1016/j.jcomdis.2013.02.003>. PMID:23628222.
2. Strand EA, McCauley RJ, Weigand SD, Stoeckel RE, Baas BS. A motor speech assessment for children with severe speech disorders: reliability and validity evidence. *J Speech Lang Hear Res*. 2013;56(2):505-20. [http://dx.doi.org/10.1044/1092-4388\(2012/12-0094\)](http://dx.doi.org/10.1044/1092-4388(2012/12-0094)). PMID:23275421.
3. Grigos MI, Moss A, Lu Y. Oral articulatory control in childhood apraxia of speech. *J Speech Lang Hear Res*. 2015;58(4):1103-18. http://dx.doi.org/10.1044/2015_JSLHR-S-13-0221. PMID:25951237.
4. Sugden E, Baker E, Munro N, Williams AL. Involvement of parents in intervention for childhood speech sound disorders: a review of the evidence. *Int J Lang Commun Disord*. 2016;51(6):597-625. <http://dx.doi.org/10.1111/1460-6984.12247>. PMID:27017993.
5. Müürsepp I, Aibast H, Gapeyeva H, Pääsuke M. Motor skills, haptic perception and social abilities in children with mild speech disorders. *Brain Dev*. 2012;34(2):128-32. <http://dx.doi.org/10.1016/j.braindev.2011.02.002>. PMID:21376491.
6. Lousada M, Jesus LMT, Hall A, Joffe V. Intelligibility as a clinical outcome measure following intervention with children with phonologically based

- speech-sound disorders. *Int J Lang Commun Disord.* 2014;49(5):584-601. <http://dx.doi.org/10.1111/1460-6984.12095>. PMID:24861159.
7. Barrozo TF, Pagan-Neves LO, Vilela N, Carvalho RM, Wertzner HF. The influence of (central) auditory processing disorder in speech sound disorders. *Braz J Otorhinolaryngol.* 2016;82(1):56-64. <http://dx.doi.org/10.1016/j.bjorl.2015.01.008>. PMID:26612604.
 8. Vick JC, Campbell TF, Shriberg LD, Green JR, Abdi H, Rusiewicz HL, et al. Distinct developmental profiles in typical speech acquisition. *J Neurophysiol.* 2012;107(10):2885-900. <http://dx.doi.org/10.1152/jn.00337.2010>. PMID:22357794.
 9. Farquharson K. Language or motor: reviewing categorical etiologies of speech sound disorders. *Front Psychol.* 2015;6:1708. <http://dx.doi.org/10.3389/fpsyg.2015.01708>. PMID: 26583005.
 10. Goldman R, Fristoe M. Goldman-Fristoe Test of Articulation 2 - Test manual. Circle Pines, MN: American Guidance Services Inc; 2000.
 11. Second WA, Donohue JS. CAAP - Clinical Assessment of Articulation and Phonology – Supplemental Examiner’s Manual. USA: Super Duper®Publications; 2002.
 12. Yavas M, Hernandez CLM, Lamprecht RR. Avaliação Fonológica da Criança. Porto Alegre: Artes Médicas; 2002.
 13. Wertzner HF. Fonologia (Parte A). In: Andrade CRF, Befi-Lopes DM, Fernandes FDM, Wertzner HF, editors. Teste de linguagem infantil nas áreas de fonologia, vocabulário, fluência e pragmática. Carapicuíba: Pró-Fono; 2004. Vol. 2, 98 p.
 14. Kirk C, Vigeland L. A psychometric review of norm-referenced tests used to assess phonological error patterns. *Lang Speech Hear Serv Sch.* 2014;45(4):365-77. http://dx.doi.org/10.1044/2014_LSHSS-13-0053. PMID:25091265.
 15. McLeod S, Verdon S. A review of 30 speech assessments in 19 languages other than English. *Am J Speech Lang Pathol.* 2014;23(4):708-23. http://dx.doi.org/10.1044/2014_AJSLP-13-0066. PMID:24700105.
 16. Cunha VLO, Capellini AS. Construção e validação de instrumento de avaliação da compreensão de leitura para escolares do terceiro ao quinto ano do Ensino Fundamental. *CoDAS.* 2014;26(1):28-37. PMID:24714856.
 17. Flipsen P Jr, Ogiela DA. Psychometric characteristics of single-word tests of children’s speech sound production. *Lang Speech Hear Serv Sch.* 2015;46(2):166-78. http://dx.doi.org/10.1044/2015_LSHSS-14-0055. PMID:25654693.
 18. Friberg JC. Considerations for test selection: How do validity and reliability impact diagnostic decisions? *Child Lang Teach Ther.* 2010;26(1):77-92. <http://dx.doi.org/10.1177/0265659009349972>.
 19. McCauley RJ, Swisher L. Use and misuse of norm-referenced tests in clinical assessment: a hypothetical case. *J Speech Hear Disord.* 1984;49(4):338-48. <http://dx.doi.org/10.1044/jshd.4904.338>. PMID:6389982.
 20. Fonseca RP, Parente MAMP, Côté H, Ska B, Joannette Y. Apresentando um instrumento de avaliação da comunicação à Fonoaudiologia Brasileira: bateria MAC. *Pro Fono.* 2008;20(4):285-91. <http://dx.doi.org/10.1590/S0104-56872008000400014>. PMID:19142474.
 21. Pagliarini KC, Ortiz KZ, Parente MAMP, Arteche A, Joannette Y, Nespolous JL, et al. Montreal-Toulouse Language Assessment Battery for aphasia: Validity and reliability evidence. *NeuroRehabilitation.* 2014;34(3):463-71. PMID:24473247.
 22. Silva MK, Ferrante C, Borsari JV, Pereira MMB. Aquisição fonológica do Português Brasileiro em crianças do Rio de Janeiro. *J Soc Bras Fonoaudiol.* 2012;24(3):248-54.
 23. Ceron MI, Gubiani MB, Oliveira CR, Keske-Soares M. Factors influencing consonant acquisition in Brazilian Portuguese-Speaking Children. *J Speech Lang Hear Res.* 2017;60:759-771. http://dx.doi.org/10.1044/2016_JSLHR-S-15-0208.
 24. Savoldi A, Ceron MI, Keske-Soares M. Quais são as melhores palavras para compor um instrumento de avaliação fonológica? *Audiol Commun Res.* 2013;18(3):194-202. <http://dx.doi.org/10.1590/S2317-64312013000300009>.
 25. Ceron MI. Instrumento de Avaliação Fonológica (INFONO): Desenvolvimento e estudos psicométricos. Universidade Federal de Santa Maria, Santa Maria; 2015.
 26. Pavez MM, Maggiolo M, Coloma CJ. Test para evaluar procesos de implificación fonológica - TEPROSIF-R. Santiago: Pontificia Universidad Católica de Chile; 2009.
 27. Gubiani MB, Keske-Soares M. Evolução fonológica de crianças com desvio fonológico submetidas a diferentes abordagens terapêuticas. *Rev CEFAC.* 2014;16(2):663-71. <http://dx.doi.org/10.1590/S1516-18462012005000071>.
 28. Giacchini V, Mota HB, Mezzomo CL. Variáveis relevantes no processo terapêutico para a aquisição do onset complexo na fala de crianças com desvio fonológico. *Rev CEFAC.* 2015;17(1, Supl 1):17-26. <http://dx.doi.org/10.1590/s1516-18462012005000049>.
 29. Murphy CFB, Pagan-Neves LO, Wertzner HF, Schochat E. Children with speech sound disorder: comparing a non-linguistic auditory approach with a phonological intervention approach to improve phonological skills. *Front Psychol.* 2015;6:64. <http://dx.doi.org/10.3389/fpsyg.2015.00064>. PMID:25698997.
 30. Lousada M, Mendes AP, Valente AR, Hall A. Standardization of a phonetic-phonological test for European-Portuguese children. *Folia Phoniatr Logop.* 2012;64(3):151-6. <http://dx.doi.org/10.1159/000264712>. PMID:22965068.

Contribuição dos autores

MIC foi responsável pela concepção do estudo, coleta, análise e interpretação dos dados; MBG foi responsável pela coleta e revisão do artigo; CRO colaborou com a interpretação dos dados e revisão do artigo; MKS foi responsável pelo delineamento do estudo e aprovação final da versão a ser publicada.