MEASURING SPEECH INTELLIGIBILITY IN BACKGROUND NOISE BY USING PSYCHOPHYSICAL JUDGMENTS IN DIFFERENT GROUPS

Sueli Aparecida Caporali [1] and José Aparecido Da Silva [2]
1) Department of Speech Pathology of UNIMEP and UNAERP
2) University São Paulo, Ribeirão Preto, São Paulo - Brazil

Abstract

The speech perception in background noise was investigated through psychophysics judgments in groups of young listeners with normal hearing, adults and elderly with hearing loss. Ratio and interval scales were used to estimate the intelligibility of everyday sentences presented with cocktail noise at three different signal-to-noise ratios. According to the results, the intelligibility of sentences improves as the signal-to-noise ratio increases, as observed in the three groups, although differences among groups have been found in the intelligibility judgments. Both psychophysical scaling methods were valid and reliable to perform this evaluation which showed to be effective to measure speech intelligibility. However, further studies are desirable before clinical usage can be done with this evaluation method.

The speech perception in background noise has been investigated by several researchers. Since the psychophysical scaling tests can be used to evaluate an observer's impression of the relationship between properties of a physical stimulus and the sensation it produces (Barry & Kidd, 1981), the psychophysical method has been used to measure perceptual attributes of speech (Fucci et al, 1994). According to Cox and McDaniel, 1984; Nakatani and Dukes, 1973, several studies have shown that intelligibility judgments are monotonically related to signal-to-noise ratio and to the amount of degradation of the speech signal by filtering.

The purpose of this investigation was to examine the abilities of young listeners with normal hearing, and young and elderly listeners with hearing loss to perform intelligibility judgments using ratio and interval scales for speech presented in noise. In this study speech intelligibility was defined as "how well the speech is understood" reported numerically by the listener using CD samples of everyday sentences with "cocktail noise" at three different signal-to-noise ratios.

Method

Subjects

Sixty subjects were given pure tone (250 – 8000 Hz) and speech audiometric evaluations before performing psychophysical judgments. The subjects are distributed in three groups according to age and hearing.

Group 1

Twenty young women, were randomly chosen to serve as subjects. Their ages ranging from 21 to 38 (M age = 23,30) and having normal hearing.

Group 2
Twenty adults within 29 – 50 years of age (M age = 40,45), eighteen men and two women, were chosen to serve as subjects, since they had hearing loss only at high frequencies (above Hz).
Group 3
Twenty elderly within 60 –77 years of age (M age = 66,85), ten men and ten women, were chosen since they had hearing loss only in high frequencies (above 3000 Hz). Each subject in group 3 was selected on the basis of hearing sensitivity that was matched reasonably well to that of a subject in group 2.

Stimulus

The stimuli used were 22 everyday sentences in Portuguese (Oliveira, 1992). See annex 1. The original sentences and "cocktail noise" were recorded separately by PC in a professional studio. The speaker was a male adult who spoke standard Portuguese. The recording level was controlled to assure that the speech level and noise were constant during all recordings. After the original recording, the samples were edited and recorded on a compact disk and subsequently played by a CD player. Three series of sentences were produced in randomized order. The output of the CD player was routed to the input of a two-channel speech audiometer (AC 30 - Kamplex) located in the anechoic chamber; the output of the audiometer was routed through the wall of the chamber to a pairs of TDH-39 earphones. The first stimulus was preceded by one sentence: "Please, listen carefully to all sentences!" Afterwards, the first sentence was presented and followed by an 8 - to second silent response interval before the next sentences.

Psychophysical judgments

The three groups were divided in two subgroups to perform different tasks, such as, category scaling with numbers among 1- 9 (N=10) and magnitude estimation without modulus (N=10). They were instructed to make judgments to each sentence in background noise according to the psychophysical tests proposed by Stevens (1975). The subjects that performed ratio scaling were trained before they started the listening tasks.

The subjects were tested individually in an anechoic chamber. The everyday sentences were presented to the subjects only on the right ear. Subjects were informed that they should evaluated the sentences in terms of intelligibility. They made scores to sentences three times, once for each signal-to-noise ratio different.

Experiment

The sentences were presented to the subjects with "cocktail" noise in randomized order. The presentation level of sentences was constant at 40 dB sensation level relative to each subject's average hearing threshold level for 500, 1000 and 2000 Hz at three different signal-to-noise ratios, i.e. the noise was varied .
Condition 1: Minus five (- 5) dB signal-to-noise ratio.
Condition 2: Zero (0) dB signal-to-noise ratio.
Condition 3: Plus five (+ 5) dB signal-to-noise ratio.

Results and Discussion

For data analysis the geometric means of magnitude estimation and arithmetic means of category scaling were obtained. The results showed that all psychophysics scaling were in agreement in relation to speech intelligibility judgments. In Figures 1 A, B and C, the means of category scaling for the sentences for three groups at three different signal-to-noise ratios

were presented. The Figures 2 A, B and C show the magnitude estimates for each sentence with "cocktail noise" at the same signal to noise ratios.

The Pearson correlation coefficients between magnitude estimation and category scaling were significant, especially at two of the signal-to-noise ratios (-5 and 0 dB) in all groups. This suggests good reliability of the intelligibility estimates. However, at plus five dB (+5 dB) signal-to-noise ratio the obtained coefficients showed low level of reliability between ratio and interval scaling. See Table 1:

Table 1 – Correlation coefficients between psychophysical tasks of the three groups for each signal-to-noise ratio. EM – Magnitude Estimation ; EC- Category Scaling. * $p \leq 0,01$

Sentences	Signal-to-noise Ratio	Group 1	Group 2	Group 3
EC X EM	- 5 dB	0,87 *	0,95 *	0,87 *
EC X EM	0 dB	0,85 *	0,84 *	0,80 *
EC X EM	+ 5 dB	0,50	0,14	0,79 *

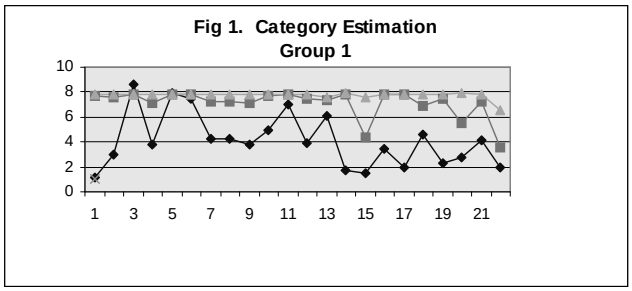
The majority of the results are in agreement with Purdy and Pavlovic (1992) and Mantelatto and Da Silva (2000 a), who found similar results comparing sensitivity in psychophysical procedures. Possibly the low coefficients found in the + 5 signal-to-noise ratio is due to the fact that subjects who performed category rating attributed high scores to all sentences, and the numbers used were close to the scaling limits. The same did not occur with the groups performing magnitude estimation.

The judgements to both scales showed that the speech intelligibility improves as the signal-to-noise ratio increases. The three groups consistently used smaller numbers in – 5 dB signal-to-noise ratio, indicating a decrease of speech intelligibility. Curiously, the scores used by young people (groups 1 and 2) in 0 and + 5 dB signal-to-noise ratios to all sentences presented were close, mostly for the groups that performed category scaling. These results were in agreement with Cox and McDaniel, 1984; Nakatani and Dukes, 1973, whose studies have shown that intelligibility judgments are monotonically related to signal-to-noise ratio. Our data are also in agreement to Fucci *et al* (1994) who recommend the use of psychophysical procedures to measure of the inherent subjective parameters involved in speech perception.

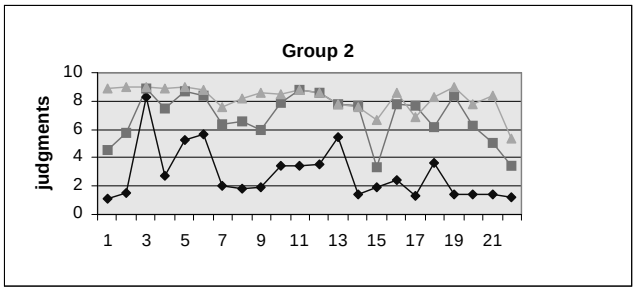
The analysis of variance (ANOVA) of the scores by subjects that performed category scaling showed statistically significant difference among groups $\{F(2,27) = 6,361; p < 0,01\}$ and signal-to-noise conditions $\{F(1,27) = 67, 67; p < 0,01\}$. However, for the magnitude estimates there were not statistically significant differences among groups neither to the three different signal-to-noise ratios ($p > 0,05$). These results are not in accordance with other findings obtained by Mantelatto and Da Silva (2000 a) where the statistically analysis were in agreement to the three psychophysical scaling methods. Probably, the lack of modulus can have influenced on the results, since in that previous study the authors used magnitude estimation with modulus.

The Figures 1 A, B, C and 2 A, B, C show the judgments performed by the three groups. The scores used by group 1 were higher compared to the groups with hearing losses (2 and 3). These data are in agreement with Pekkarmem, Salmivalli and Suonppa (1990), who reported that the speech perception depends on much more of masking stimulus in relation of spectrum, frequency, in proportion to speech signal presented and audiogram configuration.

A



B



C

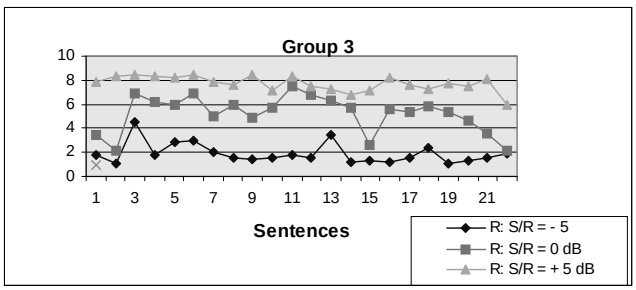
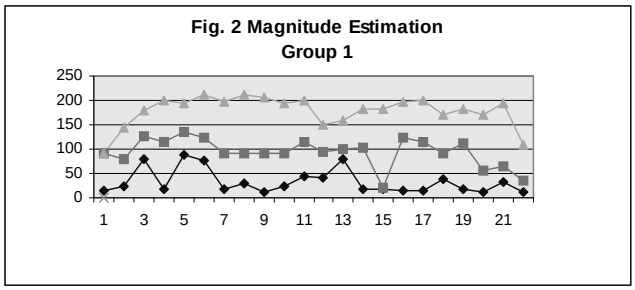
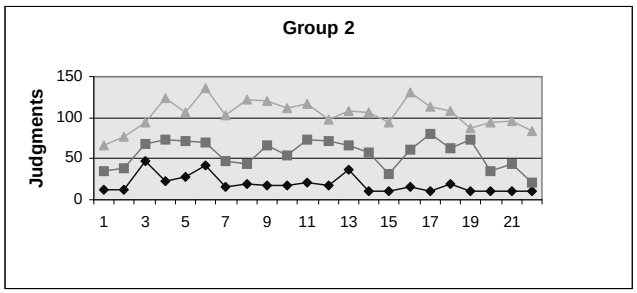


Figure 1: **Comparison among three groups using category scaling of Speech Intelligibility in "cocktail noise" at three different signal-to-noise ratios.** Data are expressed as by arithmetic means ten subjects per group. Each bar on the abscissa represents one different sentence. A - Group 1; B – Group 2; C – Group 3.

A



B



C

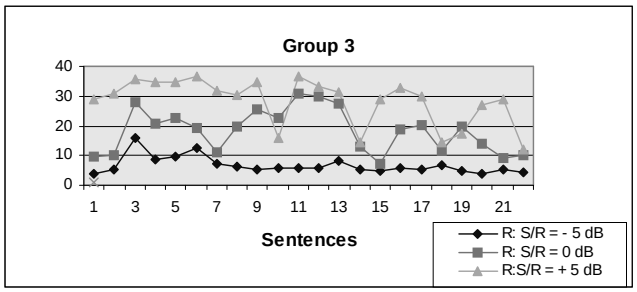


Figure 2: **Comparison among three groups using magnitude estimation of Speech Intelligibility in "cocktail noise" at three different signal-to-noise ratios.** Data are expressed as by geometric means ten subjects per group. Each bar on the abscissa represents one different sentence. A - Group 1; B – Group 2; C – Group 3.

We can also conclude that speech intelligibility scores were affected by age, since there were differences among the scores of group 2 and 3, with similar hearing losses. Dubno, Dirks and Morgan (1984) and Gordon-Salant and Fitzgibbons (1995a) also found that the elderly listeners perform more poorly than younger listeners for speech presented in noise, however they used other kind of measures with speech recognition scores in their studies.

These data also show that the psychophysical scaling are reliable and valid to measure the speech presented in noise. The speech intelligibility judgments can be affected by factors such as hearing loss and age. Further studies are required to examine the potential of psychophysical scaling to measure speech intelligibility in noise, including the effects of age and hearing loss to speech perception.

References

BARRY, S. K. & KIDD, G. (1981) Psychophysical scaling of distorced speech. *Journal of Speech and Hearing Research*, 46, 44-47.

DUBNO, J.R., DIRKS, D.D. & MORGAN, D.E. (1984) Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of América*, 76: 87 - 96.

COX, R. M. & MCDANIEL, D. M. (1984) Intelligibility ratings of continuo discourse: Application to hearing aid selection. *Journal of the Acoustical Society of America*, 76, 758-766.

FUCCI, D.; DOMYAN, S.; ELLIS, L. & HARRIS, D. (l994) Magnitude-Estimation: An effective method for the measurement of the quality of filtered speech. *Perceptual and Motor Skills, 78, 348-350.*

MANTELATTO, S.A.C. & DA SILVA, J.A**.** (2000 a) Inteligibilidade de Fala e Ruído: Um estudo com sentenças do dia-a-dia. Pró Fono, 12: 48-55.

NAKATAMI, L. H. & DUKES, K. D**.** (1973). A sensitive test of speech communication quality. *Journal Acoustical Society American, 53,* 1083 - 1092.

OLIVEIRA, S. T. (1992) Avaliação da percepção de fala utilizando sentenças do dia a dia. São Paulo - PUC/SP - 83p. (Dissertação de Mestrado).

PEKKARMEN, E.; SALMIVALLI, A. & SUONPAA, J. (1990) Effect of Noise on Word Discrimination of Subjects with Impaired Hearing Compared with those with Nourmal Hearing, *Scandinavian Audiology*, 36, 19 –31.

PURDY, S. C. & PAVLOVIC, C. V. (1992) Reliability, Sensitivity and Validity of Magnitude Estimation, Category scaling and Paired-Comparison Judgments of Speech Intelligibility by Older Listeners. *Audiology, 31,* 254 - 271.

STEVENS, S. S. (1975) *Psychophysics: introduction to its perceptual, neural, and social prospects.* New York: Wiley.

Appendix 1 - Everyday Sentences (Oliveira, 1992)

1. Poderia passar a manteiga.
2. Os ratos se escondem dos gatos.
3. Amanhã sairemos sem falta.
4. A tia dele foi nos visitar no Sábado.
5. Pensei que você tivesse ido embora.
6. Fui ao cinema depois do trabalho.
7. A televisão quebrou no meio do filme.
8. Eu não sei se isto é possível.
9. Tomei um copo de suco de limão.
10. Tocamos violão a noite inteira.
11. Preciso telefonar para ela.
12. Todo mundo sabia que ele era assim.
13. Tenho saudades dos velhos tempos.
14. Eles chegaram muito tarde.
15. Coloquei todas as minhas roupas no armário.
16. Pedi uma pizza para viagem.
17. Pegue a caneta para mim.
18. O ônibus parou três vezes durante a viagem.
19. A professora passou a lição de casa para os alunos.
20. O cachorro latiu o dia inteiro.
21. O pedaço de pão que você não comeu está no prato.
22. Perguntei para você onde tinha ido.