

LETTERS TO THE EDITOR

This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

Across-frequency interaction in lateralization of complex binaural stimuli

Constantine Trahiotis

Center for Neurological Sciences and Department of Surgery (Otolaryngology), University of Connecticut Health Center, Farmington, Connecticut 06032

Richard M. Stern

Department of Electrical and Computer Engineering and Biomedical Engineering Program, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

(Received 25 April 1994; accepted for publication 10 August 1994)

This communication concerns the nature of the mechanisms by which the binaural auditory system combines, across frequency, interaural timing information. New observations are presented which indicate that the integration of such information is *not* due to a simple averaging across frequency. Instead, the new observations underscore the importance of the monaural time structure of the stimuli. These observations reinforce our postulation of a mechanism which relies on temporally coincident activity across frequency channels that monitor the same interaural delays.

PACS numbers: 43.66.Ba, 43.66.Nm, 43.66.Pn, 43.66.Qp [HSC]

This letter addresses the ways in which the binaural system combines, across frequency, information in lateralization tasks. Stern *et al.* (1988) and Trahiotis and Stern (1989) argued that lateralization is dominated by the components of the stimulus that produce consistent interaural timing information over a range of frequencies. Shackleton *et al.* (1992) provide a different point of view, arguing that the same data can be predicted by simply *averaging* the interaural timing information over frequency, *without* any explicit mechanism that weights more heavily interaural information that is consistent.

There are other differences between the models of Stern *et al.* and Shackleton *et al.* besides the use of consistency of interaural delay over frequency. For example, the model of Shackleton *et al.* uses a Gaussian-shaped pulse with a 0.6-ms standard deviation to specify the distributions of fiber pairs with respect to internal delay. In contrast, Stern *et al.* utilize a function that is approximately constant for 0.2 ms, and which then decreases exponentially as internal delay increases. Their function is a modification of the function $p(\tau)$ introduced by Colburn (1973). Details concerning the formulation of the function $p(\tau)$ can be found in Stern and Shear (1994). Another difference is that Shackleton *et al.* make use of a more detailed computational description of the auditory-nerve response to the stimuli based on the work of Meddis *et al.* (1990). Stern and his colleagues have used analytical

characterizations that are simpler but that provide a less detailed description of the physiological data. We do not believe that these differences (which are independent of consistency of interaural delay over frequency) are important for the purposes of this discussion.

We recently described a mechanism (Stern and Trahiotis, 1991) which accounts for the effects of interaural consistency over frequency within the context of the physiologically motivated models of binaural processing developed by Colburn (1973, 1977) and Stern and Colburn (1978). In addition to the Jeffress/Colburn display of interaural timing information as a joint function of internal delay and frequency (cf. Stern and Trahiotis, 1994), we also assume a second level of coincidence-counting units that take as their inputs the outputs from a small number of the coincidence-counting units proposed by Jeffress (1948). Each set of inputs is assumed to come from coincidence counters representing a (small) range of characteristic frequencies, but with a common internal delay. This manner of weighting straightness also sharpens the ridges of the two-dimensional cross-correlation function along the internal-delay axis without an explicit mechanism for inhibition such as those postulated by Blauert and Cobben (1978) and Lindemann (1986).

It is helpful to bear in mind that information is combined in different ways at different levels of our model of the binaural system. The original level of Jeffress–Colburn-type

units record coincidences in firing times of auditory-nerve fibers within a small time interval that is probably on the order of 100 μ s or so. The second-level coincidence units record coincidences in activity over frequency. The time interval over which these coincidences are recorded primarily affects the ways in which the response to spectrally nonstationary stimuli is affected by temporal coherence in the across-frequency patterns. Although this time interval has not been precisely specified, we assume that it is roughly on the order of 30 ms. We also assume that the resulting display of information *after* the second level of coincidences is averaged over running time over a time interval of approximately 200 ms, in the manner suggested by the results of Grantham and Wightman (1978). Considering the overall system response to sound as a running cross-correlation function, we note that the effects of temporal interaction at the three levels are different. The time interval over which coincidences are recorded at the first level affects the resolution of the systems's representation of sounds with respect to the *internal-delay axis* of the running cross-correlation function. On the other hand, the temporal integration proposed by Grantham and Wightman (i.e., the third level of temporal averaging) affects resolution with respect to *running time*.

Support for a second-level coincidence mechanism is provided by physiological findings of Takahashi and Konishi (1986). They compared the responses of interaural-time-sensitive units in the inferior colliculus in the barn owl to a single tone at the best frequency, to a tone of a second frequency, and to the two tones presented simultaneously. Takahashi and Konishi found units that were tuned to the same interaural delay over a range of stimulus frequencies. Some of these units produced responses to the simultaneous presentation of the two tones that were greater than the sum of the responses to each of the tones presented in isolation. In other words, the response to a tone presented at the best frequency with the "best" ITD is facilitated by the presentation of a second tone with that same ITD. This is exactly the type of response that would be predicted by mechanisms like the second-level coincidences across frequency that we proposed.

We believe that the experimental data of Stern *et al.* (1988) are more accurately described by the predictions of the model which explicitly includes weighting by a second layer of coincidence counters (Stern and Trahiotis, 1991, Fig. 6) than by the predictions of the model of Shackleton *et al.* [1992, Fig. 2(a) and (c)]. In our opinion, the predictions obtained by Stern and Trahiotis without straightness weighting (1991, Fig. 6) are not unlike those generated by the model of Shackleton *et al.*

We recognize that our beliefs, although based on what we believe to be an objective assessment of comparison of predictions and data, cannot be the basis of a scientifically acceptable argument favoring one model over the other. Therefore, we strived to conceive of binaural stimuli that could be used to help choose between the two manners of integrating binaural information across frequency. Specifically, we searched for sets of stimuli for which the putative response of the binaural system would be the same when averaged over frequency and over time (i.e., over a 100-ms

duration) in the manner proposed by Shackleton *et al.* (1992), but which should produce binaural images with *different* spatial properties if consistency of interaural timing information is critical. We realized that sinusoidally amplitude modulated tones with sufficiently low rates of modulation to preclude resolution of the sidebands could be used for this purpose. Our approach was to compare binaural images produced when the modulation of the tones was "monaurally" in phase with binaural images produced when the SAM tones were monaurally out of phase. The crux of the argument concerns whether across-frequency averaging is sufficient to describe the phenomena or whether the ongoing time structure of the stimuli can determine binaural images by favoring temporally coincident neural activity across frequency channels.

Consider the following two sets of binaural stimuli:

SET 1:

$$x_L(t) = [1 + \cos(\omega_m t)] \cos(\omega_1 t) + [1 + \cos(\omega_m t)] \times \cos(\omega_2 t) + [1 + \cos(\omega_m t)] \cos(\omega_3 t),$$

$$x_R(t) = [1 + \cos(\omega_m t)] \cos[\omega_1(t - T_s)] + [1 + \cos(\omega_m t)] \times \cos[\omega_2(t - T_s)] + [1 + \cos(\omega_m t)] \cos[\omega_3(t - T_s)].$$

SET 2:

$$x_L(t) = [1 - \cos(\omega_m t)] \cos(\omega_1 t) + [1 + \cos(\omega_m t)] \times \cos(\omega_2 t) + [1 - \cos(\omega_m t)] \cos(\omega_3 t),$$

$$x_R(t) = [1 - \cos(\omega_m t)] \cos[\omega_1(t - T_s)] + [1 + \cos(\omega_m t)] \times \cos[\omega_2(t - T_s)] + [1 - \cos(\omega_m t)] \cos[\omega_3(t - T_s)].$$

Note that the two sets of stimuli differ only in terms of the phase of amplitude modulation applied to the carrier frequencies ω_1 and ω_3 , and that all three components of both sets of stimuli have exactly the same interaural time differences. Now, let us consider the dynamic properties of the short-term interaural cross correlation of the stimuli. The stimuli of set 1 have the property that the peaks of the amplitude modulation at each of the three carrier frequencies occur simultaneously. Hence the responses of the interaural coincidence-counting units at each of the three frequencies would be expected to occur more or less simultaneously. Because of this, the second layer of coincidence-counting units which signal consistency of interaural delay across frequency would be activated. On the other hand, the stimuli of set 2 would be expected to produce approximately simultaneous responses of the coincidence-counting units only for the two carrier frequencies that are modulated in phase (ω_1 and ω_3). The peaks of modulation of the remaining carrier frequency (ω_2) occur at the valleys of modulation of ω_1 and ω_2 and *vice versa*. As a result, there should be substantially less activity at the second level of coincidence.

If across-frequency coincidence of neural activity of the interaural timing information is salient, these two sets of stimuli should produce *different* binaural images. Alternatively, if the binaural lateralization mechanism simply aver-

ages over running time (e.g., Grantham and Wightman, 1978), or frequency, then the stimuli of set 1 and set 2 should have the same spatial qualities.

We listened to a number of such stimuli. We chose values of the stimulus parameters for which the amplitude modulation was sufficiently slow to ensure that the components at each of the three carrier frequencies remained unresolved. A typical modulation frequency was 20–25 Hz. The three carrier frequencies were chosen to be sufficiently separated to preclude interaction of their respective sidebands, yet be close enough to become fused into a single binaural image when presented as in set 1. Typical carrier frequencies included the harmonically related set of 300, 500, and 700 Hz, as well as several inharmonically related sets of frequencies containing similar frequencies. The interaural delay for stimuli comprising both set 1 and set 2 was always 1500 μ s. This value was chosen because it produces a delay of three-quarters of a period for the carrier of the 500-Hz SAM tone, a value used in our previous studies.

We listened to ongoing repetitions of three presentations of the stimuli of set 1 followed by three presentations of the stimuli of set 2, repeated for about one minute. Each presentation of the stimuli was 500 ms in duration, with a silent interval of 200 ms between each presentation, and a larger silent interval of 500 ms between each set of three presentations.

All of the stimuli from set 1, for which the peaks of the amplitude modulation at each carrier frequency were always in phase, were heard as a single, compact, binaural image that was well-lateralized toward the ear that received the signal that was leading in time. The compactness and location of the binaural image indicates that some type of across-frequency integration of the internal response according to the actual interaural time delays was taking place. The binaural spatial properties of the stimuli from set 2 were quite different and can be described in two manners. For some combinations of carrier frequency and frequency of modulation, stimuli from set 2 were heard as two distinct images. The images corresponding to the outer two carrier frequencies (ω_1 and ω_3) were heard toward the ear receiving the leading signal. *At the same time*, however, the image corresponding to ω_2 was heard toward the opposite ear (which received the signal lagging in time). For other combinations, stimuli from set 2 produced a diffuse hollow-sounding image that frequently “filled the head.” We found that adding a 1- or 2-dB interaural intensity difference favoring the signal that was lagging in time (i.e., the signal presented to the right ear) facilitated the dissolution of the images of set 2, while having *absolutely no such effect* on the images of set 1.

These qualitative observations were quite reliable, and were reported by several others who listened to similar demonstrations. Although verbal descriptions of the stimuli depended upon the actual values of the parameters chosen for the demonstration, the differences between the binaural spa-

tial properties of stimuli from set 1 and set 2 did not. Dr. Trevor Shackleton, who served as a reviewer of this letter, graciously verified that he observed the same phenomena.

In conclusion, we believe that the across-frequency integration of binaural information is not due to a simple averaging in time of the internal response to binaural stimuli. Instead, it appears that the mechanism of spectral integration depends upon the extent to which responses occurring within each frequency channel are temporally proximate. That is the defining feature of the second level of coincidence postulated by Stern and Trahiotis (1991).

ACKNOWLEDGMENTS

This work has been supported by NIH Grant DC-00234 and AFOSR Grant 89-0030 to Constantine Trahiotis and by NSF Grant IBN 90-22080 to Richard Stern. We thank Dr. Trevor Shackleton and two anonymous reviewers along with the editor, Dr. H. Steven Colburn, for their comments which helped us clarify certain details central to this work.

- Blauert, J., and Cobben, W. (1978). “Some Consideration of Binaural Cross-Correlation Analysis,” *Acustica* **39**, 96–103.
- Colburn, H. S. (1973). “Theory of Binaural Interaction Based on Auditory-Nerve Data. I. General Strategy and Preliminary Results on Interaural Discrimination,” *J. Acoust. Soc. Am.* **54** 1458–1470.
- Colburn, H. S. (1977). “Theory of Binaural Interaction Based on Auditory-Nerve Data. II. Detection of Tones in Noise,” *J. Acoust. Soc. Am.* **61**, 525–533.
- Grantham, D. W., and Wightman, F. L. (1978). “Detectability of Varying Interaural Temporal Differences,” *J. Acoust. Soc. Am.* **63**, 511–523.
- Jeffress, L. A. (1948). “A Place Theory of Sound Localization,” *J. Comp. Physiol. Psychol.* **41**, 35–39.
- Lindemann, W. (1986). “Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals,” *J. Acoust. Soc. Am.* **80**, 1608–1622.
- Meddis, R., Hewitt, M. J., and Shackleton, T. M. (1990). “Implementation Details of a Computational Model of the Inner Hair Cell/Auditory-Nerve Synapse,” *J. Acoust. Soc. Am.* **87**, 1813–1818.
- Shackleton, T. M., Meddis, R., and Hewitt, M. J. (1992). “Across Frequency Integration in a Model of Lateralization,” *J. Acoust. Soc. Am.* **91**, 2276–2279 (L).
- Stern, R. M., Jr., and Colburn, H. S. (1978). “Theory of Binaural Interaction Based on Auditory-Nerve Data. IV. A Model for Subjective Lateral Position,” *J. Acoust. Soc. Am.* **64**, 127–140.
- Stern, R. M., and Shear, G. D. (1994). “Lateralization and Detection of Low-Frequency Binaural Stimuli: Effects of Distribution of Internal Delay,” *J. Acoust. Soc. Am.* (in revision).
- Stern, R. M., and Trahiotis, C. (1991). “The Role of Consistency of Interaural Timing over Frequency in Binaural Lateralization,” in *Proc. of the Ninth International Symposium on Auditory Physiology and Perception, Carcans, France*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford).
- 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.
- Stern, R. M., and Trahiotis, C. (1994). *Handbook of Perception and Cognition. Vol. VI. Hearing: Models of Binaural Interaction*, edited by B. C. J. Moore (Academic, New York).
- Takahashi, T. A., and Konishi, M. (1986). “Selectivity for Interaural Time Differences in the Owl’s Midbrain,” *J. Neurosci.* **6**, 3413–3422.
- Trahiotis, C., and Stern, R. M. (1989). “Lateralization of Bands of Noise: Effects of Bandwidth and Differences of Interaural Time and Intensity,” *J. Acoust. Soc. Am.* **86**, 1285–1293.