R3 Periodicity and aperiodicity in the perception of speech in both steady-state and fluctuating maskers

Kurt Steinmetzger & Stuart Rosen
UCL Speech, Hearing & Phonetic Sciences, London UK

Introduction

Users of cochlear implants (CIs) experience great difficulties in speech perception in all types of background noise, and show little benefit from fluctuations in the masker. One popular (if partial) explanation for these difficulties proposes a key role for temporal fine structure (TFS) cues which are severely limited by CI speech processing. However, there is controversy over whether TFS has a special role in allowing fluctuating masker benefit or whether its contribution to speech perception is just as important for steady maskers. Here, as a precursor to studies with CI listeners, we investigate the role of periodicity in targets and maskers for normal-hearing listeners.

Methods

Speech Reception Thresholds (SRTs) were measured adaptively for sentences processed to change their source characteristics (and hence their periodicity) in various ways.

Stimuli

Speech targets Based on IEEE sentences from an adult male talker, using two vocoding methods to manipulate the voice source. These vocoders differ substantially in the way in which they estimate spectral envelopes, and we wanted to be sure that any of our findings were not dependent on a particular method.

- a standard 24-channel vocoder (Dudley, 1939)
- TANDEM-STRAIGHT (Kawahara et al., 2008)

With excitation sources

- FxNx – pulses which track the speech F0 when it is voiced, with noise otherwise
- Nx – noise always
- Fx – pulses which track the speech F0, interpolated through periods of silence and voicelessness to make a continuous contour

Resulting in unprocessed speech plus 5 processed conditions:

- Nx-vocode – standard noise vocoding; envelope extraction by full-wave rectification and lowpass filtering at 30 Hz
- Nx-STRAIGHT – similar to Nx-vocode
- FxNx-vocode – standard channel vocoding which is highly intelligible, preserving periodic/aperiodic distinction except for mixed excitation
- FxNx-STRAIGHT – similar to FxNx-vocode, but with preservation of mixed excitation, and a more natural quality
- Fx-STRAIGHT – spectral dynamics as for FxNx-STRAIGHT, but with all excitation periodic

Maskers 2 periodicities x 2 envelopes, all shaped in spectrum to match that of the talker

Masker periodicity

- **Noise maskers:** speech-spectrum shaped noise
- **Periodic maskers:** harmonic complexes with dynamic F0 as determined from 1 of 16 talkers (chosen randomly on each trial), interpolated through periods of silence and voicelessness to make a continuous contour. A simplified version of a single talker.

Masker envelopes

- steady-state
- 100% sinusoidal amplitude modulation at 10 Hz

Listeners

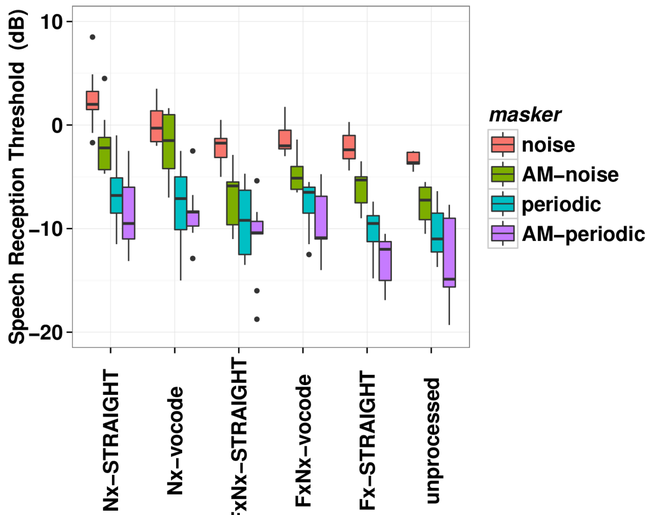
9 normal hearing young adults, native speakers of British English

Acknowledgements

Supported by The Medical Research Council (UK) and the EU ITN International Training Network INSPIRE. Thanks to Tim Green for useful comments.

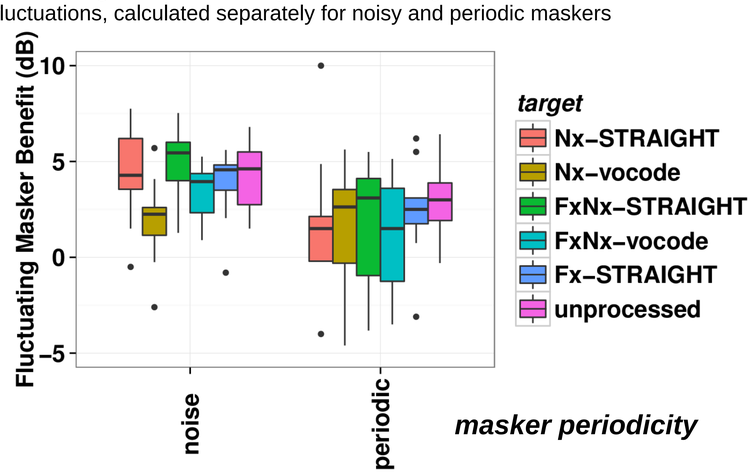
MRC Medical Research Council

Results



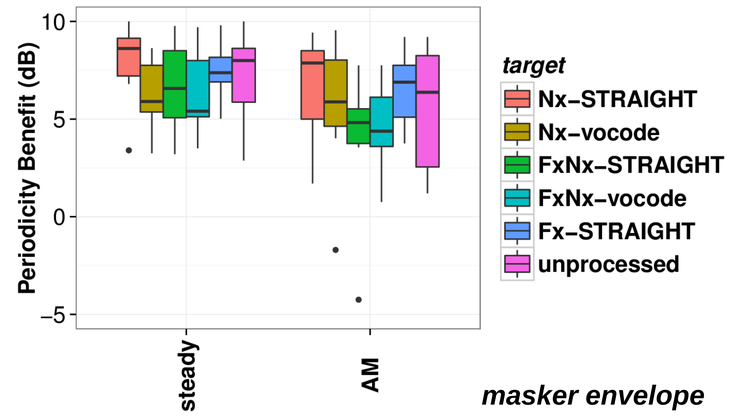
- ***Targets differ in susceptibility to masking:*** Average effect of *target* (~5 dB range) may be linked to target periodicity, but perhaps also to inherent intelligibility (although all targets were 100% intelligible in quiet)
- ***Masker periodicity more important than masker fluctuations:*** Average effect of *masker envelope is* ~3 dB whereas the average effect of *masker periodicity is* ~7 dB

Fluctuating Masker Benefit (FMB): the change in SRT due to masker fluctuations, calculated separately for noisy and periodic maskers



- ***FMB is greater for noisy maskers than periodic ones (~4 dB vs ~2 dB):*** as evidenced by a significant *masker envelope* x *masker periodicity* interaction. Perhaps the dynamic F0 contour in periodic maskers reduces FMB?
- ***Smaller FMB for Nx-vocode than Nx-STRAIGHT:*** Needs investigation!

Periodicity Benefit: the change in SRT due to masker periodicity, calculated separately for steady-state and fluctuating maskers



Periodicity benefit is greater for steady-state maskers than fluctuating ones (~8 dB vs ~6 dB): as evidenced by a significant *masker envelope* x *masker periodicity* interaction

Discussion

Periodicity in the masker strongly reduced its masking effectiveness, supporting the notion of harmonic cancellation (de Cheveigné et al., 1995). But periodicity in the target may also improve SRTs (see also Vestergaard & Patterson, 2009)

Periodicity in the masker had a larger effect in steady-state maskers than in those that fluctuate in amplitude. Thus, it appears unlikely that sensitivity to TFS has a special role in glimpsing (see also Moore, 2011)

Final remarks

Obs! These are not CI simulations, even if noisy targets in noisy maskers are very similar to those. Envelope information concerning target and masker are here kept independent, but are applied to a single carrier in simulations.

Inability to exploit periodicity in the masker may be a more important limitation to CI speech perception in backgrounds of other voices than an inability to glimpse.

References

de Cheveigné, A., McAdams, S., Laroche, J., and Rosenberg, M. (1995). "Identification of Concurrent Harmonic and Inharmonic Vowels - A Test of the Theory of Harmonic Cancellation and Enhancement," J. Acoust. Soc. Am. 97, 3736-3748.

Dudley, H. (1939). "Remaking speech," J. Acoust. Soc. Am. 11, 169-177.

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

Moore, B. C. J. (2011). "The importance of temporal fine structure for the intelligibility of speech in complex backgrounds," in *Speech Perception and Auditory Disorders*, T. Dau, J. Dalsgaard, M. Jepsen & T. Poulsen (eds): Centertryk A/S, Denmark), pp. 21-32.

Vestergaard, M. D., and Patterson, R. D. (2009). "Effects of voicing in the recognition of concurrent syllables," J. Acoust. Soc. Am. 126, 2860-2863.