

Speech recognition under adverse listening conditions in young NH listeners

V. Stenbäck^{1,3}, M. Hällgren^{3,4}, B. Lyxell^{2,3}, B. Larsby^{1,3}

¹ Technical Audiology/IKE, Linköping University, Sweden

² Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden.

³ Linnaeus Centre HEAD, Linköping University, Linköping, Sweden.

⁴ Department of Otorhinolaryngology/Section of Audiology, Linköping University Hospital, Linköping, Sweden

Introduction

- Disruption in the intelligibility of speech may occur due to physical characteristics of noise (energetic masking), or due to the informational content of the masker (informational masking).
- Various maskers have different effect on speech recognition, and some properties of a masker can facilitate speech recognition, while other properties hinder recognition.

Purpose

Examine the differences in performance in young normally-hearing (NH) individuals on Hagerman's speech-recognition-in-noise test when both energetic and informational maskers are used to mask target sentences.

Participants

16 young NH individuals (M age= 26,8 years (SD=5)).

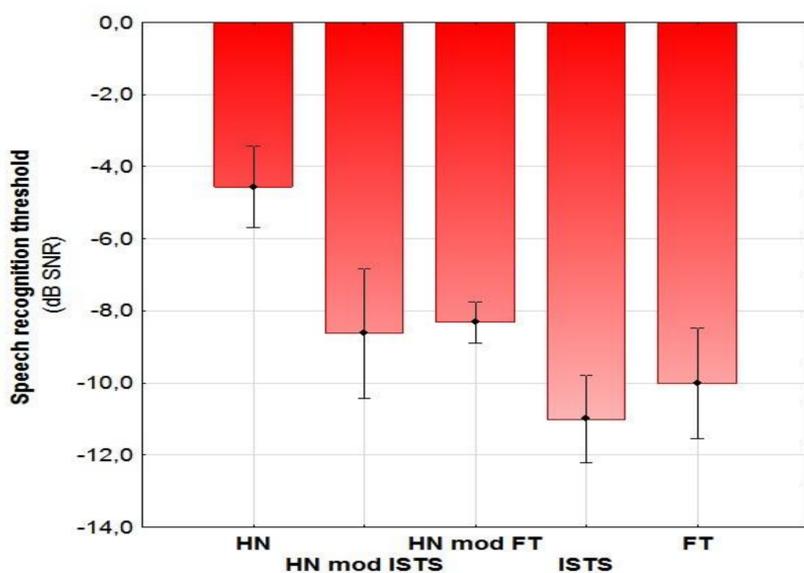
Material

- **The Hagerman speech-in-noise test** – The Hagerman sentences are low-redundancy five-word sentences with same syntactic structure. The sentences begin with a name, followed by a verb, number, adjective and lastly a noun (example: Jonas gave nine black rings). The original Hagerman noise is a slightly (10%) modulated speech-shaped noise (HN). We masked the sentences with 5 different maskers and obtained SNRs for 80% word recognition. The different maskers used were Hagerman's original noise (HN), the International Speech Test Signal (ISTS), a single female talker reading a section in a Swedish newspaper (FT), and two speech-shaped noises modulated with the envelopes of the ISTS (HN mod ISTS) and the female talker (HN mod FT), containing no speech or semantic information.

Results

An ANOVA revealed a main effect of noise type $F(4, 60) = 23.43$, $p < .001$, where HN required more favourable SNRs than the other noise types used as maskers.

A paired samples t-test was conducted to compare FT to HN mod FT, and ISTS to HN mod ISTS. There was a significant difference in speech recognition performance between FT (M=-10, SD=2.9), and HN mod FT (M=-8.3, SD=1.1), $t(15) = -2.6$, $p = .018$. There was also a significant difference between ISTS (M=-8.6, SD=3.3), and HN mod ISTS (M=-11, SD=2.3), $t(15) = -2.3$, $p = .039$.



The figure shows a significant difference between HN and the other maskers, $p < .001$, and between the informational maskers and the energetic maskers. Error bars denote 95% confidence intervals.

Discussion and Conclusions

- Temporal and spectral gaps, and the informational content of the masker influence speech recognition differently. The results indicate that when modulating HN with the envelopes of the ISTS (HN mod ISTS) or the female talker (HN mod FT), the participants exhibited greater difficulties in segregating the target speech from the masker, than when the single female talker or the ISTS were used to mask the target sentences. HN had the greatest masking effect.
- An explanation to the findings is that young NH individuals are quite successful in segregating target speech from the masking speech, hence knowing which speech stream to focus on. When the maskers contain temporal dips, the participants listen in the dips, but if a word is masked by the noise, it becomes more difficult to decipher than if the target word is masked by another word. HN had the greatest masking effect, which can be explained by the lack of temporal and spectral dips, hence no pauses to fill in obscured information.
- Our conclusion is that temporal gaps may offer release from informational maskers for young NH individuals.