The Role of Periodicity in Perceiving Speech in Quiet and in Background Noise with Simulated Cochlear Implants

Kurt Steinmetzger [1], *Stuart Rosen*[1]

[1] Department of Speech, Hearing and Phonetic Sciences, University College London, UK

kurt.steinmetzger.12@ucl.ac.uk, s.rosen@ucl.ac.uk

Index Terms: speech, noise, cochlear implants, periodicity

1. Introduction

Previously we have investigated the ability of normal-hearing listeners to perceive sentences in a variety of conditions involving the presence and absence of periodicity in both target speech and masker. Listeners were found to substantially benefit from masker periodicity, while there was no effect of periodicity in the target speech. Moreover, the masker-periodicity benefit (MPB) was substantially larger than the fluctuating-masker benefit (FMB) obtained from sinusoidal 10 Hz modulations. Here we used a similar design and tested to what extent MBP and FMB were maintained in simulations of cochlear implants (CIs).

Factors that are thought to explain the periodicity benefit include the ability to cancel out harmonic maskers from the signal mixture [1], the possibility to glimpse sections of the target speech in between the individual harmonics [2], and the sparse modulation spectrum [3]. Access to spectral information is severely limited with current CIs [4], limiting the role of the former two factors, but we still expected listeners to benefit from temporal regularity and reduced modulation masking.

Most CI users also show little or no FMB [5], which can at least in part be attributed to the elevated Speech Reception Thresholds (SRTs) [6]. Kwon and colleagues [7] recently introduced a masker that maximises glimpsing opportunities by reducing the temporal overlap with the speech. Quite surprisingly, even in the absence of energetic masking, few of their listeners showed substantial FMBs, suggesting that CI users may find it particularly difficult to identify the segmental boundaries between speech and noise. In order to further investigate this idea, the current study included this type of masker in addition to steady and modulated interferers.

2. Methods

Twelve normal-hearing listeners were tested (18–21 years). Targets were recordings of the BEL sentences [8] spoken by a male Southern British English talker, processed to have an aperiodic, mixed, or periodic source excitation using TANDEM -STRAIGHT [9]. Harmonic complexes based on the F0-contours of real speech (interpolated through silence and voice-lessness) were used as periodic maskers, while speech-shaped noise was used as aperiodic masker. Masker envelopes were either steady, 10-Hz modulated, or adjusted to be the inverse of the target sentence envelope in 50 ms steps (+MR) [7]. The signal mixture was 8-channel noise-vocoded to simulate CI processing (equal basilar membrane distance, frequency range 70 Hz–4 kHz, 400 Hz envelope low-pass filter). Participants listened to 20 sentences in each of the 18 conditions (3 targets x 6 maskers) and SRTs were measured by tracking 50% correct.

3. Results

Listeners benefitted substantially from masker periodicity and to a lesser extent also from sinusoidal modulations. However, SRTs with the +MR maskers were similar to those with steady maskers. A mixed effects model returned significant main effects of masker periodicity ($p < .001$) and masker envelope ($p < .001$), but neither the main effect of target periodicity ($p = .29$), nor any of the interactions were significant ($p \geq .41$).

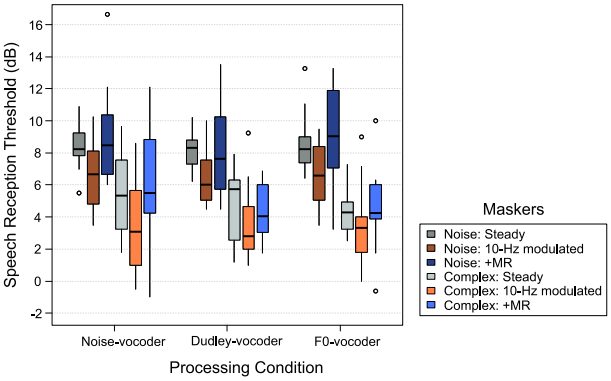

Figure 1: *Boxplots of the Speech Reception Thresholds.*

4. Discussion

Compared to normal hearing, MPB and FMB were reduced to about half their size with simulated CIs (on average from about 9 to 4 dB and 4 to 2 dB). However, even in the absence of salient pitch cues, the MPB was still twice as large as the FMB. In addition, although they did not energetically mask the targets, the +MR maskers with their speech-like envelopes led to similar SRTs as steady interferers. Because spectral cues that aid stream segregation are largely unavailable, CI processing thus appears to have a particularly detrimental effect on the listeners' ability to distinguish segments of target speech and masker.

5. Conclusions

Although its perception is restricted by CI signal processing, masker periodicity was found to aid speech perception to a greater extent than superimposed masker envelope fluctuations.

6. Acknowledgements

This project has been funded with support from the European Commission under Contract FP7-PEOPLE-2011-290000 and the MRC, UK (Grant G1001255).

7. References

[1] A. de Cheveigné, "Cancellation model of pitch perception", *Journal of the Acoustical Society of America*, vol. 103, no. 3, pp. 1261-1271, 1998.

[2] M. L. Deroche, J. F. Culling, M. Chatterjee, and C. J. Limb, "Roles of the target and masker fundamental frequencies in voice segregation", *Journal of the Acoustical Society of America*, vol. 136, no. 3, pp. 1225-1236, 2014.

[3] M. A. Stone, C. Füllgrabe, and B. C. Moore, "Notionally steady background noise acts primarily as a modulation masker of speech", *Journal of the Acoustical Society of America*, vol. 132, no. 1, pp. 317-326, 2012.

[4] T. Green, A. Faulkner, and S. Rosen, "Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants", *Journal of the Acoustical Society of America*, vol. 112, no. 5, pp. 2155-2164, 2002.

[5] H. E. Cullington and F.-G. Zeng, "Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects", *Journal of the Acoustical Society of America*, vol. 123, no. 1, pp. 450-461, 2008.

[6] J. G. Bernstein and K. W. Grant, "Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners", *Journal of the Acoustical Society of America*, vol. 125, no. 5, pp. 3358-3372, 2009.

[7] B. J. Kwon, T. T. Perry, and C. L. Wilhelm, "Sentence recognition in noise promoting or suppressing masking release by normal-hearing and cochlear-implant listeners", *Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 3111-3119, 2012.

[8] L. Calandruccio, and R. Smiljanic, "New sentence recognition materials developed using a basic non-native English lexicon", *Journal of Speech, Language, and Hearing Research*, vol. 55, pp. 1342-1355, 2012.

[9] H. Kawahara, M. Morise, T. Takahashi, R. Nisimura, T. Irino, and H. Banno, "TANDEM-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation", *Proceedings of ICASSP*, pp. 3933-3936, 2008.