



The critical role of temporal envelopes in sound source segregation

Michael Stone

Manchester Centre for Audiology and Deafness, School of Health Sciences, Faculty of Biology, Medicine and Health, UK

The measurement of human auditory modulation processing capabilities employs a signal comprising an extrinsic modulation impressed onto a carrier. If the carrier is sinusoidal, then other, temporal, cues aid detection (Strickland & Viemeister, 1997). If the carrier is broadband then its own, intrinsic, fluctuations influence the detection threshold of the carrier (Dau et al., 1997).

The measurement of spectro-temporal modulation transfer functions (STMTFs) has to use a broadband carrier in order to carry the spectral component of the modulation. Historically, this carrier has been either a noise, or closely-spaced sinusoids. A (peripheral) auditory filter captures the extrinsic modulation, along with the interaction of the broadband components which produce intrinsic modulations. However, Stone et al. (2012) showed that these intrinsic modulations strongly impaired (speech) discrimination. They reduced the influence of the intrinsic modulations by use of a dichotic presentation of spectrally interleaved modulation carriers. The best results were achieved with 1-ERBn spaced sinusoids as carriers. A central recombination process appeared to re-generate the extrinsic modulations, but less so for the intrinsic modulations, resulting in good discrimination of the target signal.

This paper compares measurements of STMTFs using the 1-ERBn-spaced sinusoidal-carrier dichotic-presentation technique against a diotic-presentation technique using 0.33-ERBn-spaced sinusoidal carriers, similar to that of Chi et al. (1999). Temporal modulation rates were 0 Hz, and well as octave-spaced from 1 to 16 Hz. The dichotic presentation technique limited the testing of spectral modulation rates up to 0.28 ripples/ERBn, equivalent to around 1.5 cycles/octave at mid-audio frequencies, although higher rates are capable of being discriminated by the human auditory system. The results suggest a need to re-visit the understanding of what the historic measures of STMTFs show.

References

- Chi T., Gao Y., Guyton M.C., Ru P., Shamma S. (1999). Spectro-temporal modulation transfer functions and speech intelligibility. *J. Acoust. Soc. Am.* 106, 2719-2731.
- Dau T., Verhey J., Kohlrausch A. (1997) Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers. *J. Acoust. Soc. Am.* 106, 2752-2760.
- Stone M.A., Füllgrabe C., Moore B.C.J. (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *J. Acoust. Soc. Am.* 132, 317-326.
- Strickland E.A. Neal F. Viemeister N.F. (1997) The effects of frequency region and bandwidth on the temporal modulation transfer function. *J. Acoust. Soc. Am.* 102, 1799-1810.