AD-A224 147

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE May 24, 1990	3. REPORT TYPE AND DATES COVERED Annual 01 May 89 to 30 Apr 90

4. TITLE AND SUBTITLE

Auditory Processing of Complex Sounds

5. FUNDING NUMBERS

G - AFOSR-89-0335
PE - 61102F
PR - 2313
TA - A6

6. AUTHOR(S)

Dr William P Shofner

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Loyola University
Parmly Hearing Inst
6525 N. Sheridan
Chicago, IL 60626

AFOSR-TR   90 0754

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Dr John F Tagney
AFOSR/NL
Building 410
Bolling AFB, DC 20332

DTIC
ELECTE
JUL 20 1990
S E

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

unlimited

12b. DISTRIBUTION CODE

Approved for public release;
distribution unlimited.

13. ABSTRACT (Maximum 200 words)

Neurophysiological experiments have been directed at gaining an understanding of how auditory neurons encode pitch related information in the temporal properties of discharge. In general, all physiological neuronal types recorded to date in the chinchilla cochlear nucleua can show periodicities in their discharges that are related to the pitch of harmonic tone complexes, but only those neurons that show phase-locking at best frequency can encode the pitch related information in cos+ rippled noise. The results of binaural psychophysical experiments suggest (1) that spectrally synthetic binaural processing is the rule when the number of components in the tone complex are relatively few (less than 10) and there are no dynamic binural cues to aid segregation of the target from the background, and (2) that waveforms having large effective envelope depths are on the average more easily lateralized than those having small effective envelope depths.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT (U)	18. SECURITY CLASSIFICATION OF THIS PAGE (U)	19. SECURITY CLASSIFICATION OF ABSTRACT (U)	20. LIMITATION OF ABSTRACT (U)

Auditory Processing of Complex Sounds Across Frequency Channels (AFOSR-89-0335)

Principal Investigator: William P. Shofner, Ph.D.
Co-PI: Raymond H. Dye, Ph.D.
Co-PI: William A. Yost, Ph.D.
Investigator: Stanley Sheft, Ph.D.

Parmly Hearing Institute
Loyola University of Chicago
6525 N. Sheridan Rd.
Chicago, IL 60626
312-508-2710

Accession For		
NTIS GRA&I		X
DTIC TAB		
Unannounced		☐
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A-1		

ANNUAL TECHNICAL REPORT   (May 1, 1989 to May 1, 1990)

SECTION I. Auditory Physiology
William P. Shofner

Neurophysiological experiments during the last year have been directed at gaining an understanding of how auditory neurons encode stimulus information found in the time domain of complex sounds, particularly those complex sounds which can generate the perception of pitch. Information in the time domain can be found in the waveform fine structure and envelope. Rippled noise is a broadband stimulus which produces the perception of pitch, yet is aperiodic in the time domain. Cos+ rippled noise is generated when a broadband noise is delayed and then added to the undelayed noise. The resulting stimulus has a power spectrum that varies in a cosinusoidal fashion in which the peaks are separated by 1/delay. The autocorrelation function of waveform fine structure of cos+ noise has a single peak at the delay of the noise. Thus, unlike wideband noise which is aperiodic and has a flat autocorrelation function, *rippled noise is an aperiodic stimulus that does not have a flat autocorrelation function.* Moreover, the autocorrelation function of the envelope of cos+ noise also has a single peak at the delay. Experiments over the last year have investigated the responses of cochlear nucleus neurons in the chinchilla to tone complexes and cos+ rippled noise. Specifically, responses based on the temporal discharge of cochlear nucleus neurons were studied. These temporal properties are those based on the time intervals between individual spikes.

In general, all physiological neuronal types recorded can show periodicities in their discharge in response to tone complexes that are related to the fundamental frequency. In contrast, only those neurons that show phase-locking at best frequency have autocorrelation functions based on the temporal discharge in response to rippled noise that show a peak at the delay. The strength of synchrony in response to cos+ noise is currently being measured as the height of the peak in the neuron's autocorrelation function. Neuronal autocorrelation functions are based on first-order interspike intervals as well as all higher-order intervals. However, correlations that may exist among higher-order interspike intervals will contribute to the height of this peak. In order to evaluate possible contributions of higher-order intervals have on the peak in the autocorrelation function, a new neuronal spike train is generated by randomly shuffling the interspike intervals. Random shuffling of the intervals removes correlations among higher-order intervals, but does not affect the first-order interval distribution. For neurons that show a peak in the autocorrelation function in response to cos+ rippled noise, a slightly reduced peak remains in the autocorrelation functions computed from the shuffled intervals. The presence of peaks in the shuffled autocorrelation functions which are only slightly reduced suggests that the heights of these peaks in the unshuffled autocorrelation functions predominately reflect contributions of first-order intervals, but correlations among higher-order intervals can also contribute to the measured synchrony.

Experiments planned in the coming year include a comparison of neuronal responses to cos+ noise with responses to cos- noise. In contrast to cos+ rippled noise, cos-noise is generated when a broadband noise is delayed and subtracted from the undelayed noise. The autocorrelation function of cos- noise has a single trough at the delay, while the autocorrelation function of the envelope of cos- noise has a single peak at the delay. Comparison of responses

1

to cos+ and cos- noise will provide data as to whether neurons extract the delay from the waveform fine structure or from the envelope.

Presented Research

Shofner, W.P. (1990) Temporal responses of cochlear nucleus units in the chinchilla to low frequency tones, tone complexes and rippled noise. Abstr. of the 13th Midwinter Meeting, Assoc. Res. Otolaryngol. p. 400.

Shofner, W.P. (1990) Measurement of the strength of synchrony of chinchilla cochlear nucleus units in response to cos+ noise. Submitted to 1990 Annual Meeting, Society for Neuroscience.

SECTION II. Psychophysics-Binaural Processing
        Raymond H. Dye, Jr. (with Mark Stellmack)

During the past year or so our work has continued along two avenues. First, **Binaural Processing of Spectrally Incoherent Signals** (signals with interaural delays that varying across the spectrum) is an attempt to gain insight into the processes by which the binaural system combines interaural information across the frequency domain to localize sound sources. Past experiments in our lab with 3-and 5-component complexes show the binaural auditory system to be spectrally synthetic in that information is pooled between low-frequency components and there is no access to information regarding which component was delayed.

Our recent work attempts to test the generality of spectral synthesis in binaural processing. First we measured interaural time thresholds for 3, 5, 7, and 9 component complexes centered at 753 Hz as a function of frequency spacing (10 Hz to 200 Hz) to see if spectral synthesis might break down as the spectrum was made richer and more dense. A single interval task in which the first interval always contained a diotic 753 Hz component and the second interval contained either a diotic n-component complex or a complex with a 753 Hz component delayed to the right ear (lateralized to the left). Worst performance was obtained for frequency differences in the 25-50 Hz range, with performance improving as Δfs approached 10 Hz. For all frequency spacings and all numbers of components, binaural interference was present. As Δf was decreased to 10 Hz, subjects actually improved in their ability to detect delays of the middle component. Subjects reported that the interval containing the dichotic stimulus was associated with a broader, more diffuse intracranial image when the spacing was 10 Hz, and that this cue was used as a basis for performing the task.

Several of our colleagues have suggested that these results might be due to the fact that all of our signal components were gated on and off at the same time, and Gestalt principles of perceptual organization would argue that this would lead to the judgement that all components had a common source. To test this notion, we measured interaural time thresholds for conditions in which the background was turned on prior to the target. This onset asynchrony varied from 25 to 200 ms, and the total duration of the signal was 200 ms (the target

duration) plus the duration of the asynchrony. For one subject, the onset asynchrony had virtually no effect, while it hurt the performance of the other three subjects. Another suggestion has been that our past work employed signals having only ongoing interaural delays with no onset delay (signals at the two ears were gated simultaneously), and that onset delays might be important for allowing the target to be heard at a different interaural position than the background. Conditions were run for 7-component complexes (Δf = 10, 50, or 100 Hz) in which onset and ongoing delays were present, and our results show ΔIDTs to be as high as those obtained with only ongoing delays. Finally, it has been suggested that harmonic complexes are special, and that making the target have a harmonic relation with the background might lead to more interference than if the components were presented in other configurations. To test this notion, ΔIDTs were measured for 3-tone complexes that were harmonic, inharmonic, or had a mistuned target (center frequency). This last condition should maximize the ability to observers to segregate the background from the target, since the background formed a harmonic complex. Again, no release from interference was found. These results suggest that spectrally synthetic binaural processing is the rule, rather than the exception, when the number of components in the complex is relatively small (fewer than 10) and there are no dynamic binaural cues to aid segregation of the target from the background (e. g., varying the delay of the target over time, or varying which component is delayed over time).

Our work on **Envelope Extraction in Binaural Processing** during the past year has extended some of our past work to envelope-delayed bands of noise. The focus has been to examine the explanatory power of envelope-detection schemes as mechanisms by which the binaural auditory system localizes high-frequency signals on the basis of envelope delay. Our strategy has been to manipulate the **effective** envelope depth without altering the spectrum of the signals, typically by varying the starting phases of the signal components. Our past work has shown the effective envelope of multicomponent complexes is a good predictor of the ease with which interaural envelope delays are detected. To characterize the effective envelope of these waveforms, the standard deviation of the instantaneous power in units of average power are computed (defined as z), and we find that waveforms of higher z are more easily lateralized. This finding is consistent with envelope-extraction schemes that consist of a nonlinearity followed by a low-pass filter.

Threshold interaural envelope delays were measured in a 2-AFC task for narrow bands of noise whose center frequency was fixed at 4000 Hz. The bandwidth of the noise was set to 50, 100, or 200 Hz. All components of the noise were equal amplitude, and different noise samples were generated by randomizing the starting phases of the components. The signal duration was 200 ms with 10-ms linear rise/decay times. The standard deviation of the instantaneous power in units of average power was computed for a large number of noise waveforms. For this study the waveforms were classified by into five ranges of z-values and efforts were made to correlate performance with z. Data were gathered using blocked trials (all 100 trials in a run employing waveforms within the same range of z's) and mixed trials (all range of z's run in a 100-trial block). While performance with high effective envelope depths was superior to that with low envelope depths on the average, it appears that relationship between z and threshold envelope delays is not simple. This is especially true of the data gathered with a 50-Hz wide band of noise. Part of the problem concerns the fact

3

that the location of the major peak in the temporal waveform varies from one waveform to the next when the bandwidths are relatively narrow. Waveforms with early peaks appear to be easier to lateralize than those with later peaks. As such, the data gathered at a single restricted range of z-values tend to be quite variable. We are currently examining signal generation schemes that will allow us to generate waveforms of variable effective envelope depth without altering the location of the prominent peak in the temporal waveform.

Presented Research
    Stellmack, M. A. and Dye, R. H. (1989). "The Effect of Onset Asynchonies between Components on the Ability to Detect a Delayed Component Embedded in a Diotic Complex," J. Acoust. Soc. Am. 86, S11.

SECTION III. Psychophysics-Auditory Object Perception
            William A. Yost and Stanley Sheft

    We have been busy with a number of studies investigating the inability of listeners to attend to the temporal modulation (usually sinusoidal amplitude modulation) of components in one spectral region when components in another spectral region(s) are also temporally modulated. According to the classic notions of the critical band theory of hearing, listeners should be able to attend to one region of the spectrum in the presence of components in remote spectral regions. In our work, listeners can attend to different spectral regions when the components do not share a common pattern of temporal modulation. However, when the components are coherently modulated, processing of the modulation in any one region of the spectrum is very difficult.

    We have called this inability to independently process the modulation of a single spectral region Modulation Detection Interference (MDI). When the temporal modulation is sinusoidal amplitude modulation, we have demonstrated MDI for the detection of modulation, the discrimination of modulation rate, and the discrimination of depth of modulation. We have shown that the amount of MDI decreases when the modulation pattern in different regions of the spectrum are different (e.g., two spectral regions modulated at different rates).

    In order to explain MDI, we feel we must understand how the modulation pattern of a temporally modulated stimulus is extracted. We have been using sinusoidally amplitude-modulated narrow-band noises and complex patterns of modulation of tonal carriers to explore different candidate mechanisms for extracting the temporal pattern of modulation. At present, we find that a transduction-like process, such as suggested by the hair-cell transducer model of Ray Meddis, provides a good way to extract the envelope of temporal modulation. The autocorrelation functions of these temporal envelopes appear (so far) to provide a decent account of the data involving the modulation of narrow-band noises and the complex modulation of tonal carriers.

    Both the empirical results and our theoretical work seem to support the contention that slow temporal modulation of different spectral components can be used by the auditory system to fuse these components into one auditory image. As such, slow temporal modulation is probably one of the cues used by the

AFOSR-89-0335

auditory system to identify sound sources in a complex (multiple sound source) acoustic environment.

Published and Presented Research

Yost, W.A. and Sheft, S. (1989). "Across-critical-band processing of amplitude-modulated tones," Journal of the Acoustical Society of America (85), 848-857.

Yost, W.A., Sheft, S., and Opie, J. (1989)." Modulation interference in detection and discrimination of amplitude modulation," Journal of the Acoustical Society of America (86), 2138-2147.

Yost, W.A. and Sheft, S. (1990). "A comparison among three measures of cross-spectral processing of amplitude modulation with tonal signals," Journal of the Acoustical Society of America, (87), 897-900.

Sheft, S. and Yost, W.A. (1990). "Temporal integration in amplitude modulation detection," Journal of the Acoustical Society of America, in press.

Yost, W.A. (1990). "Thresholds for segregating a narrow band from a broad-band noise," Journal of the Acoustical Society of America, under review.

Sheft, S. and Yost, W.A. (1989). "Spectral fusion based on coherence of amplitude modulation," Journal of the Acoustical Society of America (86), S10-S11.

Yost, W.A. (1989). "Recognition memory for arbitrary, complex waveforms," Journal of the Acoustical Society of America (86), S23.

Yost, W.A. and Sheft, S. (1989). "Temporal modulation transfer functions for pure tones, Journal of the Acoustical Society of America (85), S123.