FIST: A French Sentence Test for Speech Intelligibility in Noise


Heleen Luts

Ellen Boon

Jocelyne Wable
Saint-Mandé, France

Jan Wouters

Corresponding author:
Heleen Luts
Herestraat 49/721
B-3000 Leuven
Belgium
Tel: +32 16 33 04 77
Fax: +32 16 33 04 86
E-mail: heleen.luts@med.kuleuven.be
Introduction

Recently, the Hearing in Noise Test (HINT) was developed for Canadian Francophone listeners (Vaillancourt et al., 2005). Because of the large differences between Canadian and European French, the French Intelligibility Sentence Test (FIST) was developed using procedures similar to the HINT for adaptive testing of speech intelligibility in noise in Europe. Reference data were obtained for normal-hearing listeners and the performance of Francophone listeners from France and Belgium was compared.

Methods and Results

Preparation of test materials

The initial test material for the FIST was developed by Wable (2001). Sentences with various syntactic structures were created using a database with typical French words (Content et al., 1990). The written sentences were judged by an informal jury according to difficulty, level of abstraction, naturalness and word content. The jury consisted of several French speakers from France who differed in age and educational degree. Sentences judged inadequate by at least one member were eliminated. A total of 582 sentences were selected. These were read by a French male speaker and recorded.

Equalization of sentence difficulty

Unlike the HINT materials, sentence intelligibility of materials in the present study was equalized without modifying the sentence level, but by selecting a subset of sentences with equal intelligibility scores in noise for hearing-impaired subjects. In total, 433 French subjects with pure-tone averages ranging from 20 to 83 dB HL participated. In the first phase, the sentences were presented at a signal to noise ratio (S/N ratio) of 0 dB and 200 sentences with 67% to 87% correct responses were selected. In the second phase, sentences were presented at S/N ratios of 0 dB and -2 dB. The percentage of correct responses for both presentation levels had to differ by at least 25% to make sure there was an appreciable noise effect on the sentence difficulty. A final selection of 160 sentences remained. The number of syllables per sentence ranged from 6 to 15, with an average of 10.

Formation of sentence lists

These 160 sentences were evaluated at S/N ratios of 0 dB, -5 dB and -10 dB on 12 normal-hearing Francophone Belgian subjects, with hearing thresholds of 15 dB HL or better at all octave frequencies between 125 and 8000 Hz. Exact repetition of the sentences was required. Based on these results, the 140 best sentences were combined into 14 test lists of 10 sentences, with the aim to obtain the most homogeneous set of test lists with regard to speech reception thresholds (SRTs) and the slope of the performance-intensity (PI) function. The remaining 20 sentences were used to create a training list.

Development of norms and reliability, and estimation of the PI function

SRTs were obtained from 20 Francophone normal hearing adults. Ten subjects were from France and the other 10 were from Belgium. Their hearing thresholds were equal to or better than 25 dB HL at all octave frequencies between 125 and 8000 Hz. The sentence material was presented through headphones using an adaptive test procedure in a stationary speech-
weighted noise that has the same long-term average spectrum as the sentences at a fixed level of 65 dB SPL.

Average SRTs are shown in Table 1. There is a small, but significant difference in SRT of 0.7 dB between the two groups of listeners \[ t(18) = -2.4, \ p < .05 \]. SRTs measured using individual lists deviate from the overall mean by at the most 1.1 dB. The inter-list variability, calculated as the quadratically averaged within-subject standard deviations of repeated measurements, was 1.0 dB and 1.1 dB for French and Belgian listeners respectively. The mean within-list variability, as defined in the introduction paper earlier, was 1.6 dB (SD = 0.4 dB) for both subject groups.

The PI function is estimated based on the scores at the different presentation levels within the adaptive procedure. The slopes at 50% intelligibility are calculated based on non-linear regression to a logistic function of the PI curve of each subject (see Table 1). The slope of the PI function was not significantly different between the two subject groups \[ t(18) = 1.3, \ p > .05 \].

**Discussion and conclusions**

Although the number of sentence lists is rather limited, the PI function of the test material has a very steep slope and can provide accurate speech intelligibility scores in noise. There is a small difference in SRT between the French and Belgian listeners. This has two possible explanations. The speech material is uttered by a French male voice, which could be slightly more difficult for the Belgian listeners. A second explanation could be the fact that 6 out of 10 Belgian subjects were simultaneously bilingual, meaning they learned two first languages in early childhood, and they might have more difficulty with speech recognition in noise than monolingual listeners (Grosjean, 1989).

The French test battery for the assessment of speech intelligibility in noise will soon be completed with a closed-set sentence test and a digit triplet test suited for telephone screening, in addition to the material discussed in this paper.

**Acknowledgements**

This work was supported by grants from the European Union FP6, Project 004171 HearCom. Heleen Luts is a postdoctoral researcher funded by the Research Fund K.U.Leuven. The authors thank Marieke Delfreux for her assistance in data collection.
References


Table 1: The average SRTs and standard deviations (SD) across subjects are shown. The SRT is expressed as the S/N ratio required for 50% speech intelligibility. It is defined as the average of the presentation levels of the following sentence for the 5th to the 10th sentence in the list, in an adaptive test procedure. The slope of the PI function is calculated based on fitted data. The precision is expressed as the quadratically averaged standard errors (SE) of the fit to the data of each subject.

<table>
<thead>
<tr>
<th></th>
<th>SRT (dB S/N ratio)</th>
<th>Slope (%/dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>France</td>
<td>-7.8</td>
<td>0.6</td>
</tr>
<tr>
<td>Belgium</td>
<td>-7.1</td>
<td>0.7</td>
</tr>
<tr>
<td>All subjects</td>
<td>-7.4</td>
<td>0.7</td>
</tr>
</tbody>
</table>