Binaural unmasking of speech with small interaural time differences and its relationship to binaural temporal fine structure processing

Gusztáv Lőcsei[1], *Sébastien Santurette*[1], *Torsten Dau*[1], *Ewen N. MacDonald*[1]

[1] Hearing Systems, Department of Electrical Engineering, Technical University of Denmark

guloc@elektro.dtu.dk

Index Terms: speech recognition, temporal fine structure, spatial release from masking, binaural unmasking

1. Introduction

In everyday communication scenarios sound sources are often spatially distributed in the auditory scene. When interferers get spatially separated from a target stream, speech recognition performance improves. This phenomenon is called spatial release from masking (SRM). It is well established that hearing impaired listeners (HI) have difficulties when following a single talker in the presence of other, interfering talkers. These difficulties manifest themselves both in increased SRTs and in a decreased SRM [1,2].

SRM in the horizontal plane is partly facilitated by binaural unmasking [1]: the hearing system is able to aid speech recognition by detecting and exploiting disparities in the interaural timing differences (ITDs) of the arriving sound waves between the ears, associated with each sound source separately. Binaural unmasking is thought to be mainly facilitated by differences in ITDs at low frequencies [3]. It has been shown that both hearing loss and aging can be independently associated with reduced binaural temporal fine structure (TFS) sensitivity. This can e.g. negatively affect the ability to detect interaural time differences (or interchangeably interaural phase differences, IPDs) in pure-tone carriers at low frequencies [4,5,6] or the upper frequency limit at which listeners are sensitive to such differences [7,8]. Despite these findings, several studies have suggested that SRM due to binaural unmasking is in the same range for most of the HI listeners with a symmetrical impairment as for the normal hearing listeners (NH) [1,9]. Still, assuming that SRM due to binaural unmasking is indeed mainly facilitated by low-frequency ITDs, it is reasonable to expect a reduction in SRM for listeners who have a reduced sensitivity to detect ITDs. Furthermore, the putative effect of reduced sensitivity to ITDs is likely to be pronounced in listening scenarios where SRM is triggered by ITDs in the range or below the individual ITD discrimination thresholds.

In the current study binaural TFS coding abilities at low frequencies and the amount of SRM due to binaural unmasking are being measured in 10 young NH listeners and 10 middle aged or elderly listeners with normal hearing below but with a symmetrical hearing loss above 1.5 kHz. It is hypothesized that a reduced sensitivity to detect ITDs has a more pronounced effect on SRM due to binaural unmasking once the spatial separation between target and maskers is small, i.e. triggered by ITDs in the range of where they are just detectable for the individuals.

2. Methods

2.1. Binaural TFS coding

To assess the robustness of binaural TFS coding, interaural phase discrimination thresholds (IPDTs) for pure tones are measured for a range of frequencies below 1.5 kHz. A weighted up-down, 3I-3AFC method is used to track the 70.7% point on the psychometric function. First, the upper frequency limit at which subjects are able to detect a 180° phase shift between the ears is assessed ($IPDT_{FR}$). After this, IPDT thresholds are evaluated starting at 250 Hz, and then at increasing frequencies in 250 Hz steps, until the $IPDT_{FR}$ is reached. Stimuli are presented in a similar way to [4].

2.2. Speech intelligibility and SRM

The stimuli in the speech intelligibility experiments are delivered via headphones. SRTs are measured using the Danish DAT corpus as target sentence material [10]. The target sentences are embedded into a stream of sentences uttered by the two other talkers from the same corpus or into speech shaped noise (D2 or SSN conditions, respectively). The target is always presented diotically, while maskers are presented either diotically (reference condition) or dichotically, steering their lateral position to the side using ITDs only. When large ITDs (680 µs) are used in the dichotic masker conditions, SRMs for both noise types are calculated as the difference between SRTs in the reference and spatial conditions. In the D2 masker condition, the effect of spatial separation due to small ITDs is assessed by measuring the amount of ITD needed to yield a fixed 3 dB SRM. As the amount of SRM with SSN is typically just around 3 dB, the effect of spatial separation due to small ITDs in the SSN condition is assessed by measuring SRMs for fixed small ITDs (270 µs).

3. Preliminary results

According to preliminary data collected on 10 NH and 6 HI listeners, it appears that processing of ITDs in this HI group is only affected at frequencies above about 750 Hz. In the speech intelligibility experiments HI listeners show elevated SRTs but similar SRM values as the NH listeners. The data of the HI group suggest that the $IPDT_{FR}$ values might be predictive of SRM in stationary background noise only, but not in fluctuating two-talker babble. One possible explanation could be that in the latter case susceptibility to informational masking gets confounded with abilities related to binaural processing.

4. References

[1] A. Bronkhorst, "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acustica*, vol. 86, pp. 117–128, 2000.

[2] J. Peissig and B. Kollmeier. "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners", *J. Acoust. Soc. Am.*, vol. 101, no. 3, pp. 1660–1670, 1997.

[3] A. W. Bronkhorst and R. Plomp, "The effect of head-induced interaural time and level differences on speech intelligibility in noise", *J. Acoust. Soc. Am.*, vol. 83, no. 4, pp. 1508-1516, 1988.

[4] K. Hopkins and B. C. J. Moore, "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise", *J. Acoust Soc. Am.*, vol. 130, no. 1, pp. 334–349, 2011.

[5] T. Neher, T. Lunner, K. Hopkins and B. C. J. Moore, "Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners", *J. Acoust. Soc. Am.*, vol. 131, no. 4, pp. 2561–2564, 2012.

[6] A. King, K. Hopkins and C. J. Plack, "The effects of age and hearing loss on interaural phase difference discrimination", *J. Acoust. Soc. Am.*, vol. 135, no. 1, pp. 342–351, 2014.

[7] B. Ross, T. Fujioka, K. L. Tremblay and T. W. Picton, "Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase", *J. Neurosci.*, vol. 27, no. 42, pp. 11172–11178, 2007.

[8] T. Neher, S. Laugesen, N. S. Jensen and L. Kragelund, "Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities?", *J. Acoust. Soc. Am.*, vol. 130, no. 3, pp. 1542–1558, 2011.

[9] A. W. Bronkhorst and R. Plomp, "Binaural speech intelligibility in noise for hearing-impaired listeners", *J. Acoust. Soc. Am.*, vol. 86, no. 4, pp. 1374–1383, 1989.

[10] J. Nielsen, T. Dau and T. Neher, "A Danish open-set speech corpus for competing-speech studies," *J. Acoust. Soc. Am.*, vol. 135, pp. 407–420, 2014.