Masking Effects on Perceptual Cues in Hearing-Impaired Ears

Siyan Guo[1], Jont Allen[2], Cliston Cole[3]

[1]University of Illinois, Urbana-Champaign, Email: sguo16@illinois.edu
[2,3]University of Illinois, Urbana-Champaign

Objectives

Current hearing-aids mainly work by improving audibility. However a common complaint is "I can hear the speech but I can't understand it." This has led us to investigate the strategy of the hearing-impaired (HI) ear in listening to speech. In this study, effects of varying masking on primary cues in HI ears are analyzed, with and without the presence of conflicting cues [1]. Using an error pattern analysis, the results provide insights into the HI ear's listening strategy.

Introduction

One of the common problem of the current signal processing techniques for hearing aids is the methodology of the amplification strategy. For example, some strategies use a formula that amplifies high, mid, and low frequencies (i.e. National Acoustic laboratory Revised (NALR)). A hearing impairment with only high-frequency hearing loss necessitates amplification only at the high frequencies and not the mid or low. Therefore, for hearing aids to better enhance speech perception, it is essential that signal processing engineers understand the necessary and sufficient perceptual cues that an HI ear uses for correct recognition. We shall show that the failure of past strategies is due to inadequate metrics of performance because of a poor understanding of critical speech cues. We shall also show that HI ears are using the same primary cue as NH ears. The reason why the HI cannot understand speech in noise as NH ears is because they cannot hear these critical speech cues, due to a combination of hearing loss and masking effects of the noise.

Perceptual Cues

The following figure shows an example of perceptual cues within a spoken utterance in the form of AI-gram:

- Time-frequency energy patterns of spoken utterance /ka/
- Box 1 indicates the *primary cue* /k/
- Box 2 indicates the *conflicting cue* at higher frequency /t/
- Box 3 indicates the conflicting cue at lower frequency /p/

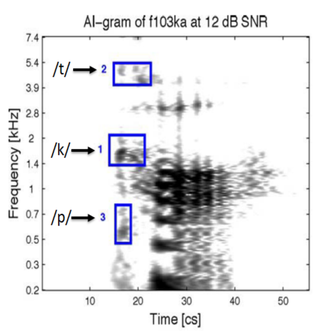

Figure 1: Example of utterance /ka/ spoken by female talker 103 at 12 [dB] SNR

Design

Two consonant-vowel (CV) identification experiments were conducted on 5 NH and 10 HI subjects (17 ears). Four plosive consonants (tokens) /t,k,d,g/ paired with vowel /a/, in CV context, were used as target stimuli. The CVs were presented at signal-to-noise ratios (SNRs) of 0, 9, and 18 dB. In experiment *one* the primary cue for each CV was presented 5 ways: removed, attenuated by 6dB, unmodified, and amplified by 6dB and 12dB. Experiment *two* was similar to experiment one except the conflicting cues were removed. Subjects' responses were analyzed based on average probability of error over SNR.

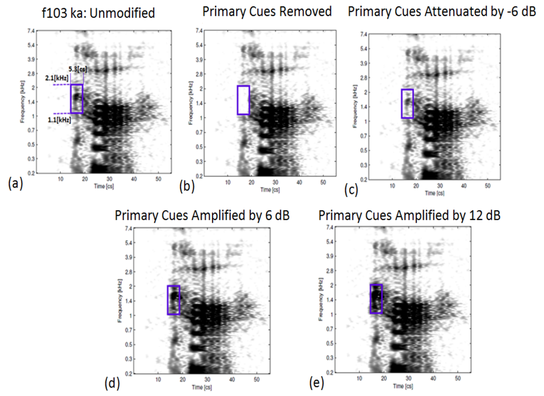

Figure 2: Varying the masking of the prmary cue for female talker f103: (a) Unmodified version. (b) Primary cue removed. (c) Primary cue attenuated. (d) Primary cue amplified by 6 [dB]. (e) Primary cue amplified by 12 [dB].

Audiogram

All HI subjects were given an audiological evaluation. As shown in the figures, HI subjects have been separated into a low-error group (LEG) and a high-error group (HEG). The method used to separate them is described in the results section.

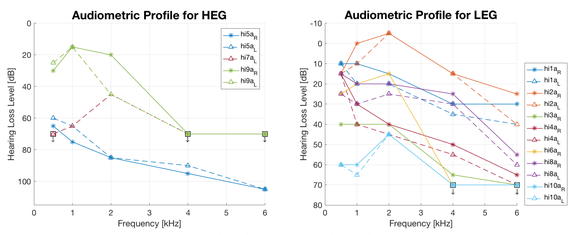

Figure 3: Audiogram for HEG (left) and LEG (right): The black arrow indicate ears that have HL greater than 70 [dB] (e.g. severe). HI5a from HEG has profound HL (i.e. $\geq$ 90 [dB]) for both ears at 4 and 6 [kHz]

Results

Here we study the result of how masking the plosive's primary cues affect the intelligibility for both NH and HI ears. The performance of NH subjects are consistent, with lower error.

Results (Cont'd)

As shown in figure 4, average normal hearing (ANH) group has 0% error for all tokens for 9 and 18 dB SNR. Two tokens shows ∼10% error at 0 dB SNR. We observe from the figure that when the primary cue is removed, the performance of the ANH group significantly degraded. For most of the tokens, the error increases to over 90%, with a few at 50% (figure 4c). This pattern shows ANH depends on primary cue for speech perception.

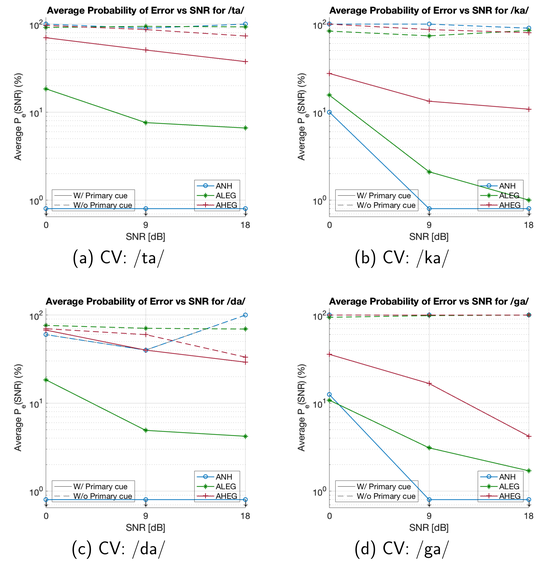

Figure 4: Average probability of error as a function of SNR for both conditions: primary cue present (solid) and primary cue removed (dashed). The black arrow indicates 0% error. Token /da/ has the lowest error for all 3 groups when primary cue is removed. ANH: Average normal hearing group. ALEG: Average low-error group. AHEG: Average high-error group

The errors of unmodified tokens for all 3 SNR levels for HI ears naturally divide into two groups: low-error group (LEG) and high-error group (HEG). The HI ears in the LEG tend to make fewer errors than the HI ears in the HEG. Additionally, the audiometric profiles from each HI group show a significantly different hearing loss (HL), mainly in the high frequency region at or above 4 kHz.

The averaged probability of error for both groups are shown in figure 4. The LEG data is similar to the ANH group, but with slightly higher error. Subjects in the HEG exhibit much higher error under all circumstances when compare to ANH and LEG, especially for high frequency sounds such as /ta/ and /da/, due to the severe high frequency HL.

To determine the role of primary cue for HI ears, we remove the primary cue and study the errors. Both HEG and LEG showed dramatic degradation in the performance when the primary cue is removed. This demonstrates that HI ears are using the same primary cue as NH for speech perception.

Results (Cont'd)

The effect of removing conflicting cues is not significant for ANH and LEG, which reinforces the observation that the primary cue is the most important speech feature. On the other hand, HEG subjects show increased sensitivity to the absence of conflicting cues. Their performance is either improved or degraded depending on the subject. Some subjects in the HEG may even be using conflicting cues.

Summary

- The primary cue is critical and vital information for correct identification of plosive consonants, for both NH and HI ears
- LEG subjects are more likely to benefit from hearing-aids than HEG subjects

Conclusion

Our analysis shows that HI listeners are using the same primary cue as NH ears. The strength of the primary cue is a critical quality for low-error HI speech perception, especially important in the presence of noise.

Overall, the LEG showed results similar to the average normal hearing (ANH) group, for both experiments, but with slightly higher error. The removal of the conflicting cues did not have notable impact on their performance. However, the subjects in the HEG exhibit sensitive reaction towards the presence of conflicting cues. For the HEG, the removal of conflicting cues can either enhance or diminish speech perception. Interestingly, a few ears from HEG depend on conflicting cues for correct recognition.

Our results suggest that the LEG are more likely to benefit from hearing-aids. Thus it would be clinically useful for classifying subjects, to predict success with their aided condition, as well as to improve the fitting algorithm.

References

[1] F. Li and J.B. Allen. Manipulation of consonants in natural speech. *IEEE Trans. Audio, Speech, Lang. Process.*, 19(3):496-504, 2011.

[2] A. Kapoor and J.B. Allen. Perceptual effects of plosive feature modification. *The Journal of the Acoustical Society of America*, 131(1):478-491, 2012.

[3] C. Cole. *On the Effects of Masking of Perceptual Cues in Hearing-Impaired Ears*. PhD Dissertation, Univ. of Illinois, Urbana-Champaign, 2017.

ILLINOIS
Electrical & Computer Engineering
COLLEGE OF ENGINEERING