

The development and evaluation of the Finnish digit triplet test

Tytti Willberg, Michael Buschermöhle, Ville Sivonen, Antti A. Aarnisalo, Heikki Löppönen, Birger Kollmeier & Aarno Dietz

To cite this article: Tytti Willberg, Michael Buschermöhle, Ville Sivonen, Antti A. Aarnisalo, Heikki Löppönen, Birger Kollmeier & Aarno Dietz (2016) The development and evaluation of the Finnish digit triplet test, Acta Oto-Laryngologica, 136:10, 1035-1040, DOI: [10.1080/00016489.2016.1175662](https://doi.org/10.1080/00016489.2016.1175662)

To link to this article: <http://dx.doi.org/10.1080/00016489.2016.1175662>



Published online: 28 Apr 2016.



Submit your article to this journal [↗](#)



Article views: 72



View related articles [↗](#)



View Crossmark data [↗](#)

ORIGINAL ARTICLE

The development and evaluation of the Finnish digit triplet test

Tytti Willberg^{a,b}, Michael Buschermöhle^{c,d}, Ville Sivonen^e, Antti A. Aarnisalo^e, Heikki Löppönen^{a,b}, Birger Kollmeier^{c,d,f} and Aarno Dietz^{a,b}

^aDepartment of Otorhinolaryngology, Kuopio University Hospital, Kuopio, Finland; ^bInstitute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland; ^cHörTech gGmbH, Oldenburg, Germany; ^dCluster of Excellence, 'Hearing4all', Oldenburg, Germany; ^eDepartment of Otorhinolaryngology, Helsinki University Central Hospital, Helsinki, Finland; ^fDepartment for Medical Physics and Acoustics, Carl Von Ossietzky Universität Oldenburg, Oldenburg, Germany

ABSTRACT

Objectives: The aim of the study was to develop a reliable and easily accessible screening test for primary detection of hearing impairment.

Methods: Digits 0–9 were used to form quasirandom digit triplets. First, digit specific intelligibility functions and speech recognition thresholds (SRTs) were determined. To homogenize the test material digits with steep intelligibility function slopes were chosen and level correction up to ± 2 dB were applied to the digits as needed. Evaluation measurements were performed to check for systematic differences in intelligibility between the test lists and to obtain normative reference function for normal-hearing listeners.

Results: The mean SRT and the final slope of the test lists were -10.8 ± 0.1 dB signal-to-noise ratio (SNR) and $21.7 \pm 1.8\%/dB$, respectively (measurements at constant level; inter-list variability). The mean SRT and slope of the test subjects were -10.8 ± 0.5 dB SNR and $23.4 \pm 5.2\%/dB$ (measurements at constant level; inter-subject variability). The mean SRT for normal-hearing young adults for a single adaptive measurement is -9.8 ± 0.9 dB SNR.

Conclusion: The Finnish digit triplet test is the first self-screening hearing test in the Finnish language. It was developed according to current standards, and it provides reliable and internationally comparable speech intelligibility measurements.

ARTICLE HISTORY

Received 12 January 2016
Revised 24 March 2016
Accepted 28 March 2016
Published online 27 April 2016

KEYWORDS

Digit triplet test (DTT); speech audiometry; speech test in noise; speech recognition threshold (SRT); hearing screening; signal-to-noise ratio (SNR)

Introduction

Age-related hearing impairment among older adults is common, and especially the eldest patients tend to underestimate its severity [1,2]. Since the proportion of the elderly is increasing, the number of patients with age-related sensorineural hearing impairment can also be expected to rise. An affordable, efficient, and easily accessible test for the screening of hearing impairment is needed to meet the growing need for the primary detection of hearing impairment.

The difficulty of hearing in noise is often the first sign of sensorineural hearing impairment. Pure-tone audiograms alone do not provide sufficient estimates of the speech recognition abilities in noisy everyday situations [3]. Various types of speech recognition tests in noise have been developed for a more accurate assessment of the speech reception impairment caused by sensorineural hearing impairment [4]. Most of the different tests for speech recognition in noise were developed for diagnostics in clinical or research applications. However, the basic features of speech tests in noise make them also well suited for screening purposes since: (1) The stimuli are presented at suprathreshold level, which means less demands for the testing environment. (2) For a large range of intensities, the speech recognition threshold (SRT,

i.e. the signal-to-noise ratio that yields 50% speech intelligibility) is not influenced by the absolute presentation level [5] and, therefore, the requirements for the calibration of the screening equipment are less stringent. (3) With a closed set of speech material the test can be implemented as an automatic screening test. (4) The SRT can be accurately determined by an adaptive procedure within a few minutes time [6].

Guidelines for internationally comparable speech tests in noise were produced by the HearCom project [7]. Recently, the International Collegium of Rehabilitative Audiology (ICRA) also published their recommendations for constructing multilingual speech tests [8]. In both of these recommendations the digit triplet test (DTT) was the choice for an internationally comparable screening test for hearing impairment. The DTT was originally developed in the Netherlands as a self-screening test by telephone [9]. Nowadays the test can be implemented as a mobile application.

The basic concept of the DTT makes use of digit triplets (for example, 2-4-7, pronounced *two-four-seven*) as speech material and adaptive up-and-down tracking for the SRT determination. The interfering noise matches the average long-term spectrum of the speech material, and the noise level is kept constant throughout the test. The responses are

given with a keypad and the presentation level of the triplets is varied in 2 dB steps based on the responses: an incorrect response increases the presentation level, and after a correct response the presentation level is decreased. A list of digit triplets (comprising 27–30 triplets per list depending on the test language) is presented, and the SRT is calculated by averaging the SNRs from a pre-defined triplet onwards. Typically, it takes 3–4 min to complete the test.

The success of the early versions of the DTT has led to the development of similar tests in different languages [7,8]. It has been shown that these tests can accurately discriminate between normal-hearing and hearing-impaired individuals, and their results correlate well with other speech tests in noise that utilize more complex speech material [9,10].

DTTs are increasingly used for screening for hearing impairment. The original Dutch telephone test received over 65 000 calls during the first 4 months [2]. As a quick and reliable self-test the DTT is also suitable for large-scale cohort studies; the British English version of the DTT is used as a part of the UK Biobank project with already over 160 000 subjects having completed the test [11]. The German DTT is one of the tests used in the German National Cohort, which is another large-scale cohort study [12].

The aim of the present study was to develop a Finnish version of the DTT to provide an efficient screening test for the primary detection of hearing impairment. This article describes the development, optimization, and evaluation procedures of the Finnish DTT. During the development, current guidelines [7] were followed throughout the study to ensure adequate test reliability and international comparability. Finally, the results were compared with those of the recently developed Finnish matrix test [13] and with the results of DTTs in other languages [7,14,15].

Materials and methods

Recording of the speech material

The speech material of the Finnish DTT consists of the digits 0–9 combined into triplets (e.g. 7-4-6). In the Finnish language the digits 0–6 have two syllables, and the digits 7–9 have three syllables. The speech material was recorded in the House of Hearing in Oldenburg, Germany. The speaker was a trained native Finnish female speaker who is working as a news anchor for Finland's national public service broadcasting company YLE. She was also the speaker for the Finnish matrix test [13]. For the Finnish DTT she spoke the numbers in standard Finnish dialect and with a standard pronunciation. A limited set of digit triplets were formed and arranged into two lists in random order. Each digit occurred twice on each position within a triplet, and each list was recorded twice. A carrier phrase '*Numerot (The digits):...*' preceded every triplet during the recording.

The recording took place in a sound insulated room. The set-up met the ISO 8253-3:2012 requirements for recording speech material for speech tests, and it was identical to the set-up used for the recording of the speech material for the Finnish matrix test (described in more detail in Dietz et al. [13]). The speaker was instructed to use natural speech rate

and speaking effort, and to maintain a constant distance from the microphone during the recordings.

Cutting the speech material and re-synthesizing the triplets

After recording, high pass filtering at 50 Hz was used to reduce low frequency noise. The sound files were cut into individual triplets, and the digit root mean square levels (RMS) of each triplet were equalized to eliminate any long-term trend in speaking effort. Next, the recorded triplets were manually cut into individual digits with a pre- and post-word flanking of 5 ms. Digits starting with a plosive (digits 2, 3, 6, 8) were cut 10–20 ms before the beginning of the plosive. All other digits were cut as close as possible to their beginnings and ends.

For each digit at each position in the triplet two of the most natural sounding versions (version 1 and version 2) of the recordings were chosen for optimization. Test lists for the optimization measurements were created by combining the sound files containing single digits and their ramps into digit triplets in a way that preserved the individual digit's original position in the triplet. This method allowed the preservation of prosody when single digits were combined to form triplets. The same method was also used for the German DTT [7] and for the Finnish matrix test [13]. Six test lists of 30 triplets were formed for both versions of the digits, resulting in 12 test lists in total. In each test list, each digit occurred three times in each position. A single triplet occurred only once within the six test lists for both versions 1 and 2.

For the carrier phrase only one version of the recordings was used. The RMS of the announcement phrase was set at 2 dB higher than the RMS of the triplets since especially at poor signal-to-noise ratios it is beneficial for the carrier phrase to be slightly more audible than the triplets [7].

Development of the masking noise

In accordance with the current recommendations [7,8], the masking noise for the Finnish DTT is quasi-stationary masking noise that has a long-term spectrum corresponding to the long-term spectrum of the triplets without fluctuations in level. It was generated by superposing all individual digit files 30-fold using variable and random delays before the start of the sound files and between the sound files. The recordings of the carrier phrase were not used to create the noise.

Participants

Both the optimization and evaluation measurements took place at Kuopio University Hospital. Sixteen native Finnish speakers aged from 20–30 years (mean = 23.1 years) participated in the optimization measurements, and 19 native Finnish speakers aged from 18–34 years (mean = 23.2 years), who had not participated in the optimization measurements, participated in the evaluation measurements. All the participants had normal hearing confirmed by pure-tone

audiometry at the beginning of the session (pure-tone threshold ≤ 15 dB HL between 0.125–8 kHz). All measurements were performed monaurally on the better ear. Informed consent was obtained from all participants. The study was approved by the Research Ethics Committee of the Northern Savo Hospital District.

Optimization

The standard deviation of an SRT estimate can be modeled as being inversely proportional to the slope of the intelligibility function [16]. As shown earlier, the slope of the intelligibility function depends on the slope of the individual test items and their variation in intelligibility [17]. Therefore, both these variables were optimized to obtain homogenous speech material and to increase the precision of the Finnish DTT.

The measurements were done in a sound-attenuated booth using free field equalized Sennheiser HDA200 headphones (Sennheiser Electronics GmbH & Co. KG, Wedemark-Wennebostel, Germany). The equipment and set-up were the same as for the optimization of the speech material for the Finnish matrix sentence test [13].

Subjects were told that all triplet combinations of digits 0–9 were possible, and they were instructed to repeat the presented three digits in the correct order. The experimenter entered the repeated triplet into the software. If a digit could not be heard, the subjects were allowed to say so, and it was, thus, assumed that the guess rate was 0. The subjects received no feedback on their responses during the test.

To randomize the presentation of the test lists, the subjects were divided into two groups. Both groups initially performed two training lists at constant SNR of 0 and -2 dB. After training both groups performed eight test lists in random order at the following constant SNRs: -21.0 , -18.5 , -16.0 , -13.5 , -11.0 , -8.5 , -6.0 , and -3.5 dB. The noise level was held constant at 65 dB SPL. Group 1 used lists from recording version 1 and group 2 used lists from recording version 2. This way each digit at each position from the selected recording version was presented to the same test subject at each SNR exactly three times. For each digit at each position in the triplet, intelligibility scores were determined for every SNR measured. The psychometric function for each individual digit realization was obtained by performing a maximum likelihood fit using the raw data and the equation

$$I(L) = \frac{1}{1 + e^{4s(L_{50}-L)}} \quad (1)$$

where I is the intelligibility of the digit, L is the level (given here as signal to noise ratio), s is the slope of the psychometric function, and L_{50} is the SRT of the digit. The fitted parameters for each digit were s and L_{50} .

The goal of optimization was to obtain test material where individual digits have psychometric functions with steep slopes and intelligibility scores that are close to each other. Therefore, the level of each digit was adjusted to bring it as close as possible to the average L_{50} of all digits. The amount of level correction needed was determined by comparing each individual digit's L_{50} to the average L_{50} of all digits. The

level differences were within 2 dB in all but five digits for which the calculated correction needed was 2.1 or 2.2 dB. However, it was decided to limit the level corrections to 2 dB to avoid unnatural sounding sound level changes within a triplet. The slopes of the psychometric functions of recording versions (version 1 and version 2) of digits were compared and the version with the steeper slope was chosen for each digit separately.

The optimized digits were arranged into triplets to create the final test lists. Altogether six lists each comprising 30 triplets were formed. Within each test list one digit occurs at each position exactly three times. A single digit may occur twice in one triplet but never sequentially, i.e. triplet 4-6-4 is allowed, but triplet 2-2-3 is not. No triplet occurs twice in the test lists. The background noise file and the carrier phrase remained the same as for the optimization measurements.

Evaluation measurements

The goal of the evaluation measurements was to obtain a normative reference function for normal-hearing listeners, and to verify that there is no systematic difference in intelligibility between the final test lists.

For the evaluation measurements we used the same apparatus and set-up as described for the optimization measurements. In the beginning of the session each subject performed two training lists in an adaptive procedure. The measurements started at 0 dB SNR. Triplet scoring was used, i.e. all digits had to be identified correctly in the correct order. Based on the given answer the speech level of the next triplet was adjusted with an automatic adaptive up-and-down procedure with a step size of 2 dB. The noise level was held constant at 65 dB SPL. The lists used were selected randomly from the six test lists. This same test protocol will be used for the final test application.

Subsequently, each subject was tested with all the six test lists at three different constant SNRs (18 measurements in total per subject). The SNRs used were -14.0 dB SNR, -12.5 dB SNR, and -11.0 dB SNR, since they were expected to yield intelligibilities of $\sim 20\%$, 50% , and 80% , respectively. The order of the test lists and SNRs used was randomized to minimize training or fatigue effect on the results.

Results

Results from the optimization measurements

Before optimization the mean SRT for all digits was -12.3 ± 1.7 dB SNR, and the mean slope of the intelligibility function was $20.8 \pm 5.8\%/dB$. As can be seen in Figure 1 the optimization procedure increased considerably the homogeneity of the test material, and the SRT for the remaining final digits was expected to be -12.3 ± 0.1 dB SNR, and the expected slope was $22.9 \pm 6.4\%/dB$.

Results from the evaluation measurements

The normative reference function and the SRT for normal-hearing listeners were determined by pooling the data from

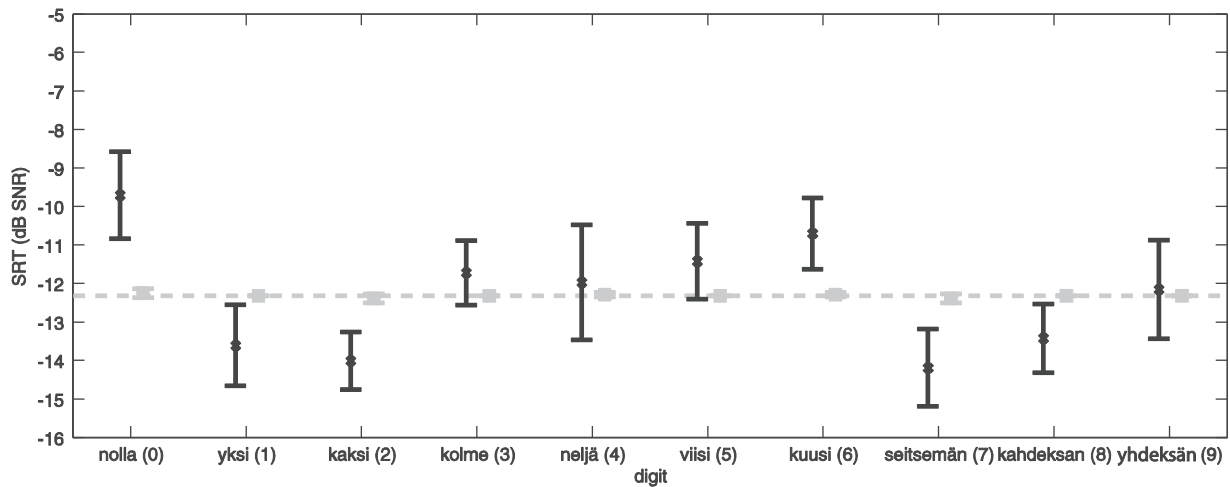


Figure 1. SRTs of the individual digits before and after optimization. Error bars indicate standard deviations. Before optimization, the standard deviation is calculated across the two versions and the three positions within a triplet. After optimization, the standard deviation is calculated across the three positions within a triplet (data as predicted from optimization measurements with application of level correction).

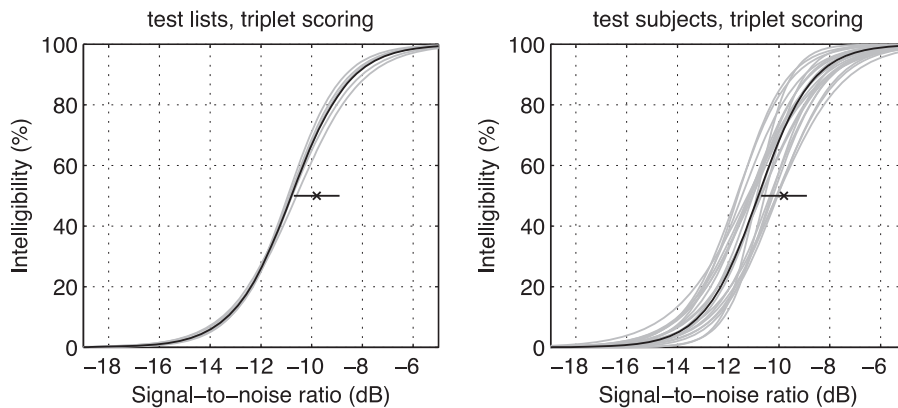


Figure 2. Psychometric functions of the Finnish Digit Triplets Test with triplet scoring. Left: grey lines indicate the psychometric functions of the six different test lists. Right: grey lines indicate the psychometric functions of the 20 test subjects. In both panels, the solid black line indicates the average psychometric function (left: averaged across test lists, right: averaged across test subjects). The cross and solid horizontal black lines indicate the mean plus/minus one standard deviation of the adaptively determined SRT.

all test lists measured at constant levels for one subject and by fitting the logistic model function to these data. When using triplet scoring (i.e. every single digit in a triplet had to be correctly recognized and repeated in the correct order) the guess rate is 0.001, which means that a chance level of 0% can be assumed. Therefore, we used the same logistic function (Equation 1) as for the data from the optimization measurements. Averaged across all test subjects the mean SRT was -10.8 ± 0.5 dB SNR when triplet scoring was used. With triplet scoring the mean slope of the reference function of the Finnish DTT was $23.4 \pm 5.2\%/dB$.

List-specific speech intelligibility functions were determined by pooling the data from all subjects for one test list and by fitting the logistic model function (Equation 1) to these data. With triplet scoring the mean slope for all test lists was $21.7 \pm 1.8\%/dB$ and the mean SRT for all test lists was -10.8 ± 0.1 dB SNR. Because the standard deviation of the SRT for all the test lists was only 0.1 dB SNR, while the standard deviation of the SRT for all test subjects was 0.5 dB, the lists can be considered equal in intelligibility.

The results from the adaptive measurements that were conducted in the beginning of the evaluation measurements were analyzed as well. Only triplet scoring was used for this.

The SRT was determined by averaging the SRTs from the fifth triplet on. For the first adaptive measurement the average SRT was -9.8 ± 0.9 dB SNR. For the second measurement it was -10.1 ± 0.8 dB SNR. All results determined with triplet scoring are summarized in Figure 2.

The triplet scoring method described above will be used for the final test. For evaluation purposes we analyzed the data from the measurements with constant SNRs also with digit scoring. If digit scoring is used, the assumed guess rate is 0.1, which leads to 10% chance level. To take this into account when analyzing the results we used the psychometric function

$$I(L) = 0.1 + \frac{0.9}{1 + e^{4s(L_{50}-L)}} \quad (2)$$

The variables used in this equation are the same as in Equation (1).

As expected digit scoring resulted in a slightly shallower mean slope of $20.2 \pm 4.9\%/dB$ for all the test subjects. The test lists also had a shallower mean slope of $18.9 \pm 1.2\%/dB$. With digit scoring the SRT for all test subjects decreased to -12.3 ± 0.6 dB SNR and the SRT for all test lists decreased to

-12.3 ± 0.02 dB SNR. A summary of the results of the evaluation measurements can be found in Table 1.

Discussion

The Finnish DTT is the first self-screening hearing test developed for the Finnish language. The first DTT was implemented in the Netherlands more than 10 years ago [9], and since then DTTs in multiple languages have successfully been used as screening tools for hearing impairment [2,11,15].

The development of the Finnish DTT followed the guidelines [7] that were set to assure that new DTTs provide results that are both reliable and internationally comparable. Previously there have been some slight differences between the development procedures of different DTTs, for example in regards of the optimization of the triplets [9]. Since the German and French DTTs [7,14,15] were developed according to the same guidelines as the Finnish DTT, we focused our comparison of the test results to these DTTs. The comparison shows that the steep slope of the Finnish DTT lines up well with the slopes of the other DTTs (see Table 2). Also the SRT of the Finnish DTT corresponds well to the results from the other two DTTs (see Table 2).

The evaluation measurements of the Finnish DTT showed that the test lists had very homogenous list-specific SRTs, with a standard deviation of only 0.1 dB between the lists. We, therefore, expect that the Finnish DTT will provide as reliable and accurate results as the other DTTs, and that also the Finnish version of the DTT can be implemented in screening for hearing impairment.

The speech material for the Finnish DTT uses all the digits from 0–9. In the Finnish language there are no monosyllabic digits; digits 0–6 have two syllables and the digits 7–9 have three syllables. In the Dutch and German versions bisyllabic digits were omitted because it was suspected that the few bisyllabic digits might stand out and be more easily recognized, which might have a negative effect on the discrimination function [7,9]. For the Finnish version all digits 0–9 were used because, after optimization, no major differences in intelligibility could be measured (see Figure 1).

A fair correlation between the results of DTTs and speech tests in noise with more complex speech material, such as the matrix test, has been shown [9,10]. For most languages DTT’s reference SRT values are lower (i.e. better) than the corresponding reference SRT values for the matrix test (see Table 3). This has largely been attributed to digit triplets being very limited and easy speech material [10]. For the Finnish tests, however, the difference between the reference SRTs is very small (see Table 3). Whereas the reference SRT for the Finnish DTT is very similar to other DTTs (see Table 3), the reference SRT for the Finnish matrix test has been shown to be considerably lower than for most other matrix tests [13].

The reference SRT of a speech test in noise is a test-specific parameter that depends on multiple test-specific variables such as speech material, scoring method, speaker, and speech rate [4,19]. The comparison between the Finnish tests is interesting because, unlike with many other DTTs and matrix tests, the speech material for both tests was recorded during the same recording session and with the same trained speaker using a constant speech rate. Also, for both tests, a similar procedure to create the masking noise was used.

One possible explanation for the small difference between the reference SRTs might be the phonological nature of the Finnish language. There are few published studies on how the phonemic characteristics of different languages affect the speech intelligibility in noise. However, compared to many other languages the Finnish language has rather straightforward phonological characteristics and the use of vowel harmony makes the occurrence of vowels within a word somewhat predictable [20]. This predictability may contribute to the seemingly better intelligibility (i.e. lower SRT) of the

Table 3. Comparison of reference parameters for Matrix Sentence Tests and Digit Triplets Tests in different languages. All data are provided for monaural headphone measurements with normal hearing test subjects, measured with constant levels.

	SRT (dB SNR)		Slope (%/dB)	
	Matrix	DTT	Matrix	DTT
Finnish (this publication [13])	-10.1	-10.8	16.7	23.4
German [7,18]	-7.1	-9.3	17.1	19.6
French [10,15]	-6.0	-10.5	14.0	27.1

Table 1. Summary of results of the evaluation measurements.

	Triplet scoring	Digit scoring
SRT, adaptive, across subjects	-9.8 ± 0.9 dB SNR (test) -10.1 ± 0.8 dB SNR (retest)	Not available
SRT, constant levels, across subjects	-10.8 ± 0.5 dB SNR	-12.3 ± 0.6 dB SNR
SRT, constant levels, across lists	-10.8 ± 0.1 dB SNR	-12.3 ± 0.02 dB SNR
Slope, constant levels, across subjects	23.4 ± 5.2%/dB	20.2 ± 4.9%/dB
Slope, constant levels, across lists	21.7 ± 1.8%/dB	18.9 ± 1.2%/dB

Table 2. Comparison of key parameters of the Finnish DTT with other DTTs. The data is summarized for triplet scoring and the broadband version of the test (monaural presentation for normal hearing listeners).

	SRT	Slope of psychometric function
Finnish (this publication)	-10.8 ± 0.5 dB SNR (constant levels, across subjects) -9.8 ± 0.9 dB SNR (adaptive, across subjects)	23.4 ± 5.2%/dB (constant levels, across subjects)
German [7,14]	-9.3 ± 0.2 dB SNR (constant levels, across lists) -9.4 ± 0.4 dB SNR (adaptive, across subjects)	19.6 ± 2.2%/dB (constant levels, across lists)
French [15]	-10.5 ± 0.3 dB SNR (constant levels, across subjects) -10.2 ± 0.5 dB SNR (adaptive, across subjects)	27.1 ± 3.0%/dB (constant levels, across subjects)

Finnish version of the matrix test, and explain the similarity of the reference SRTs for the Finnish DTT and the matrix test. Further validation studies will show whether this effect is seen also with hearing impairment.

So far in Finland no objective self-test has been available for people to assess their hearing. The primary diagnostics of hearing impairment takes place mostly at public healthcare centers from where patients with hearing impairments are referred to public ENT clinics for a more precise assessment and rehabilitation if necessary. At present in Finland pure-tone audiometry is the only way to objectively assess hearing at primary healthcare level. These audiograms can give a good estimate of patients' pure tone thresholds, but they are often not precise or reliable enough for clinical diagnostics, since many primary healthcare centers lack the trained professionals and precisely calibrated equipment. When considering a referral for hearing rehabilitation, an automated screening test that can reliably identify patients who are likely to benefit from hearing rehabilitation could be a more efficient and cost-effective option in many cases. An easily accessible auditory screening test has also the potential to increase people's awareness of their hearing impairment and the likelihood of seeking proper rehabilitation.

Conclusions

The Finnish DTT is the first self-screening hearing test in the Finnish language. The development of the test followed the current international guidelines in order to establish speech audiometric tests comparable across different languages, and the Finnish DTT provides reliable results that are similar to other DTTs [7].

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

- [1] Roth TN, Hanebuth D, Probst R. Prevalence of age-related hearing loss in Europe: a review. *Eur Arch Otorhinolaryngol* 2011;268:1101–7.
- [2] Smits C, Houtgast T. Results from the Dutch speech-in-noise screening test by telephone. *Ear Hear* 2005;26:89–95.
- [3] Plomp R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J Acoust Soc Am* 1978;63:533.
- [4] Theunissen M, Swanepoel DW, Hanekom J. Sentence recognition in noise: Variables in compilation and interpretation of tests. *Int J Audiol* 2009;48:743–57.
- [5] Wagener KC, Brand T. Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: influence of measurement procedure and masking parameters. *Int J Audiol* 2005;44:144–56.
- [6] Smits C, Houtgast T. Measurements and calculations on the simple up-down adaptive procedure for speech-in-noise tests. *J Acoust Soc Am* 2006;120:1608–21.
- [7] Zokoll MA, Wagener KC, Brand T, Buschermöhle M, Kollmeier B. Internationally comparable screening tests for listening in noise in several European languages: the German digit triplet test as an optimization prototype. *Int J Audiol* 2012;51:697–707.
- [8] Akeroyd MA, Arlinger S, Bentler RA, Boothroyd A, Dillier N, Dreschler WA, et al. International Collegium of Rehabilitative Audiology (ICRA) recommendations for the construction of multilingual speech tests. *Int J Audiol* 2015;54:17–22.
- [9] Smits C, Kapteyn TS, Houtgast T. Development and validation of an automatic speech-in-noise screening test by telephone. *Int J Audiol* 2004;43:15–28.
- [10] Jansen S, Luts H, Wagener KC, Kollmeier B, Del Rio M, Dauman R, et al. Comparison of three types of French speech-in-noise tests: A multi-center study. *Int J Audiol* 2012;51:164–73.
- [11] Moore DR. Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. *PLoS One* 2014;9:e107720.
- [12] Ahrens W, Hoffmann W, Jöckel KH, Kaaks R, Gromer B, Greiser KH, et al. The German National Cohort: Aims, study design and organization. *Eur J Epidemiol* 2014;29:371–82.
- [13] Dietz A, Buschermöhle M, Aarnisalo AA, Vanhanen A, Hyyrynen T, Aaltonen O, et al. The development and evaluation of the Finnish matrix sentence test for speech intelligibility assessment. *Acta Otolaryngol* 2014;134:728–37.
- [14] Buschermöhle M, Wagener KC, Berg D, Meis M, Kollmeier B. The German Digit Triplets Test (Part I): implementations for telephone, internet and mobile devices. *Z Audiol* 2014;53:139–45.
- [15] Jansen S, Luts H, Wagener KC, Frachet B, Wouters J. The french digit triplet test: a hearing screening tool for speech intelligibility in noise. *Int J Audiol* 2010;49:378–87.
- [16] Brand T, Kollmeier B. Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *J Acoust Soc Am* 2002;111:2801–10.
- [17] Kollmeier B. *Meßmethodik, Modellierung und Verbesserung der Verständlichkeit von Sprache*. Göttingen, Germany: Georg-August-Universität; 1990.
- [18] Wagener KC, Brand T, Kollmeier B. Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil III: evaluation des Oldenburger Satztests. *Z Audiol* 1999;38:86–95.
- [19] Macpherson A, Akeroyd MA. Variations in the slope of the psychometric functions for speech intelligibility: a systematic survey. *Trends Hear* 2014;18:1–26.
- [20] Eerola O, Savela J, Laaksonen J, Aaltonen O. The effect of duration on vowel categorization and perceptual prototypes in a quantity language. *J Phon* 2012;40:315–28.