# COMPARISON OF THE DURATION OF PHONES AND WORDS IN THE SPEECH OF NORMAL-HEARING AND HEARING-IMPAIRED INDIVIDUALS

Nina Alarotu and Mietta Lennes
*University of Helsinki, Finland*

**ABSTRACT**

The aim of this study was to confirm previously reported differences in the durations of phones and words spoken by normal-hearing and hearing-impaired individuals. For this purpose, utterances of 400 isolated Finnish words were recorded from eight hearing-impaired and four normal-hearing adult subjects. The utterances were manually labeled by two phoneticians and collected into a searchable speech database. Durations of phones and words were measured for each subject. The mean durations of both words (N=4800) and phones (N=27478) were longer for the hearing-impaired subject group. The two groups differed significantly regarding the durations of the words and phones. The results are important regarding both speech recognition technology and clinical methods.

## 1. INTRODUCTION

The hearing-impaired constitute a special target group for applications using automatic speech recognition. The aim of this study was to confirm the durational difference between normal-hearing and hearing-impaired subjects that has previously been reported.

### 1.1 Previous studies

Finnish is known to use phonemic quantity oppositions. This suggests that errors in producing the durational cues for phonemic contrasts of Finnish may affect the intelligibility of speech. Several studies on other languages have shown that hearing-impaired individuals have difficulties in the temporal aspects of their speech production [4, 5, 6, 10]. As for Finnish, it has been reported that the durations of phones of hearing-impaired speakers are generally longer than for normal-hearing speakers [7, 3]. However, although Palomaa [7] collected a large amount of data for his study, he used only one normal-hearing and one hearing-impaired subject. Hurme and Sonninen [3] reported results from 20 hearing-impaired and 21 normal-hearing children along with ten normal-hearing adults but did not include data from hearing-impaired adults. Also, their data consisted of utterances of 13 isolated words only. Therefore, it is desirable to obtain further evidence for this phenomenon.

### 1.2 Present study

Due to technological and methodological development, highly accurate analyses can now be performed on large amounts of speech data. This is possible by building a *speech database* [9], in which different transcription levels and the corresponding speech signals are linked allowing for repetitive searches with different criteria. Nevertheless, the labeling of the database (the segmentation and the linking of transcriptions with segments) must still be done manually in case accurate segment boundaries are required.

This study provides results on the durations of phones and words in a large speech database that has been collected in order to compare the phonetic properties that are typical to the speech of normal-hearing and hearing-impaired people.

## 2. METHODS

### 2.1. Subjects

Eight hearing-impaired (3 males) and four normal-hearing (1 male) adult Finnish subjects participated in the experiment.

Two of the hearing-impaired subjects were profoundly deaf and six severely hearing-impaired. One male subject was hearing-impaired from birth, others postlingually.

### 2.2. Word list

The word list (400 words) was built considering the natural frequencies of words in Finnish speech. First, the 200 most frequently used words were picked out from previously recorded spontaneous speech data spoken by 11 normal-hearing Finnish speakers. This list was supplemented with an additional 200 words from the frequency dictionary of Finnish [8]. A few words were replaced to cover as many diphones as possible without affecting naturalness.

### 2.3. Recordings

The recordings were performed in a meeting room with minor background noise. A Sennheiser high-quality miniature microphone attached to a head set was used for recording and placed a few centimeters to the right from the right corner of the subject's mouth. An identical, randomized list of the 400 words was presented to each subject with a computer system, which also recorded the sound signals straight to the computer's hard disk.

Subjects were asked to read each word appearing on the screen and to pronounce it three times consecutively in their usual speech rate, pitch, and level. They were also instructed to keep a small pause between each utterance to reduce the effect of coarticulation. Subjects were allowed to have short breaks during the recording session.

### 2.4. Building and searching the database

The second utterances (4800 in total) of all recorded word triplets were manually labeled (segmented and transcribed) by two phoneticians. All occurrences of phones were segmented from the speech signal regardless of the phonemic structure of the original stimulus word. A database of the samples was built [9] with the QuickSig object-oriented signal analysis system.

## 3. RESULTS

### 3.1. Durations of words

Durations of all 4800 word occurrences (400 for each subject) were measured as well as the durations of all phones (8853 occurences for the normal-hearing subjects and 18625 for the hearing-impaired). This data was averaged for the two subject groups.

The two subject groups differed significantly ($p<0.001$) with regard to the durations of both phones and words. The average duration of phones and words was longer for the hearing-impaired subjects (Figure 1).

Since the speech material was not labeled phonemically but phonetically, the average number of phones per word could be measured individually and was found to differ between subjects. The hearing-impaired subjects had approximately 5.8 phones per word, whereas the corresponding number for normal-hearing subjects was 5.5.

The standard deviation for the duration of words was 149 ms for the normal and 244 ms for the hearing-impaired subjects (Figure 2). For phones, the standard deviation was 52 ms for the normal-hearing group and 78 ms for the hearing-impaired group (Figure 3).

Factors such as age, amount of hearing loss, and elapsed time from loss of hearing had no significant effect on the durations.

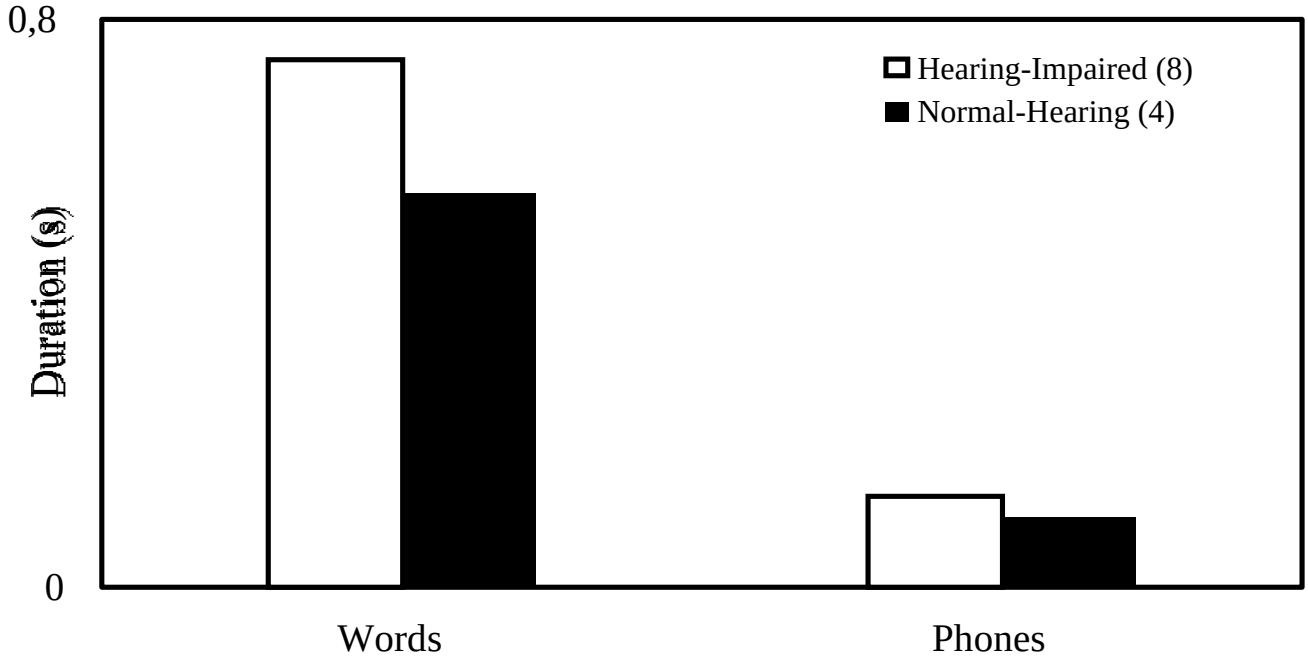


Figure 1. Mean durations of isolated words and phones for the hearing-impaired and the normal-hearing subject groups.

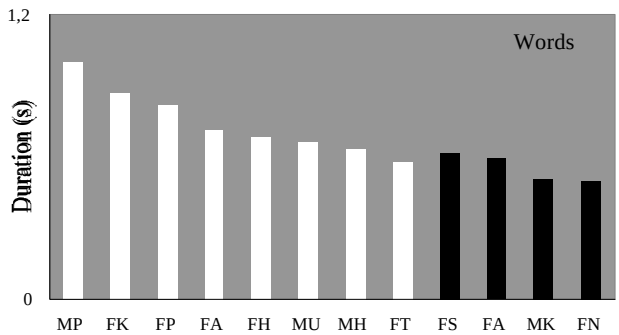


Figure 2. Mean durations of words for each hearing-impaired (white) and normal-hearing (black) subject. The first letter in the subject's name stands for sex (M= male, F= female), the second one is an ID letter.

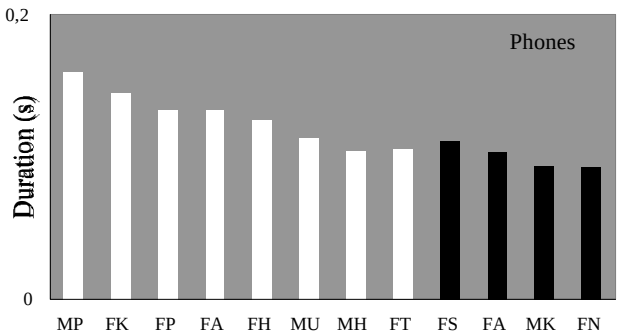


Figure 3. Mean durations of phones for each hearing-impaired (white) and normal-hearing (black) subject. The first letter in the subject's name stands for sex (M= male, F= female), the second one is an ID letter.

## 4. DISCUSSION

As suggested by earlier studies [3, 7], Finnish-speaking hearing-impaired individuals were conclusively shown to produce longer phones and words in comparison to normal-hearing subjects, although the groups did not differ in this respect as dramatically as reported by Palomaa [7] and Hurme and Sonninen [3].

There was a slight difference between the average number of phones per word in the two subject groups, suggesting that hearing-impaired subjects may have a tendency to produce "additional" phones in their speech. This result is in accordance with previous studies (e.g. [2]).

A preliminary auditory analysis [1] of this data has been reported earlier by the authors. Since some minimal pairs based on phonemic quantity seem to have been confused by the hearing-impaired subjects, it seems necessary to look more closely into the relative durations of phones within words.

One of the difficulties this experiment is facing is the fact that only utterances of isolated words were studied. It is not clear whether the observed tendency would be present in running speech samples.

The results will be of great importance in designing and testing automatic speech recognizers as applications for hearing-impaired individuals.

Since the data was collected to a speech database, it may be further analysed for possible spectral differences between the speech sounds of normal-hearing and hearing-impaired speakers.

### REFERENCES


[1] Alarotu, N. and Lennes, M. 1998. Variability of Finnish speech spoken by hearing-impaired individuals. *Proceedings of the Finnic Phonetics Symposium, August 11-14, 1998, Pärnu, Estonia. Linguistica Uralica*, 34, 295-300.

[2] Hudgins, C. and Numbers, F. 1942. An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs*, 25, 289–392.

[3] Hurme, P. and Sonninen, A. 1982. Normaalikuuloisten lasten ja aikuisten sekä kuulovammaisten ja lasten tuottamien KVKV- ja KVKKV-sanojen kestohahmoista. (English summary: Duration patterns in Finnish CVCV and CVCCV words as produced by children and adults with normal hearing and by hearing-impaired children.) *11th Meeting of Finnish Phoneticians – Helsinki. 1982* (ed. A. Iivonen and H. Kaskinen). Publications of the Department of Phonetics, University of Helsinki, 35, 39–53.

[4] McGarr, N. S. and Löfqvist, A. 1982. Obstruent production by hearing-impaired speakers: interarticulator timing and acoustics. *Journal of the Acoustical Society of America*, 72 (1), 34–42.

[5] Metz, D. E., Samar, V. J., Schiavetti, N., Sitler, R. W. and Whitehead, R. L. 1985. Acoustic dimensions of hearing-impaired speakers' intelligibility. *Journal of Speech and Hearing Research*, 28, 345–355.

[6] Nejime, Y. and Moore, B. 1998. Evaluation of the effect of speech-rate slowing on speech intelligibility in noise using a simulation of cochlear hearing loss. *Journal of the Acoustical Society of America*, 103 (1), 572–576.

[7] Palomaa, J. K. 1946. *Suomen kielen äännekestoista puhumaan oppineen kuuromykän ja kuulevan henkilön ääntämisessä. Vertaileva kokeellisfoneettinen tutkimus*. Publicationes Instituti Phonetici Universitatis Helsingforsiensis, 8.

[8] Saukkonen, P., Haipus, M., Niemikorpi, A. and Sulkala, H. 1979. *A Frequency Dictionary of Finnish*. Porvoo: WSOY.

[9] Vainio, M., Altosaar, T., Karjalainen, M., and Iivonen, A. 1995. Experiments with an object-oriented speech database. *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm, Sweden*, 3, 334–337.

[10] Yeni-Komshian, G. H. and Bunnell, H. T. 1998. Perceptual evaluations of spectral and temporal modifications of deaf speech. *Journal of the Acoustical Society of America*, 104 (2), 637–647.