Validação de itens para uma escala de avaliação da inteligibilidade de fala****

Validation of items for a speech intelligibility assessment scale

Ana Paula Ramos de Souza*
Jair Mendes Marques**
Lisiane Collares Scott***

*Fonoaudióloga. Doutora em Linguística Aplicada pela Pontifícia Universidade Católica do Rio Grande do Sul (PUC - RS). Professor-Adjunto do Departamento de Otorrino-Fonoaudiologia da Universidade Federal de Santa Maria. Endereço para correspondência: R. Raposo Tavares, 134 - Apto. 401 - Santa Maria - RS CEP 97015-560. (ramos1964@uol.com.br).

**Matemático e Engenheiro Químico. Doutor em Ciências Geodésicas pela Universidade Federal do Paraná. Docente de Estatística do Curso de Mestrado em Distúrbios da Comunicação da Unversidade Tuiuti do Paraná (UTP).

***Fonoaudióloga. Mestre em Distúrbios da Comunicação pela UTP.

****Trabalho Realizado na UTP.

Abstract

Background: speech intelligibility evaluation. Aim: to develop and validate items for a speech intelligibility assessment scale based on the speech of individuals with phonological disorder (PD) who present common phonological repair strategies (RS) in the Brazilian Portuguese language (BP), through the testing of its effectiveness in classifying the speech of these subjects. Also to observe the speech intelligibility generated by the use of RS and possible interference of factors such as gender, age, literacy and contact of the judges with children. Method: spontaneous narratives of five children with typical PD and one control subject were recorded in a compact disc (CD) and presented to 103 judges, with normal hearing abilities, with ages ranging from 18 to 39 years and with elementary, high-school and undergraduate backgrounds. Results: results indicate the statistical validity of the scale and that the type of RS if fundamental for speech intelligibility. There was no statistical interference of the investigated variables of gender, age, literacy and contact with children in the performed judgments. Conclusion: the scale items were validated and demonstrated efficacy in the assessment of speech intelligibility of the studied cases.

Key Words: Speech; Speech Intelligibility; Speech Disorder.

Resumo

Tema: avaliação da inteligibilidade de fala. Objetivo: esta pesquisa objetivou desenvolver e validar itens para uma escala de inteligibilidade de fala a partir da fala de sujeitos com distúrbios fonológicos (DF) que apresentassem estratégias de reparo (ER) frequentes em Português Brasileiro (PB), através da testagem de sua eficácia para classificar a fala desses sujeitos. Também observou a inteligibilidade de fala gerada pelo uso das distintas ER e a possível interferência de variáveis como sexo, idade, escolaridade e contato com crianças entre os julgadores das amostras de fala. Método: assim, narrativas espontâneas de cinco crianças cujas falas representassem casos clínicos típicos e a de um sujeito controle foram apresentadas em *compact disc* (CD) a 103 juízes adultos, com habilidades auditivas normais, entre 18 e 39 anos de idade, com escolaridade fundamental, média e superior. Resultados: os resultados demonstraram a validade estatística dos itens da escala e que o tipo de ER é fundamental no processo de inteligibilidade. Não houve interferência estatística das variáveis sexo, idade, escolaridade ou contato com crianças nos julgamentos realizados. Conclusão: os itens da escala foram validados e demonstraram eficácia na avaliação da inteligibilidade de fala dos casos estudados.

Palavras-Chave: Fala; Inteligibilidade; Transtorno de Fala.

Introduction

Communication disorders widely affect children and adolescents and may represent a prevalence of 29.1% of speech disorders[1]. Some studies include familial profile[2], genetic aspects[3], impact on quality of life and on relationship with family members, as well as participation in the therapeutic process4,5,6,7,8,9. Other studies analyze several clinical characteristics such as neuropsychologica10,11,12,13,14,15 and audiological16,17 aspects but few investigate the consequence of speech disorder on speech intelligibility given its difficult measument18. Among the variables that affect speech intelligibility are the type and frequency of phonemic errors, the similarity or not of the segment produced with the target phoneme in the speech samples. Regarding the raters who evaluate the speech samples, their experience in analyzing children's speech (typical or atypical), their knowledge about the speaker and the production context18 have been reported to affect speech intelligibility ratings.

Several studies on the Brazilian Portuguese (BP) language have addressed the issue through the analysis of speech samples carried out by Speech-Language Pathologist19-21 or by layman22-23 raters. It is common, in all such studies, the analysis of the phonological processes impact on speech intelligibility as well as the evaluation of scales measuring intelligibility24. In the current study, the validation of items for the speech intelligibility scale carried out on the study by Collares25 is described.

From such theoretical considerations, the main purposes of the study were to develop and validate items for a speech intelligibility scale and to test the effectiveness of the items of the scale in classifying the speech of these subjects. In relation to the raters, the possible interference of variables such as gender, age, education and contact with children from the sample was analyzed.

Method

This study is part of the research project entitled: Phonological Disorders: characterization, assessment and therapy, under responsibility of Dr. Ana Paula Ramos de Souza, approved by the Institutional Ethics Committee - protocol number 2002-107.

Inclusion and exclusion criteria of subjects whose speeches were rated

Six subjects participated in the study: one child with typical phonological acquisition and five children with Phonological Disorder (PD) aged between four and nine years of age. All subjects were monolingual native speakers of BP. The subjects with PD were patients of the school clinic of Porto Alegre and the control-subject was from the city of Porto Alegre. The selection of subjects with PD was made based on phonological assessment and selection of common ER in phonological acquisition of BP based on studies on the prevalence of ER in this population 22,23. The control-subject was a convenience sample, i.e. a child known to one of the authors.

Parents of subjects authorized their participations in the study and signed a consent form allowing data collection with guaranteed confidentiality and volunteerism right.

The method of selection of subjects was based on the results of the childhood phonological assessment6, i.e.; speech samples of the children were elicited by pictures and were evocatively recorded, phonetically transcribed, and were analyzed for contrast target in PB and repair strategies presented. The selected subjects were distributed regarding the ER:

. subject 1 - Devoicing: This subject exclusively presented ER in 100% of the production of plosives, fricatives and sound affricates as voiceless.
. subject 2 - Velar Plosives Anteriorization: ER in which / k / and / g / were substituted with / t / e / d / respectively.
. subject 3 - Stopping, Deletion and Semi-vocalization of Initial, Final, and Intervocalic Liquids, and Reduction of Complex Onset:

Stopping was observed in 65% of the production of fricatives and affricates. The remaining 35% were distributed among deletions, or affrication or nasalization of fricatives. The subject deleted the liquids in 50% of the possibilities, presented semi-vocalization of liquids in 36%, and 14% of the time the subject replaced the non-lateral liquid by the lateral one. The complex onsets were reduced in 100% of the possibilities. The use of metathesis was occasionally observed.

. subject 4 - Completed Phonological Acquisition, without ER.
. subject 5 - Palatal Fricative Anteriorization - the palatal fricatives were anteriorized in 100% of possibilities for the alveolar / s / and / z /.
. subject 6 - Deletion of non-lateral Alveolar Liquid in the Medial Onset and Coda, and Reduction of the Complex Onset: These processes occurred in 100% of the possibilities.

From such sample selection, procedures were used to collect the speech sample to be analyzed by the raters.

Procedures and Materials for the Organization of Speech Samples to be Assessed

The speech samples were collected after the selection of participants according to the phonological assessment and after effective authorization by their parents or guardians. To collect the speech of the six subjects, four logical sequences with activities of daily life of a boy (sleeping, waking, eating, going to school) and a girl (playing chase butterflies or playing a bone to the dog) were used. When necessary, the researcher gave an example of reporting in order to present the children with a model. Questions that helped the subjects to organize the report were also allowed. This procedure was accomplished on a speech sample pre-recording training from each subject. The recording was conducted when the subject reported being able to produce a continuous report for recording.

The recordings were made with a Sony MiniDisc model MZ-R70, a Sony MD Digital Audio Recorder of 74 minutes and a Leson MP68 microphone type electret cardioid, with key, cable and low impedance.

Preparation of Speech Samples for Assessment

After the reports were recorded on Minidisc, the speech sample from each subject was listened and transcribed in order to determine whether there was or not a percentage of consonants correct necessary for the balancing of samples among the subjects regarding the occurrence of ER - i.e. in order to balance the same number of words with ER for all subjects. The percentage of consonants correct (PCC) was not used as a parameter because subject 3 - who presented more ERs - would naturally have a smaller percentage of PCC. The percentage was balanced and the message was maintained when comparing subjects 1, 2, 5 and 6.

The speech samples from six subjects were saved via Sound Forge 6.0 software and a Toshiba Satellite notebook model 1805-S207. The samples were edited to remove background noise and to unite speech in order to compose equal hearing durations of 40 seconds each with a 20 seconds interval between them. After this editing, data were transferred to a CD.

The reports of the subjects with PD demanded the use of one additional logical sequence in order to produce connected speech. This was possibly due to their poor vocabulary and creativity. In contrast, the report of the control subject composed the required number of words from the report of only one logical sequence. All speech samples contained about 85 words, whilst the need to maintain the meaning.

The CD was recorded with two sequences of presentation for the test and retest in order to not induce similar responses by the raters. In the first month, the raters heard the sequence a) S1, S2, S3, S4, S5, S6. For the retest (three months later) the sequence b) S3, S6, S5, S1, S4 and S2 was presented to the raters.

Procedures for Construction of Items for the Intelligibility Scale

The procedures for construction of items for the intelligibility scale were: consulting the literature on the Likert scales and selection of terms that are commonly observable or accessible in ordinary dialogue concerning the intelligibility of speech. Therefore, the word "understand" was used as definition of the scale item and the term understandable was preferred to intelligible on the items of the scale.

Based on such selections, the Likert scale consisted of the five items displayed in Figure 1:

FIGURE 1. Items of the Intelligibility Scale.

Scale Item	Definition for rater reading
Non-understandable	I do not understand any of the words or the message he/she is saying
Hardly understandable	I understand a few words with difficulty, but not the meaning of the message.
Understandable	I understand some words which allows the understanding of great part of the message
Very understandable	It is possible to understand the majority of the words and message
Totally understandable	It is possible to understand all words and the entire message.

Sample and inclusion and exclusion criteria for selection of the sample of raters

Considering the aspects pointed out on the literature as influential regarding the raters, a questionnaire with identification information, occupation, number of children and their ages was applied to the raters of the current study.

The sample of raters was comprised of adults between 18 and 39 years with normal hearing and intellect; elementary school, secondary school and university students. The average age was 26.78 with a standard deviation of 6.44.

Therefore, the following inclusion criteria were considered: no presence of presbycusis, no suspected hearing loss or cognitive or mental disorder, and not being a professional who directly works with children - individually or in groups. These were observed on informal interview conducted by the researcher as well as through information provided by their teachers and professors. The volunteers who reported hearing difficulties or demonstrated difficulties during the initial interview were excluded from the sample. Those who had evident cognitive or psychological difficulties were also excluded.

From such profile, 52 men (50.5%) and 51 women (49.5%) were selected. The sample totalized 103 raters, native speakers of BP who were not professionals with direct contact with children or deviant speech such as Pedagogy and Speech-Language Pathology. About 40% of the raters were in elementary school, 34% in secondary school and 26% were university students. In relation to marital status, 40% were married, 49% single, 6% divorced, and 5% reported other status. Almost half the sample had no children (48%). The remaining had only one (28%), two (17%), three (6%) or four (1%) children. Of these, about 91% had children between three and nine years old, with the greatest concentration between three and six years (60%) - a period during wich the phonological acquisition is being completed.

Collection procedures with the raters

Sony Discman model D-172CK, Sony headphones and a CD were used to present the recordings of the first - test - and second - retest - sequences to the raters.

Initially, the raters were invited at their classrooms to participate in the study. Those who agreed to participate signed a consent form. After such agreement, they responded to the initial questionnaire and then began to listen to the speech samples, subject by subject. The raters assigned a value from 1 to 5 on the Likert scale for each sample. The researcher marked the responses. It was possible to read and reread the items of the scale before and after listening to each speech sample as the items remained visually available during all rating process.

Testing was conducted during the first month of collection and retest was conducted three months later.

In the test condition, 125 raters were present. For the retest, there was a loss of 22 raters and, therefore, 103 raters completed the analyses on the retest condition. The instructions provided were the following: "You will hear six children and after hearing each child you will assign one value as you agree on the following scale". The five previously mentioned scale items - which remained visible at each listening on a card - were then read out loud to the rater.

Methods of data analysis

The quantitative tests used were the Chi-square test, which examines the dependence among the questions on the scale, and the ?(alpha) Cronbach's coefficient, which accesses the significance of the retest. Such tests were applied to the responses of raters in order to verify the agreement between the 103 responses obtained in the test condition and the 103 responses obtained in the retest condition. These tests were also applied to analyze the interference or not of variables related to the raters such as gender, age, and contact with small children on their analyses.

As the purpose of this study was to validate the items and not to generalize conclusions regarding the type of ER verified in the speech of children, the analysis of the responses considering each type of ERs was qualitatively performed. The sample of six subjects for composition of speech samples do not reflect a statistical reality of the population from the city where the study was conducted.

It was possible to qualitatively interpret the phonological data by comparing it with data from the literature due to the high classification agreement between the test and retest conditions. However, generalizations about the classification of each ER require a larger number of subjects assessed by the scale. Furthermore, the sample size should be calculated from the exact prevalence of ER and the prevalence of PD in the population of

the city where the study is conducted. Such generalization will only be possible from the validation of the items of the scale and should be accomplished as a continuity of the current study.

Results

Table 1 presents the results of the statistical test applied to the classification agreement of scale items between the test and retest conditions for each subject.

The first analytical procedure was calculated through the Spearman correlation between test and retest of speech of each subject. Such procedure revealed a high correlation between the two conditions as, for an $? = 0.05$ (5%), $p < ?$ was verified. Both on test and retest conditions, the ? value was higher than 0.7, which demonstrated that the retest nearly doubles the test. The Chi-square test showed high significance ($p < ?$) for all items tested, demonstrating the validity of the scale items in the assessment of the speech samples. Subject four was not included in these analyses as the subject did not show any variation between test and retest conditions. These results are summarized in Table 1.

The responses attributed to each subject on the test and retest conditions are distributed in Table 2, which shows almost complete agreement between test and retest conditions. Importantly, S3 - the subject with the highest number of ERs - was the only one to receive the classification 1 of the scale (Non-understandable) from 79 of the 103 raters and hardly understandable from the other 24 raters. This subject presents ERs covering stopping and liquid simplifications. Already the subjects 2, 5 and 6 received the classification 3 - understandable - as their baseline. The subject 1 received the intelligibility attribution between 3 and 4 on the Likert table, corresponding to understandable and very understandable. Therefore, all subjects with PD, except subject 3, were classified in a range which speech is understandable or very understandable. However, none of these subjects were classified as completely understandable, as the subject who has completed phonological acquisition.

Spearman's correlation revealed no statistical significance regarding the association between gender, education and contact with children and the intelligibility classification provided by the raters when considering $? = 0.05$ (5%).

TABLE 1. Analyses of validity of each item for all subjects according to test/retest reliability

Item –subject	Sperman's Correlation T	Sperman's Correlation R	P	Chi-Square Test	P
S1	21,85300	0,908530	0,0000	80,00	0,0001*
S2	22,05635	0,909989	0,0000	86,20	0,0001*
S3	18,60072	0,879797	0,0000	76,80	0,0001*
S4	-	-	-	200,00	0,0001*
S5	28,65561	0,943649	0,0000	84,80	0,0001*
S6	22,31028	0,911765	0,0000	52,80	0,0001*
Total of valid n	Mean Test= 14,49	Mean Retest=14,46			
103	SD= 1,93	SD=1,90			
	Cronbach's α =0,7399	Cronbach's α=0,7275			

TABLE 2. Classifications attributed on the scale for each subject on test and retest conditions

SCALE	QUESTIONS											
	S1		S2		S3		S4		S5		S6	
	T	R	T	R	T	R	T	R	T	R	T	R
Non-understandable	-	-	-	-	79	79	-	-	-	-	-	-
Hardly understandable	02	02	09	09	24	24	-	-	03	03	11	11
Understandable	51	51	64	62	-	-	-	-	63	62	57	56
Very understandable	50	50	30	32	-	-	-	-	37	38	35	36
Totally understandable	-	-	-	-	-	-	103	103	-	-	-	-

T = Test; R = Retest – highet response concentrations for each subject are highlighetd in gray

Discussion

As shown in Table 1, the proposed items for the classification of speech intelligibility were valid considering the high significance observed on the comparsion between the test and retest conditions .

The validation of the scale however, should occur from the application of the scale before a larger number of raters and with a larger number of speech samples of individuals with PD and other speech disorders.

The results displayed in Table 2 indicate the following order of intelligibility regarding the analyzed ERs: intelligibility with stopping and liquid simplification < anteriorization of velar = anteriorization of palatal fricative = liquid simplification < devoicing. These data only indicates what needs to be confirmed with a larger sample of subjects with PD. The classification values of each ER also need statistical confirmation on samples calculated from the prevalence of each ER in this population. Furthermore, the prevalence of each ER needs to be defined from the prevalence of PD in the city where the study is conducted.

The present data suggest that the subject with the largest number of processes (subject 3) was the subject with poorer intelligibility. Although we have not analyzed the percentage of consonants correct (PCC) as a variable of analysis, this finding is in agreement with other studies20,21 that suggest a direct relationship between severity measured by PCC and intelligibility.

Regarding the speech profile, subjects 5 and 6 fit the profile described as phonological delay23. However, there was no confirmation that such profile would cause less impact on intelligibility due to the ERs being evolutionarily deleyed22. In contrast, subjects 1 and 2, who presented ERs that are usually eliminated earlier, were assigned intelligibility classifications equal to (S2) or higher than (S1) subjects 5 and 6. The analysis shows that only the presence of a higher number of ERs has materialized into a major intelligibility alteration. Such finding adds to the ER effects that some authors have reported as impacting intelligibility 20, 21.

Regarding the variables gender, education and contact with children, it is clear that they have no interference on the classifications of speech samples. This fact demonstrates that the rater, except the rater who lives with a child with PD, may not be interfered from such variables in a speech listening situation without context influence18. Thus, one can affirm that the linguistic variables identified by the authors were more relevant in this study.

The clinical implications of this study are the resizing of the intelligibility concept in terms of both assessment and clinical intervention.

Regarding the assessment, it was reported by Ramos et al.[22] that subject 3 - with stoping and liquid alteration - is a subject whose speech is not so frequently observed in the profile of subjects with PD (maximum 30% of group). Therefore, when attributing the unintelligible label to the group of subjects with PD in general, one can incur a major error that does not reflect the reality experienced by the individual.

The data also suggest the need for the intelligibility scale to compose the evaluation of individuals with PD, particularly to assess familial reactions to the speech of the child. A more understandable speech to the family may indicate less disruption in communication in order to motivate the individual with PD to change their speech patterns.

The implication in terms of clinical intervention appears to be under the motivation for change in the speech itself. The present study suggests that 70% of subjects with PD - considering the frequency of ERs in larger groups of PD [22] - should not feel their speech as little intelligible to the interlocutor. This is justified by the finding that subjects with more frequent ERs had their speech samples classified as understandable and very understandable. Therefore, the fact that the interlocutor is possibly not signaling

communication disruptions may be the resistance nature of some individuals in adhering the Phonological therapy.

Conclusion

In relation to the purpose of developing and validating the items of the intelligibility scale, it is considered that the results were positive as the items created were effective in assessing the intelligibility of the analyzed speech samples. The analyzed items were compatible both in test and in the retest conditions.

The variables gender, education and frequent contact with children were not significant in the analysis of speech samples of the group studied.

Considering the speech intelligibility of individuals with PD, it was observed that the scale presented can assess speech intelligibility and differentiate ERs.

References

1. Papp AC, Wertzner HF. O aspecto familial e o transtorno fonológico. Pró-Fono. 2006;18(20):151-60.

2. Broomfield J, Dodd B. Children with speech and language disability: caseload characteristics. Int J Lang Commun Disord. 2004;39(3):303-24.

3. Stein CM, Millard C, Kluge A, Miscimarra LE, Cartier KC, Freebairn LA, Hansene AJ, Shriberg LD, Taylor HG, Lewis BA, Iyengar SK. Speech sound disorder influenced by a locus in 15q14 region.

4. McLeod S, McCormack J. Application of the ICF and ICF - Children and youth in children with speech impaiment. Semin Speech Lang. 2007;28(4):254-64.

5. Markham C, Dean T. Parent´sand professional´s preceptions of quality of life in children with speech and language difficulty. Int J Lang Commun Disord. 2006; 41(2):189-212.

6. Rudolph M, Kummer P, Eysholdt U, Rosanowski F. Quality of life in mothers of speech impaired children. Logoped Phoniatr Vocol. 2005;30(1):3-8.

7. Wink M, Rosanowski F, Hoppe U, Eysholdt U, Grässel E. Subjective burden in mothers of speech-impaired children. Folia Phoniatr Logop. 2007;59(5):268-272.

8. Rudolph M, Rosanowski F, Eysholdt U, Kummer P. Anxiety and depression in mothers of speech impaired children. Int J Pediatr Otorhinolaryngol. 2003; 67(12):1337-41.

9. Pappas NW, McLeod S, McAllister L, McKinnon HD. Parental involvement in speech intervention: a national survey. Clin Linguist Phon. 2008; 22(4-5):335-44.

10. Sutherland D, Gillon GT. Development of phonological representations and phonological awareness in children with speech impairment. Int J Lang Commun Disord. 2007; 42(2):229-50.

11. Mainela-Arnold E, Evans JL, Alibali MW. Understanding conservation delays in children with specific language impairment: task representations revealed in speech and gesture. J Speech Lang Hear Res. 2006; 49(6):1267-79.

12. Gillon GT. Facilitating phoneme awareness development in 3- and 4-year-old children with speech impairment. Lang Speech Hear Serv Sch. 2005;36(4):308-24.

13. Bernhardt B, Major E. Speech, language and literacy skills 3 years later: a follow-up study of early phonological and metaphonological intervention. Int J Lang Commun Disord. 2005;40(1):1-27.

14. Raitano NA, Pennington BF, Tunick RA, Boada R, Shriberg LD. Pre-literacy skills of subgroups of children with speech sound disorders. J Child Psychol Psychiatry. 2004;45(4):821-35.

15. Blischak DM, Shah SD, Lombardino LJ, Chiarella K. Effects of phonemic awareness instruction on the encoding skills of children with severe speech impairment. Disabil Rehabil. 2004;26(21-22):1295-304.

16. Psillas G, Psifidis A, Antoniadou-Hitoglou M, Kouloulas A. Hearing assessment in pre-school children with speech delay. Auris Nasus Larynx. 2006;33 (3):259-63.

17. Buller N, Ptok M. Is there a correlation between low-level auditory processing and phonological processing in preschool children? HNO. 2006;54(9):715-20.

18. Yavas M, Lamprecht R. Os processos e a inteligibilidade na fonologia com desvios. In: Yavas M. Desvios fonológicos em crianças: teoria, pesquisa e tratamento. Porto Alegre: Mercado Aberto; 1990. p. 231-256.

19. Hodson B, Paden E. Targeting intelligible speech: a phonological approach to remediation. San Diego: College-Hill Press; 1983.

20. Wertzner HF. O distúrbio fonológico em crianças falantes do português: descrição e medidas de severidade. [tese]. São Paulo (SP): Faculdade de Medicina da Universidade de São Paulo; 2002, 228f.

21. Wertzner HF, Papp ACCS, Amaro L, Galea DES. Relação entre processos fonológicos e classificação perceptiva de inteligibilidade de fala no transtorno fonológico. Revista da Sociedade Brasileira de Fonoaudiologia. 2005;10(4):193-200.

22. Ramos APF, Pergher GL, Marques J, Collares LM, Carreirão L. Distúrbio fonológico: perfil e inteligibilidade de fala. Cadernos de pesquisas em lingüística. 2005;1(1):67-79.

23. Lamprecht RR. A aquisição fonológica normal e com desvios fonológicos evolutivos: aspectos quanto à natureza da diferença. Letras de Hoje. 1995;30(4):117-25.

24. Huttunen K, Sorri M. Methodological aspects of assessing speech intelligibility among children with impaired hearing. Acta Otolaryngol. 2004;124(4):490-4.

25. COLLARES LM. Avaliação da inteligibilidade de fala em crianças com distúrbios fonológicos: criação de itens para uma escala de avaliação [Dissertação de Mestrado]. Curitiba (PR): Universidade Tuiuti do Paraná; 2003.

26. Yavas MS, Hernandorena CLM, Lamprecht RR. Avaliação Fonológica da criança. Porto Algre: Artmed;