Speech intelligibility tests for the Turkish language

Melanie A. Zokoll1, Sabine Hochmuth1, Dilek Fidan2, Kirsten C. Wagener3, İclal Ergenç4, Birger Kollmeier,1,3,5

1Medical Physics, Institut of Physics, Carl von Ossietzky University Oldenburg, Oldenburg, Germany
2Department of Turkish Language Teaching, Kocaeli University, Kocaeli, Turkey
3Hörzentrum Oldenburg GmbH, Oldenburg, Germany
4Department of Linguistics, Ankara University, Ankara, Turkey
5HörTech gGmbH, Oldenburg, Germany

Key words: hearing screening, speech-in-noise test, digit triplet test, Matrix test, Turkish

Introduction

Professional hearing diagnostics and verification of technical hearing aids often fails because of speech barriers: speech intelligibility tests can usually only be used if both, testing person and patient have the same language competence and results are not comparable internationally. If additionally the tests are measured in noise to obtain results in situations close to reality, non-native listeners perform significantly worse in comparison to native listeners with the same hearing ability (e.g., Warzybok et al., 2010, van Wijngaard et al., 2002). Multilingual versions of tests provide the possibility to test non-native listeners more exactly. Thus, two projects (HearCom and its partial follow-up HurDig) are concerned with the selection and if necessary the development, optimization and evaluation of internationally comparable speech intelligibility tests for several languages. A main focus is laid on the internationalization of the digit triplet test (Smits et al., 2004, Wagener et al., 2005, 2006) and the Oldenburg sentence test (OLSA, Wagener et al., 1999 a,b,c, see also Hagerman, 1984) as a matrix test.

Meanwhile both tests are available in several languages (British English, Danish, German, French, Polish, Swedish, Spanish and soon also in Russian and American English). This contribution reports the development and optimization of the Turkish digit triplet and Matrix test as an extension of the existing test battery. For Germany this language is very important since about 2 Million Turkish native speakers live in Germany (status 2009).

Speech material

The speech material for the Turkish digit triplet test as well as the Turkish Matrix test was developed by linguists of the University of Ankara and Kocaeli which are also co-authors of this paper. For the digit triplet test, all Turkish digits between zero and nine were selected [sıfır (0), bir (1), iki (2), üç (3), dört (4), beş (5), altı (6), yedi (7), sekiz (8), dokuz (9)]. Six different test lists were prepared. Each list contained 30 digit triplets. In each list a certain digit appeared three times at each position in a triplet (1st, 2nd, and 3rd position). In accordance with the procedure applied for the German digit triplet test, the noise background was generated by 30-fold superimposition of the speech material (i.e., the digits and the announcement “sayılar”). For the Matrix test, a basis list of ten sentences à five words with the same sentence structure was prepared (Tab.1).
The words were selected according to their frequency in the Turkish language. By pseudo-randomly selecting words within the single word groups of this basis list new sentences were composed and combined in 30 test lists à ten sentences each. For the optimization measurements described below, succeeding test lists were combined to ten test lists à 30 sentences. The test-specific noise was generated like for the digit triplet test (see Wagener et al., 1999a, 2003 for more details).

<table>
<thead>
<tr>
<th>Name</th>
<th>Numeral</th>
<th>Adjective</th>
<th>Object</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gönül</td>
<td>yedi</td>
<td>mavi</td>
<td>sepet</td>
<td>haketmiş Gönül deserved seven blue baskets.</td>
</tr>
<tr>
<td>Zuhal</td>
<td>bir</td>
<td>yeni</td>
<td>kılım</td>
<td>verdi Zuhal gave single new rug.</td>
</tr>
<tr>
<td>Frat</td>
<td>küçük</td>
<td>beyaz</td>
<td>yatak</td>
<td>satmış Frat sold seven white beds.</td>
</tr>
<tr>
<td>Hikmet</td>
<td>üç</td>
<td>küçük</td>
<td>çatal</td>
<td>getirdi Hikmet brought three small forks.</td>
</tr>
<tr>
<td>Tuncay</td>
<td>altı</td>
<td>yeşil</td>
<td>cmzik</td>
<td>bulmuş Tuncay found six green tweezers.</td>
</tr>
<tr>
<td>Nurşen</td>
<td>boş</td>
<td>temiz</td>
<td>gömlek</td>
<td>çıktı Nurşen drew five clean shirts.</td>
</tr>
<tr>
<td>Poyraz</td>
<td>dokuz</td>
<td>renkli</td>
<td>balon</td>
<td>fiatmuş Poyraz threw nine colorful balloons.</td>
</tr>
<tr>
<td>Seyhan</td>
<td>on</td>
<td>bordo</td>
<td>minder</td>
<td>gördü Seyhan saw ten maroon cushions.</td>
</tr>
<tr>
<td>Meltem</td>
<td>iki</td>
<td>güzel</td>
<td>terlik</td>
<td>kazanmış Meltem won two nice slippers.</td>
</tr>
<tr>
<td>Dilek</td>
<td>dört</td>
<td>siyah</td>
<td>fincan</td>
<td>yolladı Dilek sent four black cups.</td>
</tr>
</tbody>
</table>

Table 1. Basis list of the Turkish Matrix test.

Optimization of the speech material

Testing in noise requires higher precision of the SRT estimation than testing in quiet, since the SRT of normal hearing listeners is raised by the masking noise level and differences in speech intelligibility between normal hearing and hearing impaired become smaller. The speech intelligibility function of a sentence test can be described by the following logistic model function:

\[
SI(SNR) = \frac{1}{1 + e^{-4s_{50}(SRT-SNR)}}
\]

\(SI\) is the speech intelligibility and \(s_{50}\) the slope of the intelligibility function at the \(SRT\). Homogeneous word-specific speech intelligibility is supposed to result in a steep intelligibility function of the overall speech material and therefore in an exact estimation of the SRT (Kollmeier 1990). Thus, the optimization of the speech test material targets to homogenize the speech intelligibility of the single words by adequate selection of the material and level adjustments. To obtain the word-specific intelligibility, optimization measurements with twelve normal hearing Turkish native speakers (between 21 and 36 years of age, mean age 28 years) were conducted. All subjects grew up and visited school in Turkey. At time of testing, their residence in Germany was less than five years.

Word-specific intelligibility was measured presenting the test lists (six test lists à 30 digit triplets for the digit triplet test, and ten test lists à 30 sentences for the Matrix test, respectively) over headphones (Sennheiser HDA200) at fixed SNRs [range from –2 to –16 dB SNR (step size 2 dB SNR) for the digit triplet test, and 5.5 to –22 dB SNR (step size 2.5 dB SNR) for the Matrix test]. The level of the noise was 65 dB SPL. Responses were gathered with word scoring, i.e., every word was counted separately.

The mean word-specific SRT was calculated and the intelligibility of all words was adjusted to this value (either by increasing or decreasing the level). The level adjustments were limited to 3 dB for both tests. For the Matrix test, this was predominantly done to preserve the natural sound impression of the sentences. Matrix test words with a very shallow slope of the word-specific speech intelligibility function were excluded.

Results

Turkish digit triplet test

The mean digit-specific speech intelligibility function before optimization has a SRT of \(-10.0 \pm 2.2\) dB SNR and a slope of 21.9 %/db. The resulting slope of the overall speech material, estimated according to the probabilistic model of Kollmeier (1990), is 13.9 %/db. Fig. 1 shows the expected shift in SRT for the different digits independent of their position in the triplet. The optimization procedure resulted in a computed mean SRT of \(-10.1 \pm 0.5\) dB SNR, whereas the slope stays the same. According to the probabilistic model, decreasing the standard deviation of the SRTs from 2.2 to 0.5 dB SNR increases the slope of the overall material (\(s_{50\text{mat}}\)) from 13.9 to 21.0 %/db.
Fig. 1. Intelligibility functions for the different digits independent of their position in the triplet before (left graph) and after optimization (right graph).

Turkish Matrix test

Fig. 2 shows the distribution of the word-specific SRTs of the speech material before (left graph) and after optimization (right graph). The mean SRT of all words is -9.4 ± 2.6 dB SNR before and -9.3 ± 0.8 dB SNR after optimization. The median slope of the word-specific speech intelligibility function is 17.8 %/dB. The resulting slope of the overall speech material (s50test) could be increased from 11.6 to 16.8 %/dB due to optimization.

Fig. 2. Distribution of the word-specific SRTs for the speech material of the Matrix test before (left graph) and after optimization (right graph).

Discussion

The goal of the study was to develop Turkish speech in noise tests that are internationally comparable and multilingual applicable. For both tests, suitable speech material was generated and optimized, i.e., homogenized in intelligibility. This is shown by the decreasing distribution of the SRT values (Fig. 1 and 2) and the steeper slope of the overall material (s50test) in comparison to the non-optimized test material.

In comparison to the test-specific speech intelligibility functions obtained for other languages with tests of the same format in headphone measurements (e.g., Wagener et al., 2009, see Fig. 3), the Turkish tests show similar test-specific SRTs and slopes.
Fig. 3. Test-specific speech intelligibility functions for digit triplet and Matrix tests in several languages. Speech intelligibility functions of the Turkish Matrix test base on optimization data. For the British English Matrix test, data were obtained with sentences scoring (i.e., complete sentences were counted as either correct or incorrect).

Acknowledgements
This work has been done in the context of the European Regional Development Fund (ERDF) funded project “Network for multilingual hearing and speech intelligibility diagnostics” (HurDig). Special thanks to Müge Kaya who helped in judging the sound impression of the sentences.

References


