

Artículo Original

# Estudio de la validez y fiabilidad del Test de Repetición Fonético-Fonológica (TREFF) como instrumento para evaluar el desempeño fonético-fonológico en niños chilenos

Sandra Vásquez Gajardo <sup>a,\*</sup>, Hernán León Valdés <sup>b</sup>, Jaime Soto-Barba <sup>a</sup>, Katia Sáez-Carrillo <sup>c</sup>

<sup>a</sup> *Departamento de Español, Facultad de Humanidades y Arte, Universidad de Concepción, Chile*

<sup>b</sup> *Departamento de Fonoaudiología, Facultad de Medicina, Universidad de Concepción, Chile*

<sup>c</sup> *Departamento de Estadística, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Chile*

## RESUMEN

En la actualidad, se requiere un mayor repertorio de pruebas válidas y confiables que evalúen el desempeño fonético-fonológico en niños chilenos. El objetivo de este estudio es determinar la validez y fiabilidad del Test de Repetición Fonético-Fonológica (TREFF). La validación fue llevada a cabo mediante un juicio de expertos y la fiabilidad fue determinada mediante el análisis estadístico de la estabilidad temporal y la concordancia intra- e interjueces. Para este proceso, se aplicó el instrumento en una población infantil perteneciente al nivel sociocultural medio alto, con un desarrollo fonético-fonológico acorde a lo esperado para su edad. El análisis de las respuestas fue realizado por tres jueces evaluadores, quienes determinaron si el niño o la niña lograba producir los elementos evaluados. Respecto a la validación, el instrumento fue modificado considerando las sugerencias de los expertos. En cuanto a la fiabilidad, los resultados obtenidos demostraron estabilidad temporal y un alto grado de concordancia a nivel intra- e interjueces. El estudio permite concluir que el test TREFF cumple con los criterios de validez y fiabilidad.

## Palabras clave:

Validez; Validación; Fiabilidad; Evaluación fonético-fonológica

## Study of the validity and reliability of the Phonetic-Phonological Repetition Test (TREFF) as an instrument to evaluate the phonetic-phonological performance in Chilean children

## ABSTRACT

At present, a greater repertoire of valid and reliable tests is required to evaluate the phonetic-phonological performance in Chilean children. The objective of this study is to determine the validity and reliability of the Phonetic-Phonological Repetition Test (TREFF). Validation was carried out by expert judgment and reliability was determined by statistical analysis of temporal stability and intra and inter-judge agreement. For this process, the instrument was applied in a medium-high sociocultural level child population, with a phonetic-phonological development according to what was expected for their age. The analysis of the responses was carried out by three evaluating judges, who determined if the boy or girl achieved to produce the elements evaluated. In terms of validation, the instrument was modified considering the experts' suggestions. Concerning to reliability, the results obtained demonstrated temporal stability and a high degree of concordance at the intra-judge and inter-judge levels. The study allows us to conclude that the TREFF test meets the validity and reliability criteria.

## Keywords:

Validity; Validation; Reliability; Phonetic-phonological evaluation

\* *Autor/a correspondiente:* Sandra Vásquez Gajardo  
*Email:* [flgasandravasquezgajardo@gmail.com](mailto:flgasandravasquezgajardo@gmail.com)

Recibido: 26-10-2020  
Aceptado: 17-11-2021  
Publicado: 22-11-2021

## INTRODUCCIÓN

En toda lengua resulta relevante poder determinar los comportamientos fonético-fonológicos del habla infantil. Específicamente, conocer los patrones y tendencias de adquisición del repertorio fonológico según la edad de niños y niñas, el desempeño típico de estos procesos, la alteración del comportamiento fonético-fonológico en determinados grupos de la población infantil, entre otros. Sin embargo, la tarea presenta importantes desafíos, sobre todo porque se necesita contar con un instrumento que recoja muestras de habla que permitan hacer comparaciones entre sujetos, que elicite diferentes producciones enmarcadas en el repertorio y la fonotaxis y que permita una aplicación que evalúe las poblaciones requeridas en plazos razonables. Todo esto con el fin de observar el comportamiento oral infantil de manera efectiva de acuerdo con los criterios y variables que se quieran considerar. Debido a que este instrumento debe recabar información fidedigna, es imprescindible que cumpla con determinadas propiedades psicométricas, es decir, presenten validez y fiabilidad o confiabilidad. Lo anterior implica, por lo tanto, someter la prueba utilizada a un estudio que comprenda las fases de validación y a la determinación de su fiabilidad.

La validez consiste en el grado en que un instrumento mide verdaderamente aquello que busca medir o funciona para el objetivo para el cual fue diseñado (Arribas, 2004). En ella, encontramos diversos tipos, tales como la validez de expertos, la validez de contenido, la validez de criterio y la validez de constructo. La validez de expertos consiste en la selección de personas con experticia en el tema que pueden entregar información, evidencia, juicios y valoraciones respecto del instrumento que se revisa (Escobar Pérez & Cuervo Martínez, 2008). Por su parte, la validez de contenido indica si el instrumento mide todos o la mayoría de los componentes del dominio de contenido de las variables a medir (Hernández & Mendoza, 2018). La validez de criterio alude a que el instrumento funciona similar a otros instrumentos (Sánchez & Echeverry, 2004) y, la de constructo, evalúa el nivel en que el instrumento refleja la teoría del concepto que mide.

La fiabilidad o confiabilidad consiste en el grado en que un instrumento mide con precisión y con un mínimo de error aquello que se pretende observar (Arribas, 2004), es decir, que su aplicación repetida al mismo sujeto produce los mismos resultados (Hernández et al., 2014).

En el ámbito de la observación oral infantil en Chile, uno de los pocos instrumentos válidos y confiables para evaluar el

desempeño fonológico es el TEPROSIF-R, el cual se basa en la teoría de la fonología natural (Ingram, 1983). Este test se encuentra estandarizado y, por lo mismo, ha sido ampliamente utilizado en Chile, tanto en investigación como en clínica. Sin embargo, dado que no considera todo el repertorio fonológico del Español de Chile, ni todas las posibilidades de producción fonética dependientes de la fonotaxis y de la acentuación, se necesita disponer de otras alternativas que evalúen el componente oral de la manera más amplia y completa posibles, cumpliendo, además, con la exigencia de ser opciones que presenten validez y confiabilidad.

Hamdan-Rosales et al. (2020), en el marco de una investigación en la cual se relaciona el comportamiento fonético-fonológico y léxico-semántico en una población infantil determinada, crean el Test de Repetición Fonético-Fonológica (TREFF). Este test se diseña con el objetivo de obtener un corpus de producciones para el Español de Chile que permita un análisis fonético y/o fonológico, con el enfoque que cada investigador requiere. Se opta por utilizar el método de la repetición inmediata de palabras, pues permite una aplicación en tiempos acotados a muestras de participantes en cantidades suficientes como para obtener resultados significativos. Por otra parte, se descarta la utilización de la nominación de láminas, debido a que el evaluado se puede enfrentar a posibles barreras léxicas al no lograr acceder a la palabra o al no conocerla, lo que imposibilitaría obtener el enunciado requerido. Además, no se dispone de láminas que representen todos los contextos fonéticos que se espera evaluar con TREFF. A su vez, tampoco se considera la evaluación del habla espontánea, pues si bien presenta muchas ventajas, también tiene inconvenientes en tareas de investigación como la gran cantidad de tiempo requerido para realizar la transcripción y el análisis. Otros inconvenientes son que, en el caso en que el habla del niño es ininteligible, no se puede llevar a cabo la transcripción y que el sujeto examinado no produzca todos los aspectos del habla que se buscan evaluar (Vásquez, 2020).

La versión original de TREFF consta de una lista de 104 palabras que permite evaluar la producción del repertorio fonológico del español de Chile. Se encuentra subdividida en 7 partes: Primero, los 17 fonemas del Español de Chile los cuales son evaluados a través de la repetición de palabras bisilábicas y trisilábicas, ordenándose por zona, de anterior a posterior (bilabiales /p-b-m/, labiodental /f/, postdentales /d̪-t̪/, alveolares /s-n-l-r-r/, palatales /d̺-ɲ-t͡ʃ/ y velares /k-g-x/), en las posiciones inicial, medial y final de la palabra (se descartan bilabiales, labiodentales y palatales) y en posición trabante (se descartan postdentales y palatales); segundo, los diptongos semiconsonánticos y semivocálicos; tercero, los hiatos; cuarto, los grupos consonánticos /pl-pr/, /bl-

br/, /tr/, /dr/, /kl-kr/, /gl-gr/, /fl-fr/ en posición inicial (se descarta /tl/) e interior (se descarta /fl/); quinto, trisílabos; sexto, cuatrísílabos y séptimo, pentasílabos. En todos los casos, se prioriza la utilización de palabras cercanas al léxico infantil preescolar y escolar, específicamente para el rango de 3 a 7 años.

Además de la aplicación del test TREFF por Hamdan-Rosales et al. (2020), la herramienta fue utilizada en otras investigaciones en las que también logró obtener un corpus de habla que permitió efectuar análisis del desempeño fonético-fonológico infantil. La información obtenida concuerda con los datos aportados por otras investigaciones, tanto desde el punto de vista descriptivo mediante el análisis con la pauta CLAFF, como al observar, con una mirada evolutiva, el repertorio fonológico detectado (Alarcón, 2019; Fuica & Soto-Barba, 2014; León Valdés et al., 2019; Soto-Barba et al., 2011; J. C. Torres et al., 2018; V. Torres & Soto-Barba, 2016). Si bien la prueba logra el objetivo para el cual fue diseñada, aún necesita completar un proceso de estandarización que incluya tanto su validez como su confiabilidad.

En el marco expuesto, este estudio tiene como objetivo presentar el proceso de validez y confiabilidad del instrumento TREFF. Además, se muestran determinadas adecuaciones que otorgan a esta prueba una mayor precisión en la investigación de fenómenos fonético-fonológicos que se manifiestan en el habla infantil de niños chilenos.

## METODOLOGÍA

El presente estudio se organiza, principalmente, en dos etapas. La primera etapa consiste en la validación de TREFF por juicio de expertos y, la segunda, en la determinación de la fiabilidad con el análisis de concordancia intra- e interjueces y la estabilidad temporal. La segunda etapa se logra mediante la aplicación del instrumento a una población infantil, en dos oportunidades diferidas, en un periodo de 7 a 14 días.

### Diseño del estudio

El estudio presenta un diseño de investigación con un enfoque metodológico cuantitativo no experimental transversal de tipo descriptivo.

## Proceso de validación de TREFF por juicio de expertos

### Participantes

Se seleccionaron 5 jueces expertos quienes debían cumplir con las características de poseer formación de postgrado en lingüística con orientación hacia las ciencias del habla y familiarizados con el ámbito fonético-fonológico de la producción oral infantil. Con estos criterios, se escogieron tres expertos con grado de doctor y dos con grado de magíster en el área de ciencias del habla, los que debían analizar si cada uno de los componentes de la prueba medía lo que pretendía medir.

### Materiales y procedimiento

Se confeccionó y utilizó una planilla denominada “Evaluación de Juicio de Expertos” (ver Anexo 1), que mostraba la prueba TREFF bajo un formato que permitió a cada experto valorar sus componentes. Para organizar de mejor forma la información presentada, la prueba consideró la existencia de dimensiones e ítems. Las dimensiones informaban sobre el elemento evaluado, por ejemplo, la zona en que se articula el fonema y su posición dentro de la palabra; y los ítems correspondían a las palabras estímulo que el niño debía repetir.

Al inicio, se instruyó a cada juez sobre el objetivo del test, la metodología de aplicación y de registro, y el rol que como expertos les correspondía realizar en esta tarea. Se les indicó que debían analizar si cada uno de los ítems evaluaba congruentemente las dimensiones, valorándolos en casillas enumeradas del 1 al 3. Si el ítem evaluaba claramente la dimensión esperada y no requería modificaciones, debían elegir el número 3; si el ítem evaluaba la dimensión esperada, pero requería modificaciones, el número 2; si el ítem no evaluaba la dimensión esperada y requería reformularse, el número 1. Además, cada evaluador contó con un espacio para realizar observaciones por ítem o dimensión. A su vez, tenía un espacio para realizar comentarios generales del instrumento. En la Figura 1, se muestra un ejemplo de la planilla utilizada.

Ítem	Dimensiones			Evaluación			Observaciones
	Zona	Fonema	Posición	1	2	3	
/pa.to/	Bilabiales	/p/	Fonema inicial				
/ma.pa/			Fonema interior				
/ap.to/			Coda interior				

Figura 1. Imagen de un aspecto de la planilla utilizada por los jueces expertos.

Finalmente, se recibieron las respuestas de los jueces, las que se analizaron y organizaron, para determinar los ajustes pertinentes a la prueba definitiva que sería sometida a un análisis de fiabilidad.

### Ajustes de TREFF

Las observaciones efectuadas por los expertos indicaron que era apropiado utilizar de modo prioritario en la prueba, desde la perspectiva léxica, sustantivos concretos y comunes al léxico infantil. Desde la perspectiva fonético-fonológica, sugirieron sistematicidad en métrica y patrón acentual en los ítems, utilizar palabras que aseguraran que el sonido evaluado esté libre de variaciones inducidas por el contexto fonético y descartar la evaluación de sonidos en coda en la cual se produzcan

neutralizaciones que generen archifonemas al momento de la articulación; esto último teniendo en cuenta que se podrían generar variaciones en la percepción del fono producido por parte de quienes apliquen el test TREFF en alguna investigación.

Considerando la opinión de los expertos, se ordena y segmenta la prueba en seis módulos, realizando ajustes en ciertas dimensiones e ítems (ver resumen en Tabla 1). Para las dimensiones, se modifica la dimensión según posición dentro de la palabra por la dimensión, según la acentuación. Es decir, el elemento que se evaluaba en posición inicial, interior, coda interior y coda final dentro de la palabra, actualmente se evalúa en sílaba acentuada o no acentuada de la palabra. Las codas interiores y finales son evaluadas en un módulo diferente (Vásquez, 2020).

Tabla 1. Ajustes del test TREFF.

A. Ajustes en la distribución del contenido: considerando la evaluación de los expertos, se segmenta la prueba en seis módulos, realizando ajustes en ciertas dimensiones e ítems.

Partes de versión original		Módulos de la versión ajustada	
1.	Evaluación de producción de fonema según posición dentro de la palabra.	1.	Evaluación de producción de fonema según acentuación.
2.	Evaluación de producción de diptongos semivocálicos y semiconsonánticos.	2.	Evaluación de producción de fonemas en posición codal interior y/o final.
3.	Evaluación de producción de Hiatos	3.	Evaluación de producción de diptongos semivocálicos y semiconsonánticos.
4.	Evaluación de producción de ataques complejos según posición de la sílaba.	4.	Evaluación de producción de Hiatos.
5.	Evaluación de producción de palabras de distinta métrica: trisílabos, cuatrísílabos y pentasílabos.	5.	Evaluación de producción de ataques complejos según acentuación.
		6.	Evaluación de producción de palabras de distinta métrica: trisílabos, cuatrísílabos y pentasílabos.

B. Ajustes en dimensiones: se modifica la dimensión “posición dentro de la palabra” a “acentuación”

Dimensión original	Dimensión con ajustes
Evaluación del elemento según “posición dentro de la palabra”: El elemento es evaluado en posición inicial, medial, en coda interior y en coda final.	Evaluación del elemento según “acentuación”: El elemento es evaluado en sílaba tónica y en sílaba átona.

C. Ajustes en los ítems

Elemento evaluado	Ítem original	ítem ajustado	Motivo
/f/	/ka.'fè/	/'xe.fe/	Ajuste del ítem por cambio de dimensión: Debido a que se realizó el cambio de dimensión “posición dentro de la palabra” a “acentuación” se ajustaron aquellos ítems que no coincidían con la nueva dimensión.
/t/	/pa.'tè/	/'pa.ʈa/	
/g/	/'xu.go/	/'ma.go/	

/gl/	/i.'glu/	/'re.gla/	
/tʁ/	/a.'tʁas/	/'pos.tʁe/	
/ɲ/	/ɲaɲ.'ðu/	/'na.tø/	
/x/	/xo.'se/	/'xu.go/	
/dʁ/	/dʁa.'gon/	/'dʁo.ga/	
/iu/	/siu.'dʁad/	/'biu.dʁa/	Ajuste del ítem para lograr sistematicidad en acentuación: Los expertos sugirieron que, en el caso del diptongo, el elemento evaluado se debía encontrar en la sílaba tónica, por lo que se ajustaron los ítems que no coincidían con tal requerimiento.
/ua/	/'a.gua/	/'gua.tʁa/	
/f/	/'fo.ka/	/'fa.ro/	Ajuste del ítem por posibles cambios en la articulación: Se realizaron ajustes a los ítems para evitar posibles cambios en la articulación debido a su contexto fonético (como velarización y/o dentalización del elemento evaluado).
/n/	/'kaɲ.tʁo/	/'pan.sa/	
/l/	/'pal.tʁa/	/'bol.sa/	
/kl/	/'an.kla/	/'tʁe.kla/	
/fl/	/'flan/	/'fla.ko/	Ajuste del ítem para lograr sistematicidad en métrica: Los expertos sugirieron el uso sistemático de bisílabos a lo largo de la prueba. Cabe considerar que la evaluación de diferentes métricas de palabra (trisílabos, polisílabos y pentasílabos) es efectuada en un módulo aparte.
/tʁ/	/'tʁen/	/'tʁo.pa/	
/ei/	/'rei/	/'rei.na/	
/oi/	/'boi/	/'boi.na/	
/ie/	/'pie/	/'sie.lo/	
Cuatrisílabo	/tʁe.'le.fo.no/	/tʁa.dʒa.'ri.nes/	Ajustes del ítem para lograr sistematicidad en organización métrica en pies: Se utiliza el mismo pie métrico en todos los estímulos del módulo de métrica para homologar la estructura acentual.
Pentasílabo	/re.fri.xe.ra.'dʁor/	/ka.pe.ru.'si.tʁa/	
Pentasílabo	/au.tʁo.'ma.tʁi.ko/	/es.ka.ra.'ba.xo/	
Pentasílabo	/e.li.'kop.tʁe.ro/	/na.tʁu.ra.'le.sa/	

## Proceso de determinación de la fiabilidad de TREFF

### Participantes

La muestra, determinada por conveniencia, se constituyó por 31 participantes de 4.0 a 5.6 años de niveles medio mayor y prekínder, pertenecientes al nivel sociocultural medio alto.

Para el proceso de selección de los participantes se determinó, en primer lugar, que fueran niños y/o niñas que se encontraran en una etapa del desarrollo fonético-fonológico medianamente desarrollada, por lo que el estudio se concentró en abordar edades entre los 4.0 a 5.6 años. Además, su desempeño fonético-fonológico debía ser acorde a lo esperado para su edad, lo que requirió descartar alteraciones del habla y del lenguaje. Cabe mencionar que la variable sexo no fue considerada en esta investigación.

Se constató que los informantes poseyeran órganos fonoarticulatorios anatómica y funcionalmente indemnes, evaluándolos mediante una pauta de evaluación confeccionada de manera ad hoc para el proceso. Además, se aplicó el TEPROSIF-R con el fin de descartar alteraciones en el desempeño fonológico.

Todas las aplicaciones fueron registradas en audio con una grabadora TASCAM DR-40 para analizar el desempeño. Asimismo, a través de la información otorgada por las educadoras de párvulos, se descartaron a aquellos niños con diagnóstico de trastornos del desarrollo del lenguaje o trastornos del espectro autista, así como de otras condiciones que pudieran afectar el desempeño fonético-fonológico. Por otra parte, todos los niños debían poseer un consentimiento informado, firmado por los padres. Por su parte, los niños debían manifestar su asentimiento verbal antes de participar en las evaluaciones (ver Anexo 2).

### Procedimiento

Se realizó la aplicación del instrumento a cada uno de los niños seleccionados, siguiendo la misma modalidad. Cada niño debía sentarse frente al evaluador, quien debía tapar sus labios para evitar el posible apoyo visual en la articulación de las palabras. El evaluador comenzaba dando las instrucciones, haciendo énfasis en que el participante debía escuchar atentamente y repetir las palabras entregadas verbalmente. Se iniciaba con estímulos de ejemplo para corroborar que el niño comprendiera la tarea. Una vez comprendida la actividad, se comenzaba la evaluación

siguiendo el orden de los módulos y solicitando la repetición de los estímulos de izquierda a derecha. Cabe mencionar que, para mantener la atención del niño durante la prueba, se le motivaba en contexto de juego, diciéndole que se irían pasando etapas hasta llegar a la meta, reforzándolo positivamente cada vez que finalizaba un módulo. Si se observaba cansancio en el niño, se disponía de un espacio de tiempo de 10 a 20 segundos para, luego, retomar el módulo siguiente.

La prueba TREFF se aplicó en una primera instancia (aplicación n°1) y, luego, dos semanas después, se procedió a aplicar el mismo instrumento por segunda vez a la misma muestra, bajo las mismas condiciones (aplicación n°2). Todas las aplicaciones fueron registradas en audio con una grabadora TASCAM DR-40.

Las audiciones de las grabaciones de TREFF fueron realizadas por tres evaluadores en dos oportunidades diferentes (primera audición y segunda audición), quienes también tenían formación de postgrado en el área de ciencias del habla. Una vez efectuado el juicio auditivo, se realizó la transcripción fonológica de cada palabra producida por los niños.

Después de que las respuestas de los estudiantes fueron transcritas y analizadas, se procedió a determinar si cada ítem era logrado o no logrado. Un ítem se consideró logrado si el niño producía el elemento fonológico, de acuerdo con lo esperado, en la dimensión evaluada. Por el contrario, un ítem se consideró no logrado si el niño no era capaz de producir el elemento fonológico, de acuerdo con lo esperado, en la dimensión evaluada. Este criterio se aplicó con todos los ítems de los módulos del uno al cinco, los cuales evaluaron la producción de fonemas consonánticos en sílaba tónica, átona, coda, diptongos, hiatos y ataques complejos. Sin embargo, al evaluar la metría, en el módulo seis, se consideró como logrado si se obtenía la metría esperada y la mantención de la secuencia silábica y fonemas constituyentes de las sílabas. Por el contrario, se consideró como no logrado aquel ítem que no conservó la metría o la secuencia silábica o la secuencia fonémica al interior de cada sílaba. Para realizar el análisis estadístico, el desempeño fue codificado numéricamente, asignando un puntaje de 1 a los ítems logrados y un puntaje de 0 a los ítems no logrados.

Los resultados de todo el proceso fueron sometidos, en primer lugar, a un análisis de concordancia intrajuez, que medía la consistencia de un mismo evaluador entre la primera y la segunda audición de una misma aplicación; en segundo lugar, a un análisis de concordancia interjueces, que determinaba qué variabilidad había entre las percepciones de los distintos evaluadores para una misma audición. En tercer lugar, a un análisis de estabilidad

temporal, cuyo resultado esperado era que no existieran diferencias significativas, entre las respuestas de la primera y segunda aplicación de la prueba.

Los tres evaluadores, descritos como Juez 1, Juez 2 y Juez 3, debieron oír la grabación de la Aplicación n°1 del instrumento, en dos instancias, con un intervalo de tiempo de 7 a 14 días. Todas las audiciones fueron inmediatamente transcritas en las hojas de registro de TREFF, las cuales se analizaron posteriormente. En una segunda instancia, los mismos evaluadores debieron oír la grabación de la segunda aplicación del instrumento, en dos instancias, con el mismo intervalo de tiempo anteriormente mencionado. Esta tarea fue realizada con el fin de determinar si existía concordancia en los registros de cada evaluador, es decir, si las respuestas de la primera y segunda audición, de la misma aplicación, eran iguales o existía algún grado de variación. Además, la información recolectada permitió estudiar el grado de concordancia entre los evaluadores, a través de la comparación de sus resultados. Finalmente, se realizó el estudio de la estabilidad temporal del instrumento, en el que se compararon, en cada juez, los resultados de la Aplicación n°1 con los resultados de la Aplicación °2, considerando los resultados totales de la prueba.

Para la obtención de los resultados a nivel estadístico, las concordancias y la estabilidad temporal fueron analizadas mediante el Coeficiente de Correlación Intraclass (CCI) y el Coeficiente de Correlación de Concordancia de Lin (CCC). Se decidió utilizar dos coeficientes con el fin de contrastarlos y corroborar el grado de confiabilidad. Los coeficientes de fiabilidad pueden oscilar entre cero y uno, en los cuales, cero indica nula confiabilidad y uno significa lo máximo de fiabilidad (Hernández & Mendoza, 2018).

## **RESULTADOS**

Para determinar la concordancia intrajuez, primero se realizó el análisis de los registros correspondientes a las respuestas de la primera y segunda audición de la aplicación N°1 de cada evaluador. Para ello, se consideraron los resultados totales del instrumento. Los resultados de correlación obtenidos de los coeficientes CCI y CCC muestran, para el Juez 1, 2 y 3, un puntaje sobre 0.9, lo que refleja un alto grado de concordancia. Esto indica que las evaluaciones de la primera y segunda audición de la aplicación N°1 poseen una mínima variabilidad entre las respuestas, lo que demuestra fiabilidad del instrumento (ver Tabla 2).

**Tabla 2.** Análisis de concordancia intrajuez entre la primera y segunda audición de la aplicación N°1.

	Coeficiente de correlación intraclase	Límites		Coeficiente de Lin	Límites	
		Inferior	Superior		Inferior	Superior
Juez 1	.998	.996	.999	.996	.991	.998
Juez 2	.984	.967	.992	.967	.935	.984
Juez 3	.968	.901	.987	.936	.874	.968

Posteriormente, se realizó el estudio de análisis de concordancia intrajuez para medir la concordancia de las respuestas entre la primera y segunda audición de la aplicación N°2, considerando

los resultados totales del instrumento. Los resultados obtenidos para el Juez 1, Juez 2 y Juez 3, son sobre 0.9, lo que indica un alto grado de fiabilidad (ver Tabla 3).

**Tabla 3.** Análisis de concordancia intrajuez entre la primera y segunda audición de la aplicación N°2.

	Coeficiente de correlación intraclase	Límites		Coeficiente de Lin	Límites	
		Inferior	Superior		Inferior	Superior
Juez 1	.994	.987	.997	.987	.974	.994
Juez 2	.99	.967	.996	.981	.962	.99
Juez 3	.959	.917	.98	.92	.841	.96

Para determinar la concordancia entre los evaluadores, se consideraron los resultados totales de la prueba. En primer lugar, se realizó la comparación de los datos entre los jueces para la aplicación N°1, contrastando las evaluaciones del Juez 1 con el Juez 2; luego, el Juez 1 con el Juez 3; y, finalmente, el Juez 2 con

el Juez 3. Los métodos estadísticos utilizados fueron los mismos que se aplicaron para analizar la concordancia intrajuez. Los resultados obtenidos para la concordancia interjueces en la aplicación N°1 demuestran un grado de concordancia favorable (ver Tabla 4).

**Tabla 4.** Análisis de concordancia interjueces de la primera audición de la aplicación N° 1.

	Coeficiente de correlación intraclase	Límites		Coeficiente de Lin	Límites	
		Inferior	Superior		Inferior	Superior
Jueces	.975	.954	.988	---	---	---
Juez 1- Juez 2	.978	.954	.989	.955	.91	.978
Juez 1- Juez 3	.959	.891	.983	.92	.848	.959
Juez 2- Juez 3	.952	.899	.977	.906	.817	.952

El análisis de concordancia entre jueces para la aplicación N°2, también consideró los resultados totales de la prueba. En CCI y CCC, los puntajes para las comparaciones de Juez 1 con Juez 2,

Juez 1 con Juez 3 y Juez 2 con Juez 3 son sobre 0.9, indicando un alto grado de concordancia (ver Tabla 5).

**Tabla 5.** Análisis de concordancia interjueces de la primera audición de la aplicación N°2.

	Coeficiente de correlación intraclase	Límites		Coeficiente de Lin	Límites	
		Inferior	Superior		Inferior	Superior
Jueces	.987	.975	.994	---	---	---
Juez 1- Juez 2	.987	.97	.994	.974	.948	.987
Juez 1- Juez 3	.976	.928	.99	.953	.912	.975
Juez 2- Juez 3	.979	.957	.99	.958	.917	.979

Para el análisis de la estabilidad temporal, se compararon los resultados de la aplicación N°1 con los resultados de la aplicación N°2, por cada juez, considerando los resultados totales de la prueba. Los datos obtenidos son sobre 0.9 tanto para el Juez 1, como para el Juez 2 y Juez 3. Esto señala que no hubo cambios

significativos entre las respuestas de la primera aplicación y la segunda aplicación, las cuales fueron llevadas a cabo con un intervalo de tiempo de 7 a 14 días. El instrumento demostró ser estable en el tiempo, lo que denota fiabilidad (ver Tabla 6).

**Tabla 6.** Análisis de estabilidad temporal entre la aplicación N°1 y aplicación N°2.

	Coeficiente de correlación intraclase	Límites		Coeficiente de Lin	Límites	
		Inferior	Superior		Inferior	Superior
Juez 1	.978	.956	.99	.956	.913	.978
Juez 2	.98	.958	.991	.961	.921	.981
Juez 3	.956	.909	.979	.913	.828	.957

## DISCUSIÓN

El presente estudio permitió cumplir con los objetivos planteados inicialmente, esto es, conseguir la validación de TREFF, junto con lograr determinar que cumple con las propiedades de fiabilidad necesarias. La validez fue lograda mediante el proceso de juicio de expertos, en el que se consideraron las sugerencias de los jueces para llevar a cabo las modificaciones que aseguraron que el instrumento medía los aspectos que declaraba evaluar. En cuanto a la fiabilidad, el test TREFF demostró ser un instrumento estable en el tiempo, con un alto grado de concordancia intrajuez e interjueces.

Esta investigación, al aportar un nuevo instrumento con estudios de confiabilidad y validez, contribuye significativamente al ámbito lingüístico y fonoaudiológico, principalmente para fines de investigación, en el ámbito de la producción del habla infantil. Obtener el desempeño fonético-fonológico de niñas y niños chilenos a partir de un instrumento fidedigno permitirá profundizar los análisis de fenómenos lingüísticos de tipo dialectal e identificar un comportamiento típico o presencia de alteraciones fonético-fonológicas.

Como proyección de este trabajo, se espera utilizar TREFF como instrumento para concretar otros procesos formales, como lo es la obtención de perfiles de comportamiento típico en diferentes rangos etarios infantiles y en distintos niveles socioeconómicos, que podrán constituir valores normativos para este desempeño. Debe tenerse presente que las manifestaciones del habla infantil son dinámicas y están sujetas a cambios por razones sociales y culturales, entre otras, por lo se requiere actualizar las evaluaciones. Igualmente se podrían realizar adaptaciones de la prueba para estudiar y describir otros grupos chilenos como minorías étnicas, poblaciones rurales, entre otros. Se proyecta, además, que al contar con la mencionada descripción de desempeño típico de niños y niñas se podrá aplicar esta prueba a otros grupos infantiles atípicos, en diversos rangos de edad. Resulta importante comentar que, tratándose de un instrumento que utiliza la metodología de la imitación directa de palabras para recoger su muestra, en algunos tipos de investigaciones se podría efectuar una complementación con evaluaciones de habla espontánea o semiespontánea en los mismos informantes, para obtener información del desempeño de habla en contextos fonético-fonológicos diferentes a los que evalúa TREFF.

Con el fin de orientar el uso de este instrumento a la clínica fonoaudiológica, se proyecta la posibilidad de efectuar los análisis necesarios para determinar su especificidad, sensibilidad y validez predictiva. Por último, con la información obtenida en los procesos mencionados, se podrían establecer rangos etarios en los cuales se apliquen determinados módulos del test, transformando la prueba en una herramienta adaptable y flexible a diversos requerimientos de observación e investigación del habla infantil.

## FINANCIAMIENTO

Proyecto VRID Asociativo, código 218.083.036-1.0.

## REFERENCIAS

- Alarcón, E. (2019). *Ajustes Fonético-Fonológicos en niños de primer año de educación básica provenientes del sistema educativo público y privado de la comuna de San Pedro de la Paz. Incidencia del sexo y el tipo de establecimiento educacional al que asisten*. [Tesis de Magister, Universidad de Concepción]. <http://repositorio.udec.cl/jspui/handle/11594/701>
- Arribas, M. C. M. (2004). *Diseño y validación de cuestionarios*. 5(17), 23–29. [https://www.enferpro.com/documentos/validacion\\_cuestionarios.pdf](https://www.enferpro.com/documentos/validacion_cuestionarios.pdf)
- Escobar Pérez, J., & Cuervo Martínez, Á. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en medición*, 6(1), 27–36. <https://dialnet.unirioja.es/servlet/articulo?codigo=2981181>
- Fuica, M. A., & Soto-Barba, J. (2014). Ajustes fonético-fonológicos en niños(as) de 4 a 5 años que aprenden inglés como segunda lengua. *Literatura y lingüística*, 30, 258–283. <https://doi.org/10.4067/S0716-58112014000200014>
- Hamdan-Rosales, N., Soto-Barba, J., Sáez-Carrillo, K., & Riffó, B. (2020). Desempeño fonético-fonológico y léxico-semántico en un grupo de preescolares de sectores vulnerables de la provincia de Concepción: Incidencia del sexo y relación entre los niveles lingüísticos. *Nueva revista del Pacífico*, 73, 356–377. <https://doi.org/10.4067/S0719-51762020000200356>
- Hernández, R., Fernández, C., & Baptista, M. P. (2014). *Metodología de la investigación* (6ta ed.). McGraw-Hill.
- Hernández, R., & Mendoza, C. (2018). *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta*. Mc Graw Hill Education.
- Ingram, D. (1983). *Trastornos fonológicos en el niño*. Médica y Técnica.
- León Valdés, H., Ñanculeo Raguileo, M., & Soto-Barba, J. (2019). Ajustes fonético-fonológicos en niños de habla típica entre 3 y 4, y entre 4 y 5 años de edad. *Onomázein*, 43, 200–222. <https://doi.org/10.7764/onomazein.43.05>
- Sánchez, R., & Echeverry, J. (2004). Validación de escalas de medición en salud. *Revista de Salud Pública*, 6, 302–318. <https://doi.org/10.1590/S0124-00642004000300006>
- Soto-Barba, J., León, H., & Torres, V. (2011). Una propuesta para la clasificación de los ajustes fonético-fonológicos del habla infantil (CLAFF). *Onomázein*, 23(1), 69–79.
- Torres, J. C., León, H., & Figueroa, M. (2018). Desempeño fonético-fonológico en niños con Trastorno del Espectro Autista (TEA) de pre-kínder a tercero básico. *Revista Chilena de Fonoaudiología*, 17, 1–18. <https://doi.org/10.5354/0719-4692.2018.51611>
- Torres, V., & Soto-Barba, J. (2016). Ajustes fonético-fonológicos en niños con trastornos específicos del lenguaje mixto (TEL Mixto). *Onomázein*, 33, 69–87. <https://doi.org/10.7764/onomazein.33.7>
- Vásquez, S. (2020). *Determinación de la validez y fiabilidad del test de repetición fonético-fonológica (TREFF) como instrumento para evaluar el desempeño fonético-fonológico en niños chilenos*. [Tesis de Magister, Universidad de Concepción]. <http://repositorio.udec.cl/jspui/handle/11594/4679>

ANEXO 1. Parte 1.

Evaluación de Expertos TREFF

Evaluador: \_\_\_\_\_  
Instrumento: Test de Repetición Fonético-Fonológica (TREFF)

Instrucción:

Se solicita a usted participar en calidad de experto en la valoración de un instrumento que busca evaluar el desempeño fonético-fonológico en niños preescolares y escolares. Dicho instrumento consiste en un listado de palabras para repetir en forma directa, es decir, el evaluador lee al niño cada uno de los ítems, tal como lo indican las instrucciones del test, y el niño debe repetir. Esta tarea es grabada en audio.

Una vez aplicado y grabado el test TREFF, el evaluador procederá a transcribir fonéticamente las producciones del niño evaluado para luego determinar su desempeño en la Hoja de registro, mediante los indicadores Logrado o No logrado, los que a su vez atribuirán un puntaje al desempeño.

En su calidad de experto, le solicitamos que analice si cada uno de los ítems evalúa congruentemente las dimensiones “zona”, “fonema” y “posición de palabra” como se indica.

Para realizar esta tarea, se le solicita que, por cada ítem, marque con una equis (“x”) la alternativa que mejor represente su opinión, utilizando una de estas opciones:

1	2	3
El ítem <b>no evalúa</b> la dimensión esperada, debe reformularse.	El ítem <b>evalúa</b> la dimensión esperada, sin embargo requiere <b>modificaciones</b> .	El ítem <b>evalúa claramente</b> la dimensión esperada y no requiere modificaciones.

Adicionalmente, en aquellos casos en que ud. califique el ítem con la opción 2, podrá hacer observaciones y sugerencias acerca de la selección de éste o proponer otro que le parezca más adecuado en la sección “observaciones” de la presente pauta de evaluación ubicada junto a la evaluación de cada ítem. Otros comentarios generales podrán ser señalados en la sección ubicada al final del documento, en la cual además podrá opinar sobre la hoja de registro.

ANEXO 1. Parte 2.

Ítem	Dimensiones			Evaluación			Observaciones
	Zona	Fonema	Posición	1	2	3	
'pa.to/ 'ma.pa/ 'ap.to/	Bilabiales	/p/	Fonema inicial				
			Fonema interior				
			Coda interior				
'bo.te/ 'nu.be/ 'ob.xe.to/		/b/	Fonema inicial				
			Fonema interior				
			Coda interior				
'ma.no/ 'ka.ma/ 'kam.po/		/m/	Fonema inicial				
			Fonema interior				
			Coda interior				
'fo.ka/ 'ka.fe/ 'af.ta/	Labio dentales	/f/	Fonema inicial				
			Fonema interior				
			Coda interior				
'de.do/ 'ko.do/ 'pa.'red/	Post dentales	/d/	Fonema inicial				
			Fonema interior				
			Coda final				
'ta.sa/ 'pa.te/		/t/	Fonema inicial				
			Fonema interior				
'sa.po/ 'pa.sa/ 'pas.to/ 'bus/	Alveolares	/s/	Fonema inicial				
			Fonema interior				
			Coda interior				
			Coda final				
'ni.do/ 'mo.no/ 'kan.to/ 'pan/		/n/	Fonema inicial				
			Fonema interior				
			Coda interior				
			Coda final				
'lu.na/ 'pa.la/ 'pal.ta/ 'sol/ 'sol/		/l/	Fonema inicial				
			Fonema interior				
			Coda interior				
			Coda final				
'pe.ra/ 'par.te/ 'mar/	/r/	Fonema inicial					
		Fonema interior					
		Coda final					
'ro.sa/ 'pe.ro/	/r/	Fonema inicial					
		Fonema interior					
'd̥ʒa.be/ 'po.d̥ʒo/	Palatales	/d̥ʒ/	Fonema inicial				
			Fonema interior				
'nan.'d̥u/ 'ni.no/		/n/	Fonema inicial				
			Fonema interior				
'tʃi.ko/ 'le.tje/		/tʃ/	Fonema inicial				
			Fonema interior				
'ka.sa/ 'bo.ka/ 'kak.tus/	Velares	/k/	Fonema inicial				
			Fonema interior				
			Coda interior				
'ga.ʔo/ 'xu.go/ 'sig.no/		/g/	Fonema inicial				
			Fonema interior				
			Coda interior				
'xo.'se/ 'ro.xo/ 're.'lox/		/x/	Fonema inicial				
			Fonema interior				
			Coda final				

ANEXO 1. Parte 3.

Ítem	Dimensiones		Evaluación			Observaciones
	Ataque complejo	Posición	1	2	3	
/'plu.ma/ /so'pla/	/pl/	Inicial				
		Interior				
/'blu.sa/ /'ka.ble/	/bl/	Inicial				
		Interior				
/'flia/	/fl/	Inicial				
		Interior				
/a.'tle.ʧa/	/tl/	Inicial				
		Interior				
/'kla.bo/ /'an.kla/	/kl/	Inicial				
		Interior				
/'glo.bo/ /i.'glu/	/gl/	Inicial				
		Interior				
/'pre.so/ /'kom.pra/	/pr/	Inicial				
		Interior				
/'bra.so/ /'po.bre/	/br/	Inicial				
		Interior				
/'fru.ʧa/ /'ko.fre/	/fr/	Inicial				
		Interior				
/ʧren/ /a.'ʧras/	/ʧr/	Inicial				
		Interior				
/'gra.'gon/ /'la.ɣra/	/dʒr/	Inicial				
		Interior				
/'kre.ma/ /'mi.kro/	/kr/	Inicial				
		Interior				
/'gri.ʧo/ /'ʧi.gre/	/gr/	Inicial				
		Interior				

Ítem	Dimensión		Evaluación			Observaciones
	Diptongo semiconsonántico		1	2	3	
/'pie/	/ie/					
/siu.'dad/	/iu/					
/'pio.xo/	/io/					
/'ue.bo/	/ue/					
/'pia.no/	/ia/					
/'kuo.ta/	/uo/					
/'a.gua/	/ua/					

Ítem	Dimensión		Evaluación			Observaciones
	Diptongo semivocálico		1	2	3	
/'rei/	/ei/					
/'au.to/	/au/					
/'ai.re/	/ai/					
/'peu.mo/	/eu/					
/'kui.da/	/ui/					
/'boʏ/	/oi/					

Ítem	Dimensión		Evaluación			Observaciones
	Hiato		1	2	3	
/'ʧi.o/	/'i.o/					
/ba.'ul/	/a.'u/					
/go.'ʧe.a/	/'e.a/					
/'pu.a/	/'u.a/					
/re.'i/	/e.'i/					
/a.ʧa.'ud/	/a.'u/					

ANEXO 1. Parte 4.

Ítem	Dimensión	Evaluación			Observaciones
	Metría	1	2	3	
/t̥o.ˈma.ʧe/	trisílabo				
/pa.ˈle.ʧa/	trisílabo				
/ko.ˈne.xo/	trisílabo				
/xi.ˈra.fa/	trisílabo				
/ma.ri.ˈpo.sa/	cuatrísílabo				
/mo.ʧo.ˈne.ʧa/	cuatrísílabo				
/ʧe.ˈle.fo.no	cuatrísílabo				
/ʧa.dʒa.ˈri.nes/	cuatrísílabo				
/ka.ra.bi.ˈne.ro/	pentasílabo				
/re.fri.xe.ra.ˈdor/	pentasílabo				
/au.ʧo.ˈma.ʧi.ko/	pentasílabo				
/e.li.ˈkop.te.ro/	pentasílabo				

Comentarios