Revista Chilena de Fonoaudiología 20 (2021), 1-13
https://revfono.uchile.cl/
ISSN 0719-4692

Original Article

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

Sandra Vásquez Gajardo [a, *], Hernán León Valdés [b], Jaime Soto-Barba [a], Katia Sáez-Carrillo [c]

[a] *Department of Spanish Language, Faculty of Arts, University of Concepción, Chile*
[b] *Department of Communication Sciences and Disorders, Faculty of Medicine, University of Concepción, Chile*
[c] *Department of Statistics, Faculty of Physics and Mathematics, University of Concepción, Chile*

ABSTRACT

Currently, a greater repertoire of valid and reliable tests is needed to evaluate the phonetic-phonological performance in Chilean children. The objective of this study was to determine the validity and reliability of the Phonetic-Phonological Repetition Test (TREFF). Validation was performed using expert judgment and reliability was determined using a statistical analysis of temporal stability, as well as intra- and inter-rater agreement. For this process, the instrument was applied to a population of children belonging to the upper-middle sociocultural level, with a phonetic-phonological development expected for their age. The analysis of the responses was carried out by three judges, who determined whether the child succeeded in producing the evaluated elements. Regarding the validation, the instrument was modified considering the experts' suggestions. As to reliability, the results showed temporal stability and a high degree of intra-rater and inter-rate agreement. This allows us to conclude that the TREFF test meets the validity and reliability criteria.

Keywords:
Validity; Validation; Reliability; Phonetic-phonological evaluation

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

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

* *Corresponding Author:* Sandra Vásquez Gajardo
Email: flgasandravasquezgajardo@gmail.com

Received: 26-10-2020
Accepted: 17-11-2021
Published: 22-11-2021

https://doi.org/10.5354/0719-4692.2021.59250

INTRODUCTION

The ability to determine the phonetic-phonological behaviors in children's speech is something relevant to every language, especially when considering the developmental patterns in the acquisition of the phonological repertoire, the normative performance of these processes, and disturbances in the phonetic-phonological behaviors in some groups of children. However, this task is a challenging one, since it is necessary to count with an instrument that allows collecting speech samples to compare subjects' performances, that elicits different productions within the repertoire and phonotactics, and that can be applied to populations in a reasonable amount of time. All this, in order to observe children's speech effectively, according to specific criteria and variables. Because said instrument must collect reliable information, it should meet certain psychometric characteristics, that is, it must have validity and reliability. Therefore, the test should be submitted to a study that includes the phases of validation, and the confirmation of its reliability.

Validity is the degree to which an instrument truly measures what it seeks to measure, or achieves the objective for which it was designed (Arribas, 2004). There are various types of validity, such as expert judgment validity, content validity, criterion-related validity, and construct validity. Expert judgment validity consists of the selection of experts in the studied subject, who provide information, evidence, judgments, and evaluations regarding the instrument (Escobar Pérez & Cuervo Martínez, 2008). On the other hand, content validity indicates whether the instrument measures all or most of the components of the content of the measured variables (Hernández & Mendoza, 2018). Criterion-related validity measures whether the instrument works similarly to other instruments (Sánchez & Echeverry, 2004). Lastly, construct validity assesses whether the instrument reflects the theory behind what it measures.

Reliability is the degree to which an instrument measures accurately and with a minimum of error what it is intended to observe (Arribas, 2004). This means that its repeated application to the same subject produces the same results (Hernández et al., 2014).

In the field of child speech and language in Chile, one of the few valid and reliable instruments to evaluate phonological performance is TEPROSIF-R, which is based on the theory of Natural Phonology (Ingram, 1983). This test is standardized and, therefore, it has been widely used in Chile both in research and in clinical practice. However, since it does not consider the entire phonological repertoire of Chilean Spanish, nor all the possibilities of phonetic production depending on phonotactics and accentuation, it is necessary to count with alternatives that evaluate the verbal aspect in the broadest and most complete way possible, and that meet the requirements of validity and reliability.

Hamdan-Rosales et al. (2020) created the Phonetic-Phonological Repetition Test (TREFF) in the context of a study that relates the phonetic-phonological and lexical-semantic behavior in a given child population. This test is designed to obtain a corpus of speech productions in Chilean Spanish that allows a phonetic and/or phonological analysis, with the focus that each researcher requires. The direct word repetition method is used, since it allows applying the test in a short amount of time to a sufficient number of participants to obtain significant results. Picture naming is excluded because the participant may face lexical barriers if they cannot access or do not know the word, which would make it impossible to obtain the desired utterance. Moreover, there is not any set of pictures that represent all the phonetic contexts to be evaluated with TREFF. In turn, a spontaneous speech evaluation is not considered because, although it presents many advantages, it also has some disadvantages in research, such as the amount of time needed to transcribe and analyze samples. Other inconveniences are the possibility of having an unintelligible sample where the transcription cannot be done, and that the subject does not produce all the aspects of speech that need to be assessed (Vásquez, 2020).

The original version of TREFF has a list of 104 words that allows assessing the production of the phonetic repertoire in Chilean Spanish. It is subdivided into 7 sections: First the 17 phonemes of Chilean Spanish, which are evaluated through repetition of disyllabic and trisyllabic words, sorted by articulation point: anterior to posterior (bilabial /p-b-m/, labio-dental /f/, dental /d̪-t̪/, alveolar /s-n-l-ɾ-r/, palato-alveolar /d͡ʒ-ɲ-t͡ʃ/ and velar /k-g-x/), at the beginning, middle, and end of a word (bilabials, labiodental, and palate-alveolar are discarded in this case), and in coda position (dentals and palate-alveolar are discarded); second, semi-consonant and semi-vowel diphthongs; third, hiatuses; fourth, consonant clusters /pl-pɾ/, /bl-bɾ/, /t̪ɾ/, /d̪ɾ/, /kl-kɾ/, /gl-gɾ/, /fl-fɾ/ in initial (/tl/ is discarded) and inner positions (/fl/ is discarded); fifth, trisyllables; sixth, four-syllable words, and seventh, five-syllable words. In all cases, the use of words close to the lexicon of preschool and school children is prioritized, specifically for the age range of 3 to 7 years.

Besides the application of the TREFF test by Hamdan-Rosales et al. (2020), the instrument has been used in other research to obtain a language sample and analyze the phonetic-phonological performance of children. The information agrees with the data

provided by other research from a descriptive point of view, using the CLAFF guideline for analysis (Classification of phonetic-phonological adjustments in child language), as well as from a developmental perspective, observing the detected phonological repertoire (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). While the test achieves its intended purpose, it still needs to complete a standardization process that includes both validity and reliability.

Considering the aforementioned, this study aims to present the process of validation and reliability assessment of the TREFF test. In addition, certain adjustments are described that make the test more precise for the research of phonetic-phonological phenomena in the speech of Chilean children.

METHODOLOGY

The present study was organized mainly in two stages. The first stage consisted of the validation of TREFF by expert judgment, and the second of the determination of its reliability, analyzing intra- and inter-rater agreement and temporal stability. The second stage was carried out by applying the instrument to a child population, twice in a period of 7 to 14 days.

Study design

This study presents a descriptive, non-experimental, cross-sectional, quantitative design.

Validation process of the TREFF by expert judgment

Participants

Five expert judges were selected who had to meet the criteria of having postgraduate training in linguistics with an orientation towards communication sciences, as well as being familiar with the phonetic-phonological aspect of children's oral language. Three experts were chosen who had a doctorate and two who had a master's degree in the field of communication sciences, who analyzed whether each of the components of the test measured what it was intended to measure.

Materials and procedure

A spreadsheet titled "*Evaluación de Expertos TREFF*" [TREFF Expert Evaluation] was used (see Appendix 1), which presented the test in a format that allowed experts to evaluate each element. To better organize the information, the test was divided into dimensions and items. The dimensions referred to the evaluated element, for example, the point of articulation of the phoneme and its position within the word; and the items corresponded to the stimulus words that the child had to repeat.

In the beginning, each judge was informed about the objective of the test, the application and recording methods, and the role that they had in this task. They were instructed to analyze whether the items congruently evaluated each of the dimensions, scoring them using boxes numbered from 1 to 3. If the item clearly evaluated the expected dimension and did not require modifications, they should choose number 3; if the item evaluated the expected dimension but required modifications, number 2; if the item did not evaluate the expected dimension and required reformulation, number 1. In addition, each evaluator had a space to make observations for each item or dimension, and one to write general remarks about the instrument. An example of the template is shown in Figure 1.

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

Figure 1. Image of one of the sections of the form used by the expert judges.

Finally, the responses were analyzed and organized to determine the pertinent adjustments to be made, after which the test would be subjected to a reliability analysis.

Adjustments to the TREFF test

The experts' indicated that from a lexical perspective it would be adequate to prioritize the use of concrete nouns, common in children's vocabulary. From a phonetic-phonological perspective,

they suggested consistency in length and stress patterns, the use of words that ensure each sound is free of variations due to phonetic context, and to discard sounds in coda position, in which neutralizations that occur and generate archiphonemes. The latter considers the possibility that the sound is perceived differently by those who apply the TREFF test in any study.

Considering the judges' opinion the test was organized and divided into six modules, adjusting certain dimensions and items (see summary in Table 1). The dimension "Position within the word" is replaced with the dimension "Stress". That is, the element that was originally evaluated in initial, middle, inner coda, and final coda position within the word is now evaluated in stressed or unstressed syllables. The inner and final coda positions are evaluated in a different module (Vásquez, 2020).

Table 1. Adjustments to the TREFF test.

A. Adjustments to the distribution of content: considering the assessment by experts, the test is divided into six modules, making adjustments in certain dimensions and items.

Segments of the original version	Modules in the adjusted version
1. Evaluation of phoneme production according to its position within the word.	1. Evaluation of the production of phonemes according to stress.
2. Evaluation of semi-vowel and semi-consonant diphthongs.	2. Evaluation of the production of phonemes in inner and final coda position.
3. Evaluation of the production of Hiatuses.	3. Evaluation of the production of semi-vowel and semi-consonant diphthongs.
4. Evaluation of the production of complex onsets according to syllable position.	4. Evaluation of the production of hiatuses.
5. Evaluation of the production of words with different lengths: trisyllables, four-syllable words and pentasyllables.	5. Evaluation of the production of complex onsets according to stress.
	6. Evaluation of the production of words with different lengths: trisyllables, four-syllable words and pentasyllables.

B. Adjustments to the dimensions: "position within the word" is changed to "stress"

Original dimension	Adjusted dimension
Evaluation of the element according to "position within the word": the element is evaluated in initial, medial, inner coda, and final coda positions.	Evaluation of the element according to "stress": the element is evaluated in stressed and unstressed syllables.

C. Adjustments to the items

Evaluated element	Original item	Adjusted item	Justification
/f/	/ka.'fe/	/'xe.fe/	Item adjusted due to a change in the dimension: Because the dimension was changed from "position within the word" to "stress", items that did not agree with the new dimension were adjusted.
/t/	/pa.'t̪e/	/'pa.t̪a/	
/g/	/'xu.go/	/'ma.go/	
/gl/	/i.'glu/	/'re.gla/	
/t̪r/	/a.'t̪ras/	/'pos.t̪re/	
/ɲ/	/ɲaɲ.'d̪u/	/'ɲa.t̪o/	
/x/	/xo.'se/	/'xu.go/	
/d̪ɾ/	/d̪ra.'gon/	/'d̪ɾo.ga/	
/iu/	/siu.'d̪ad̪/	/'biu.d̪a/	Adjustments to the item to achieve consistency in stress: the experts suggested that in the case of diphthongs the evaluated element should be in the stressed syllable. Consequently, the items that did not meet this requirement were adjusted.
/ua/	/'a.gua/	/'gua.t̪a/	
/f/	/'fo.ka/	/'fa.ɾo/	Adjustments to the item due to possible changes in articulation: adjustments were made to the items to avoid possible changes in articulation due to their phonetic context (such as velarization or dentalization of the evaluated element)
/n/	/'kaɲ.t̪o/	/'pan.sa/	
/l/	/'pal.t̪a/	/'bol.sa/	

/kl/	/'an.kla/	/'ʈe.kla/	
/fl/	/'flan/	/'fla.ko/	Adjustments to the item to achieve consistency in length: the experts suggested a consistent use of disyllables throughout the test. It is important to mention that the evaluation of different word lengths (trisyllables, polysyllables, and pentasyllabic words) is carried out in a different module.
/ʈr/	/'ʈren/	/'ʈro.pa/	
/ei/	/'rei/	/'rei.na/	
/oi/	/'boi/	/'boi.na/	
/ie/	/'pie/	/'sie.lo/	
Four-syllable word	/ʈe.'le.fo.no/	/ʈa.d͡ʒa.'ri.nes/	Adjustments to the item to achieve consistency in the metric organization: The same metric foot is used for every stimulus in the word length module, in order to standardize the accentuation structure.
Pentasyllable	/re.fri.xe.ɾa.'ɖoɾ/	/ka.pe.ru.'si.ʈa/	
Pentasyllable	/au.ʈo.'ma.ʈi.ko/	/es.ka.ɾa.'ba.xo/	
Pentasyllable	/e.li.'kop.ʈe.ro/	/na.ʈu.ɾa.'le.sa/	

Process for determining the reliability of the TREFF

Participants

The sample, selected by convenience, consisted of 31 participants between 4.0 and 5.6 years old, in transitional and prekindergarten levels, belonging to an upper-middle sociocultural level.

For the selection process, it was determined that the participants should be children in a moderately developed stage of phonetic-phonological development, therefore the study focused on ages between 4.0 and 5.6 years. In addition, their phonetic-phonological performance had to be in line with what is expected for their age, which required ruling out speech and language disturbances. It is worth mentioning that the gender variable was not considered for this research.

The children were first assessed using an evaluation guideline specially created for the process, to confirm they did not have any anatomical or functional disturbance in their speech organs. In addition, the TEPROSIF-R was applied to rule out phonological disorders. All the evaluations were audio-recorded using a TASCAM DR-40 to analyze performance. Furthermore, based on the information provided by educators, children diagnosed with developmental language disorder or autism spectrum disorders, as well as other conditions that could affect their phonetic-phonological performance, were excluded. Lastly, all children were required to have an informed consent signed by their parents, and they had to give verbal consent before participating in the evaluations (see Appendix 2).

Procedures

The instrument was applied to each of the selected children using the same modality. Each child sat in front of the evaluator, who covered their lips while uttering the words to avoid visual support. The evaluator began by giving the instructions, emphasizing that the participant should listen carefully and repeat the words verbally. They then proceeded to provide example stimuli to confirm that the child understood the task. Once the activity was understood the evaluation began, following the given order of the modules and requesting the repetition of stimuli from left to right. It is worth mentioning that to keep the child's attention the test was carried out in a playful context, telling them that they would go through stages until they reached the goal, and giving them positive reinforcement each time they finished a module. If fatigue was observed the child had a period of 10 to 20 seconds of rest, to then continue with the next module.

The TREFF test was applied a first time (application N°1) and two weeks later a second time to the same sample, under the same conditions (application N° 2). All applications were recorded in audio with a TASCAM DR-40 recorder.

Three evaluators with postgraduate studies in communication sciences listened to the TREFF recordings on two different occasions (first listening and second listening). Once the listening judgment was finished, a phonological transcription of each word produced by the children was carried out.

After the responses were transcribed and analyzed, it was determined whether each item was achieved or not. An item was considered achieved if the child produced the phonological element according to what was expected in the evaluated dimension. On the other hand, an item was considered not achieved if the child was not capable of producing the phonological element. This criterion was applied to all the items in modules one to five, which evaluate the production of consonant phonemes in stressed and unstressed syllables, in coda position, diphthongs, hiatuses, and complex onsets. However, module six was considered as achieved if the expected length was

produced along with maintaining the syllabic sequence and the phonemes that constitute the syllables. In contrast, items where length was not maintained or where the syllabic or phonemic sequence was unsuccessfully produced were considered as not achieved. In order to carry out the statistical analysis, performance was coded numerically, assigning a score of 1 to the achieved items and 0 to the ones not achieved.

The results of the entire process were first subjected to an intra-rater agreement analysis, which measured the consistency of each evaluator's assessment between the first and second listening of each application of the test. An inter-rater agreement analysis was then performed, determining the degree of variability between the perceptions of different evaluators for the same recording. Finally, the results were subjected to a temporal stability analysis, where the expected result was that there were no significant differences between the responses of the first and second application of the test.

The three evaluators, described as Judge 1, Judge 2, and Judge 3, listened to the application Nº 1 of the instrument twice, with an interval of 7 to 14 days between each time. The responses were immediately transcribed into the TREFF record sheets, which were subsequently analyzed. The evaluators then listened to the recordings of the second application of the instrument, repeating the same conditions. This was done to determine the agreement between the records of each evaluator, that is if the responses of the first and second listening of one application were the same or there was some degree of variation. In addition, the information collected allowed studying the level of agreement between evaluators, by comparing their results. Finally, a study of the temporal stability of the instrument was carried out, in which the results of Application Nº 1 were compared with the results of Application Nº 2, considering the global results of the test.

To obtain statistical results, the level of agreement and the temporal stability were analyzed using the Intraclass Correlation Coefficient (ICC) and Lin's Concordance Correlation Coefficient (CCC). Two coefficients were used in order to contrast them and verify the degree of reliability. Reliability coefficients can range between zero and one, zero indicating null reliability and one meaning maximum reliability (Hernández & Mendoza, 2018).

RESULTS

In order to determine the intra-rater agreement, an analysis was first carried out of the records of each evaluator for the first and second listening of application Nº 1. For this, the global results of the test were used. The correlation results obtained from the ICC and CCC coefficients show a score higher than 0.9 for the three judges, which reflects a high degree of agreement. This indicates that there is minimal variability between answers in the evaluations of the first and second listening of application Nº 1, which shows the level of agreement of the instrument (see Table 2).

Table 2. Analysis of intra-rater agreement between the first and second listening of application N°1.

		Limits			Limits	
	Intraclass Correlation Coefficient	Inferior	Superior	Lin's Coefficient	Inferior	Superior
Judge 1	.998	.996	.999	.996	.991	.998
Judge 2	.984	.967	.992	.967	.935	.984
Judge 3	.968	.901	.987	.936	.874	.968

Subsequently, an analysis of intra-rater agreement was carried out to measure the agreement of the responses between the first and second listening of application Nº 2, considering the global results of the instrument. The results obtained for Judge 1, Judge 2, and Judge 3 are higher than 0.9, which indicates a high degree of agreement (see Table 3).

Table 3. Analysis of intra-rater agreement between the first and second listening of application N°2.

		Limits			Limits	
	Intraclass Correlation Coefficient	Inferior	Superior	Lin's Coefficient	Inferior	Superior
Judge 1	.994	.987	.997	.987	.974	.994
Judge 2	.99	.967	.996	.981	.962	.99
Judge 3	.959	.917	.98	.92	.841	.96

To determine the inter-rater agreement, the total test results were used. First, the data were compared between judges for application Nº 1, contrasting the evaluations of Judge 1 and Judge 2; then Judge 1 and Judge 3; and, finally, Judge 2 and Judge 3. The statistical methods were the same as those used to analyze the intra-rater agreement. The results obtained for inter-rater agreement in application No. 1 show a high degree of agreement (see Table 4).

Table 4. Analysis of the inter-rater agreement for the first listening of application N°1.

		Limits			Limits	
	Intraclass Correlation Coefficient	Inferior	Superior	Lin's Coefficient	Inferior	Superior
Judges	.975	.954	.988	---	---	---
Judge 1- Judge 2	.978	.954	.989	.955	.91	.978
Judge 1- Judge 3	.959	.891	.983	.92	.848	.959
Judge 2- Judge 3	.952	.899	.977	.906	,817	.952

The inter-rater agreement for application Nº 2 also considered the total results of the test. In ICC and CCC, the scores for the comparisons between Judge 1 and Judge 2, Judge 1 and Judge 3, and Judge 2 and Judge 3 were above 0.9, indicating a high degree of agreement (see Table 5).

Table 5. Analysis of the inter-rater agreement for the first listening of application N°2.

		Limits			Limits	
	Intraclass Correlation Coefficient	Inferior	Superior	Lin's Coefficient	Inferior	Superior
Judges	.987	.975	.994	---	---	---
Judge 1- Judge 2	.987	.97	.994	.974	.948	.987
Judge 1- Judge 3	.976	.928	.99	.953	.912	.975
Judge 2- Judge 3	.979	.957	.99	.958	.917	.979

For the analysis of temporal stability, the results of application Nº 1 and the results of application Nº 2 were compared considering each judge, using the global results of the test. The results were over 0.9 for Judge 1, Judge 2, and Judge 3. This indicates that there were no significant changes between responses in the first and second applications, considering the 7-14 interval between them. The instrument proved to be stable over time, which denotes reliability (see Table 6).

Table 6. Analysis of temporal stability between application Nº 1 and Nº 2.

		Limits			Limits	
	Intraclass Correlation Coefficient	Inferior	Superior	Lin's Coefficient	Inferior	Superior
Judge 1	.978	.956	.99	.956	.913	.978
Judge 2	.98	.958	.991	.961	.921	.981
Judge 3	.956	.909	.979	.913	.828	.957

DISCUSSION

The present study was able to reach its objectives, which are to achieve the validation of the TREFF and to determine its reliability. Validity was achieved through expert judgment, in which adjustments were made following the judges' suggestions, to ensure the instrument measures what it is intended to measure. Regarding reliability, the test proved to be stable over time, with a high degree of intra-rater and inter-rater agreement.

By providing a new valid and reliable instrument this work makes a significant contribution, mainly to research in the fields of child linguistics and speech-language pathology. Observing the phonetic-phonological performance of Chilean children using a reliable instrument will allow deepening the analysis of dialectal linguistic phenomena, thus identifying typical behaviors or phonetic-phonological disturbances.

As a projection, it is expected that TREFF is used to help formal processes such as obtaining profiles of typical phonetic-phonological performance in different age ranges and socioeconomic status, and thus counting with normative values. It should be borne in mind that the manifestations of children's speech are dynamic and subject to change for social and cultural reasons, hence it is necessary to constantly update the data. Similarly, the test could be adapted to study and describe the performance of other groups in Chile, such as ethnic minorities and rural populations, among others. An additional projection is that by counting with a description of the typical performance of children, this test could be applied to atypical child groups in different age ranges. It is important to mention that, because this instrument uses direct word repetition to collect its sample, it should be complemented in some types of research with spontaneous or semi-spontaneous speech samples. This is to observe linguistic performance in phonetic-phonological contexts different from those that the TREFF evaluates.

In order to adapt this instrument to the speech-language therapy clinical practice, the possibility of determining its specificity, sensitivity, and predictive validity is projected. Finally, with the information obtained from the aforementioned processes, certain modules of the test could be applied to established age ranges, making the test an adaptable and flexible tool used in diverse observation and research contexts, in the field of children's speech and language.

FUNDING

Associative VRID Project, code 218.083.036-1.0.

REFERENCES

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

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., & Riffo, 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

APPENDIX 1. Part 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 **no evalúa** la dimensión esperada, debe reformularse.	El ítem **evalúa** la dimensión esperada, sin embargo requiere **modificaciones.**	El ítem **evalúa claramente** 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.

APPENDIX 1. Part 2.

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

APPENDIX 1. Part 3.

	Dimensiones		Evaluación			Observaciones
Ítem	Ataque complejo	Posición	1	2	3	
/'plu.ma/	/pl/	Inicial				
/so'pla/		Interior				
/'blu.sa/	/bl/	Inicial				
/'ka.ble/		Interior				
/flan/	/fl/	Inicial				
/'in.fla/		Interior				
	/tl/					
/a.'t̪le.t̪a/		Interior				
/'kla.bo/	/kl/	Inicial				
/'an.kla/		Interior				
/'glo.bo/	/gl/	Inicial				
/i.'glu/		Interior				
/'pre.so/	/pr/	Inicial				
/'kom.pra/		Interior				
/'bra.so/	/br/	Inicial				
/'po. bre/		Interior				
/'fru.t̪a/	/fr/	Inicial				
/'ko.fre/		Interior				
/t̪ren/	/tr/	Inicial				
/a.'t̪ras/		Interior				
/d̪ra.'gon/	/dr/	Inicial				
/'la.d̪ra/		Interior				
/'kre.ma/	/kr/	Inicial				
/'mi.kro/		Interior				
/'gri.t̪o/	/gr/	Inicial				
/'t̪i.gre/		Interior				

	Dimensión		Evaluación			Observaciones
Ítem	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/					

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

	Dimensión		Evaluación			Observaciones
Ítem	Hiato		1	2	3	
/'t̪i.o/	/'i.o/					
/ba.'ul/	/a.'u/					
/go.'t̪e. a/	/'e.a/					
/'pu.a/	/'u.a/					
/re.'i/	/e.'i/					
/a.t̪a.'ud̪/	/a.'u/					

APPENDIX 1. Part 4.

Ítem	Dimensión Metría		Evaluación 1	 2	 3	Observaciones
/t̪o.ˈma.t̪e/	trisílabo					
/pa.ˈle.t̪a/	trisílabo					
/ko.ˈne.xo/	trisílabo					
/xi.ˈra.fa/	trisílabo					
/ma.ri.ˈpo.sa/	cuatrisílabo					
/mo.t̪o.ˈne.t̪a/	cuatrisílabo					
/t̪e.ˈle.fo.no	cuatrisílabo					
/t̪a.d͡ʒa.ˈri.nes/	cuatrisílabo					
/ka.ra.bi.ˈne.ro/	pentasílabo					
/re.fri.xe.ra.ˈdor/	pentasílabo					
/au.t̪o.ˈma.t̪i.ko/	pentasílabo					
/e.li.ˈkop.te.ro/	pentasílabo					

Comentarios