

FP6-004171 HEARCOM
Hearing in the Communication Society

INTEGRATED PROJECT
Information Society Technologies

**D-1-11: Telephone implementation of
Polish Digit Triplet Test**

Contractual Date of Delivery:	Extra report
Actual Date of Submission:	24-02-2009
Editor:	E. Ozimek
Sub-Project/Work-Package:	SP1/WP1
Version:	1.1
Total number of pages:	32

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	
Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006) This information is confidential and may be used only for information purposes by Community Institutions to whom the Commission has supplied it		

VERSION DETAILS

Version:	1.1
Date:	24-02-2009 / 22-02-2010 (public)
Status:	Submitted

CONTRIBUTOR(S) to DELIVERABLE

Partner	Name
PL - IAM	Edward Ozimek
PL - IAM	Andrzej Wicher
PL - IAM	Anna Warzybok
PL - IAM	Dariusz Kutzner
PL - IAM	Paweł Libiszewski

DOCUMENT HISTORY

Version	Date	Responsible	Description
1.0	14-10-2008	Edward Ozimek	First draft
1.1	24-02-2009	Edward Ozimek	Final version after review
1.1	22-02-2010	M. Vlaming	Changed into public report

DELIVERABLE REVIEW

Version	Date	Reviewed by	Conclusion*
1.0	21-01-2009	Mark Lutman	Accept
1.0	03-02-2009	Michael Buschermöhle	Accept with modifications

* e.g. Accept, Develop, Modify, Rework, Update

Table of Contents

Pre-Amble	6
1 Executive Summary	7
2 Introduction.....	8
3 Laboratory version of PDTT	10
3.1 <i>Selection of digit triplets and recording session.....</i>	<i>10</i>
3.2 <i>Measurements of triplet-specific intelligibility functions</i>	<i>10</i>
3.3 <i>Composition of final digit triplet test</i>	<i>11</i>
3.4 <i>Retest measurements.....</i>	<i>13</i>
3.5 <i>Across-language comparison of digit triplet tests</i>	<i>14</i>
4 Telephone version of PDTT	16
4.1 <i>Implementation of PDTT in the telephone system</i>	<i>16</i>
4.1.1 Set-up for speech-in-noise measurement via telephone.....	16
4.1.2 Procedure of automatic speech-in-noise measurement	17
4.1.3 Website working with telephone version of PDTT.....	23
4.1.4 Database structure	24
4.2 <i>Telephone version of PDTT – evaluation measurements.....</i>	<i>26</i>
4.3 <i>Comparison of laboratory and telephone versions of PDTT.....</i>	<i>28</i>
5 Conclusions and future plans.....	29
6 References	31

List of Figures

Fig. 3.1 <i>List-specific intelligibility functions for PDTT</i>	12
Fig. 3.2 <i>Comparison of the mean list-specific intelligibility functions for digit triplet tests across languages</i>	15
Fig. 4.1. <i>A block diagram of the prototype system used in the telephone version of PDTT</i>	17
Fig. 4.2. <i>A procedure of automatic speech-in-noise measurement used in a telephone version of PDTT – flow chart</i>	22
Fig. 4.3. <i>Website working with the telephone version of PDTT – log-in screen</i>	23
Fig. 4.4. <i>Example of a graphic presentation of the data obtained with the staircase procedure</i>	24
Fig. 4.5. <i>The mean SRTs (averaged across telephones and callers) for 4 lists and corresponding standard deviations (error bars).</i>	27
Fig. 4.6. <i>The mean SRTs (averaged across lists and callers) for 4 telephones and corresponding standard deviations (error bars).</i>	27

List of Tables

Tab. 4.1. <i>Characterization of the laboratory and the telephone versions of PDTT – comparison</i>	28
---	----

Acknowledgement

Supported by grants from the European Union FP6, Project 004171 HEARCOM. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

Pre-Amble

This deliverable is devoted to implementation of the Polish Digit Triplet Test (PDTT) for auditory self-screening via a telephone. The deliverable is composed of two parts. The first one outlines development and evaluation of a laboratory version of Polish Digit Triplet Test. The second one addresses implementation of the test into a telephone system. A prototype system, test procedure and results of evaluation measurements of the telephone version of PDTT are given. The digit triplet test is available for several languages (D-1-8) and is one of the major type of screening test developed within a framework of the HearCom project. The present test reveals a synergy with other Deliverables (D-1-6, D-1-8) and tests developed within the HearCom project.

1 Executive Summary

This deliverable describes implementation of the Polish Digit Triplet Test (PDTT) in the telephone system. PDTT is an intelligibility screening test based on digit triplets composed of limited vocabulary (ten digits). During the test the subject is presented with speech signals at a background of masking noise, at different signal-to-noise ratios (SNR). Measurement result is expressed as speech reception threshold (SRT) i.e. SNR yielding 50% speech intelligibility.

The deliverable is composed of two parts. The first part deals with the preparation and evaluation of the laboratory version of PDTT. This is actually a concise description of the test preparation. Detailed report on the PDTT development can be found in Ozimek et al., (2008).

The second part addresses the implementation of PDTT into the telephone system. This part describes a prototype system designed to realize intelligibility testing via the telephone and presents preliminary data and general evaluation of the telephone version of PDTT. After dialling, the caller is presented with a voice menu and basic information on the caller is collected. Subsequently, the caller is presented with one of 4 PDTT lists and SRT is determined. Finally, the caller is informed about his/her intelligibility data related to the normative data. Also a web site has been developed at which the results of self-screening are presented in a graphic form.

The data showed that the telephone SRTs were somewhat higher than those obtained in the laboratory conditions, as expected. Furthermore, the speech intelligibility did not depend on the telephone model used in the measurements. Hence, it is justifiable to say that the telephone version of PDTT has been validated successfully and can be used in the programme of national speech intelligibility screening via the telephone system.

Proposals for improvement of the telephone version of PDTT as well as a campaign encouraging people to consider intelligibility self-testing are outlined in this deliverable.

2 Introduction

Digits and numbers as a test material have been used in speech perception measurements for a long time. They have been used for speech intelligibility measurements and for clinical purposes (Fletcher, 1929; Pruszewicz et al., 1994a; Pruszewicz et al., 1994b), intelligibility and auditory screening (Smits, 2005; Wagener et al., 2005), in studies on the influence of speech context on intelligibility (Miller et al., 1951; Kalikow et al., 1977; Ozimek et al., 2006), on pronunciation differences between native and non-native American English speakers (Schmidt-Nielsen, 1989), in evaluation of central auditory processing (Bellis, 1996), in experiments on dichotic speech intelligibility for various signal bandwidths (Strouse et al., 2000), in a study of digit recognition in the presence of different kinds of masking signals (Smits et al., 2007), etc. More recently, digits have been used in a study concerning the intelligibility of speech presented in a multi-talker noise for normally-hearing and hearing-impaired subjects (Wilson et al., 2005).

A special category of digit test material comprises complexes composed of different digits like digit pairs, digit triplets etc. Like standard word or sentence tests, these complexes are usually presented to the subject against interfering noise at various signal-to-noise ratios (SNR). On the basis of the test results, the speech-reception-threshold (SRT), i.e. SNR yielding 50% speech intelligibility, can be estimated. Digit complexes like triplets have several advantages over single digits. First of all, they produce steep intelligibility functions, so the SRT estimate is characterized by a relatively low standard deviation (Smits et al., 2004). Digit triplets were proved to provide accurate SRT values, without making demands on the patient's cognitive abilities (memory) (Smits et al., 2004). Besides, since digit triplets have no context, it is hard to learn them all by heart, especially when they are presented in a random order. Although SRTs obtained for the digit triplet test are lower than those for the sentence test, the results of both tests are correlated (Smits et al., 2004). Furthermore, since measurements of the SRT can be made via telephone or the Internet, the digit triplet test can be widely used for extensive screening of speech intelligibility. Digit triplet tests have been so far developed for Canadian English (Rudmin, 1987), Danish (Elberling et al., 1989), American English (Wilson et al., 2004), English (Hall, 2006), Dutch (Smits, 2005) and German (Wagener et al., 2005).

The telephone version of triple-digit tests, and publicity campaigns promoting auditory self-screening, have been so far introduced in the Netherlands, England, Germany and France. The campaigns generally aim at increasing the awareness in society of hearing impairment by means of

encouraging people to perform intelligibility self-testing. Apart from obvious benefits for the general population, the data gathered during measurements provide insight into the hearing abilities of specific groups of the population.

As far as the Polish language is concerned, the Polish Digit Triplet Test has been recently developed within the framework of the HearCom project. However, the test has been available so far only in the laboratory version. This derivable focuses on the implementation of PDTT in the telephone system, to be useful for 'remote' self-screening by the Polish population.

3 Laboratory version of PDTT

This chapter addresses the development of the laboratory version of PDTT, which progressed through the following stages: initial selection of digit triplets, recording session, triplet intelligibility measurements, selection of a homogeneous group of triplets and composition of statistically and phonemically equivalent lists containing different digit complexes. Below only a concise review of development of the PDTT is presented while details on the experimental procedure, measurements and final results of the PDTT are given in deliverable D-1-6a (Ozimek et al., 2007) and in the paper by Ozimek et al. (2008).

3.1 Selection of digit triplets and recording session

The digit triplet test contains sequences of three digits from 0, 1 to 9 spoken separately¹. From all possible 10^3 combinations of triplets, only 160 triplets were selected in such a way that repeated digits were excluded and for each digit the probability of occurrence in any position in triplets was approximately equal. The 160 triplets were read out in a radio studio by a male Polish native speaker. The speaker was asked to keep a natural intonation, approximately the same loudness level and vocal effort over time. After the recording all the signals were edited and stored on a PC hard disc as separate sound files (*.wav).

3.2 Measurements of triplet-specific intelligibility functions

For each triplet, speech intelligibility was measured by means of the constant stimuli paradigm. Each of 160 triplets was presented to the subject at 7 values of SNR: -14.5; -13.0; -11.5; -10.0; -8.5; -7.0 and -5.5 dB. The noise sound pressure level was kept constant at 70 dB SPL, thus SNR value was determined entirely by the speech signal sound pressure level. Both, the order of the triplet presentation as well as that of SNRs were randomized. The triplets were mixed digitally with interfering noise and presented monaurally via the Sennheiser HD-580 headphones to the subjects. The signals were generated by means of the TDT-3 equipment (the 24-bit real time signal processor RP2, the headphone

buffer HB7). The interfering noise was the so-called digit triplet noise, i.e. the noise generated by means of superposition of all the recorded triplets.

Fifty normal-hearing subjects took part in the experiments. During the intelligibility measurements the subject was seated in an acoustically-insulated booth and asked to type on a keyboard what he/she had just heard. The subject's response was scored 1 if the entire triplet was repeated correctly, otherwise the response was scored 0 (the so-called triplet scoring). The subject had to confirm his/her response by means of typing it again. The 'double response' method reduces lapses (i.e. for example typing errors made to audible stimuli) and consequently spread of data is decreased. Each subject was presented with 1,120 triplets (7 SNRs*160 units). The total number of collected and analyzed responses was 56,000.

In the next step, the intelligibility data were determined for each triplet, i.e. the proportions of correct responses obtained for the respective SNRs and for all subjects were computed. Cumulative distribution functions (CDF) were fitted to the intelligibility data using the maximum likelihood criterion. The intelligibility function was characterized by two main parameters, namely: SRT (i.e. the SNR value yielding 50%-probability of correct responses) and S_{50} (i.e. the slope of the function at the SRT point). Finally, 160 triplet-specific intelligibility functions and 160 corresponding SRT and S_{50} values were determined.

3.3 Composition of final digit triplet test

It is widely known that reliable tests for speech intelligibility measurements must be composed of statistically homogenous speech material; i.e. characterized by similar and steep psychometric functions. Therefore, it was decided that the 'optimal' digit triplets should meet the following criteria: SRT values should fall into the range of ± 1.5 dB with respect to the average SRT obtained for 160 triplets; S_{50} values should be at least 13 %/dB. As a result, 100 triplets fulfilling the above conditions were chosen. Mean SRT and mean S_{50} for the selected 100 triple digit complexes were -9.4 dB and 21.4 %/dB, respectively. On the basis of the 100 triplets selected, 4 triplet lists containing 25 different triplets were composed. A dedicated algorithm was used to generate statistically equivalent (i.e. yielding similar SRT and $S_{50\text{list}}$) and phonemically equivalent (i.e. yielding comparable phoneme distributions) lists. The algorithm performed the following operations:

- Generation of a random permutation of triplet indexes and formation of 4 preliminary lists, each comprising 25 triplets;
- Analysis of the mean SRT and S_{50} characterizing each list;
- Analysis of phonemic content of the lists.

These operations were repeated until 4 lists, composed of different digit triplets, met the following conditions:

- The mean SRT and $S_{50\text{mean}}$ characterizing each list fell into the range of ± 0.1 dB and ± 1 %/dB, respectively, of the mean SRT and mean S_{50} for the 100 selected triplets, i.e. -9.4 dB and 21.4 %/dB, respectively (i.e. the lists are composed of different triplets, but yielding similar intelligibility);
- For each list, a distribution of each phoneme fell into the range of ± 1.5 percentage points with respect to the reference phoneme distribution characterizing phonemic content of all the recorded triplets (i.e. the lists are composed of different triplets, but revealing comparable phoneme distributions).

The lists meeting these conditions revealed very close phonemic content and produced similar list-specific intelligibility functions; i.e. they might be regarded as statistically and phonemically equivalent. Although phoneme distributions are close across lists, it should be stressed that the lists are not phonemically representative for the Polish language since the test comprises strongly limited lexicon (only ten words). Fig. 3.1 presents list-specific intelligibility functions for PDTT.

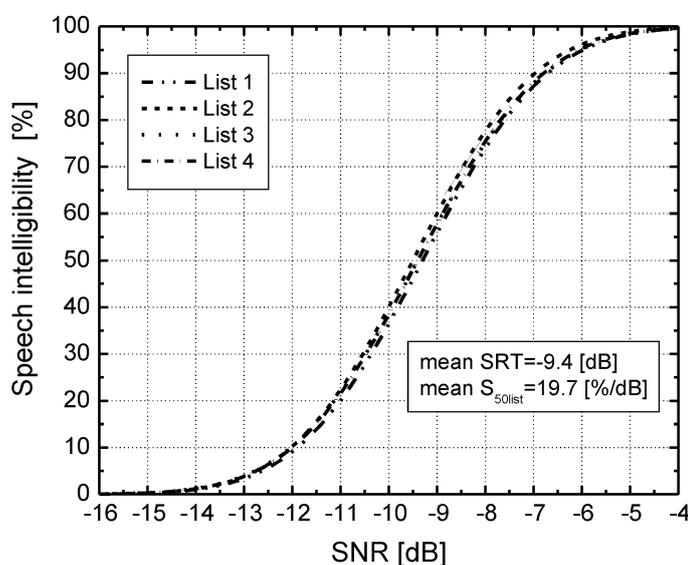


Fig. 3.1 List-specific intelligibility functions for PDTT

As can be seen, the lists produce very similar SRT values. Standard deviation across SRT values in each list does not exceed 1 dB. The statistical equivalence of the respective lists was confirmed by the results of ANOVA test applied to the baseline data. The results of the test proved that SRT did not depend statistically on the list index $\{F(3,99)=0.19, p=0.9\}$.

3.4 Retest measurements

In order to verify the statistical equivalence of PDTT lists, three retest experiments have been carried out. Another group of twenty 'fresh' normal-hearing listeners took part in each experiment. In the first one, individual intelligibility functions were derived for each list using a standard constant stimuli paradigm (i.e. intelligibility scores were measured, and SRT and S_{50} parameters were determined by means of fitting the cumulative distribution function to the intelligibility data). In the second one, statistical parameters characterizing the respective lists were analyzed by application of the adaptive staircase procedure with the *1-up/1-down* decision rule, and SRTs of the lists were compared. Both the first and second test-retest measurements were performed using the same apparatus as in the main investigation. Since, in the Internet hearing screening scenario, measurements were carried out via commonplace equipment (i.e. PCs equipped with a standard sound card applying the adaptive *up/down* method), the third retest experiment was conducted using a standard PC notebook computer in non-laboratory conditions (office).

In the first retest experiment, the triplets for respective lists were presented at the following SNR values: -13.0; -11.5; -10.0; -8.5 and -7.0 dB. The order of the lists, SNRs and triplets presented were randomized. The SRTs obtained for the respective lists and subjects were very close; i.e. the results of the measurements confirmed the expected reliability of the lists developed. Statistical analysis was performed to examine the equivalence of the four lists constituting the Polish digit triplet test. ANOVA results revealed that triplet intelligibility score was determined entirely by SNR value and did not depend on the list index used in the measurements. Thus, the lists may be considered as fully equivalent, yielding similar intelligibility data.

In the second and third retest experiment, an adaptive procedure with *1-up/1-down* decision rule was employed to determine speech intelligibility. The initial value of SNR was -3 dB, which was far above the expected SRT; i.e. the speech was perfectly understood at the beginning of the measurement. In order to make the procedure converge to the SRT point as quickly as possible, the initial step was set to 2 dB and was reduced to 1 dB after the third incorrect response. During measurements, 25 triplets were presented, while SRT was calculated as a mean of the last 15 SNRs at which the signals were presented, including the so-called 'virtual SNR' (i.e. SNR at which the 26th digit triplet would be presented, if it existed).

In measurements with laboratory equipment, like in the case of retest measurements by the constant stimuli method, the SRT values obtained for the adaptive procedure were close to the expected values, i.e. the

differences between 'expected' SRT and a result of the retest measurement did not exceed ± 0.3 dB. The obtained SRTs were pooled across listeners and subjected to one-way within-subject ANOVA (with individual SRTs as repeated measurements) with respect to the 'list' factor, which revealed that the differences in SRT across the lists were not statistically significant $\{F(3,79)=0.77, p=0.51\}$.

In the retest experiment with PC and standard sound card, prior to the measurement session, the listeners were allowed to adjust loudness of the sound to their comfort. The SRTs collected were pooled across subjects and analyzed by means of one-way within-subject ANOVA (with individual SRTs as repeated measurements). Again, it turned out that the 'list' factor was not statistically significant $\{F(3,79)=0.71, p=0.55\}$. What is more, measurements in an ordinary room by using a PC equipped with a standard sound card yielded SRTs close to those obtained in laboratory conditions (i.e. the differences between 'expected' SRT and a result of the measurement for non-professional equipment did not exceed ± 0.3 dB).

3.5 Across-language comparison of digit triplet tests

Fig. 3.2 presents a comparison of the mean list-specific intelligibility functions for the Polish digit triplet test (solid lines) with respect to the mean list-specific intelligibility functions for digit triplet tests developed for several European languages, some within HearCom. It should be emphasized that due to certain differences in linguistic structure and measurement methods used in different laboratories, only a general comparison is possible. All the intelligibility functions were obtained in measurements via headphones.

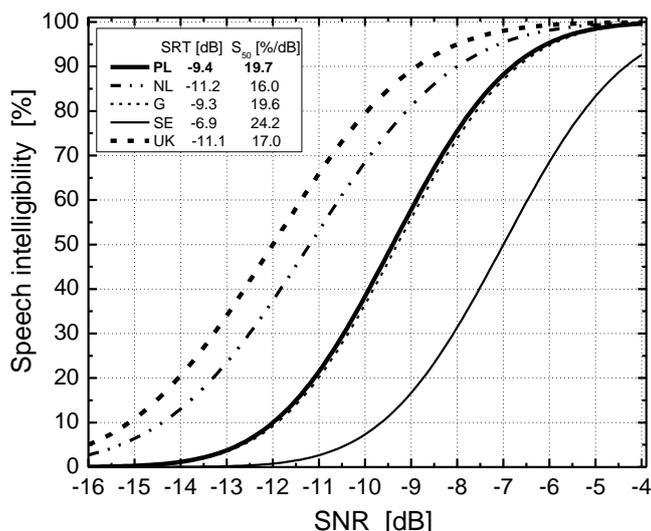


Fig. 3.2 Comparison of the mean list-specific intelligibility functions for digit triplet tests across languages

As can be seen, SRT varies across languages from -11.2 dB (the Dutch test) to -6.9 dB (Swedish test); the shallowest slope is 16.0 %/dB (the Dutch test), while the steepest slope is observed for the Swedish materials. The functions characterizing the German and Polish tests are almost identical. Factors affecting differences between parameters of the intelligibility functions across languages might result from: different experimental procedures, differences in linguistic structure of languages, differences across speakers, type of masker used etc.

4 Telephone version of PDTT

The second part of the project concerns the implementation of the PDTT in the Polish telephone system to be used by people interested in intelligibility self-testing. This chapter describes a prototype system designed to realize such testing via telephone and the procedure of automatic speech-in-noise measurement. Also preliminary data as well as results of evaluation of measurements are presented.

4.1 Implementation of PDTT in the telephone system

4.1.1 Set-up for speech-in-noise measurement via telephone

To carry out digit triplet intelligibility test via a telephone line, a PC server (the processor Intel Core DUO 2 GHz, 2 GB RAM, 250 GB HDD) equipped with the telephonic card DIVA SERVER 4BRI was connected to the telephone line. The prototype system described in this chapter enables 4 simultaneous independent callers to be tested (with the possibility of extension in the future of the number of connections). The testing is controlled by dedicated software implemented in the Axxium 6.0 environment. The signals are processed at 16-bit resolution and at a sampling rate of 8 kHz. Prior to the implementation, the original .wav files (i.e. laboratory version) were re-sampled and stored at the hard disc as .cal files that are standard files used in telecommunication (CAPI A-law files). The data collected are stored in an MS SQL data base. Furthermore, a special website, which offers the callers access to graphic representation of individual results has been developed. Fig. 4.1 presents a general scheme of the system realizing the telephone version of PDTT.

Subsequently, meaningful and short comments for a voice menu were written and recorded in the same radio studio and by the same speaker who participated in the recording of the triplets (section 3.1). The recorded utterances were also converted to .cal files.

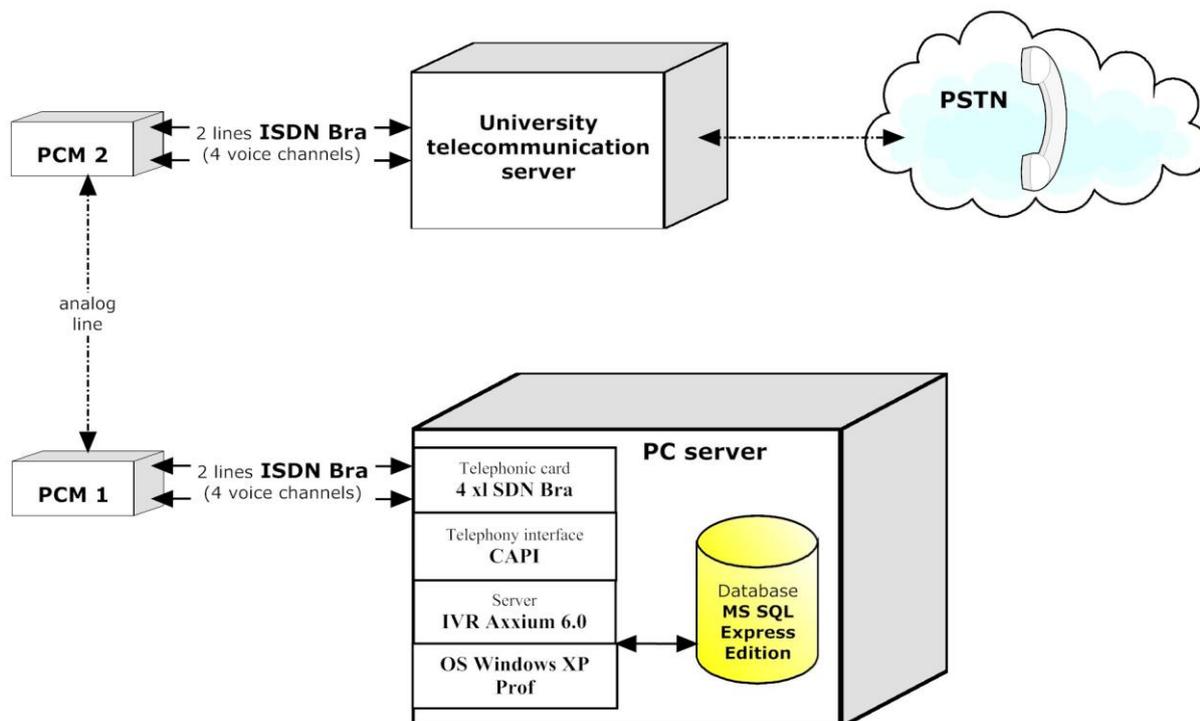


Fig. 4.1. A block diagram of the prototype system used in the telephone version of PDTT

4.1.2 Procedure of automatic speech-in-noise measurement

An integral part of PDTT is a voice menu that supports measurement and gives the caller the detailed instructions on the test. The voice menu is composed of different parts, which are presented below (Polish version – italic-blue font and the corresponding English translation - italic-red font). Expressions in brackets [] describe actions performed by the system in response to the caller’s performance.

The procedure for automatic speech-in-noise measurement via the telephone is as follows. Firstly, the caller is welcomed and he or she is presented with basic information about the test (i.e. purpose, duration, a recommendation of using only land-line phones).

Witamy w teście zrozumiałości mowy opracowanym w Instytucie Akustyki Uniwersytetu imienia Adama Mickiewicza w Poznaniu, w ramach projektu europejskiego HearCom. Test umożliwia ocenę zrozumiałości mowy prezentowanej na tle szumu i będzie trwać około pięciu minut. Zaleca się przeprowadzenie testu za pomocą telefonu stacjonarnego. Stosowanie telefonu komórkowego, może prowadzić do błędnych wyników.

Welcome to a speech intelligibility test developed at the Institute of Acoustics, Adam Mickiewicz University in Poznan within the framework of the European HearCom project. The test allows measurement of intelligibility of speech presented at a background of noise and lasts about five minutes. It is recommended to use a landline phone for the test performance. If a mobile phone is used, the results might be erroneous.

After that the caller is asked to type his/her age using a telephone keypad.

Podaj ile masz lat, używając klawiatury telefonu i naciśnij gwiazdkę [czekanie na odpowiedź, jeśli jej brak]: Brak odpowiedzi. Dziękujemy za udział w badaniu.

Please provide your age using the keypad and press the star key [waiting for the answer, if there is not any]: No answer. Thank you for your participation in the test

The next question is related to the sex of the caller.

Jeśli jesteś kobietą wybierz 1.

If you are a woman press 1.

Jeśli jesteś mężczyzną wybierz 2.

If you are a man press 2.

The information concerning caller's sex influences the grammatical structure of utterances presented in the voice menu. This is due to the fact that in Polish language some expressions in the menu have to be gender-specific, for instance the English comment "You have not given response" is translated into Polish as "*Nie udzielieś odpowiedzi*" for a man, but "*Nie udzieliłaś odpowiedzi*" when the caller is a woman.

If no response is given by the caller, the test procedure stops.

Nie udzielono odpowiedzi. Dziękujemy za udział w teście.

No answer. Thank you for your participation in the test.

Next the caller is presented with a precise instruction about the test (characteristics of the stimuli, task, a way of giving response etc.) and the actual measurement begins. Round brackets contain gender-specific suffixes ("aś" female suffix, "eś" male suffix).

Za chwilę usłyszysz zestawy trzech cyfr, których zrozumiałość będzie zmieniać się w czasie trwania testu. Po prezentacji każdego zestawu wybierz na klawiaturze telefonu cyfry, które usłyszałaś (-eś) i naciśnij

gwiazdkę. Cyfry wpisz w takiej kolejności, w jakiej usłyszałaś (-eś). W przypadku, gdy nie rozumiałaś (-eś) żadnej cyfry, również naciśnij gwiazdkę. Uwaga, rozpoczynamy test.

[w przypadku braku odpowiedzi]

Wpisz cyfry w takiej kolejności, jakiej usłyszałaś (-eś).

You will hear sequences of three digits. The intelligibility of the speech will be varying during the test. Once each sequence has been presented, press the number keys corresponding to digits you have just heard, and press the star key. Enter the digits in the same order as you have heard them. If you did not understand any digit at all, also press the star key. Attention please, we are now starting the test

[if there is no answer within one minute from presenting speech]

Press the number keys corresponding to digits you have just heard, and press the star key

The test procedure is very similar to that of the laboratory measurements (an adaptive procedure with the 1-up/1-down decision rule, step decreased from 2 dB to 1 dB after the first incorrect response, triplet scoring), except for the presentation mode (a telephone receiver) and loudness level (might be adjusted by the caller). One of four digit triplet lists is randomly selected (according to uniform distribution, i.e. a probability of choosing of each lists is 0.25). During the test, 25 digit triplets are presented in a random order. After typing digits in an appropriate order, the caller is asked to accept the response by pressing the "*" key (functioning as the 'enter' key).

If the caller still does not give any answer for 60 seconds, the test stops and the information on the test interruption is presented and, consequently, the test is aborted.

Nie udzieliłaś(-eś) odpowiedzi. Brak Twojej odpowiedzi uniemożliwia kontynuację testu. Dziękujemy za udział w badaniu.

You have not given any answer. Therefore it is impossible to continue with the test. Thank you for your participation in the test.

If the caller's performance is totally unsatisfactory, i.e. if nine out of ten first answers are wrong, further measurement is pointless, because the SNR constantly increases instead of converging to 50%-intelligibility value. In this case the test stops and the information of test termination is given and the line is disconnected.

Z uwagi na niewystarczającą powtarzalność Twoich odpowiedzi, proponujemy ponowne przeprowadzenie testu. Dziękujemy.

Because your response have turned out to be incoherent, it is not possible to continue the test. We recommend trying again. Thank you.

Having completed responses to all the triplets, the caller is informed that the test procedure is finished.

To był ostatni zestaw cyfr.

That has been the last sequence of digits.

Subsequently, an individual SRT is computed as a mean of 16 last SNRs (i.e. ten first responses are rejected) and compared to the normative data (confidence interval derived for measurement via a telephone for normal-hearing listeners).

If SRT is lower than the upper limit of the 95% confidence interval (-5.4 dB), the information about normal speech intelligibility is generated. If the SRT is larger than -5.4 dB, the caller is informed about his or her poor performance. A recommendation of visiting an ENT doctor is presented to the caller, when needed.

Wynik badania wskazuje, że Twoja zrozumiałość mowy jest prawidłowa. Jeśli jednak uważasz, że masz problemy ze słuchem zgłoś się do laryngologa lub protetyka słuchu.

Your performance has turned out to be satisfactory. However, if you still feel that you have some problems with hearing or ears, please visit an ENT doctor or a hearing care professional.

...or

Wynik badania wskazuje, że masz problemy ze zrozumiałością mowy. Zalecamy wizytę u laryngologa lub protetyka słuchu.

The test result has showed that your performance was poor. We recommend visiting an ENT doctor or a hearing care professional for further tests.

After measurements, the caller is informed about the website at which the result of measurement is presented in a graphic form. The caller is given a three-digit code which in a conjunction with their first six digits of telephone number enables logging to the unique listener's account.

Na stronie internetowej www.ucho.amu.edu.pl możesz zobaczyć szczegółowy wynik twojego badania. Powtarzamy adres strony: www.ucho.amu.edu.pl. Po wejściu na stronę wprowadź numer telefonu, z którego dzwonisz oraz wprowadź następujący kod ... Powtarzamy kod...

You can view detailed results of your test at www.ucho.amu.edu.pl. We repeat the website address www.ucho.amu.edu.pl. Upon accessing the website, enter the first six digits of your phone number you were calling from and type in the following code... We repeat the code...

At the end, the caller's participation is acknowledged.

Dziękujemy za udział w badaniu.

Thank you for your participation in the test.

The test procedure lasts approximately 5 minutes per ear. Fig. 4.2 presents a flow chart of the automatic speech-in-noise intelligibility measurement used in the telephone version of PDTT.

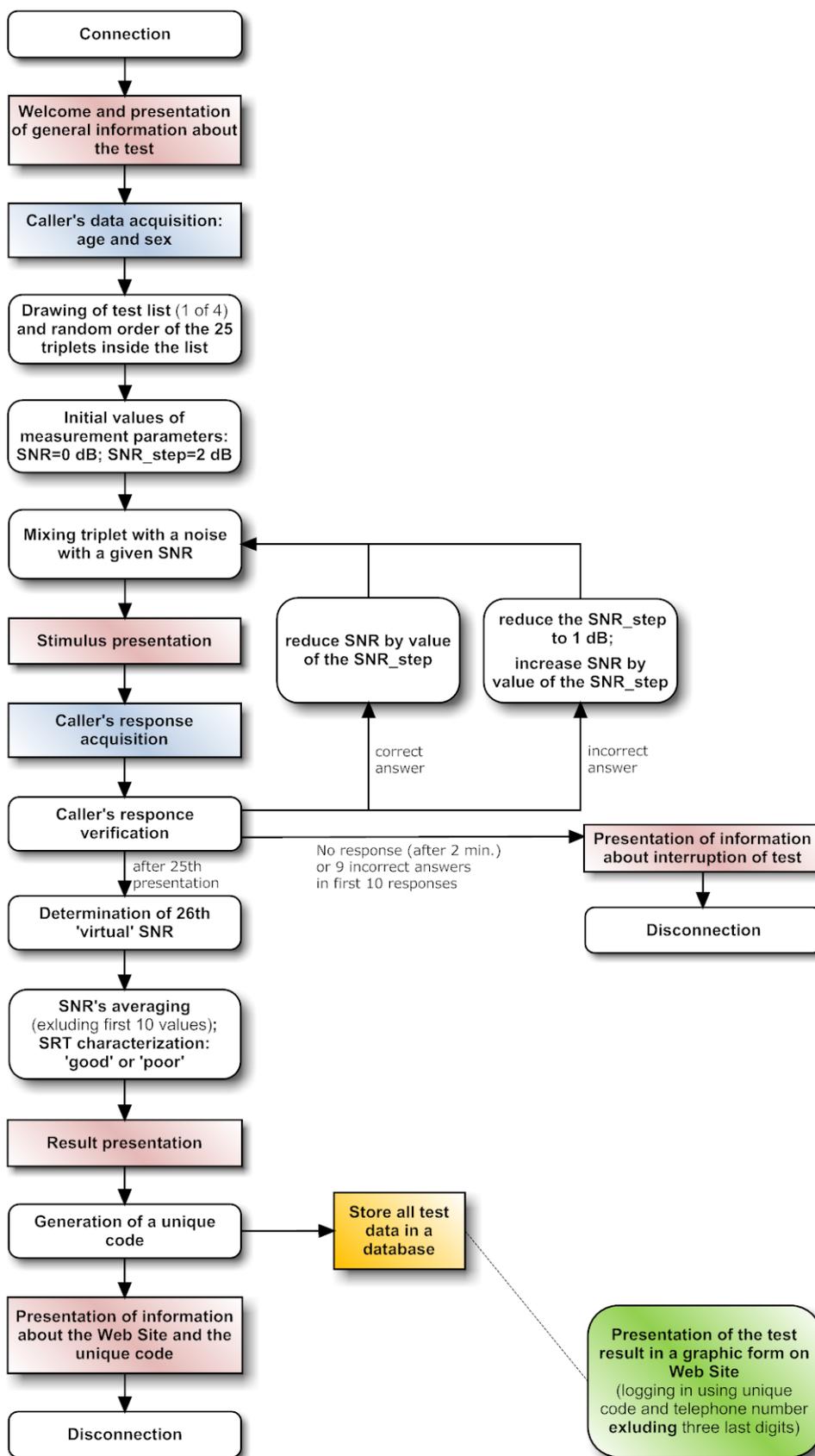
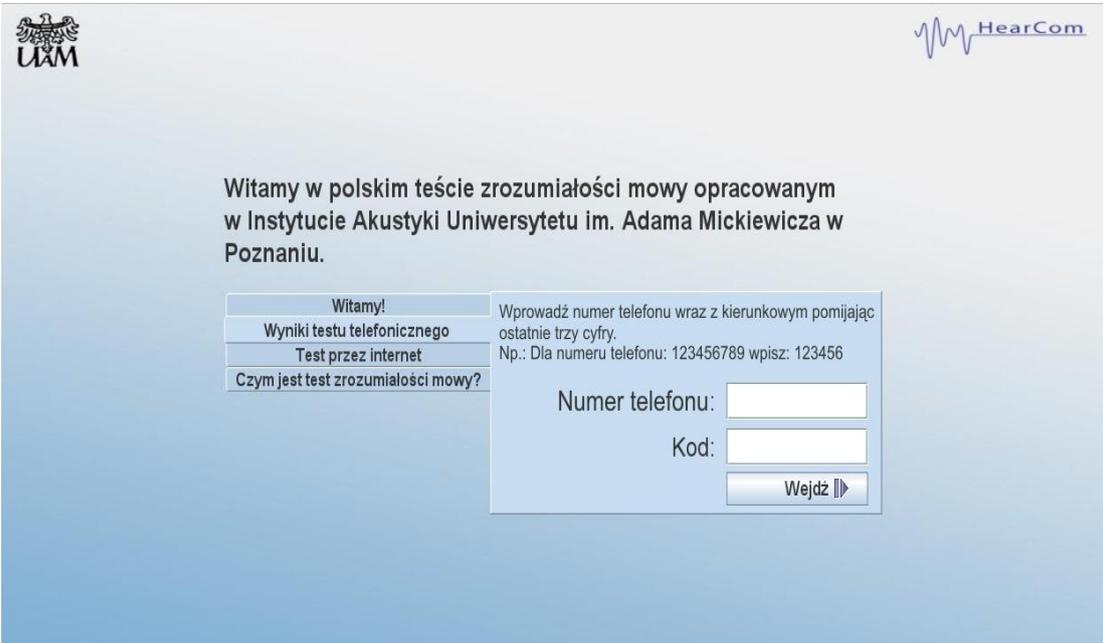


Fig. 4.2. A procedure of automatic speech-in-noise measurement used in a telephone version of PDTT – flow chart

4.1.3 Website working with telephone version of PDTT

As mentioned above, a special website was designed that offers the callers a possibility to gain insight into details of the test performed. The website might be, therefore, very useful for the callers interested in further details of the self-testing in which they have taken part. Furthermore, the results can be printed out by the callers themselves and presented to an ENT doctor or a hearing aid professional. What is more, the Internet version of PDTT is also available at the website for the users interested in self-testing using a home PC (to be evaluated in the future).

After typing the appropriate address (www.ucho.amu.edu.pl) into the Internet browser, the user is presented with a logging screen and asked to type their first six digits of telephone number and the test code. Fig. 4.3 depicts a logging screen presented to the user.



Witamy w polskim teście zrozumiałości mowy opracowanym w Instytucie Akustyki Uniwersytetu im. Adama Mickiewicza w Poznaniu.

Witamy!	Wprowadź numer telefonu wraz z kierunkowym pomijając ostatnie trzy cyfry. Np.: Dla numeru telefonu: 123456789 wpisz: 123456
Wyniki testu telefonicznego	
Test przez internet	
Czym jest test zrozumiałości mowy?	

Numer telefonu:

Kod:

Wejdz 

Fig. 4.3. Website working with the telephone version of PDTT – log-in screen

The log-in menu has 4 tabs:

- “Witamy!” (“Welcome!”) this a default tab presented to user at the beginning
- “Wyniki testu telefonicznego” (“Results of telephone testing”) – access to outcomes of the previous measurement carried out by telephone (Fig. 4.3). The user is asked to type in first six digits of the telephone (“Telefon”) number and the corresponding code (“Kod”)

- “Test przez Internet” (“Testing via Internet”) – access to the Internet version of PDTT (to be evaluated, normative data to be determined)
- “Czym jest zrozumiałość mowy” (“Importance of speech intelligibility”) – brief text concerning speech intelligibility and common disorders of speech perception.

Having typed the required data (the tab “Wyniki testu telefonicznego”), the user is presented with a graphic form of the staircase procedure (Fig. 4.4).

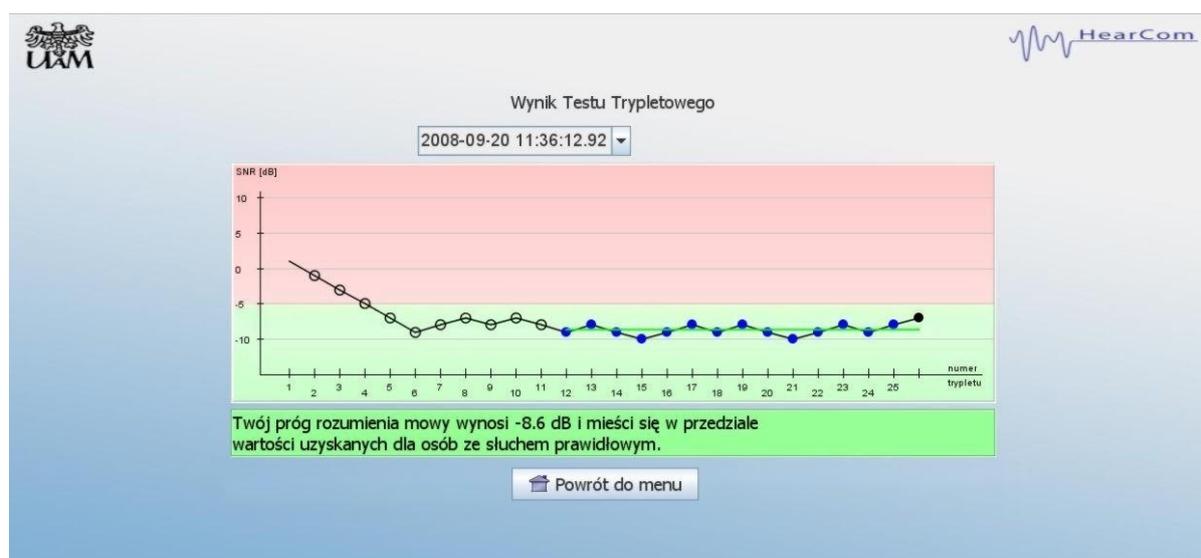


Fig. 4.4. Example of a graphic presentation of the data obtained with the staircase procedure

Blue circles present data to be averaged, the line represents the mean value; the green area presents the range of normal hearing. The comment below the figure contains the information about consistency or inconsistency of the caller’s SRT with the normative data.

4.1.4 Database structure

The database contains data concerning the caller and results of the measurements. The database was designed to be useful for further data analysis as well as to be compatible with the website. Each record in the database contains data related to a single call. Below is a list of variables in each record:

- Abbreviated caller’s telephone number

'nr_tel' – 6-number parameter containing a directional number and the caller's telephone number excluding **3 last digits** due to the privacy reasons. For example: if the telephone number is 61829 6383, only the sequence 618296xxx is stored in the database. This parameter is set to 0 in the case of an ex-directory number

- Date of measurement

'date' – in format: rrrr:mm:dd gg:mm:ss; where rrr is year, mm-month, dd-day, gg-hour, mm-minute, ss-second.

- Code of measurement

'kod' – tree-digit code

- Caller's sex

'pleć' – 'K' for woman, 'M' for man

- Caller's age

'wiek' – three-digit variable

- Presented triplets

'triplet_out(1)', 'triplet_out(2)', ... , 'triplet_out(25)' – 25 three-digit variables

- Collected responses

'triplet_in(1)', 'triplet_in(2)', ... , 'triplet_in(25)' – 25 three-digit variables

- SNR of presentation and the so-called 'virtual' SNR

'SNR(1)', 'SNR(2)', ... , 'SNR(25)', 'SNR(26)' – 26 variables in the format $\pm xx$;

- Mean SNR (excluding ten first presentations)

'SRT' – a variable in the format $\pm xx,x$;

4.2 Telephone version of PDTT – evaluation measurements

The re-test measurements described in 3.4 confirmed statistical balance of the laboratory version of PDTT and delivered normative data for laboratory measurements. In the general case, one may assume that if the lists are equivalent in laboratory conditions, there are no statistically significant differences in different measurement conditions. In order to confirm the equivalence of the lists in the telephone version of PDTT, other evaluation measurements were carried out. Another group of thirty normal-hearing subjects (callers) took part in the experiments. What is more, 4 different models of telephones were examined, including cordless telephone (wireless, but not cellular phones).

Each telephone was located in a different room (offices) at the Institute of Acoustics. Each caller was asked to 'visit' each of 4 rooms, dial the number and perform the test. To provide conditions that are comparable to final application of the test (i.e. when an anonymous caller dials the number), prior to the measurement the participants had not been given any information about the test. The callers, therefore, had to rely only on the information presented in the voice menu. During each call, a randomly chosen triplet list was presented, i.e. the callers were not presented each list once, but the probability of choosing each list was the same, i.e. 0.25. After the measurements, the caller was asked to write down their comments on the test (not obligatory).

Preliminary measurements revealed that in some cases the callers were not able to hear the initial numbers of triplets after typing their responses. This was the case for telephone models whose keypad was integrated with the receiver – since it takes some time to hold the receiver to the ear tested, the initial part of utterances could be inaudible. This in turn might lead to wrong results (i.e. higher SRTs). In accordance, the pause between accepting the response and the presentation of the next triplet was increased from 1 s to 2 s.

After measurements, the data obtained for each list and each telephone model were pooled across callers and subjected to two-way within-subject ANOVA (with individual SRTs as repeated measurements) with the 'list' and 'telephone' factor. The ANOVA revealed that both factors were statistically not significant (the differences across telephones and lists turned out to be smaller than inter-individual differences). The figure below presents mean individual SRTs and the corresponding standard deviations (error bars) for 4 lists and 4 telephone models.

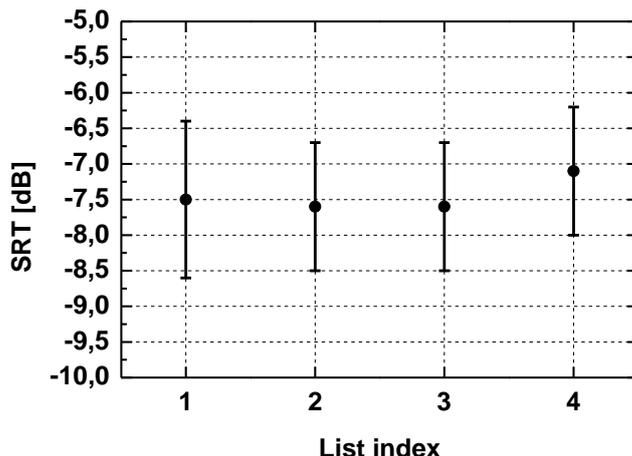


Fig. 4.5. The mean SRTs (averaged across telephones and callers) for 4 lists and corresponding standard deviations (error bars).

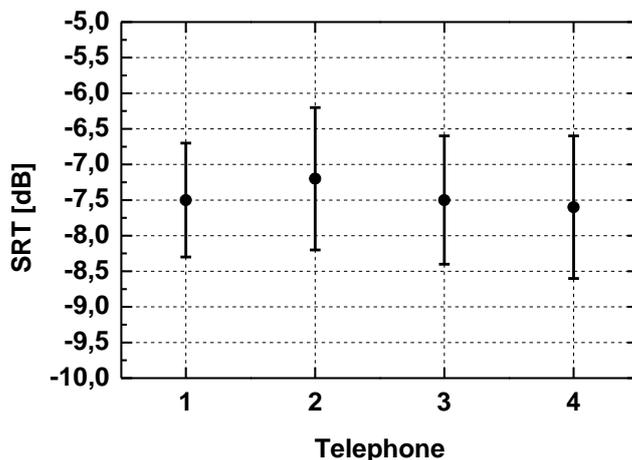


Fig. 4.6. The mean SRTs (averaged across lists and callers) for 4 telephones and corresponding standard deviations (error bars).

The mean SRT turned out to be -7.4 dB, i.e. 2.2 dB higher than SRT obtained for laboratory conditions; a higher value was expected. This result was due to distortions (compression, limitation of a frequency range, cross-talks) related to transitions of a signal through the telephone system that have some impact on a sound quality. This was also the case in the German and Dutch tests for which the difference in SRT between the laboratory and the telephone versions of the test was 2.9 dB and 4.1 dB, respectively. However, apart from the absolute SRT value, the measurements confirmed the statistical equivalence of the PDTT lists. As mentioned above, the upper limit of 95% confidence interval for SRTs for telephone version is -5.4 dB. This means that SRT higher than -5.4 dB is a result significantly poorer than the average.

4.3 Comparison of laboratory and telephone versions of PDTT

Although the laboratory and the telephone version of PDTT utilise the same speech material, there are some differences between them. These differences are due to specific equipment as well as measurement procedure used during the test. Tab. 4.1 presents a detailed comparison between the laboratory and the telephone version of PDTT.

Tab. 4.1. Characterization of the laboratory and the telephone versions of PDTT – comparison

	Laboratory version	Telephone version
Sampling rate [Hz]	48288 (TDT III)	8000
Measurement method	Constant stimulus, adaptive 1-up/1-down	adaptive 1-up/1-down
Presentation mode	headphones	telephone receiver
Response acquisition	PC keyboard	telephone keypad
Confirmation of response	Yes	No
Voice menu	No	Yes
Pause between data acquisition and speech presentation	1 s	2 s
normative SRT [dB] (SD [dB])	-9.4 (1)	-7.4 (1)
Upper limit of 95% confidence interval [dB]	-7.4	-5.4

5 Conclusions and future plans

This report describes the present status of PDTT. Two versions of PDTT are available, i.e. the laboratory and the telephone ones. The telephone version was implemented in the prototype set-up connected to the telephone system and is available via a temporary telephone number **+48 618217109**. A voice menu was designed and normative data were determined. The telephone test proved to be useful for intelligibility self-screening. The relevant website is available at: www.ucho.amu.edu.pl. Nevertheless further improvements/actions are planned to be taken:

- A modification of the 'diagnosis' block – an intermediate step between 'normal performance' and 'poor performance' will be added ('insufficient performance');
- Direct distinction of ear tested – in the end the callers will be recommended to examine the opposite ear (If they accept the proposal, another test will be carried out on the same connection.);
- An integration of the server with the dial center of the telecommunication provider and extension of the number of telephone lines;

Another challenging issue is organization of a campaign promoting the telephone intelligibility self-screening test in Polish society. This campaign is under preparation. Special care is paid to the following issues:

- Launch cooperation with Telekomunikacja Polska S.A., which is the largest and the most reliable Polish telecom company;
- Preparation of a long term business plan;
- Recruiting celebrities who will encourage self-screening;
- Organization of advertisements in the media (TV, radio, www, papers).

¹ It has been generally suggested that for the sake of homogeneity the test should contain only monosyllabic digits (Smits *et al.*, 2005). It results from the fact that when presented digits, subjects first discriminate the number of their syllables and then discriminate on the basis of phonemes. If there is a disproportion between the number of monosyllabic and disyllabic digits, certain words can deviate perceptually from others and, consequently, the homogeneity of speech material might be reduced. This must have an effect on the intelligibility function.

This problem was considered thoroughly and it was found that when proportions of disyllabic and monosyllabic digits were similar, the difference between SRTs for two types of digits was about 0.2 dB, thus indicating that the type of digit (either disyllabic or monosyllabic) did not have any significant effects on speech material homogeneity. However, if the proportion of monosyllabic and disyllabic digits is high, for example 8 to 2, the difference between SRTs for two types of digits amounts to about 1 dB. In this case it is reasonable to reject the digits that deviate perceptually from others and might reduce material homogeneity. This is why the disyllabic digit '7' and the disyllabic digits '7' and '9' were excluded from German (Wagener *et al.*, 2005) and Dutch (Smits *et al.*, 2005) digit triplet tests, respectively. However, in the Polish language as many as six out of ten digits are disyllabic. Therefore, if the disyllabic digits were to be excluded, the test would consist of four digits only. Accordingly, the prepared test has been composed of all monosyllabic as well as disyllabic digits, i.e. all digits from 0 to 9.

6 References

- Bellis, T. J. (1996). Assessment and management of central auditory processing disorders in the educational setting. San Diego, Singular Publishing.
- Elberling, C., et al. (1989). "DANTALE: a new Danish speech material." Scandinavian Audiology **18**(3): 196-75.
- Fletcher, H. (1929). Speech and hearing. New York, Van Nostrand.
- Kalikow, D. N., et al. (1977). "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability." J. Acoust. Soc. Am. **61**(1337): 1337-1351.
- Miller, G. A., et al. (1951). "The intelligibility of speech as a function of the context of the test materials." Journal of Experimental Psychology **41**: 329-335.
- Ozimek, E., et. al (2006) "The Polish sentence test for speech intelligibility evaluations." Archives of Acoustics. **31**(4): p. 431-438.
- Ozimek, E., et al. (2007). Deliverable D-1-6a: Development of Polish triple-digit test. Institute of Acoustics AMU.
- Ozimek, E., et al. (2009). "Development and evaluation of Polish digit triplet test for auditory screening." Speech Communication, **51** (4): 307-316.
- Pruszewicz, A., et al. (1994a). "New articulation lists for speech audiometry. Part I." Otolaryngol. Pol. **48**: 50-55.
- Pruszewicz, A., et al. (1994b). "New articulation lists for speech audiometry. Part II." Otolaryngol. Pol. **48**: 56-62.
- Rudmin, F. (1987). "Speech reception thresholds for digits." J. Audiol. Res. **27**: 15-21.
- Schmidt-Nielsen, A. (1989). "The intelligibility of native and non-native speakers of American English using spelling alphabet test materials." The Journal of the Acoustical Society of America **86**(S1): S76-S77.
- Smits, C. (2005). Hearing screening by telephone. Amsterdam, VU University Medical Center.

- Smits, C. and T. Houtgast (2007). "Recognition of digits in different types of noise by normal and hearing-impaired listeners." International Journal of Audiology **46**: 134-144.
- Strouse, A. and R. H. Wilson (2000). "The effect of filtering and inter-digit interval on the recognition of dichotic digits." Journal of Rehabilitation Research and Development **37**(5).
- Wagener, K., et al. (2005). Ziffer-Tripel-Test: Spracherverständlichkeitstest über das Telefon. Tagungs-CD der DGA Jahrestagung.
- Wilson, R. H., et al. (2005). "A comparison of word-recognition abilities assessed with digit pairs and digit triplets in multitalker babble." Journal of Rehabilitation Research and Development **42**(4): 499-510.
- Wilson, R. H. and D. G. Weakley (2004). "The use of digit triplets to evaluate word-recognition abilities in multitalker babble." Sem. Hear. **25**: 93-111.
-