Audio Engineering Society

Convention Paper

Presented at the 140th Convention
2016 June 4–7, Paris, France

Exploiting envelope fluctuations to enhance binaural perception

G. Christopher Stecker[1]

[1] *Department of Hearing and Speech Sciences, Vanderbilt University School of Medicine, Nashville TN USA*

Correspondence should be addressed to G. Christopher Stecker (`cstecker@spatialhearing.org`)

ABSTRACT

A review of recent and classic studies of binaural perception leads to the conclusion that envelope fluctuations, such as sound onsets, play a critical role in the sampling of spatial information from auditory stimuli. Specifically, listeners' perception of sound location corresponds with the binaural cues (interaural time and level differences) that coincide with brief increases in sound amplitude, and disregards binaural cues occurring at other times. This discrete, envelope-triggered sampling of binaural information can be exploited to enhance spatial perception of synthesized sound mixtures, or to facilitate the localization of mixture components.

1 Introduction

Spatial hearing (for example, sound localization, or segregation of multiple sources) depends on the brain's processing of multiple time-varying features ("spatial cues") of sounds. Chief among these are the binaural cues, interaural time and level differences (ITD and ILD) in the sound arriving at the two ears. Although decades of research have described human listeners' sensitivity to these cues and the initial brain mechanisms that compute them, a full understanding of how the brain represents spatial information remains elusive. Current models of spatial hearing also fail to capture human listeners' ability to accurately perceive complex spatial mixtures, particularly in acoustically complex (i.e. reverberant) scenes.

The temporal envelopes of sounds are known to profoundly affect several aspects of auditory perception, including timbre, duration, and perceptual grouping. Fluctuations of the envelope are also necessary for the conveyance of ITD in high-frequency, amplitude-modulated (AM) sound. When such fluctuations of the ongoing envelope become too weak, or the AM rate too fast, listeners lose access to envelope-ITD except at the overall sound onset, which then dominates spatial perception [1]. A surprising result of recent research is that the same thing happens to binaural cues that should not require the occurrence of envelope fluctuations, such as ILD [2, 3], or the ITD carried in the temporal fine-structure of low-frequency sounds [4, 5]. These results demonstrate that envelope fluctuations such as sound onsets are necessary for processing all aspects of binaural information in periodic sounds.

In contrast to the strong onset dominance observed for periodic tones and AM sounds, studies of binaural sensitivity for noise [6, 7] and for aperiodic AM sounds [8] suggest nearly the opposite result: greater sensitivity to binaural cues in the ongoing waveform than at sound onset. Those results suggest that binaural sensitivity is gated by envelope fluctuations occurring within, rather than across, auditory critical bands. Whether those fluctuations arise intrinsically within the auditory filter (in the case of noise) or due to the overall temporal envelope (for periodic sounds), the mechanism—and its consequence—is the same.

In this paper, we review the literature on this topic and develop a theoretical framework for the critical role of envelope fluctuations in binaural processing. We end with a discussion of future applications in the domains of machine listening and spatial audio synthesis.

2 Methods

Three main approaches have been used to measure the relative influence of binaural information available during sound onsets versus later portions of a sound. We refer to these as the *Temporal Integration* approach, the *Dynamic Cues* approach, and the direct *Temporal Weighting* approach.

2.1 Temporal Integration

In the temporal integration approach, binaural sensitivity (e.g. binaural discrimination threshold) is measured as a function of sound duration. The idea is that longer sounds provide proportionally more binaural information, from which an optimal listener would benefit. Assuming stationarity of binaural information (i.e., the cue to be discriminated is constant over the duration) and temporally independent internal noise, optimal thresholds decreas in proportion to the square root of duration Hafter and Dye [1]. A plot of log threshold versus log duration thus appears linear, with a slope of -0.5. A shallower slope would indicate suboptimal integration of binaural information over time, for example if judgments are dominated by sound onset and make minimal use of post-onset information. This approach has been used (and suboptimal slopes reported) for discrimination of ITD in bands of low-frequency noise [9], trains of narrowband high-frequency impulses [1], and pure tones ranging 250–1000 Hz [10].

2.2 Dynamic Cues

An alternative approach is to measure the discrimination of binaural cues applied strategically to coincide with different temporal portions of a sound. For example, Stecker and Brown [11] and Stecker and Bibee [5] presented listeners with brief target sounds whose binaural cues were confined mainly to sound onset or offset in different conditions. Specifically, ITD or ILD progressed linearly from 0 at sound onset to favor the right ear at sound offset (condition 0R) or from a maximal value at onset to 0 at sound offset (condition R0). In a third condition (RR), ITD or ILD was constant over the sound duration. Listeners' detection of ITD was similar in conditions RR and R0, but significantly impaired in condition 0R, suggesting that listeners relied mainly on sound onset cues to perform the task.

2.3 Temporal Weighting Functions

A third approach to estimating the temporal weighting of binaural information is to derive temporal weighting functions directly. In this approach, listeners judge the spatial location of sounds whose binaural cues vary randomly over time (e.g. click trains in which each click is presented from a slightly different location). Multiple regression is used to relate the judgments, across a few hundred trials, to the cue values presented at each temporal position (e.g. click number). The regression weights obtained from this analysis provide an estimate of the relative influence of each temporal position (click) on a listener's judgment. Plotted as a function of time, the weights define a temporal weighting function. Onset dominance is evidenced by greater weight on the first position than on later positions. This approach has been used for discrimination [12, 13] as well as lateralization [3] of stimuli presented over headphones, as well as localization of sounds presented in the free field [14].

3 Results

3.1 Rate-dependent onset dominance for ITD in periodically modulated high-frequency tones

Using the temporal-integration approach, Hafter and Dye [1] measured ITD thresholds for trains of narrowband-filtered clicks (approximately 1/2 octave centered at 4 kHz). When clicks repeated at an interval longer than 10 ms, thresholds improved optimally with duration. Shorter interclick intervals (ICI), however, produced shallower threshold-duration functions. For ICI of 1–2 ms, thresholds for trains of 16 or 32 clicks were hardly better than for single clicks. Hafter and Dye [1] concluded that listeners localize high-rate (short ICI) stimuli on the basis of the overall onset. At slower rates, the ongoing information contributes more. That finding has since been replicated using essentially identical stimuli via the dynamic-cues [11] and temporal-weighting approaches [12, 13, 3, 15]. Temporal weighting functions, in particular, showed that onset dominance enhances the weight of only the very first click—not of early clicks in general. That finding strongly suggests that the onset *per se*—perhaps the initial rise of the amplitude envelope—drives this effect.

Because the relevant cue for high-frequency modulated sounds is the ITD carried by the sound envelope, it stands to reason that the effective depth of modulation should constrain listeners' sensitivity to ITD. The studies mentioned above all presented click trains with 100% modulation. However, the action of the auditory periphery temporally extends the response to each click, as the basilar membrane of the inner ear "rings" (produces its impulse response) with each click. At slow rates, there is sufficient time for damping of the response between clicks, but as the rate gets faster, successive responses begin to overlap in time, reducing the effective modulation depth. This effect is particularly strong at carrier frequencies below 1-2 kHz [15], but additional rate-limiting processes in the central auditory system produce similar effects even at 4 kHz and beyond [16]. In either case, listeners lose sensitivity to the ongoing envelope ITD as the modulation rate increases, and are left with the overall onset as the only feature capable of supporting envelope-ITD sensitivity.

In other words, rate-dependent onset dominance is to be expected for ITD in periodically modulated high-frequency tones, simply because the physical cue (ongoing envelope ITD) is reduced at high rates. Such effects are expected only to affect envelope-ITD processing, however, and should not be seen for other binaural cues such as ILD or fine-structure ITD at low frequencies. These cues are considered in the following sections.

3.2 Rate-dependent onset dominance for ILD in periodically modulated high-frequency tones

Hafter et al. [2] investigated whether rate-dependent onset dominance was specific to envelope-ITD processing by replicating their previous experiment [1] with ILD as the cue to be discriminated. Other aspects of the experimental stimuli and procedure were identical. The results revealed a similar degree of onset dominance for both ITD and ILD, and a similar dependence on the click rate, leading the authors to argue against a purely peripheral account of onset dominance. Rather, it appears that onset dominance for binaural discrimination reflects a characteristic of binaural neurons in the central auditory system. Subsequent studies have confirmed the appearance of rate-dependent onset dominance for ILD using the dynamic-cues [11, 17] and temporal-weighting [13, 3, 15] approaches. These have, however, revealed a key difference between the temporal weighting of ITD and ILD cues, which is that ongoing ILD cues, and in particular late-arriving ILD cues (i.e., those occurring near sound offset) have a significant effect on localization. In many cases, this manifests as a "U-shaped" temporal weighting function for ILD [18, 17], with potent ILD at sound onset and offset but not middle portions of the sound. For ITD, in contrast, onset dominance appears more absolute.

In contrast to the case for envelope-ITD processing at high frequencies, there is little reason to expect ILD sensitivity to depend on envelope features, and thus no reason to expect rate-dependent onset dominance due to acoustical or peripheral effects. The similarity of temporal weighting for ITD and ILD cues at high frequencies suggests a common, central mechanism.

3.3 Enhanced sensitivity to low-frequency fine-structure ITD during positive envelope fluctuations

The "textbook" account of binaural sensitivity to ITD is that of a low-frequency (e.g., 500 Hz) pure tone, delayed to one ear by a small amount. Each cycle of the waveform elicits a response in the auditory nerve, and the relative timing of responses in the two ears is extracted by coincidence-detection neurons in the brainstem that effectively cross-correlate the left and right signals. Because the system is sensitive to the cycle-by-cycle phase difference of the tone (the "temporal fine structure" of the signal, as opposed to the envelope), one expects the system to integrate information across cycles in a more-or-less optimal fashion, as would occur for cross-correlation over a reasonably long term (say, a few hundred ms). In particular, there is little reason to expect onsets or other envelope fluctuations to have any substantive effect on low-frequency ITD sensitivity. Recent studies, however, reveal a temporal dynamics for low-frequency fine-structure ITD that does not markedly differ from that described above for high-frequency envelope ITD and ILD.

Houtgast and Plomp [9] used the temporal-integration approach to study ITD discrimination in narrow bands of noise centered at 500 Hz, and found shallow threshold-duration slopes almost identical to those reported by Hafter and Dye [1] for high-rate high-frequency click trains. The authors suggested that listeners made use mainly of ITD cues carried by the sound onset. That result has been replicated in our lab, using 40–640 ms pure tones of 500 Hz [5] whose envelopes were gated on and off simultaneously in the two ears so that the envelope itself did not provide an ITD cue. That study, and a followup [10], also utilized the dynamic-cues approach to compare fine-structure ITD discrimination in pure tones where the cue was available at sound onset but not offset (condition R0 as in [11]) or vice versa (condition 0R). Consistent with the results for high-rate high-frequency sounds, ITD thresholds in both studies were significantly better when the ITD cue was present at sound onset. That is, we observed strong onset dominance for pure tones, even when the onset itself was diotic. Thus, it appears that processing of the fine-structure ITD is enhanced during the early, onset, portion of the sound.

Dietz et al. [4] conducted a lateralization experiment using a periodic variant of the dynamic-cues approach. In that study, pure tones of 484 and 516 Hz were presented to the two ears, resulting in a “binaural beat” whose ITD sweeps repeatedly from -1 to +1 ms, 32 times per second. To this sound, the authors applied sinusoidal AM with a modulation frequency of 32 Hz (matching the ITD sweep rate). The relative phase of the AM and the binaural beat were varied from trial to trial, and listeners’ spatial judgments consistently reflected the ITD which coincided with the rising portion of the sinusoidal envelope rather than the peak or the falling portion. That is, the envelope fluctuations imposed by AM enhanced sensitivity to ITD only within a brief window triggered by the onset of each modulation cycle. As was the case for single tone bursts [5] and for high-rate click trains [11], ITD perception was thus mediated by the early, rising, portion of each modulation cycle. It is important to note that the relatively slow 32 Hz modulation rate is well within the range which Hafter and Dye [1] showed to exhibit optimal threshold-duration slopes (i.e. equal sensitivity to ITD in each modulation period). Thus, although the results reveal strong “onset” dominance *within* each modulation cycle, optimal integration *across* modulation cycles likely occurred, contributing to listener’s reliable perception of a single stable location.

3.4 Little to no onset dominance for “noise”

With few exceptions, the stimuli described in the previous sections were periodic: pure tones, regular AM, or click trains with constant ICI. Aperiodic stimuli, in contrast, have been shown in numerous studies to support greater sensitivity to *ongoing* binaural information than to *onset* cues. Tobias and Schubert [6], for example, asked listeners to lateralize broadband noise in which the envelope onset ITD and the ongoing fine-structure ITD led to opposite ears. For sounds longer than $\sim$ 100 ms, listeners’ judgements consistently reflected the ongoing ITD and were unaffected by the onset ITD. Similarly, Freyman et al. [7] presented trains of 1-ms noise bursts, repeating at 2 ms, in which the initial (onset) burst carried an ITD opposite to the remaining (ongoing) bursts. When the waveforms of the bursts were identical (so that the sound was in fact periodic), listeners’ judgments were sometimes consistent with the onset, and sometimes with the ongoing ITD. But when new samples of noise were presented in each burst (so that the sound was truly aperiodic), listeners consistently lateralized the sound in the direction of the ongoing ITD. Consistent with that result, temporal weighting functions reveal strong onset dominance for trains of repeating noise bursts, but much weaker onset dominance for trains of non-repeating noise bursts [19]. Importantly, the broadband envelopes of sounds presented in these studies were themselves periodic (at a rate of 500 Hz); however, within any narrow frequency band (i.e. at any single place along the basilar membrane of the inner ear), only the repeated bursts produce regular activation. Non-repeating bursts produce temporally irregular narrowband envelopes that occasionally feature large fluctuations similar to those introduced by sound onsets or slow AM. We argue that such fluctuations explain the difference in temporal weighting of binaural information for periodic sounds and for “noise:” sensitivity to ongoing cues is driven by the aperiodic nature of the narrowband envelope, not by the broadband spectrum of the noise itself.

Further consistent with that view is the effect of introducing temporal irregularity into otherwise-periodic trains of narrowband clicks. Binaurally synchronized temporal “jitter” of the ICI results in improved ITD thresholds [20] and reduced onset dominance [13]. That phenomenon has been explored, in the audiology clinic, as means to enhance ITD sensitivity for users of cochlear implants [21].

Finally, it may be considered whether the presence of masking noise could similarly influence the temporal weighting of binaural information. The study of Houtgast and Plomp [9] addressed this question via the temporal integration method, and found nearly optimal threshold-duration slopes for targets presented at a signal-to-noise ratio of just 5 dB. More recent and extensive testing in our laboratory has replicated that result [22] and shown equivalent discrimination, using the dynamic-cues approach, of ITD present early or late in the sound duration, consistent with good sensitivity to ongoing ITD (i.e. no onset dominance) for noise-band targets presented in a noisy background.

4 Discussion

The evidence cited above suggests a critical role for envelope fluctuations in the sampling of binaural information, regardless of the type of cue (ITD or ILD) or the frequency content of the sounds. The view that emerges from a complete consideration of the literature on this topic is that binaural cues are not processed except when “triggered” by envelope fluctuations in the form of sound onsets, slow periodic modulations, and temporally irregular fluctuations of the envelope within auditory bands. The neural mechanism that subserves this behavior is not clear. While some authors have argued for an active triggering and “readout” process [23], similar effects could be generated by strong response adaptation in the auditory brainstem [24], particularly given the precise nature with which auditory neurons encode the timing of amplitude increase at sound onset [25].

Whatever the mechanism, the data on rate-dependency suggest a refractory period of 2–10 ms following each triggering event, during which further triggering is inhibited. Retriggering can be achieved by periodic modulation below this rate, by intrinsic slow fluctuations in stochastic signals, or by introducing spectral or temporal changes within a more rapid stimulus sequence [26].

The results of these studies have clear implications for how the brain extracts spatial cues from naturally fluctuating sounds such as human speech, and how that process is altered by echoes, reverberation, and competing sources in real auditory scenes. In fact, they dramatically change our view of how the brain tracks objects in a spatial scene: rather than continuous processing of spatial information, it appears that sound envelopes form the basis for discrete and temporally sparse sampling of sound-source locations. For example, the components of a spatial mixture may be accurately localized on the basis of binaural cues that occur synchronously with infrequent fluctuations in the independent envelopes of each component. These may form the basis of independent quanta of binaural information, which may be assigned to other features of the sound (pitch, identity,...) and/or grouped across time and frequency [cf. 27]. Practical consequences of this concept may find application in domains such as computational scene analysis and the processing of reverberated signals. Figure 1 of Dietz et al. [4], for example, illustrates the advantage of weighting ITD by the envelope slope when binaural signals include both direct and reflected sound. Nelson and Takahashi [28, see also Mao and Carney [29]], similarly showed how auditory neurons can use envelope-slope-weighted firing to reliably code binaural information when echoes are present.

Envelope fluctuations can also be exploited in the opposite direction; that is, in applications to spatial audio synthesis. The temporal envelopes of source material can be modified to control the salience of binaural information, and thus the perceived location of source objects in a synthesized scene. Alternately, binaural differences may be applied strategically to coincide with naturally occurring fluctuations. Future work in this area may lead to applications in dynamic binaural panning and in perceptually driven data compression of spatial audio.

5 Summary

1. Binaural perception is dominated by cues that coincide with positive envelope fluctuations, such as sound onsets.

2. Onset dominance is observed consistently across binaural cue type and frequency range, affecting amplitude-modulated sounds faster than ~ 100 Hz, as well as unmodulated sounds.
3. Slowly modulated sounds (< 100 Hz AM rate), and sounds with random within-band envelope fluctuations, do not experience onset dominance.
4. Published results thus suggest that binaural information is sampled discretely, at moments of infrequent positive-going envelope fluctuations. Potential benefits of this sparse sampling could be clearer segregation of competing sounds and robust rejection of echoes and reverberation.
5. Applications to machine listening include envelope-slope-weighting of binaural information for spatial processing in reverberation and spatiotemporal grouping.
6. Applications to spatial sound synthesis utilize psychoacoustic models of dynamic binaural sensitivity to achieve robust spatial perception and data compression of spatial audio.

6 Acknowledgments

The author thanks Anna Diedesch for her work on earlier versions of this presentation. Portions of this work were presented previously to the Acoustical Society of America (2014, Indianapolis IN) and Association for Research in Otolaryngology (2015, Baltimore MD). Supported by NIH/NIDCD R01DC011548.

References

[1] Hafter, E. R. and Dye, R. H. J., “Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number,” *J. Acoust. Soc. Am.*, 73(2), pp. 644–651, 1983.

[2] Hafter, E. R., Dye, R. H. J., and Wenzel, E. M., “Detection of interaural differences of intensity in trains of high-frequency clicks as a function of interclick interval and number,” *J. Acoust. Soc. Am.*, 73, pp. 1708–1713, 1983.

[3] Stecker, G. C., Ostreicher, J. D., and Brown, A. D., “Temporal weighting functions for interaural time and level differences. III. Temporal weighting for lateral position judgments,” *J. Acoust. Soc. Am.*, 134(2), pp. 1242–52, 2013, doi: 10.1121/1.4812857.

[4] Dietz, M., Marquardt, T., Salminen, N. H., and McAlpine, D., “Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds,” *Proc Natl Acad Sci U S A*, 110(37), pp. 15151–6, 2013, doi:10.1073/pnas.1309712110.

[5] Stecker, G. C. and Bibee, J. M., “Nonuniform temporal weighting of interaural time differences in 500 Hz tones,” *J. Acoust. Soc. Am.*, 135(6), pp. 3541–3547, 2014.

[6] Tobias, J. V. and Schubert, E. R., “Effective onset duration of auditory stimuli,” *J. Acoust. Soc. Am.*, 31, pp. 1595–1605, 1959.

[7] Freyman, R. L., Zurek, P. M., Balakrishnan, U., and Chiang, Y. C., “Onset dominance in lateralization,” *J. Acoust. Soc. Am.*, 101(3), pp. 1649–1659, 1997.

[8] Brown, A. D. and Stecker, G. C., “Temporal weighting functions for interaural time and level differences. II. The effect of binaurally synchronous temporal jitter,” *J. Acoust. Soc. Am.*, 129(1), pp. 293–300, 2011, doi:10.1121/1.3514422.

[9] Houtgast, T. and Plomp, R., “Lateralization threshold of a signal in noise,” *J. Acoust. Soc. Am.*, 44(3), pp. 807–812, 1968, ISSN 0001-4966 (Print).

[10] Diedesch, A. C. and Stecker, G. C., “Temporal weighting of binaural information at low frequencies: Discrimination of dynamic interaural time and level differences,” *J Acoust Soc Am*, 138(1), pp. 125–33, 2015, doi:10.1121/1.4922327.

[11] Stecker, G. C. and Brown, A. D., “Temporal weighting of binaural cues revealed by detection of dynamic interaural differences in high-rate Gabor click trains,” *J. Acoust. Soc. Am.*, 127(5), pp. 3092–3103, 2010.

[12] Saberi, K., “Observer weighting of interaural delays in filtered impulses,” *Percept. Psychophys.*, 58(7), pp. 1037–1046, 1996.

[13] Brown, A. D. and Stecker, G. C., “Temporal weighting of interaural time and level differences in high-rate click trains,” *J. Acoust. Soc. Am.*, 128(1), pp. 332–341, 2010.

[14] Stecker, G. C. and Hafter, E. R., “Temporal weighting in sound localization,” *J. Acoust. Soc. Am.*, 112(3), pp. 1046–1057, 2002.

[15] Stecker, G. C., “Temporal weighting functions for interaural time and level differences. IV. Effects of carrier frequency,” *J Acoust Soc Am*, 136(6), p. 3221, 2014, doi:10.1121/1.4900827.

[16] Bernstein, L. R. and Trahiotis, C., “Enhancing sensitivity to interaural delays at high frequencies using “transposed stimuli”,” *J. Acoust. Soc. Am.*, 112, pp. 1026–1036, 2002.

[17] Stecker, G. C. and Brown, A. D., “Onset- and offset-specific effects in interaural level difference discrimination,” *J. Acoust. Soc. Am.*, 132(3), pp. 1573–80, 2012, doi:10.1121/1.4740496.

[18] Stecker, G. C. and Hafter, E. R., “A recency effect in sound localization?” *J. Acoust. Soc. Am.*, 125(6), pp. 3914–3924, 2009, ISSN 1520-8524 (Electronic), doi:10.1121/1.3124776.

[19] Stecker, G. C., “Effects of carrier frequency and bandwidth on temporal weighting of binaural differences,” *Proc. Meet. Acoust.*, 19, p. 050166, 2013.

[20] Goupell, M. J., Laback, B., and Majdak, P., “Enhancing sensitivity to interaural time differences at high modulation rates by introducing temporal jitter,” *J Acoust Soc Am*, 126(5), pp. 2511–21, 2009, doi:10.1121/1.3206584.

[21] Laback, B. and Majdak, P., “Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates,” *Proc. Nat. Acad. Sci. U.S.A.*, 105(2), pp. 814–817, 2008, ISSN 1091-6490 (Electronic), doi:10.1073/pnas.0709199105.

[22] Diedesch, A. C. and Stecker, G. C., “Temporal weighting of interaural time differences in low frequency noise presented at low signal-to-noise ratio,” *J. Acoust. Soc. Am.*, 135, p. 2283, 2014.

[23] Dietz, M., Marquardt, T., Stange, A., Pecka, M., Grothe, B., and McAlpine, D., “Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds II: single-neuron recordings,” *J Neurophysiol*, 111(10), pp. 1973–85, 2014, doi:10.1152/jn.00681.2013.

[24] Hafter, E. R., Buell, T. N., and Richards, V. M., “Onset-coding in lateralization: its form site, and function,” in G. M. Edelman, W. E. Gall, and W. M. Cowan, editors, *Auditory function: neurobiological bases of hearing*, pp. 647–676, John Wiley & Sons, New York, 1988.

[25] Heil, P., “Representation of sound onsets in the auditory system,” *Audiology and Neuro-Otology*, 6, pp. 167–172, 2001.

[26] Hafter, E. R. and Buell, T. N., “Restarting the adapted binaural system,” *J. Acoust. Soc. Am.*, 88(2), pp. 806–812, 1990.

[27] Darwin, C. J. and Hukin, R. W., “Auditory objects of attention: the role of interaural time differences,” *J Exp Psychol Hum Percept Perform*, 25(3), pp. 617–29, 1999.

[28] Nelson, B. S. and Takahashi, T. T., “Spatial hearing in echoic environments: the role of the envelope in owls,” *Neuron*, 67(4), pp. 643–55, 2010, doi:10.1016/j.neuron.2010.07.014.

[29] Mao, J. and Carney, L. H., “Binaural detection with narrowband and wideband reproducible noise maskers. IV. Models using interaural time, level, and envelope differences,” *J Acoust Soc Am*, 135(2), pp. 824–37, 2014, doi:10.1121/1.4861848.