ARCHIVES OF ACOUSTICS
32, 4 (Supplement), 179–185 (2007)

POLISH DIGIT TRIPLET TEST FOR AUDITORY SCREENING: DEVELOPMENT AND INITIAL EVALUATION

Edward OZIMEK, Dariusz KUTZNER,
Aleksander SĘK, Andrzej WICHER

Adam Mickiewicz University, Institute of Acoustics
Division of Room Acoustics and Psychoacoustics
Umultowska 85, 61-614 Poznań, Poland
e-mail: ozimaku@amu.edu.pl

(*received July 15, 2007; accepted November 6, 2007*)

The objective of this study was to develop and evaluate the Polish digit triplet test for speech intelligibility screening. The first stage of the work deals with the preparation of the test and the recording procedure. The second part presents results of the digit triplets intelligibility measurement and retest data. The test consists of 4 statistically equivalent and phonemically balanced lists, each containing 25 different triplets. The constant stimuli paradigm was used to determine the digit triplet psychometric functions. Fifty otologically normal subjects took part in the measurements. The mean list-specific SRT and S_{50} for Polish digit triplet test are -9.4 dB and 19.7%/dB, respectively, and correlate very well with analogous parameters obtained for German digit triplet test, i.e. -9.3 dB and 19.6%/dB. The Polish digit triplet test evaluated in retest experiments revealed high reliability and accuracy for a fast intelligibility screening.

Keywords: digit triplet test, speech intelligibility, psychometric function, masking, auditory screening.

1. Introduction

Digits as a speech material have been used for hearing screening for a long time. They have been used for basic speech intelligibility measurements and for clinical purposes [1], auditory screening [2, 3] and in studies aimed at an influence of speech context on intelligibility [4].

A special category of digit test material are complexes of different digits, namely digit pairs, like 2-5, 8-1, or digit triplets, like 1-4-3, 9-2-1, that spoken as: *two-five*, *eight-one*, or *one-four-three*, *nine-two-one*. Like other speech tests, the triplets are presented to a subject against an interfering noise at various signal-to-noise ratios (SNR). On the basis of such results, the speech-reception-threshold (SRT), i.e., SNR yielding 50% speech intelligibility can be estimated.

The digit triplets have several advantages over single digits or numbers. They produce relatively steep psychometric functions (PF), therefore the SRT is characterized by a relatively small standard deviation (SD). Although SRTs obtained for triplets are lower than those for sentence tests, the results of these tests are highly correlated [2, 3]. Thus, it is possible to predict the SRT for sentence intelligibility on the basis of SRT obtained for triplets. Furthermore, since the measurements of the SRT can be done via telephone or the Internet [2], the digit triplet test can be widely used for extensive screening. However, as the triplets are composed of single digits, they cannot serve as a reliable test for precise speech intelligibility measurements. Unfortunately, at the moment there is no digit triplet test for the Polish language. This fact prompted us to develop such a test aimed at fast speech intelligibility screening.

2. Development of the test material

A single digit triplet is composed of 3 digits of values from 0, 1... to 9. Thus a list containing the total number of possible combinations, i.e. 10^3, was prepared. It has been suggested that a triplet test should contain monosyllabic digits only [2, 3]. However, in Polish language, six out of ten digits are disyllabic ones, therefore all Polish digits were taken into account. Moreover, triplets composed of different digits were considered only. The initial set of 1000 triplets was narrowed down to 720. In the next step, triplets characterized by an approximately equal occurrence probability of numbers 1, 2, ..., 9 across the respective positions were chosen. In this way a set of 160 triplets was selected.

The triplets were read out in a radio studio by a male speaker keeping a natural tempo and intonation. Each sequence was read out and recorded at least twice with the Neumann U87 capacitor microphone. The microphone output fed the input of the Yamaha 02R mixer. The microphone signal was pre-amplified and converted into the digital domain. It was also digitally high-pass filtered at a cut-off frequency of 80 Hz.

3. Measurements of digit triplets intelligibility against noise

3.1. Equipment, procedure and subjects

The signals were generated by means of the Tucker-Davis Technologies System 3 equipment: the RP2 Real Time Processor and the HB7 headphone amplifier and presented monaurally to the subjects via the Sennheiser HD 580 headphones. The triplets were mixed up digitally with the so-called babble triplet noise, i.e. noise generated by means of superposing of all the recorded triplets. A power spectrum of the noise matched optimally a power spectrum related to presented triplets.

Each of 160 triplets was presented to a subject at 7 values of SNR: -14.5; -13.0; -11.5; -10.0; -8.5; -7.0 and -5.5 dB. The level of the noise was kept constant at 70 dB SPL, thus SNR value was determined entirely by the speech signal level. The or-

der of the triplets presentation as well as the SNRs was randomized. Each triplet was allowed to be presented to a given listener 7 times (i.e. at 7 SNRs).

During the measurements the subjects were seated in an acoustically-insulated booth and asked to type on a keyboard what they had just heard. They were asked to confirm the response by typing it once again. The response was stored if both the first and the second response were identical. The subject's response was scored 1 if the entire triplet was repeated correctly, otherwise the response was scored 0. Fifty otologically normal subjects took part in the experiments (22 females and 28 males).

3.2. Determination of the intelligibility function parameters: SRT and steepness (S_{50})

In the next step, the intelligibility data were determined for each triplet. As expected, the intelligibility score depended on SNR and increased monotonically with increasing SNR. On the basis of the intelligibility data, the PFs were fitted, using the least-mean-square (LMS) method. The PF was characterized by two parameters: SRT, i.e. the SNR value yielding 50% correct responses and S_{50}, i.e. the steepness of the PF at the SRT point.

Since both the standardized cumulative distribution function (CDF) [2, 6] and the logistic function (LF) [3, 7, 8] are used as PFs, the intelligibility data were fitted using both CDF and LF functions and, therefore, two SRT (SRT_{CDF} and SRT_{LF}) and two S_{50} (S_{50CDF} and S_{50LF}) were obtained for each triplet. It was found that SRT did not depend on the type of the intelligibility function, while there were some differences in S_{50} obtained for CDF and LF method. However, the difference between S_{50} obtained for LF and CDF functions was about 1 percentage point/dB (LF was slightly steeper). Since the CDF function is directly related to the probability density functions, only SRT_{CDF} and S_{50CDF} were taken into consideration in further analyses, and they are called SRT and S_{50}, respectively. This initial measurement showed the mean SRT $= -9.4$ dB and mean $S_{50} = 15.0$ %/dB.

4. Composition of the final triplet lists

A reliable and accurate test for speech intelligibility measurements must be composed of material characterized by similar and possibly steep PFs [2, 3, 5–8]. Therefore, a final set of triplets was selected from the set of 160 units. The chosen triplets met the following conditions:

- SRT value fell into the range of ± 1 dB with respect to the average SRT obtained for 160 triplets,
- S_{50} of PFs was at least of 13%/dB.

As a result, 100 triplets fulfilling the above conditions were chosen. Mean SRT and mean S_{50} for the selected triplets were -9.4 dB and 21.4 %/dB, respectively.

On the basis of the 100 triplets, 4 lists, containing 25 different triplets were composed by means of a specially prepared algorithm (*Matlab 7.0*). Having a limited phoneme content constituting triplets, it was only possible to generate test material that wass phonemically balanced across the respective lists and did not reflect the phoneme distribution for Polish language. The algorithm performed the following operations:

- generation of a random permutation of 100 triplets and grouping them within of 4 preliminary lists (25 triplets each),
- analysis of statistical properties of the lists,
- analysis of phonemic content of the lists.

These stages were repeated until 4 triplet lists met the following conditions:

- the mean SRT and S_{50} of each list fell into the range ± 0.1 dB and ± 1%/dB, respectively, in comparison with the mean SRT and S_{50} for all selected triplet units (i.e. -9.4 dB and 21.4%/dB),
- distribution of each phoneme fell into the range ± 1.5 percentage points within each list with respect to a reference distribution of all recorded triplets.

The lists meeting these conditions reveal very close phonemic content and produce similar psychometric functions, i.e. they might be regarded as statistically and phonemically equivalent. The left panel of Fig. 1 presents an example of the psychometric functions and list-specific psychometric function (bold line) for list no. 3, while the right panel shows a juxtaposition of the list-specific functions for all the lists. The list-specific psychometric function takes into account the mean steepness of psychometric functions as well as the spread of SRT values within the list and can be computed according to the so-called probabilistic model described by KOLLMEIER [8].

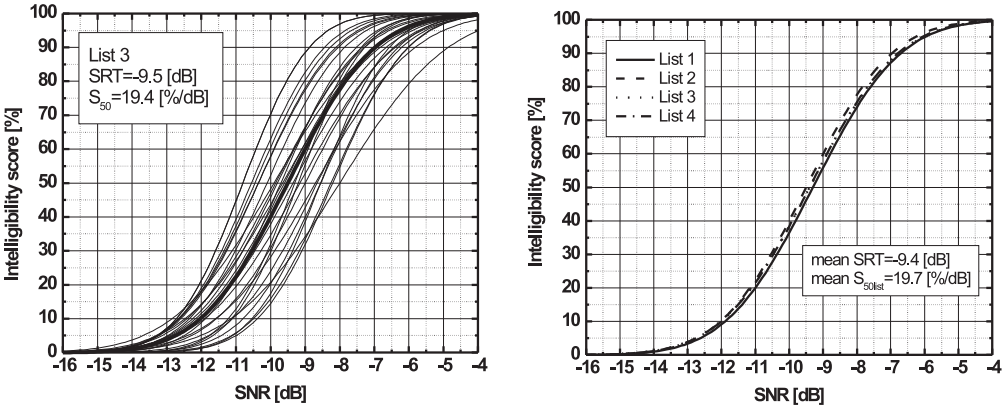

Fig. 1. Individual psychometric functions (thin lines) for the triplets in list 3 (left panel) and list-specific psychometric function (bold line) of the lists no. 3. Right panel depicts juxtaposition of the list-specific psychometric functions for respective digit triplet lists.

Distributions of the phonemes /e/, /l/, /n/, /m/, /St/ and /z/ can be regarded as comparable both for the Polish Sentence Test [5] and the triplet distribution only. However, for

19 out of 39 phonemes the differences are considerably high. This clearly results form extremely limited speech material (10 words only), which certainly does not reflect a phonemic content of everyday speech. There are some differences in SRT and slope of the psychometric functions obtained for the different languages. Only the functions obtained for the Polish and German [3] languages are quite similar.

5. Retest measurements

In order to verify statistical equivalence of the Polish digit triplet lists, two retest experiments were carried out. In the first one, for each list, psychometric functions were derived using a standard constant stimuli paradigm, i.e. intelligibility scores were measured and SRT and S_{50} parameters were determined by means of fitting CDF functions to the intelligibility data. In each retest experiment new group of 10 subjects were tested.

The triplets from respective lists were presented at the following SNR values: -13; -11.5; -10; -8.5 and -7 dB. The order of the triplets presented as well as the order of SNR values was randomized. The order of the list was intentionally kept constant to explore potential learning effect. If such an effect occurred, the lowest value of SRT would be always observed for the list presented as the first. Figure 2 (left panel) depicts intelligibility data obtained in the experiment.

The intelligibility data were analyzed by means of two-way *within subject* analysis of variance. It has turned out that the "list index factor" was statistically insignificant, while SNR was highly statistically significant. Therefore, the ANOVA results revealed that triplet intelligibility was determined only by SNR value and did not depend on the list number. Thus, the lists may be considered as equivalent yielding similar SRT and S_{50}. In this retest experiment the men SRT was equal to -9.7 dB, while $S_{50} = 21.9$ %/dB.

In the second test experiment, the lists were analyzed applying an adaptive *1-up/1-down* procedure and SRTs of the lists were compared. The adaptive procedure with this decision rule converges to the 50% correct point on the PF, i.e. SRT value. In this case, the SNR value depends on the subject's last response and it is decreasing or increasing by some value, i.e. the so-called step, if the subject's response is correct or incorrect, respectively. The initial SNR was set to 0 dB, i.e. the speech was clearly understood at the measurement beginning. The initial step was set to 2 dB and was reduced to 1 dB after the second incorrect response. The SRT was calculated as a mean of SNRs of last 15 of 25 presented triplets. Each subject was presented a given list three times. The order of the triplet list was kept constant to examine whether possible learning effect occurs, while the order of the triplets in a given list was random. An example of the subject response trace using staircase *up/down* adaptive procedure is presented in Fig. 2 (right panel).

The three way ANOVA showed that all the factors and all interactions were not statistically significant, whereas the "subject" factor was marginally significant. In other words, the SRT did not depend on the triplet list. What is more, the "position" factor

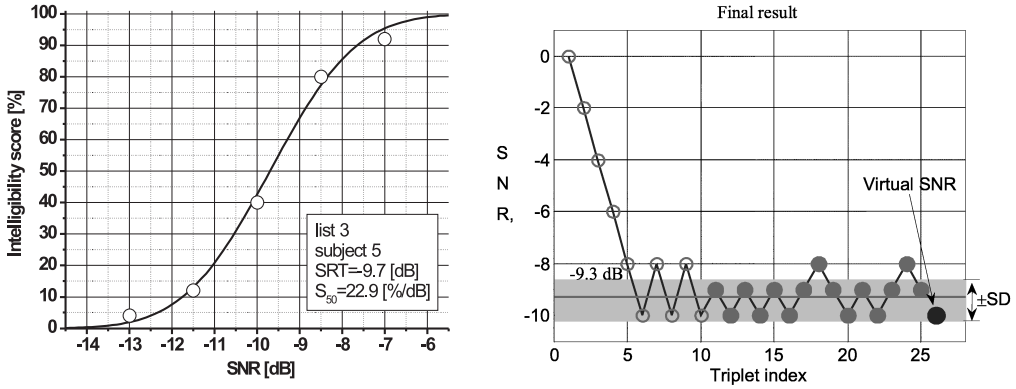

Fig. 2. An exemplary psychometric function obtained in the retest measurement for the constant stimuli paradigm (left panel) and an example of the subject response trace using staircase up/down adaptive procedure (right panel).

was found to be insignificant, i.e. the learning effect was not observed. In this retest experiment the SRT was equal to -10.0 dB, i.e. it was very close to SRT obtained by means of the constant stimuli paradigm.

6. Conclusions

The experiments performed enable formulation of the following conclusions:

- The proposed four lists of 25 triplets each lead to very similar intelligibility (SRT) and steepness (S_{50}) of the psychometric function.
- The SRT measured by means of the triplet test is approximately 3 dB lower that that measured this for the Polish Sentence Test [5]. A similar difference is observed for other languages [2].
- The proposed triplet test yields similar psychometric function to those obtained for German language [3].
- The results of the retest experiments are highly consistent with the preliminary measurements. The adaptive method turned out to be very useful and fast in determination of the SRT using the triplet test.
- The triplet test seems to be an appropriate tool for fast auditory screening that can be used not only in clinics but also via Internet and telephone line.

Acknowledgments

This work was supported by the European Union FP6, Project 004171 HEARCOM and the State Ministry of Science and Higher Education grant.

References

[1] FLETCHER H., *Speech and hearing in communication*, Van Nostrand, New York 1953.

[2] SMITS C., KAPTEYN T., HOUTGAST T., *Development and validation of an automatic speech-in-noise screening test by telephone*, International Journal of Audiology, **43**, 15–28 (2004).

[3] WAGENER K., EEENBOOM F., BRAND T., KOLLMEIER B., *Ziffer-Tripel-Test: Spracherverstandlichkeitstest uber das Telefon*, [in:] Tagungs-CD der DGA Jahrestagung, 2005.

[4] KALIKOW D. N., STEVENS K. N., ELLIOTT L. L., *Test of speech intelligibility in noise using sentences with controlled word predictability*, Journal of Acoustical Society of America, **60**, S28 (1976).

[5] OZIMEK E., KUTZNER D., SEK A., WICHER A., SZCZEPANIAK O., *The Polish sentence test for speech intelligibility evaluations*, Archives of Acoustics, **31**, 4, 431–438 (2006).

[6] VERSFELD N. J., DAALDER L., FESTEN J. M., HOUTGAST T., *Method for the selection of sentence material for efficient measurement of the speech reception threshold*, Journal of Acoustical Society of America, **107**, 1671–1684, 2000.

[7] KOLLMEIER B., WESSELKAMP M., *Development and evaluation of a sentence test for objective and subjective speech intelligibility assessment*, Journal of Acoustical Society of America, **102**, 4, 1085–1099 (1997).

[8] KOLLMEIER B., *Messmetodik, Modellierung und Verbeserung der Verstandlichkeir von Sprache*, Georg-August-Universtat, Gottingen 1990.