Interaural-time-difference thresholds for broad band-limited pulses are affected by relative bandwidth not temporal envelope sharpness

Paul G. Mayo, Philip C. Saunders, and Matthew J. Goupell
Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742, USA
paulmayo@umd.edu, psaund99@umd.edu, goupell@umd.edu

Abstract: Humans are sensitive to interaural time differences (ITDs) conveyed by slow modulations on high-frequency carrier signals. Sensitivity appears to be affected by temporal envelope sharpness, but it is unclear if there is a limit to which sharpness affects sensitivity. Pulse trains were varied in relative bandwidth (re: critical bandwidths) and center frequency. ITD sensitivity increased with increasing bandwidth. There was no effect of center frequency when relative bandwidths were analyzed, suggesting that the temporal envelope sharpness (concomitantly absolute bandwidth in Hz) did not affect performance. Rather, sensitivity was most easily explained by recruitment of additional auditory channels.

© 2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Interaural time differences (ITDs) are an important cue for sound localization in the horizontal plane (Macpherson and Middlebrooks, 2002). Humans are exquisitely sensitive to ITDs in the low-frequency (<1500 Hz) temporal fine structure, as well as for high frequencies (>1500 Hz) when there are slow (<300 Hz) envelope modulations (Hartmann, 2021). Envelope ITD sensitivity has been measured using a variety of modulators, including sinusoidal-amplitude-modulation, raised-sine envelopes, and Gaussian envelopes (Stecker et al., 2021). ITD sensitivity for such signals is affected by the modulation rate, modulation depth, and temporal envelope sharpness (Bernstein and Trahiotis, 2009, 2010; Klein-Hennig et al., 2011; Laback et al., 2011).

Band-limited acoustic pulses have been used as a simulation of the electrical pulse trains presented to cochlear-implant users, where a primary goal is to approximately match the wide bandwidths that occur with electrical stimulation and current spread (Goupell et al., 2009). An advantage of using such stimuli is that it is possible to parametrically investigate factors underlying ITD sensitivity without the variability inherent to the clinical cochlear-implant population (Kan and Litovsky, 2015). One assumption, however, is that the temporal-envelope onset of these stimuli is sufficiently sharp to approximate the instantaneous onset of an electrical pulse. The purpose of this study was to return to that assumption, as it must be met for these single acoustic pulses to be a valid simulation of electrical stimulation. Specifically, is the temporal envelope of band-limited acoustic pulses sharp enough that it does not appreciably impact ITD sensitivity and produce confounds for interpreting data?

Any change to the spectral properties of a narrowband acoustic signal will have concomitant physical changes to the temporal properties and vice versa. Larger bandwidths consequently produce sharper temporal pulse envelopes, and have been utilized in multiple studies where the bandwidth was kept approximately constant in millimeters to simulate current spread of a cochlear-implant (CI; Majdak and Laback, 2009), which corresponds to an approximately constant number of channels or equivalent-rectangular-bandwidth (ERB) rate (Moore and Glasberg, 1983). In other words, as the center frequency increased, the individual pulses became sharper (as ERBs are wider absolute bandwidths at higher center frequencies), which has the potential to improve ITD sensitivity. Likewise, Goupell et al. (2013) used stimuli that had mismatched center frequencies but kept the bandwidth constant in millimeters, meaning that the stimuli across the ears had different absolute bandwidths and thus, different amounts of temporal envelope sharpness.

Therefore, the goal of this study was to determine if ITD sensitivity could be better explained by temporal envelope sharpness or relative changes in the frequency domain, as this would help better understand the stimulus parameters that affect simulations of pulsatile electrical stimulation. To do this, we employed very low-rate [10 pulses per second.

a)Author to whom correspondence should be addressed, ORCID: 0000-0001-5751-5038.
b)ORCID: 0000-0003-2662-3057.
(pps) acoustic pulses to avoid temporal overlap and any rate effects, and then varied the bandwidth (in ERB) of these acoustic pulses for different center frequencies with different channel or critical bandwidths. We hypothesized that if temporal envelope sharpness was the driving factor behind ITD sensitivity, there would be an interaction between the physical sharpness/bandwidth scale in Hertz and center frequency. This interaction would be due to the saturation in sharpness/bandwidth that occurs from peripheral filtering for stimuli with bandwidths >1 ERB. There was also an alternative hypothesis that posited continual improvement with increasing relative bandwidth from recruitment of additional auditory channels (i.e., across-frequency combination of information).

2. Methods

2.1 Listeners
There were nine young normal-hearing adult listeners (23–38 y, average = 23.7 y) with pure-tone air-conduction thresholds ≤20 dB (re: hearing level) for octave frequencies from 0.25–8 kHz. Because some of the stimuli were presented at very high frequencies, we also tested hearing thresholds at 12 kHz. The listeners had no asymmetry >10 dB at any tested frequency.

2.2 Stimuli

The stimuli were trains of Gaussian-envelope-modulated cosine pulses, or "Gabor" pulses (Gabor, 1947; e.g., Fig. 1). The sinusoidal carrier had a frequency of 4, 8, or 12 kHz. The bandwidth of the pulses was manipulated so that the −3-dB bandwidth was 0.5, 1, 1.5, 2, 2.5, or 3 ERBs (according to Moore and Glasberg, 1983) for their respective center frequencies, making 18 total conditions. All trains were generated in the frequency domain using custom software written in MATLAB (the Mathworks, Natick, MA). The trains had a pulse rate of 10 pps, low enough to avoid any substantial pulse overlap, reduction in modulation depth, or interaction between center frequency and ITD sensitivity (Goupell et al., 2019).

Each pulse train had a duration of 300 ms and thus, consisted of three pulses. Stimuli were created at a 400-kHz sampling rate, allowing for a 2.5-μs ITD resolution, and down-sampled to 100 kHz after the ITD was applied. The left and right earphones were independently calibrated so that a long-duration, 10-pps, 8-kHz, 1-ERB stimuli measured 65 dBA; the other pulse trains were spectral peak normalized because of the desire to better understand the stimuli from Goupell et al. (2013).

Binaurally uncorrelated masking noise was presented to mask possible low-frequency distortion products. The masker was Gaussian white noise with a 1st-order Butterworth low-pass filter at 0.2 kHz, and a 3rd-order Butterworth low-pass filter at 1 kHz. The masking noise was presented at 40 dBA.

The sharpness (or attack time) of a single Gabor pulse was quantified using the \( t_{90} \) measure, the time difference between 10% and 90% of the peak temporal envelope value (Giannoulis et al., 2012). The \( t_{90} \) (in seconds) of a Gabor pulse can be precisely calculated as a function of its −3-dB bandwidth (in Hz):

\[
t_{90} = \sqrt{-2* \left( \frac{\sqrt{2} \ln(\sqrt{2})}{BW\pi} \right)^2 + \ln(10/100)} - \sqrt{-2* \left( \frac{\sqrt{2} \ln(\sqrt{2})}{BW\pi} \right)^2 + \ln(90/100)} \approx 0.447 \frac{BW}{BW}
\]

This equation can be modified to calculate the time difference between arbitrary points on the temporal envelope by substituting 10% and 90% for alternate starting and ending proportions, respectively.

2.3 Procedure

Stimuli were passed through a digital-to-analog converter (RP2.1, HB7, Tucker Davis Technologies, Alachua, FL). They were presented over circumaural headphones (Sennheiser HD650, Old Lyme, CT) to listeners seated in a double-walled, sound attenuating booth (IAC, North Aurora, IL).

Listeners performed an adaptive ITD left–right lateralization discrimination task using a custom MATLAB interface. The listeners were presented a 0-μs ITD stimulus, 300 ms of silence, then, the same stimulus with a randomly left or right leading ITD. Listeners were asked if the second sound (with the non-zero ITD applied) was perceived as left or right of the first. No correct answer feedback was given.

A 3-down-1-up adaptive procedure was used to find ITD thresholds (Levitt, 1971), with an initial ITD of 2500 μs and step factor of 2. The step factor decreased to \( \sqrt{2} \) after the second reversal. Staircases ended after 10 reversals, and thresholds were calculated as the geometric mean of the last six reversals. Each of the 18 conditions was measured three times in uniformly randomized blocks, resulting in 54 adaptive staircases per listener. The final reported threshold was the geometric mean over the three individual track measurements.

All listeners completed a training phase prior to the main experiment. Training consisted of at least three staircases with correct answer feedback using the 8-kHz center frequency and 1-ERB bandwidth stimuli. After the third staircase, a running average of all previous thresholds were compared to the most recent threshold. Training stopped when the most recent threshold was within ±15% of the moving average, indicating performance saturation. All listeners completed between three and eight staircases (mean = 4.8) before they moved on to the main experiment.
3. Results

The results of the experiment are shown in Fig. 2, where panel (A) is plotted as a function of absolute bandwidth in Hz (and therefore, also as a function of \( t_{90} \)) and panel (B) is plotted as a function of relative bandwidth in ERB. If increasing pulse sharpness influences the ITD thresholds for these already sharp stimuli, then the curves for the three different center frequencies should be aligned in Fig. 2(A), as absolute bandwidth/sharpness is a property independent of center frequency for stimuli with such a low rate, and then diverge when sharpness plateaus from peripheral filtering [see the open points in Fig. 1(C)]. No characteristics of this trend were observed. If the relative bandwidth and number of auditory channels influences the ITD thresholds, then the curves for the three different center frequencies should be aligned in Fig. 2(B), which appeared to be the case. Therefore, a two-way repeated-measures analysis of variance with factors bandwidth (in ERB) and center frequency was conducted. A Greenhouse–Geisser correction was used when the assumption of sphericity was violated. The ITD thresholds improved with increasing bandwidth \[ F(1.5, 13.0) = 30.4, \ p < 0.0001, \ \eta^2_p = 0.79 \]. Post hoc paired two-sample t-tests that were Bonferroni-corrected for 15 comparisons revealed that all bandwidths were significantly different from each other \((p < 0.05)\), except that the 1- and 1.5-ERB bandwidth were not different \((p = 0.10)\) and 2- and 2.5-ERB bandwidth stimuli were not different \((p = 0.068)\). The effect of center frequency was not significant \[ F(2, 16) = 0.68, \ p = 0.52, \ \eta^2_p = 0.08 \] and the bandwidth \times\ center frequency interaction was not significant \[ F(2.6, 21.0) = 0.92, \ p = 0.44, \ \eta^2_p = 0.10 \].

Finally, the data in Fig. 2(B) were fit to find the rate of change in ERB. We found a slope of \(-0.20\) (95% confidence interval = \([-0.22, -0.17]\), \( R^2 = 0.95 \)) for log(ITD) versus ERB and a slope of \(-0.61\) (95% confidence interval = \([-0.73, -0.50]\), \( R^2 = 0.89 \)) for log(ITD) versus log(ERB).

Fig. 2. Panel (A) shows the data as a function of absolute bandwidth in Hz (and therefore, also as a function of \( t_{90} \)) for the three different center frequencies. Panel (B) shows the data as a function of relative bandwidth in ERB.
4. Discussion

The purpose of this study was to determine if increased pulse sharpness improved ITD sensitivity for relatively sharp stimuli, often used in CI simulations (e.g., Majdak and Laback, 2009; Goupell et al., 2013), or if the increased sensitivity could be better explained by changes in the frequency domain. We found that ITD sensitivity improved with increasing bandwidth in ERB, but there was no effect of center frequency or absolute bandwidth/sharpness. Sensitivity improved even for stimuli with bandwidths >1 ERB, where sharpness plateaus due to peripheral filtering [see Fig. 1(C)]. This supports the alternative hypothesis that relative bandwidth and auditory channel recruitment, rather than pulse sharpness, affected performance. Previous studies have investigated the effect of envelope sharpness on ITD sensitivity using exponential-shaped modulation (Laback et al., 2011) and raised-sine envelope manipulations (Bernstein and Trahiotis, 2012). However, the stimuli employed here have envelopes far sharper than those of previous studies. It is possible that sharpness has a direct but bounded effect on ITD sensitivity, recruiting additional channels increases sensitivity, and the stimuli used here exceed the sharpness limit. Regardless, this interpretation supports the notion that pulse trains with relatively large bandwidths are reasonable cochlear-implant stimulations from the standpoint that they appear to have envelopes sharp enough to not have a confound compared to the instantaneous pulse onsets that occur with electrical stimulation.

The ITD thresholds measured were relatively high compared to those reported for stimuli with higher rates (Bernstein and Trahiotis, 2009, 2010), and more in line with those studies using only a few pulses (Hafter and Dye, 1983). Increasing the bandwidth of a high-frequency carrier narrowband noise (Bernstein and Trahiotis, 1994) and pulse trains (Goupell et al., 2013) improves ITD sensitivity. Binaural models that predict performance based on the normalized cross-correlation function after physiologically inspired transformations (auditory filtering, compression, rectification, low-pass filtering capturing neural refractoriness) often focus on temporal aspects of the stimuli and choose a single best channel for describing performance, which is often assumed to be at the center frequency; however, sometimes it is required to move to an off-frequency channel at a higher frequency to obtain a sharper stimulus attack for modulated tones (Bernstein and Trahiotis, 2009, 2010) or to a lower frequency to obtain a lower modulation rate for narrowband noises (Bernstein and Trahiotis, 1994). Other models use across-channel integration of correlation patterns to describe intracranial lateralization (Stern et al., 1988; Bernstein and Trahiotis, 2012), and sometimes ITD discrimination (Trahiotis et al., 2001). The current data, which adds the important parameter of center frequency, suggest that changes in ITD sensitivity with bandwidth are predicted by considering multiple channels, consistent with the latter group of studies.

If across-channel recruitment provided an ideal improvement of ITD sensitivity according to signal detection theory, the fitted slope of log(ITD) to ERB would be $-0.5$ (an improvement of $1/\sqrt{n}$, where $n$ is the number of auditory channels recruited) (Haft and Dye, 1983) compared to the observed slope of $-0.2$. One explanation for this discrepancy is that the individual Gabor pulses are continuous in the frequency domain. That is, a continuous number of overlapping channels are recruited on either side of the stimulus center frequency when the stimulus bandwidth is extended and therefore, more than one additional channel is recruited per unit change in ERB. Additionally, a requirement for the theoretically ideal slope of $-0.5$ is that both the abscissa and ordinate be logarithmic units. While ERBs are relatively logarithmically spaced as a function of frequency, a linear change to a single frequency's ERB (e.g., changing from 1–2 ERBs at 4 kHz as done in this study) is a linear change in bandwidth. A steeper slope of $-0.61$ when fitting against log(ERB) is closer to the theoretically ideal slope of $-0.5$, but the correct approach and scale are unclear. It could also be the auditory system simply does not perfectly integrate across-channel information, and that recruiting multiple additional auditory channels does not equal a unit increase in information.

There are a few possible confounds to our interpretation. First, these data do not rule out that the improved ITD thresholds were due to off-frequency listening and energy extending into broader auditory channels (which would have sharper impulse responses than the on-frequency channels). However, this does not seem likely as the absolute bandwidth of auditory channels increases as a function of frequency, so the potential increase in sharpness available from off-frequency listening would be greater for higher-frequency stimuli compared to lower-frequency stimuli. This could result in a center frequency $\times$ ERB interaction, which was not observed. Likewise, simulations of off-frequency listening reveal that there is no off-frequency filter that would cause the 4-kHz stimuli to reach the sharpness of the 12-kHz stimuli (data not shown). The only way to have equal sharpness across these stimuli is for the auditory system to listen to a lower-frequency channel with the 12-kHz stimuli, one with up to 64 dB less energy than the on-frequency channel and likely subthreshold. Further research investigating ITD sensitivity as a function of the number of auditory channels without the confound of recruiting broader channels is necessary. Second, these data cannot rule out level effects on ITD sensitivity (Dietz et al., 2013). Due to the spectral peak normalization across all bandwidths, it is possible that there was a level-based change in ITD sensitivity as a function of bandwidth. Third, we did not see a plateau in ITD sensitivity for bandwidths $\leq 1$ ERB. If the increased ITD sensitivity was due to the recruitment of additional channels, then ITD sensitivity should remain constant when the stimulus energy is contained within a single channel. However, the spectra of our stimuli were not rectangular, and it is unlikely that energy was contained within a single auditory channel, even for the narrowest bandwidths. Additional research involving multiple stimuli with bandwidths narrower than a critical bandwidth is necessary to verify this theoretical plateau and support our interpretation.
Acknowledgments

Research reported in this publication was supported by a grant awarded by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under Award No. R01DC014948 (M.J.G.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References and links

Bernstein, L. R., and Trahiotis, C. (1994). “Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise,” J. Acoust. Soc. Am. 95, 3561–3567.

Bernstein, L. R., and Trahiotis, C. (2009). “How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli,” J. Acoust. Soc. Am. 125, 3234–3242.

Bernstein, L. R., and Trahiotis, C. (2010). “Accounting quantitatively for sensitivity to envelope-based interaural temporal disparities at high frequencies,” J. Acoust. Soc. Am. 128, 1224–1234.

Bernstein, L. R., and Trahiotis, C. (2012). “Lateralization produced by interaural temporal and intensive disparities of high-frequency, raised-sine stimuli: Data and modeling,” J. Acoust. Soc. Am. 131, 409–415.

Dietz, M., Bernstein, L. R., Trahiotis, C., Ewert, S. D., and Hohmann, V. (2013). “The effect of overall level on sensitivity to interaural differences of time and level at high frequencies,” J. Acoust. Soc. Am. 134, 494–502.

Gabor, D. (1947). “Acoustical quanta and the theory of hearing,” Nature (London) 84, 591–594.

Giannoulis, D., Massberg, M., and Reiss, J. D. (2012). “Digital dynamic range compressor design—A tutorial and analysis,” J. Audio Eng. Soc. 60, 399–408, available at http://www.aes.org/e-lib/browse.cfm?elib=16354.

Goupell, M. J., Fong, S., and Stakhovskaya, O. (2019). “The effect of envelope modulations on binaural processing,” Hearing Res. 379, 117–127.

Goupell, M. J., Laback, B., and Majdak, P. (2009). “Enhancing sensitivity to interaural time differences at high modulation rates by introducing temporal jitter,” J. Acoust. Soc. Am. 126, 2511–2521.

Goupell, M. J., Stoelb, C., Kan, A., and Litovsky, R. Y. (2013). “Effect of mismatched place-of-stimulation on the salience of binaural cues in conditions that simulate bilateral cochlear-implant listening,” J. Acoust. Soc. Am. 133, 2272–2287.

Hafter, E. R., and Dye, R. H., Jr. (1983). “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, 644–651.

Hartmann, W. M. (2021). “Localization and lateralization of sound,” in Binaural Hearing, edited by R. Y. Litovsky, M. J. Goupell, A. N. Popper, and R. R. Fay (Springer International Publishing, Cham Switzerland), pp. 9–45.

Kan, A., and Litovsky, R. Y. (2015). “Binaural hearing with electrical stimulation,” Hear. Res. 322, 127–137.

Klein-Hennig, M., Dietz, M., Hohmann, V., and Ewert, S. D. (2011). “The influence of different segments of the ongoing envelope on sensitivity to interaural time delays,” J. Acoust. Soc. Am. 129, 3856–3872.

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W. D., and Pok, S. M. (2011). “Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing,” J. Acoust. Soc. Am. 130, 1515–1529.

Levitt, H. (1971). “Transformed up-down methods in psychoacoustics,” J. Acoust. Soc. Am. 49, 467–477.

Macpherson, E. A., and Middlebrooks, J. C. (2002). “Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited,” J. Acoust. Soc. Am. 111, 2219–2226.

Majdak, P., and Laback, B. (2009). “Effects of center frequency and rate on the sensitivity to interaural delay in high-frequency click trains,” J. Acoust. Soc. Am. 125, 3903–3913.

Moore, B. C., and Glasberg, B. R. (1983). “Suggested formulae for calculating auditory-filter bandwidths and excitation patterns,” J. Acoust. Soc. Am. 74, 750–753.

Stecker, G. C., Bernstein, L. E., and Brown, A. D. (2021). “Binaural hearing with temporally complex signals,” in Binaural Hearing, edited by R. Y. Litovsky, M. J. Goupell, A. N. Popper, and R. R. Fay (Springer International Publishing, Cham Switzerland), pp. 145–180.

Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (1988). “Lateralization of complex binaural stimuli: A weighted-image model,” J. Acoust. Soc. Am. 84, 156–165.

Trahiotis, C., Bernstein, L. R., and Akeroyd, M. A. (2001). “Manipulating the ‘straightness’ and ‘curvature’ of patterns of interaural cross correlation affects listeners’ sensitivity to changes in interaural delay,” J. Acoust. Soc. Am. 109, 321–330.