Influence of the linguistic complexity in relation to speech material on non-native speech perception in noise.

Anna Warzybok[1], Kirsten Wagener[2], Thomas Brand[1], Birger Kollmeier[1]
[1] *Medizinische Physik, Carl-von-Ossietzky Universität Oldenburg,*
[2] *HörTech gGmbH, Oldenburg*

Introduction

Speech recognition in noise is especially difficult in a foreign language. Speech intelligibility by non-native listeners (NN) is not only affected by masking noise, but also by other factors like experience with the vocabulary, syntax, and semantics of the second language or degree of semantic and phonological relatedness to words found in the native language. The relative contribution made by these and other factors to foreign language recognition in noise is not well understood.

The primary purpose of this study was to investigate how linguistic complexity of speech material influences speech intelligibility by NN. It was also shown how language proficiencies of the listeners can affect speech recognition by NN.

Speech material and methods

Listeners

Twenty-four normal-hearing international students participated in a listening experiment. They were divided into four groups on the basis of their German language knowledge. The listeners of the first group had a very limited German knowledge and the fourth group's listeners were rated to be the most advanced group. This rating was done by German language teachers at the University of Oldenburg. The pure-tone threshold of all the subjects at audiometric frequencies from 250 to 8000 Hz did not exceed 15 dB HL.

Procedure

To investigate the influence of the linguistic complexity on speech intelligibility by NN for the each subject speech reception threshold (SRT, speech to noise ratio yielding 50% intelligibility) was measured using three different German speech intelligibility tests: the Digit Triplets Test (DTT), the Oldenburg Sentence Test (OLSA) and the Göttingen Sentence Test (GÖSA). These tests differ in their linguistical complexity.

The DTT is a screening hearing test and has been developed especially for speech intelligibility measurements via telephone or Internet [1] [2]. It contains strongly limited speech material, namely digit triplets (complexes of three digits that are spoken separately). Each list of the DTT is composed of 27 digit triplets.

The second test used in a present study was the OLSA test [3][4] which incorporates 5-word semantically unpredictable sentences of a fixed grammar structure (*name, verb, number, adjective,* and *noun*). The utterances are composed of words from a 50-word base matrix. There are two versions of the OLSA test available: an open and a closed set version. Using an open set version, the subjects' task is to repeat words which have been understood. The experimenters' task is to compare subjects' verbal responses with the sentence displayed on a touch screen and indicate the correctly repeated words. A closed set version does not require participation of an experimenter. After presentation of a sentence to a subject a 50 word panel containing all words of the OLSA test is displayed on a touch screen. The subjects' task is to indicate words understood.

The most complex speech material used in the current study was represented by the GÖSA test [5]. This test comprises meaningful sentences of everyday speech, of a different semantical and syntactical difficulty.

In each of the tests mentioned above, during a listening session, speech was presented against a background noise at different signal to noise ratios (SNRs), by means of the adaptive procedure. The algorithm adaptively changed the SNR to a value corresponding to the 50 % speech intelligibility (SRT). For each test a masking noise was composed by superposition of all items of a particular test, so that the average spectrum of a certain masking noise matched the average spectrum of the corresponding speech material.

Speech intelligibility measurements were performed using following order:
For the DTT, one training list was presented via headphones. Next, SRT was measured using one list via telephone. For the OLSA test, training session was performed by presenting one 30 sentences list in silence with a closed set version. During the training the speech level was kept constant at 65 dB SPL. In a next step, subjects were faced with two lists in noise, each of 20 sentences, using an adaptive procedure, whereas one list was presented using an open set version and the second one with a closed set version. For the GÖSA test a short training session with presentation of ten sentences in noise was also done. After that SRT was measured using a 20 sentences list presented against a masking noise. During listening session, the order of speech tests was randomized.

Results

Measured SRTs were pooled across subjects for each test and each group of listeners and subjected to a two-way ANOVA test with respect to the 'test' and 'group' factors. The ANOVA test revealed that both factors were statistically significant ($p<0.001$). Figure 1 depicts mean SRTs obtained for different tests and different NN groups as well as

normative data for German listeners taken from the literature [5] [2] [6].

Effect of speech intelligibility test type

Mean SRT for the DTT ranged from –2.7 to -4.3 dB SNR and decreased with increasing German language skills. Similar SRTs were obtained for both versions of the OLSA test and varied from –2.3 to -4.7 dB SNR for the open set version and from –2.5 to –4.9 dB SNR for the closed set version, whereas the highest SRTs were obtained for the first group of subjects and the lowest one for the most advanced group. For the GÖSA test, SRTs fell into the range from 4.6 dB SNR for the first group of subjects to –2.1 dB SNR for the fourth group. Post hoc comparisons (with Bonferroni adjustment for multiple comparisons) showed statistically significant differences ($p<0.001$) between the GÖSA and all other speech tests used in this study.

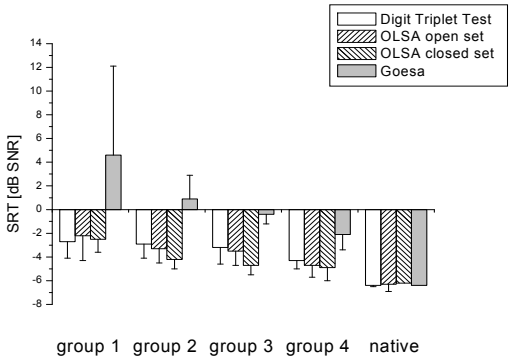

Figure 1: Mean SRTs and corresponding standard deviation for different speech intelligibility tests and different listeners' groups (with increasing German language skills with increasing group index) and reference data for native listeners.

Between-group differences for NN listeners

The smallest differences in SRT between different groups of NN were obtained when speech intelligibility was measured with tests of less linguistic complexity, namely the DTT and both versions of the OLSA test. SRTs for the GÖSA revealed significant differences between groups up to 6.7 dB SNR between the first and the fourth group and also large differences within groups, particularly in the group with only a basic German knowledge. Post hoc pairwise comparisons (Bonferroni test) showed significant differences between the first group of subjects and all other groups. According to the assumption, the biggest differences in speech recognition were observed between the first and the fourth group of NN and were larger the more complex speech material was used.

Differences in speech recognition by NN and natives

Non-native listeners' ability to understand speech in noise improved as their German language knowledge increased, but they did not perform as well as natives. Table 1 shows differences in SRT between NN and native listeners dependent on speech material and German language skills of non-natives. For relatively 'easy' speech material (DTT and OLSA test) speech intelligibility data revealed differences in SRTs between NN and natives that ranged from 2.4 dB SNR (data averaged across the 2nd, 3rd, and 4th group of NN) to 3.8 dB SNR for the first group of subjects. Significantly poorer speech intelligibility in noise for all groups of NN was observed with more complex speech material (the GÖSA test). For everyday speech, differences in SRTs between NN and natives were 5.6 dB for listeners with better German skills and 10.8 dB for subjects with limited German knowledge.

Table 1: Differences in SRT between NN and native listeners for tests with less complex speech material (DTT and OLSA test) and the linguistic more complex test (GÖSA).

	Δ SRT (Group 1-native) [dB]	**Δ SRT** (Group 2,3,4 -native) [dB]
DTT and OLSA test	3,8	2,4
GÖSA	10,8	5,6

Conclusions

The present study showed that speech recognition in noise by NN was strongly dependent on the linguistic complexity of speech material. The more complex material yielded larger differences in SRT between NN and native listeners and between different groups of NN. Speech recognition in noise was also related to German language knowledge of NN. Listeners with longer time of exposure to the second language revealed better speech in noise performance than NN with limited German language skills.

Acknowledgments

This work was carried out as a part of the ERDF project "Netzwerk für multilinguale Hör- und Sprachverständlichkeitsdiagnostik" (HurDig).

Literatur

[1] Smits, C., T. Kapteyn, and T. Houtgast,: Development and validation of an automatic speech-in-noise screening test by telephone. International Journal of Audiology (2004) 43: p. 15-28.

[2] Wagener, K.C.: Evaluation des Ziffern-Tripel-Tests über Telefon. DGA Jahrestagung. 2006.

[3] Wagener, K., T. Brandt, and B. Kollmeier: Development and evaluation of a German sentence test I:Design of the Oldenburg sentence test. Z.Audiol (1999) 38: p. 4-15.

[4] Wagener, K., T. Brandt, and B. Kollmeier: Development and evaluation of a German sentence test III: Evaluation of the Oldenburg sentence test. Z.Audiol (1999) 38: p. 86-95.

[5] Kollmeier, B. and M. Wesselkamp: Development and evaluation of a sentence test for objective and subjective speech intelligibility assessment. Journal of Acoustical Society of America (1997) 102(4): p. 1085-1099.

[6] Brand, T., et al.: Vergleich von Oldenburger Satztest und Freiburger Wörtertest als geschlossene Versionen. in DGA 7. Jahrestagung 2004, Leipzig.