Do non-native listeners benefit from speech modifications designed to promote intelligibility for native listeners?

Martin Cooke[1,2], Maria Luisa Garcia Lecumberri[2], Yan Tang[2], Mirjam Wester[3]

[1]Ikerbasque (Basque Science Foundation), Bilbao, Spain
[2]Language and Speech Laboratory, University of the Basque Country, Vitoria, Spain
[3]Centre for Speech Technology Research, University of Edinburgh, Edinburgh UK

1. Motivation

Given the increasing use of sound output technologies in cluttered acoustic environments such as transport interchanges, it is of interest to discover ways to modify generated speech which maximise the likelihood of correct message reception, without resorting to excessive sound output level or repetitive announcements. Additionally, speech modifications deployed in public spaces should ideally be beneficial to all listeners, regardless of their first language.

2. Design

We compared the effect of a number of speech modification strategies in noise on native and non-native listeners. All were based on energy reallocation across time and frequency under constant SNR and duration constraints [1, 2]. British English and Spanish listeners identified letter and digit keywords in short sentences for natural speech (**original**) and for six types of modification:

SegSNR frame-wise SNR equalisation

ChanSNR frequency channel SNRs equalised

LocalSNR SNRs equalised in time-frequency region

SelectBoost energy reallocation to boost just-audible time-frequency regions

Pausing introduction of a short pause to avoid most intense noise epoch

Combined SelectBoost + Pausing

Speech was presented mixed with either stationary speech-shaped noise (SSN) or speech envelope modulated noise (SMN) at a global SNR of -6 dB.

3. Results

While native and non-native listeners' absolute scores differed by about 14 percentage points, the pattern of scores as a function of modification type was strikingly similar, with a correlation of 0.97 (see figure 1). For example, both listener groups benefitted most from selective boosting of regions close to threshold, and both were adversely affected by pausing in the presence of stationary noise.
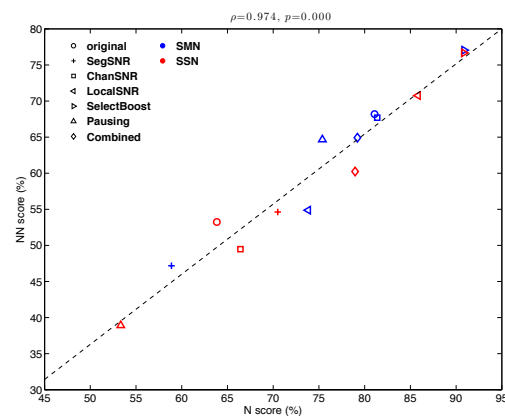

Figure 1: *Native and non-native scores for natural and modified speech in stationary and fluctuating noise.*

4. Discussion

Speech signal modifications exist which are equally effective for native and non-native listeners, suggesting that these modifications result in language-independent auditory changes which may be effectively transferred to other target languages and listener populations.

5. References

[1] Y. Tang and M. Cooke, "Energy reallocation strategies for speech enhancement in known noise conditions," in *Proc. Interspeech*, Makuhari, Japan, 2010, pp. 1636–1639.

[2] Y. Tang and M. Cooke, "Subjective and objective evaluation of speech intelligibility enhancement under constant energy and duration constraints," in *Proc. Interspeech*, Florence, Italy, 2011, pp. 345–348.