
Performance of adults with hearing loss on speech recognition index with recorded and live voice material

Desempenho de adultos com perda auditiva na pesquisa do índice de reconhecimento de fala para material de fala gravado e a viva voz

Rendimiento de adultos con pérdida auditiva en la investigación del índice de reconocimiento del habla para material grabado y a viva voz

Nanci Mecenas Lima*
Teresa Maria Momensohn-Santos**

Abstract

Introduction: Percentage of speech recognition index is a measure of the intelligibility of speech in a fixed intensity in which the individual can repeat correctly the greatest number of words. **Objective:** Compare the performance in adults with hearing loss on speech recognition index search with recorded list and with live voice list. **Material and Methods:** 14 individuals participated in the research (28 ears) with age between 50 and 85 years, showing bilateral sensorineural hearing loss, with descending audiometric configuration. All the individuals realized the research of speech recognition index with recorded list and live-voice list through a calibrated audiometer. **Results:** results showed that the percentage of adults' errors in the speech recognition index for recorded lists is smaller than for the live voice list and there is statistically significant difference for monosyllabic words in left ear ($p = 0.028$). **Conclusion:** the findings of this study show that the performance of the participants in the research of speech recognition index with recorded list is better and has less variability than that observed with the live-voice list.

Keywords: Auditory Perception; Audiometry of Speech; Hearing; Adults.

*Instituto de Estudos Avançados da Audição - IEAA – São Paulo-SP - Brazil.

**Pontifícia Universidade Católica de São Paulo – PUCSP - São Paulo-SP - Brazil.

Authors' contributions: NML Acquisition, interpretation and analysis of data; participated in the drafting and critical review of the work, with real intellectual contribution to its contents. TMMS - substantially contributed to the scientific conception and work design; participated in the writing and in the critical review of the work; final approval of content to be published.

Correspondence address: Nanci Mecenas Lima. **E-mail:** fonocenter2010@gmail.com

Received: 13/04/2016

Accepted: 07/09/2016

Resumo

Introdução: Índice percentual de reconhecimento de fala – IPRF- é a medida da inteligibilidade da fala em uma intensidade fixa na qual o indivíduo consegue repetir corretamente o maior número de palavras. **Objetivo:** Comparar o desempenho em adultos com perda auditiva na pesquisa índice de reconhecimento de fala com lista gravada e com lista a viva voz. **Método:** Participaram da pesquisa 14 indivíduos (28 orelhas) com idade entre 50 e 85 anos, com exame audiométrico apresentando perda auditiva neurosensorial bilateral, com configuração descendente. Todos realizaram a pesquisa do índice de reconhecimento de fala com lista de palavras gravadas em mídia digital e lista apresentada a viva voz, por meio de audiômetro calibrado. **Resultados:** resultados mostraram que a porcentagem de erros dos adultos no índice de Reconhecimento de Fala para listas de palavras em mídia digital é menor do que para a lista de palavras apresentadas a viva voz e que há diferença estatisticamente significativa para palavras monossilábicas na orelha esquerda ($p\text{-valor}=0,028$). **Conclusão:** os achados deste estudo mostram que o desempenho dos participantes na pesquisa do índice de reconhecimento de palavras com lista com mídia gravada é melhor e tem menor variabilidade do que o observado com a lista apresentada a viva voz.

Palavras-chave: Percepção Auditiva; Audiometria de Fala; Audição; Adulto.

Resumen

Introducción Índice percentual de reconocimiento de habla - IPRF, es la medida de la inteligibilidad de la habla fija en la cual el individuo consigue repetir correctamente el mayor número de palabras. **Objetivo:** Comparar el desempeño en adultos con pérdida auditiva en la pesquisa índice de reconocimiento de habla con lista grabada y con lista a la viva voz. **Métodos:** Participaron de la pesquisa 14 individuos (28 orejas) con edad entre 50 y 85 años, con examen audiométrico presentando pérdida auditiva neurosensorial bilateral, con configuración descendente. Todos realizaron la pesquisa del índice de reconocimiento de la habla con lista de palabras grabadas en mídia digital lista presentada a la viva voz, por medio de audiómetro calibrado. **Resultados:** resultados mostraron que el porcentaje de errores de los adultos en el índice de Reconocimiento de habla en mídias digital es discretamente menor del que para la lista de palabras presentadas a la viva voz. **Conclusión:** los resultados de este estudio muestran que el desempeño de los participantes en busca de reconocimiento de palabras lista de medios del índice registrado es mejor y tiene menos variabilidad que la observada con la lista dada el altavoz.

Palabras claves: Percepción Auditiva; Audiometría del Habla; Audición, Adulto

Introduction

Speech tests are of great importance in audiological diagnosis and aim at measuring the ability of an individual to perceive and recognize speech sounds¹. When we test human hearing function, we also evaluate their ability to communicate, that is, the possibility of comprehension and expression of speech².

Since the 1940s it has been recommended that the research on Speech Recognition Index (SRI) would be performed with recorded material instead of a live voice monitored. A pioneer in the development of speech audiometry, noticed that “the presentation through the phonograph increases the stability of the test conditions”³.

The two essential requirements to carry out the SRI tests are: the use of standard materials and the use of same procedures and stimuli for test and retest data so that they can be compared between examiners. During the process of applying these tests, it is necessary that the clinical use recorded presentation of standardized test materials⁴.

Audiometry speech or logaudiometry is a technique aimed at assessing the ability of a person to hear and understand speech. It has become a basic tool in the overall assessment of hearing loss⁵. Along with pure tone audiometry, it can help determine the degree and the type of hearing loss. The logaudiometry also brings information on degrees of discomfort or tolerance to speech

stimuli in addition to information on speech recognition skills.

Several methodological issues should be addressed when performing the measurement of speech audiometry, among them we have: the way how the word list is built; the intensity with which the speech signal is presented; if the word list is made up of familiar words, if the list is live voice presented or if it is recorded, among others. The recorded words list has been recommended because it ensures the standardization of the procedure as well as it keeps the consistency of the speech signal that is presented, which does not occur with the presentation live voice. The use of recorded materials is recommended, especially for the supra threshold speech measures⁶. It is important to remember that performance in tests of speech perception can be affected by many processes that interact with each other: sensory, perceptual and cognitive⁷.

In a basic audiological evaluation, although there are well-established relations between some thresholds for pure tones and the necessary intensity to understand speech, difficulties in understanding speech can only be displayed with speech sounds that represent a communication situation⁸.

The standardization of materials used in speech recognition tests is considered a paramount condition taking into account that the reliability of this measure can be certified. The use of recorded speech material ensures that the same procedures and the same stimuli are used, and that test and retest data can be compared across different examiners, and various institutions. Regardless of all these considerations, this practice is not found on a daily basis of audiological clinic. Speech-language pathologists report that they find it more convenient, faster and simpler to apply the monitored live voice logoaudiometric test than the recorded media. Others believe that, with live voice presentation, patients get more correct answers, because the presentation of the stimuli with the regional accent can help them better understand the word being presented⁴.

The biggest advantage of a recorded test is that it has consistency in its presentation. No speaker can replicate their standard speech articulation along different presentations of the same test material in a reasonable similar way⁹. Considering the need to develop speech recorded material for use in audiology, this research aims to compare adult performance with hearing loss in analyzing

Speech Recognition Index (SRI) with lists of words presented by means of live and recorded voice.

Methods

This study followed regulations of prospective and cross-sectional research.

The data collection was carried out in a private institution in the city of Cachoeira Paulista, state of São Paulo. All participants signed the "Informed Consent form". This project was submitted to the Research Ethics Committee and approved under number 31055414.5.0000.5493.

Casuistry

The study population was composed of 14 subjects ranging from 59 to 81 years old. Regardless of gender, they met the following inclusion criteria:

a) Symmetrical or asymmetrical audiometric curve, of mild, moderate, moderately severe and bilaterally severe¹¹ sensorineural¹⁰ type;

b) Subjects whose mother tongue was other than Brazilian Portuguese, not born in São Paulo or were diagnosed under degenerative neurological system disease or reported being under such condition were excluded from the research.

Procedures

All subjects were submitted to the following procedures:

- Visual inspection of the external auditory canal with OTOSCOPE OMNI 3000 MD;
- Anamnesis was carried out in order to obtain data on their overall health and data from otologic, audiological and professional activity history.

All the following procedures were performed in audiometric booth with supra-aural headphones and properly calibrated clinical audiometer as determined by the Federal Council of Speech-language Pathology: pure tone audiometry threshold by means of airway, in the frequencies of 250, 500, 1000, 2000, 3000, 4000, 6000, 8000 Hz; and research of Speech Recognition Threshold (SRT).

For the research of Speech Recognition Index (SRI) word lists developed by Seiva et al (2012), consisting of 25 monosyllables and 25 disyllables for each ear, were presented, on two conditions: A - recorded in digital media, B - live-voice monitored. The lists were presented at 40 dB sensation

level (dBSL) over tonal average of 500, 1000 and 2000 Hz, or in sound intensity better comfort conditions. Each subject participated in two SRI measurements in a single session, lasting about 40 minutes, including the entire procedure, namely: otoscopy, anamnesis, pure tone audiometry and speech audiometry in both conditions.

In order to avoid bias learning, familiarity with the list of words or fatigue, the application of the procedures were randomized as follows:

- SUBJECTS 1-3-5-7-9, and so forth: 1st RECORDED SRI / 2nd LIVE VOICE SRI.
- SUBJECTS 2-4-6-8 and so forth: 1st LIVE VOICE SRI / 2nd RECORDED SRI.

The lists were presented in media recorded by Compact Disc Digital Audio - Brand PB 120N - PHILCO CDR/RW. The production of the CD was carried out in recording studio located in the city of São Paulo. For the recording, a condenser microphone M-Audio Solaris positioned 6 cm from the speaker was used, the speaker was standing in an acoustic room, with audio captured by a preamplifier of Studio Projects VTB1 V Series Microphone and Audio Interface - M-Audio Firewire 1814, and recorded with Software - SONAR 8 on a computer with CPU - Pentium Dual Core with Operating System Windows XP SP 2.

Analysis criteria procedures

The SRI results were analyzed per ear, in addition, omissions of answers or any kind of sub-

stitution were considered errors. The results were expressed as percentages and, as the test carries 50 words (25 monosyllabic and 25 disyllabic), for each correct response; a 4% value was awarded. For analysis purposes, the data were presented in percentage values based on errors made by each participant in each ear

Analysis of results

Results were initially analyzed using descriptive statistics and, in order to make the comparison between the groups in each variable, the Wilcoxon test was used, since the data did not show adherence to the normal curve. The rejection level for the null hypothesis was set at a value less than or equals 0.05 (5%).

Results

Statistical analysis of the average and the median values shows that there is great variability in the percentage of errors in the study group, which allows us to say that this distribution of values is asymmetrical.

We notice that there is a statistically significant difference between the recorded situation and the live voice situation only for monosyllable in the left ear (p-value = 0.028).

Table 1. Descriptive analysis of data relating to age (years) of the sample composing this study (n = 14)

Average	SD	Median	Maxim	Minimum
70.4	6.9	68	81	59

Table 2. Descriptive analysis for the relative frequency of occurrence of values (%) for the degree of hearing loss found in the sample, classified according to Lloyd and Kaplan 1978

Loss degree	Mild	Moderate	Mod/Severe	Severe
RE	43%	21%	29%	7%
LE	36%	29%	29%	7%

Subtitle: RE - right ear; LE - left ear; mod/severe - moderately severe

Table 3. Descriptive analysis in relative frequency values of occurrence (%) for audiometric configuration of hearing loss found in the sample

Configuration of loss	Ascending	Mild Descending	Inverted U
RE	7%	86%	7%
LE	7%	86%	7%

Table 4. Descriptive analysis of the percentage of correct answers given by the evaluated subjects in the research of speech recognition index in recorded list and in live voice conditions

Subject	Monosyllabic words				Disyllabic words			
	RECORDED RE	RECORDED LE	LIVE-VOICE RE	LIVE-VOICE LE	RECORDED RE	RECORDED LE	LIVE-VOICE RE	LIVE-VOICE LE
1	68%	80%	72%	72%	76%	56%	80%	60%
2	92%	96%	36%	48%	88%	92%	64%	56%
3	100%	92%	88%	92%	100%	100%	96%	96%
4	68%	64%	100%	88%	92%	92%	100%	100%
5	80%	96%	96%	88%	72%	96%	96%	96%
6	92%	80%	80%	80%	92%	96%	96%	100%
7	92%	92%	91%	90%	100%	92%	89%	90%
8	92%	100%	92%	88%	100%	100%	96%	96%
9	92%	100%	96%	96%	100%	100%	84%	100%
10	100%	100%	88%	92%	100%	100%	88%	92%
11	76%	100%	96%	80%	100%	56%	100%	92%
12	96%	100%	84%	84%	88%	96%	92%	100%
13	100%	100%	88%	84%	100%	100%	92%	72%
14	88%	100%	84%	96%	100%	92%	88%	80%

Table 5. Descriptive analysis in the percentage of incorrect answers given by evaluated subjects in the research of speech recognition index in recorded list and in live voice conditions

%Error		Average	Standard Deviation	N	IC	P-value
Monosyllables	RE	Recorded	11.7%	11%	14	5.8%
		Live voice	14.9%	15.9%	14	8.3%
	LE	Recorded	7.1%	10.9%	14	5.7%
		Live voice	15.9%	12.3%	14	6.5%
Disyllables	RE	Recorded	6.6%	9.5%	14	5.0%
		Live voice	9.9%	9.5%	14	5.0%
	LE	Recorded	9.4%	15%	14	7.9%
		Live voice	12.1%	15%	14	7.9%

Wilcoxon test $P \leq 0,05$

Subtitle: RE - right ear; LE - left ear.

Discussion

In this study, the SRI research was carried out with the list of words presented at live voice and the list recorded in digital media. In Brazil, few speech language pathologists use the practice of performing the logaudiometric measures with recorded list. They comment that the test takes longer and that there are few national studies on logaudiometry with lists in Brazilian Portuguese. This same type of review has also been found in studies conducted in the United States.

One of the objections that the speech language pathologist uses to not use the recorded speech material is the delay; he thinks he will “waste time” to perform this procedure if he does it this way. In

2011, American researchers¹² evaluated the time difference in the management of speech audiometry tests with live voice and recorded list. They showed that the average difference administration time (for lists of 50 words recorded and live) was less than a minute (49 seconds) - which is not clinically significant. Nevertheless, they showed that patients with hearing loss had better test performance with speech recorded material than they did with the material presented at live voice. In this study, we did not measure time, however, we found that there is better performance, lower percentage of errors and lower standard deviation in both recorded lists and in both ears, even if they are not statistically significant.

These data may indicate that, for people with hearing impairment, the use of recorded media is recommended for a more reliable and accurate diagnosis, due to the possibility of reducing the speaker variability, errors in microphone placement and speaker sound calibration issues, as well as to reduce errors and articulation of inaccuracies.

The live voice test can be influenced by the intra and inter-speaker variability that is introduced on the test results, such as accent, speech rate, speech articulation pattern, intonation, among others. In order to minimize this problem, the test could be applied several times, but this would be unworkable in clinical practice. The recorded stimulus allows editing of the recording so that it approaches the recommended model for the creation of this type of material without suffering the subtle interference of speaker variability while delivering the same word at different times¹⁴.

The results showed a statistically significant difference between the results obtained with the recorded list and the live-voice list only in the left ear to monosyllable ($P = 0.028$). However, it is important to notice that, for all other conditions, the subjects' performance was always better, that is, with a lower percentage of error for the condition of lists presented in the recorded media. It is also noticed that the standard deviation is always higher at "lists presented at live voice" as they are even higher for monosyllabic words than for disyllabic words.

In our study, 13 subjects had mild, moderate or moderately severe hearing loss, while in the quoted authors, if we consider the average audiometry values for the frequencies of 500, 1000 and 2000 Hz, all subjects' audiometry would be interpreted as a normal degree. Different results from the ones found in this study were reported in another study¹⁴, however, the elderly who participated in the study in question showed different audiometric conditions from the population now studied.

The data analysis shows that the sample was predominantly composed of subjects with mild descending audiometric configuration hearing loss type. In general, individuals with this audiometric standard are likely to have worse performance in competitive communication situations such as those that occur in noisy environments.

Importantly, this study has some limitations that may have influenced the results: the group's age variability (59-81 years old); the low number

of the sample (14 subjects); and, as a research bias, the fact that it has been quoted in another study⁸, the prior knowledge of the purpose of the study may have led them to try to get more correct answers than if the collection had been made spontaneously.

In a basic audiological evaluation, although there are well-established relations between pure tone thresholds and the necessary intensity to understand speech, the real difficulties to understand speech can only be demonstrated when speech sounds that represent a communication situation are used¹⁰.

The biggest advantage of a recorded test is that it has consistency in its presentation: no speaker can replicate in a reasonably similar way their standard speech articulation along different presentations of the same test material; with the recorded speech material, the intra and inter-speakers variability can be controlled; for the same subject evaluated by different examiners, this aspect will not be a factor that may cause error in his exam. It is highly important that the possibility of error in an audiological evaluation is minimized to the lowest so that interpretations and inappropriate decisions can be avoided^{13,6}.

Regarding the usage of these tests, it is worth taking some considerations: do not take into account only the time spent in the application, but also bear in mind the reliability of speech tests applied by pre-recorded digital media; in addition, the test result is both valid and reliable while it delivers a service that is suitable to the professional responsible for the test application.

Digital media can be efficiently used in order to manage stimuli. Studies show that only 1% of speech language pathologists reported the use of digital media¹⁵.

The standardization of speech recognition materials is essential to the reliability of clinical procedures that define the information on diagnosis and the essential rehabilitation of a person's hearing ability¹².

According to these collected data, the need for studies with a larger sample is evident.

Conclusion

The results showed that the percentage of adults' errors in the Speech Recognition Index (SRI) for word lists in digital media is slightly lower than it is for the list of words showed at live

voice. We conclude that there is a statistically significant difference between the performance situation on correct rate for speech recognition in recorded material and live voice only for monosyllables in the left ear.

References

1. Seiva, AS, Momensohn-Santos T, Fortes CC, Queiroz SD. Desempenho de indivíduos no teste do índice de reconhecimento de fala à viva voz e com fala gravada. *Distúrb Comum*, São Paulo, 24 (3): 351-8,2012.
2. Frota S. Fundamentos em Fonoaudiologia. Audiologia. Ed. Guanabara Koogan. 1998.
3. Carhart R. Monitored live-voice as a test of auditory acuity. *J Acoust Soc Am*. 17: 338-49, 1946.
4. Roeser RJ, Clark JL. Live voice speech recognition audiometry - stop the madness! *Audiol Today*, 20(1); 32, 2008
5. Bess F, Humes L. Fundamentos de audiologia. 4a.ed. Editora Revinter. 2012
6. Schoepflin, J Back to Basics: Speech Audiometry. <http://www.audiologyonline.com/articles/back-to-basics-speech-audiometry-6828>. 2012. Acesso em 29/08/2015
7. Ptok M, Kiessling J. Speech perception. The basis for speech audiometry examinations. *HNO*. Sep; 52 (9): 824-30, 2004
8. Freitas CD, Lopes LFD, Costa MJ. Confiabilidade dos limiares de reconhecimento de sentenças no silêncio e no ruído. *Revista Brasileira Otorrinolaringologia* [Internet]. Oct [cited 2015 Aug 01]; 71 (5): 624-630. 2005 Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-72992005000500013&lng=en.
9. Zaboni ZC, Iório MCM. Reconhecimento de fala no nível de máximo conforto em pacientes adultos com perda auditiva neurossensorial. *Rev. Soc. Bras. Fonoaudiologia*. 14:491---7. 2009
10. Silman S, Silverman CA. Basic audiologic testing. In: Silman, S.; Silverman, C. A. *Auditory diagnosis: principles and applications*. San Diego: Singular Publishing Group; 1997. P.: 44-52
11. Lloyd LL, Kaplan H. *Audiometric interpretation: a manual of basic audiometry*, vol. 94. Baltimore: University Park Press; 1978. p. 16-7.
12. Mendel LL, Owen RS. A study of recorded versus live voice word recognition. *Int J Audiol* 50: 688-93, 2011
13. Zaboni ZC, Iório MCM. Reconhecimento de fala no nível de máximo conforto em pacientes adultos com perda auditiva neurossensorial. *Rev. Soc. Bras. Fonoaudiologia*. 14:491-7, 2009
14. Caporali, AS, Silva JA. Reconhecimento de fala no ruído em jovens e idosos com perda auditiva. *Rev. Bras. Otorr*. 70(4): 525-53. 2004
15. Martin FN, Champlin CA & Chambers JA. Seventh survey of audiometric practices in the United States. *J Am Acad. Audiol.*, 9, 95 – 104, 1998