

Influência da velocidade articulatória e da intensidade na inteligibilidade de fala***

Influence of speech rate and loudness on speech intelligibility

Simone dos Santos Barreto*
Karin Zazo Ortiz**

*Fonoaudióloga. Mestre em Distúrbios da Comunicação Humana pela Universidade Federal de São Paulo. Fonoaudióloga da Prefeitura do Rio de Janeiro. Endereço para correspondência:
Rua Botucatu, 802 - São Paulo - SP - CEP 04023-062
(simone_barret@hotmail.com)

**Fonoaudióloga. Pós-Doutorado em Neurociências pela Universidade Federal de São Paulo. Professora Adjunta do Departamento de Fonoaudiologia da Universidade Federal de São Paulo.

***Trabalho Realizado no Departamento de Fonoaudiologia da Universidade Federal de São Paulo - Escola Paulista de Medicina.

Artigo Original de Pesquisa

Artigo Submetido a Avaliação por Pares

Conflito de Interesse: não

Recebido em 28.11.2007.
Revisado em 01.03.2008; 26.05.2008.
Aceito para Publicação em 26.05.2008.

Abstract

Background: contextual cues intrinsic to speech stimuli can have an influence on speech intelligibility measurements; however, the influence of cues that depend on the acoustic signal, such as speech rate and vocal loudness, need further investigation. **Aim:** to examine whether possible reductions in the articulatory rate and increase in vocal loudness, associated to the production of different speech stimuli, can have an influence on speech intelligibility measurements. **Method:** participants of this study were thirty normal speakers and sixty normal listeners. Speakers were recorded during the repetition of three lists of speech stimuli (sentences, words and pseudo words). The averages of the articulatory rate (syllables per second) and of the vocal loudness (decibel) were calculated for each speaker according to their performance in each repetition task. Speech intelligibility was measured based on the orthographic transcription of the speech samples; the score was calculated in terms of percentage of correctly transcribed words. **Results:** it was observed that articulatory rates were statistically different between the three types of stimuli; however, the stimuli produced with the lowest articulatory rate (pseudo words followed by words) did not present higher speech intelligibility scores. Vocal loudness was statistically higher during the repetition of pseudo words; however, this increase did not have an influence on the speech intelligibility scores. **Conclusion:** the reduction of the articulatory rate or the increase of vocal loudness did not have an influence on the speech intelligibility measurements, indicating that contextual cues have a greater impact on speech intelligibility than the independent cues given by the acoustic signal.

Key Words: Speech Intelligibility; Speech Production Measurement; Speech Acoustics; Speech.

Resumo

Tema: existem evidências de que as pistas contextuais intrínsecas aos estímulos de fala elevam os escores de inteligibilidade, entretanto, a influência de pistas dependentes do sinal acústico, como a velocidade e a intensidade com as quais os diferentes estímulos são produzidos, são pouco conhecidas. **Objetivo:** investigar se a redução da velocidade articulatória e o acréscimo da intensidade da fala, em diferentes tipos de estímulos, influenciariam os escores de inteligibilidade. **Método:** participaram do estudo 30 falantes e 60 ouvintes, todos sem distúrbios da comunicação. Os falantes foram gravados durante a repetição de três listas de estímulos (frases, palavras e pseudopalavras). As médias da velocidade articulatória (sílabas por segundo) e da intensidade da fala (decibel) foram calculadas por falante, para cada lista. A inteligibilidade foi mensurada pelo método de transcrição ortográfica das amostras pelos ouvintes, sendo os escores calculados em percentagem de palavras corretamente transcritas. **Resultados:** diferenças estatisticamente significativas da velocidade articulatória foram encontradas entre os três tipos de estímulos, contudo, os estímulos produzidos com menor velocidade (pseudopalavras seguidas pelas palavras) não conduziram a escores superiores de inteligibilidade. Em relação à intensidade, apenas as pseudopalavras apresentaram valores estatisticamente superiores aos demais estímulos, porém este acréscimo também não elevou os escores de inteligibilidade da fala. **Conclusão:** nem a redução da velocidade articulatória nem o acréscimo da intensidade da fala influenciaram os escores de inteligibilidade dos sujeitos avaliados, sinalizando que as pistas contextuais exercem mais efeito sobre a inteligibilidade da fala que as informações independentes do sinal acústico.

Palavras-Chave: Inteligibilidade da Fala; Medidas de Produção da Fala; Acústica da Fala; Fala.

Referenciar este material como:



Barreto SS, Ortiz KZ. Influence of speech rate and loudness on speech intelligibility (original title: Influência da velocidade articulatória e da intensidade na inteligibilidade de fala). Pró-Fono Revista de Atualização Científica. 2008 abr-jun;20(2):87-92.

Introduction

Different types of stimuli are commonly employed for assessing intelligibility in speech disorders. Several studies have investigated the influence of stimulus type on intelligibility scores, noting that contextual cues tend to boost speech intelligibility(1-3). However, the effects of different variables, such as speech rate and intensity, are often overlooked, even though they may contribute to the differences seen in intelligibility measures when different speech stimuli are used.

The findings of a number of studies which have examined the relationship among these signal-dependent variables and speech intelligibility give credence to this notion. In terms of speech rate, studies involving dysarthric individuals have shown reduced rates to be associated to better intelligibility scores(4-7), whilst for the intensity variable, findings on articulatory dynamics in speech production at high intensity by normal individuals, such as the amplification of phonatory and articulatory force, have provided grounds for the hypothesis of improved intelligibility through raised intensity(8). Moreover, correlation has been found between increased intensity and intelligibility for specific groups of dysarthric speakers(9).

Thus, the aim of the present study was to investigate whether either a reduced articulatory rate or increased speech intensity can influence intelligibility scores for different stimulus types.

Methods

All volunteers who took part in this study were informed about the objectives of the work and the form of participation expected, and previously signed the informed term of consent. This study was approved by the Research Ethics Committee of the Federal University of São Paulo - Escola Paulista de Medicina (UNIFESP-EPM/ Report 0708/06).

Participants

A total of 30 individuals with no communication impairments were recruited from a population accompanying patients at the Acquired Neurologic Speech and Language Disorders Unit of the Speech Therapy Department of UNIFESP-EPM, or from relatives of students on the Speech Therapy course at UNIFESP-EPM. Adult native speakers of Brazilian Portuguese were selected who had no history of present or previous communication disorders, or

history of neurologic compromise (traumatic brain injury with loss of consciousness for more than 15 minutes, stroke, epilepsy, etc.), high blood pressure, use of psychotropic medication or psychiatric history. A healthy adult population was studied because they represented potential speakers with maximum efficiency in the use of speech production mechanisms. Speakers presented a mean age of 40.4 years, standard deviation of 13.2 years, and comprised 15 males and 15 females.

A further 60 subjects drawn from the population of graduate and post-graduate students from the Speech Therapy Department of UNIFESP-EPM acted as listeners of the speech samples produced. These listeners were chosen for their high level of education, in a bid to control the influence of this variable on the transcription task, and due to their familiarity with the intelligibility assessment procedure. This dispensed with the need for prior training of the listeners. This group included native speakers of Brazilian Portuguese with normal hearing (verified by basic hearing test), who had no history of disorders in language, learning or cognition, nor any relationship with the speakers and/or the speech stimuli employed, since these factors could potentially interfere in the measuring of intelligibility.

Material

Three lists of stimuli were used as speech material: sentences, words and pseudowords, where these varied in terms of the contextual cues intrinsic to each type of stimulus. Vehicle sentences were not used so as to mimic usual intelligibility assessment procedures in which separate words or sentences are recorded(1-3). Although little used, pseudowords were included to represent a stimulus with an absolute minimum level of contextual cues.

The list of sentences used(10) was selected for its phonetically balanced content, having a 99% level of correlation with the reference corpus of the Acoustic Phonetic and Experimental Psycholinguistic Laboratory of the State University of Campinas. This served as the basis for devising the other two lists of stimuli, composed of 25 sentences made up of simple periods. The sentences contained an average of five words and nine syllables and constituted an overall corpus containing 520 phonetic occurrences and 237 syllables.

The word and pseudoword lists devised, presented very strong correlations ($p \geq 0.993$ and p

< 0.001) with the list of sentences in terms of distribution of frequency of phonemes, word length and type of syllabic structure involved. Word and pseudoword lists were identical to each other in terms of spread of parameters, containing 60 stimuli each with 260 phonetic occurrences and 118 syllables.

The following equipment was used to record the speech samples: a Cyber Acoustics, model AC-100 microphone headset, a Toshiba, model Satellite L25 Notebook and the Sound Forge 4.5 program (Sonic Foundry). The Praat 4.4.13 program along with model CD-6631MV Edifier head phones were also used for sound file editing, acoustic analysis and transcription task.

Procedures

The speakers were recorded during repetition of the three lists of stimuli at natural speed and intensity. The recording was carried out in a silent environment, with the subject seated and microphone placed 5 cm from their mouth. The order of presentation of the lists and their items was performed randomly for each subject. In addition, the order of the lists was counterbalanced in the group of speakers so as to prevent interference on results due to an ordering effect. Verbal repetition was preferred over reading so as to avoid reading ability of the speaker from affecting performance.

The original sound files were edited into 145 files per speaker, based on the auditory analysis and acoustic analysis broad band spectrograms, configured according to the default values of Praat program, with the aid of the tools to display the pulses, formants and intensity curve of the speech signals. Cut-off criteria were set as the first negative peak which preceded or followed each initial and final phoneme.

The acoustic measures of intensity and articulatory speed were calculated. The intensity of each utterance was registered in decibels (dB), using the intensity function of the Praat program, after which the mean intensity of each list of stimuli was calculated. The mean number of syllables produced per second was calculated to give articulatory rate, defined as the number of speech units per time unit, excluding any pauses separating the articulatory sequences(11). The duration of each utterance in milliseconds was registered and mean velocity calculated by dividing total number of syllables in the list by pronunciation time of items.

In order to measure speech intelligibility, the item identification method was used, based on orthographic transcription of the speech samples by the listeners and quantifying of the number of correctly decoded stimuli. This model was chosen because it yields reliable scores related to the information transferred, and is the most frequently used approach(1-7,12-16).

The listeners performed orthographic transcription of the speech samples in individual sessions within a silent environment. The output audio volume of the notebook and Windows Media Player were set to a comfortable level for listeners and kept at this level throughout the entire task for all listeners. Each listener was randomly designated to transcribe the sample of one speaker only in a bid to minimize the effect of prior knowledge of stimuli on test results. Each speaker had their speech sample transcribed by two listeners in order to minimize the influence of variability of listeners on intelligibility scores. The order of presentation of the list of stimuli followed the original recording order, whereby items from each list were presented once only, one by one at intervals dictated by the listener's transcription pace.

The transcriptions were analyzed and scored according to the number of correct items per syllabic unit. Given the different number of syllables in each stimulus list, intelligibility was measured by the percentage of correctly transcribed syllables per list. Transcribed stimuli were considered correct when phonemic correspondence was observed between the orthographic transcription and the expected production of the target stimuli. However, typical variations in oral modality, represented graphically, were deemed correct, provided the word meaning was unchanged.

Data analysis involved comparisons between stimulus types for both articulatory rate measures and speech intensity measures, while any differences between their means were assessed. Comparisons between intelligibility scores by stimulus types were performed in order that this data could be analyzed together with the results of the acoustic measurements. Differences among means of continuous data were tested using parametric and non-parametric tests, showing similar results in all cases. Only parametric test results are shown. The Student t test (t) was used for dependent samples while the Wilcoxon test was employed for non-parametric data. A probability (p) of less than 0.05 was considered

statistically significant. All tests were two-tailed. Ninety five percent confidence intervals (CI) were calculated for differences between means. All analyses were performed using version 11.5.1 of the SPSS (Statistical Package for the Social Sciences) statistical package for Windows.

Results

The data for measurements of articulatory rate, intensity and intelligibility of speech, by sentence, word and pseudoword are depicted in Table 1.

Table 2 shows the results of comparisons of each measure for the three stimulus types.

With regard to articulatory rate between

sentences, words and pseudowords, the mean rate for pseudowords was found to be significantly lower than the means of the other types of stimuli. Furthermore, the mean articulatory rate of words proved lower than that of sentences. In terms of speech intensity, the mean intensity of pseudowords was found to be statistically greater than for sentences and words. Comparison of intelligibility measures by speech stimulus types revealed that intelligibility scores for sentences were significantly higher, followed by words score and lastly pseudoword score.

Analysis of the results of the comparisons revealed that speech intelligibility scores were no greater for stimulus produced at lower articulatory rates or with increased speech intensity.

TABLE 1. Descriptive analysis of articulatory rate, intensity and intelligibility of speech, by sentence, word and pseudoword.

| Speech production measure | Mean | Standard Deviation | Median | Minimum | Maximum |
|---------------------------|-------|--------------------|--------|---------|---------|
| RATE- Sentences | 4,69 | 0,49 | 4,72 | 3,58 | 5,69 |
| RATE- Words | 3,00 | 0,34 | 2,97 | 2,26 | 3,89 |
| RATE- Pseudowords | 2,89 | 0,30 | 2,87 | 2,41 | 3,49 |
| INTEN- Sentences | 60,27 | 5,71 | 62,00 | 46,00 | 69,00 |
| INTEN- Words | 60,73 | 6,05 | 62,00 | 46,00 | 71,00 |
| INTEN- Pseudowords | 61,37 | 5,95 | 63,00 | 47,00 | 70,00 |
| INTEL- Sentences | 99,28 | 1,09 | 99,79 | 95,57 | 100,00 |
| INTEL- Words | 96,90 | 2,99 | 97,67 | 88,13 | 100,00 |
| INTEL- Pseudowords | 88,46 | 7,29 | 90,25 | 71,61 | 98,30 |

RATE: rate in syllables per second; INTEN: intensity in decibels; INTEL: intelligibility in percentage of correctly transcribed syllables.

TABLE 2. Comparison among the mean articulatory rate, intensity and intelligibility of speech measures among sentences, words and pseudowords.

| Comparison | t(29) | 95%(CI) | | p |
|------------------------------|-------|---------|----------|---------|
| RATE Sentences-Words | 18,01 | 1,50 | to 1,88 | <0,001* |
| RATE Sentences-Pseudowords | 22,34 | 1,63 | to 1,96 | <0,001* |
| RATE Words-Pseudowords | 2,79 | 0,03 | to 0,18 | 0,009* |
| INTEN- Sentences-Words | -1,63 | -1,05 | to 0,12 | 0,114 |
| INTEN- Sentences-Pseudowords | -3,97 | -1,67 | to -0,53 | <0,001* |
| INTEN- Words-Pseudowords | -3,60 | -0,99 | to -0,27 | 0,001* |
| INTEL- Sentences-Words | 6,22 | 1,60 | to 3,16 | <0,001* |
| INTEL- Sentences-Pseudowords | 9,14 | 8,40 | to 13,25 | <0,001* |
| INTEL- Words-Pseudowords | 8,55 | 6,42 | to 10,47 | <0,001* |

RATE: rate in syllables per second; INTEN: intensity in decibels; INTEL: intelligibility in percentage of correctly transcribed syllables; *p < 0.05

Discussion

The result pertaining to reduced articulatory rate for pseudowords is compatible with the results of an earlier study assessing the variation in articulatory movements for speech tasks in which meaningless sentences produced by speakers with no disorders led to wider and more long-lasting articulatory movements(17).

Although pseudowords, followed by words, were produced at a lower rate, no increases in speech intelligibility were seen when using these stimuli, since intelligibility scores of pseudowords were found to be lower than the other stimuli. Similarly, speech intelligibility scores of words were lower than those of sentences. These findings were not congruent with the hypothesis, based on evidence found in the literature, postulating that intelligibility gains would occur at reduced articulatory rates. However, it is important to bear in mind that these studies differed not only in the populations studied but also in terms of their design. In the studies reviewed, the influence of variation in intra-speaker rate was analyzed using the same type of speech stimulus. In the present study, any improvement stemming from reduced rate may be irrelevant, owing to the reduction in contextual cues in words and chiefly pseudowords. In addition, it is possible that the intelligibility level of the speaker was a factor which interfered in the influence of rate, being evident only among subjects with severe speech impairment(7).

In relation to speech intensity, comparison of stimulus types revealed that only pseudowords were produced with greater intensity overall. Despite the difference found, its relevance is questionable, given that the mean intensity of each stimulus type was very similar, differing by approximately 1dB. In any event, the higher speech intensity observed in pseudowords cannot explain the differences found among intelligibility scores, as these were also found to be lower in pseudowords. The difference in sentence and word intelligibility was not influenced by intensity since

no statistically significant differences were found for this acoustic measure between the two lists. These results are in disagreement with the hypothesis, based on other studies(8,9), which holds that increased intensity in speech production can raise intelligibility. Once again, methodological differences among these studies should be taken into account with respect to the speakers assessed, as well as the use or otherwise of the same stimulus type.

To our knowledge, no other studies comparing possible interference of articulatory rate and intensity of speech on intelligibility measured using three stimulus types have been published. Also, there is no evidence in the literature that listener experience of the intelligibility assessment procedure using orthographic transcription, has an effect on speech intelligibility scores with no disorder(18), since similar findings can be found with lay listeners.

Conclusion

Based on analysis of the results obtained, it can be concluded that the variations in articulatory rate and intensity of the different stimuli produced by the speakers studied did not influence the speech intelligibility measures. Hence, these findings suggest that the signal-dependent information, such as articulatory rate and speech intensity, exert less affect on the speech intelligibility scores than the acoustic signal-independent information in the form of contextual cues, when speakers from the extreme upper end of the intelligibility spectrum are assessed.

It is possible that among subjects with speech disorders whose intelligibility scores approach those found in the normal population, similar patterns of influence from acoustic signal-dependent or independent cues may be found, whereas the same cannot be said for individuals with more severely compromised speech intelligibility.

References

1. Bain C, Ferguson A, Mathisen B. Effectiveness of the speech enhancer on intelligibility: a case study. *J Med Speech-Lang Pathol.* 2005;13(2):85-95.
2. Hustad KC. Effects of speech stimuli and dysarthria severity on intelligibility scores and listener confidence ratings for speakers with cerebral palsy. *Folia Phoniatr Logop.* 2007;59:306-17.
3. Sitler RW, Schiavetti N, Metz DE. Contextual effects in the measurement of hearing-impaired speakers' intelligibility. *J Speech Hear Res.* 1983;26:30-4.
4. Yorkston KM, Hammen VL, Beukelman DR, Traynor CD. The effect of rate control on the intelligibility and naturelness of dysarthric speech. *J Speech Hear Disord.* 1990;55:550-60.
5. Hustad KC, Jones T, Dailey S. Implementing speech supplementation strategies: effects on intelligibility and speech rate of individuals with chronic severe dysarthria. *J Speech Lang Hear Res.* 2003;46(2):462-75.
6. Hammen VL, Yorkston KM, Minifie FD. Effect of temporal alterations on speech intelligibility in Parkinson dysarthria. *J Speech Hear Res.* 1994;37:244-53.
7. Pilon MA, McIntosh KW, Thau MH. Auditory vs visual speech timing cues as external rate control to enhance verbal intelligibility in mixed spastic-ataxic dysarthric speakers: a pilot study. *Brain Inj.* 1998;12(9):793-803.
8. Schulman R. Articulatory dynamics of loud and normal speech. *J Acoust Soc Am.* 1989;85(1):295-312.
9. Tjaden K, Wilding GE. Rate and loudness manipulations in dysarthria: acoustic and perceptual findings. *J Speech Lang Hear Res.* 2004;47(4):766-84.
10. Costa MJ, Iorio MCM, Mangabeira-Albernaz PL. Reconhecimento de fala: desenvolvimento de uma lista de sentenças em português. *Acta Awho.* 1997;16(4):164-73.
11. Tsao Y, Weismer G. Interspeaker variation in habitual speaking rate: evidence for neuromuscular component. *J Speech Lang Hear Res.* 1997;40(4):858-66.
12. Kempler D, Van Lancker D. Effect of speech task on intelligibility in dysarthria: a case study of Parkinson's Disease. *Brain Lang.* 2002;80:449-64.
13. Garcia JM, Crowe LK, Redler D, Hustad K. Effects of spontaneous gestures on comprehension and intelligibility of dysarthric speech: a case report. *J Med Speech-Lang Pathol.* 2004;12(4):145-8.
14. Hanson EK, Beukelman DR. Effect of omitted cues on alphabet supplemented speech intelligibility. *J Med Speech-Lang Pathol.* 2006;14(3):185-96.
15. Hustad KC. Influence of alphabet cues on listeners ability to identify sound segments in sentences produced by speakers with moderate and severe dysarthria. *J Med Speech-Lang Pathol.* 2006;14(4):249-52.
16. Whitehill TL, Wong CC-Y. Contributing factors to listener effort for dysarthric speech. *J Med Speech-Lang Pathol.* 2006;14(4):335-41.
17. Tasko SM, McClean MD. Variation in articulatory movement with changes in speech task. *J Speech Lang Hear Res.* 2004;47(1):85-101.
18. Ellis LW, Fucci DJ. Effects of listeners' experience on two measures of intelligibility. *Percept Mot Skills.* 1992;74(3 Pt 2):1099-104.