# The contribution of intrinsic amplitude modulation to the precedence effect at high frequencies

Bernhard U. Seeber

*MRC Institute of Hearing Research, University Park, Nottingham, NG7 2RD,*
*seeber@ihr.mrc.ac.uk*

## Summary and Introduction

The precedence effect (PE) demonstrates our ability to locate sounds correctly at the source despite the presence of interfering sound reflections. It was shown to function with long duration broadband noises even when lead and lag had simultaneous onsets, i.e. with information restricted to the ongoing sound part [1].

The present study investigated if the PE can be elicited for non-transient high frequency sounds based on intrinsic envelope modulation. This is questionable because a) the main cue for localization, interaural time differences extracted from the temporal fine structure, may not be available, b) inherent amplitude modulations provide only a weak cue for sound segregation via pitch mechanisms.

In a localization dominance task participants indicated the perceived location of lead-lag stimuli presented from loudspeakers in anechoic space. Stimuli were harmonic complex tones (HCTs) and Gaussian noise bandlimited to 2.5-5.5 kHz. Results demonstrated that the PE can exist for high-frequency stimuli without explicit onsets. The depth of intrinsic modulation generally affected echo thresholds (ETs), but the difference between HCTs with Schroeder positive and negative phase was negligible. Localization dominance was weak for Gaussian noise as expressed in larger localization variance and short ETs of 2 ms, shorter than those for the HCTs (4 ms).

## Methods

### Experimental Procedures and Subjects

Localization dominance was studied for sounds presented in anechoic space from loudspeakers at ±30° [2]. In half of the trials, a leading sound was presented from +30°, while its delayed copy, the lag, was presented from -30°. In the other half of the trials the directions for lead and lag were swapped. After the sound was played a movable light spot appeared in front of the listener. Subjects used it to indicate the perceived azimuthal location of the lead-lag stimulus. If subjects perceived separate images for lead and lag they were instructed to indicate the leftmost image. Randomizing the side of lead and lag thus meant that subjects indicated the lead in one half of the trials.

Localization results were collected for the following lead-lag delays: 0, 0.5, 1, 2, 3, 4, 5, 7, 10, and 13 ms. Sound level was roved in 3 steps [-3, 0, +3 dB SPL] from a base level of 60 dB SPL. Two trials were collected for each roving level (3), sound (4), delay (10) and lead direction (2), giving a total of 480 trials per condition. Trials were administered in random order and the presentation was divided into 4 runs of about 8 min each. Brief training was given prior to data collection.

Results of six subjects with normal hearing (<20 dB HL), two male and four female (age 22-26 years, mean 23.3 years), are presented. Subjects were paid and the study protocol was approved by the ethics committee of the Psychology Department at the University of Nottingham.

### Stimuli

Four stimuli with differing intrinsic envelope modulation were used: (1) A harmonic complex tone (HCT) with zero starting phase of each harmonic (Ø), (2) a HCT with Schröder positive phase (S+) [3], (3) a HCT with Schröder negative phase (S-), (4) a Gaussian noise (N). All HCTs had a fundamental frequency of 100 Hz. Stimuli were 300 ms long. Lead and lag were simultaneously gated with a slow, 50 ms rise time to minimize onset dominance, i.e. there was no onset or offset delay between lead and lag.

In the main test condition stimuli were bandlimited to 2.5-5.5 kHz. The lack of phase locking at those frequencies prevented the extraction of interaural time differences (ITDs) from the temporal fine structure (TFS) of the sound and thus forced the auditory system to rely on ITDs and interaural level differences (ILDs) carried in the sound's envelope. The stimuli received no further amplitude modulation so that only sound-intrinsic modulation was available. Results from the high-frequency condition were contrasted to those from a broadband condition in which the same stimuli were bandlimited to 300-5500 Hz.

## Results and Discussion

Results of the localization dominance test with the lead at +30° are reported in Figure 1, top. The PE was active for all stimuli as demonstrated by the shift of median responses towards the lead. This indicates the ability to use intrinsic modulation from the envelope. However, localization dominance differed strongly between stimuli. The lead was most dominant for the zero-phase HCT (Ø); responses were closest to the lead direction, variance was small, and echo thresholds (ETs), the delay at which responses crossed to the lag side, were largest (4 ms). In contrast, median responses for the noise were only somewhat shifted to the lead direction, lower quartiles never crossed over to the lead side, and the ET was only 2 ms – too short to be of much practical use. Figure 1, bottom, highlights the differences by showing the relative proportion of responses falling on the lead side, which for the zero-phase HCT were almost 100% at 0.5-2 ms delay, thus indicating a strong precedence effect. With the noise, instead, there was no delay at which all responses were on the lead side and above 2 ms delay lead bias became non-significant, indicating very short echo thresholds.

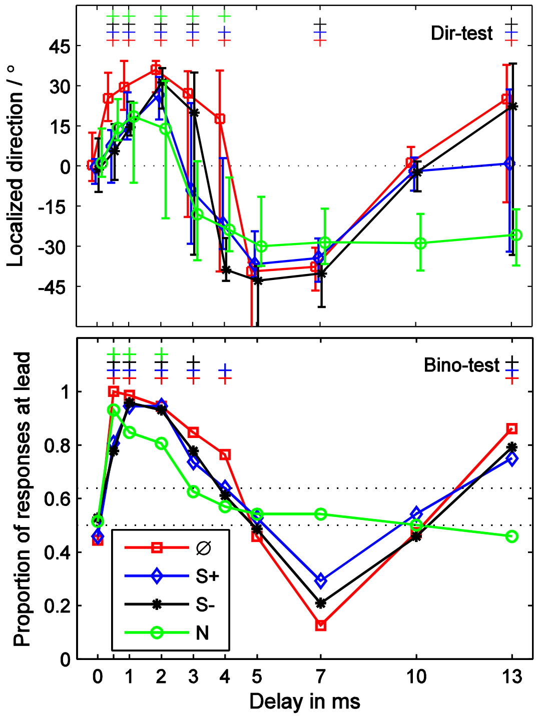

**Figure 1, *Top:* Localization results for the lead at +30° as a function of delay time. Medians of pooled responses from all listeners are connected by lines. Errorbars show interquartiles. Symbols/colours represent different stimuli (see methods). Note that the instruction was to indicate the left image when two were heard; hence responses shift to the left above the echo threshold. Symbols on the top indicate if lead direction significantly influenced the response distribution (Wilcoxon sign rank test, p<0.01).**

***Bottom:* Proportion of localization responses falling on the side of the lead. A bias to the lead was indicated if at least 46 of the 72 trials per delay (upper dotted line) were on the lead side (binomial test, p<0.01).**

The PE was further assessed for Schröder positive and negative phase HCTs. Their broadband envelope is similar and flat, while their narrow-band envelope is highly modulated (modulation depth >30 dB), but again similar. However, Schröder-positive phase HCTs are thought to evoke a peaky envelope on the basilar membrane while Schröder-negative phase HCTs should cause shallower fluctuations; thus both give rise to differing modulation amounts on the basilar membrane [3]. Results differ only slightly between them, suggesting that their envelope differences played only a minor role for the PE, and that binaural cues were similarly useful in both. Nevertheless, statistical testing indicated slightly larger ETs for the S+ HCT, in line with an increased ITD contribution from its peakier envelope.

Intrinsic modulation in high-frequency noise was not sufficient to evoke a strong, long-lasting PE. To exclude subject effects, localization dominance was further studied for broadband stimuli. Results for all broadband HCTs were very much alike, indicating that phase relationships between harmonics become irrelevant for the PE with broadband HCTs (Figure 2). This might be because redundant information for the PE can be extracted from a large number of auditory filters, leading to ceiling performance. Broadband stimuli also permit the extraction of ITDs from the TFS. ITD thresholds are generally lower at low compared to high frequencies, because the high rate of the TFS at low frequencies allows more information to be extracted. Specifically, ITD thresholds are very low even for signals with narrow bandwidth, e.g. tones. At high frequencies, instead, good ITD discrimination is possible only with wide-band stimuli. However, the usefulness of this spectral integration process may be limited in the presence of other sounds, i.e. the lag. While distinct onsets are potentially useful across auditory filters to extract ITDs from the lead, this may not be equally possible with noise. The unmodulated noise exhibits only intrinsic modulations which are only slightly correlated across auditory filters (from cross-talk). Thus, there were no common onsets across filters from which ITDs could have been extracted. Additionally, modulation depth was very small (the standard deviation of the log-envelope was around 3 dB). This will have led to poor ITD discrimination from each single auditory filter [4]. In contrast, all HCTs had a large within-filter modulation depth of at least 30 dB, sufficient for extracting ITDs [4]. While this may explain the short ET for the high-frequency noise, ETs for HCTs were larger. Envelopes of the HCTs share the same periodicity across auditory filters, thus making the evaluation of common onsets theoretically possible. The results in Figure 2 demonstrate that the PE is strong for broadband noise. For the above reasons this might entirely be based on TFS. Then, ETs become as large as 10 ms, rendering localization robust against interference from sound reflections.

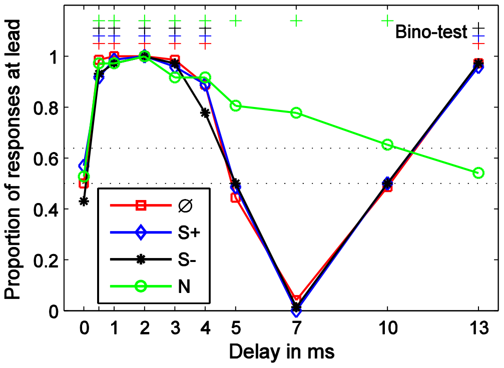

**Figure 2: Same as Figure 1, bottom, but for broadband stimuli (300 Hz - 5.5 kHz).**

## Acknowledgements

This work was supported by the Intramural Programme of the MRC.

## References

[1] Dizon, R.M., Colburn, H.S., 2006. J. Acoust. Soc. Am. 119, 2947-2964.

[2] Seeber, B.U., S. Kerber, and E.R. Hafter: Hear Res, 2010. **260**(1-2): 1-10.

[3] Moore, B., 2003. An Introduction to the Psychology of Hearing. Academic Press, San Diego, pg. 94.

[4] Monaghan, J.J.M, K. Krumbholz, and B.U. Seeber, submitted.