Naval Submarine Medical Research Laboratory

NSMRL REPORT 1171 **12 JULY 1991**

NARROWBAND AND BROADBAND ENVELOPE CUES FOR AURAL CLASSIFICATION

by
Thomas E. Hanna
and
Yvonne R. Masakowski

Released by:
R. G. Walter, CAPT, DC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS N.A.
2a. SECURITY CLASSIFICATION AUTHORITY N.A. 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE N.A.	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NSMRL Report No.	5. MONITORING ORGANIZATION REPORT NUMBER(S) Same
6a. NAME OF PERFORMING ORGANIZATION Naval Submarine Medical Research Laboratory 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Office of Naval Research
6c. ADDRESS (City, State, and ZIP Code) Naval Submarine Base New London Box 900 Groton, CT 06349-5900	7b. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 22217-5000
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable) Code 1142PS	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-88-WR24003
8c. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 22217-5000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO.: 61153N 42 PROJECT NO.: RR 04209 TASK NO.: RR0420901 WORK UNIT ACCESSION NO.: R&T4424207

11. TITLE (Include Security Classification)
(U) Narrowband and Broadband Envelope Cues for Aural Classification

12. PERSONAL AUTHOR(S)
Thomas E. Hanna and Yvonne R. Masakowski

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
	FROM ______ TO ______	19 JUL 12	14

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Auditory perception; Aural classification; Envelope features; Sonar signals

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Auditory classification of a set of brief sounds was compared to that in a number of conditions that manipulated the spectral and temporal information in the signal. Multidimensional scaling techniques identified the use of six perceptual dimensions that could be accounted for by amplitude-envelopes of critical-band filtered signals. Modulation rates above 200 Hz did not contribute much to classification although there was some suggestion that rates greater than 200 Hz were significant for one of the dimensions. These results indicate the potential importance of envelope information for auditory classification of brief sounds.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL John J. O'Hare	22b. TELEPHONE (Include Area Code) (202) 696-4502	22c. OFFICE SYMBOL Code 1142PS

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

S/N 0102-LF-014-6603

Narrowband and Broadband Envelope Cues
for Aural Classification
by

Thomas E. Hanna
and
Yvonne R. Masakowski

Naval Submarine Medical Research Laboratory
Report 1171

Office of Naval Research
Research Work Unit 61153N-RR4209.001-ONR4424207

Approved and released by

R. G. Walter

R. G. WALTER, CAPT, DC, USN
Commanding Officer
NavSubMedRschLab

SUMMARY PAGE

PROBLEM

To determine whether auditory envelope features used to classify brief sounds are derived from narrowband or broadband analyses of the acoustic signal and the extent to which higher modulation rates contribute to these envelope features.

FINDINGS

The signal envelope could account for all of the perceptual dimensions needed by listeners to classify a set of brief sounds. These perceptual dimensions were better accounted for by a narrowband analysis of the signal than by a broadband analysis. Modulation rates above 200 Hz did not contribute much to classification although there was some suggestion that rates greater than 200 Hz were significant for one of the dimensions.

APPLICATION

These results demonstrate the potential importance of envelope information for auditory classification of brief sounds. The results are consistent with a narrowband analysis of envelope information and the primary importance of modulation rates less than 200 Hz. The results suggest that good identification of brief signals can be achieved with a large reduction in the effective bandwidth of signals by using relatively low modulation rates in a small number of bands, simplifying the problem of automatic classification.

ADMINISTRATIVE INFORMATION

This investigation was conducted under Office of Naval Research Work Unit 61153N-RR04209.001-ONR4424207. The views expressed in this report are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, or the U. S. Government. It was submitted for review on 29 October 1990, approved for publication on 12 July 1991, and has been designated as Naval Submarine Medical Research Laboratory Report No. 1171.

ABSTRACT

Auditory classification of a set of brief sounds was compared to that in a number of conditions that manipulated the spectral and temporal information in the signal. Multidimensional scaling techniques identified the use of six perceptual dimensions that could be accounted for by amplitude-envelopes of critical-band filtered signals. Modulation rates above 200 Hz did not contribute much to classification although there was some suggestion that rates greater than 200 Hz were significant for one of the dimensions. These results indicate the potential importance of envelope information for auditory classification of brief sounds.

INTRODUCTION

A signal's amplitude-time variation, or envelope, contains nonspectral information that can be used by the auditory system to identify that signal. Although there is uncertainty about the mechanisms that underlie the auditory system's processing of envelope information, it is typically assumed that the signal is filtered, and then subjected to a nonlinearity such as a full-wave or half-wave rectification. Additionally, the system's limited ability to follow rapid fluctuations in the resulting signal is described by a central integrator with a time constant.

Such a model can be applied to data for modulation-rate discrimination, modulation detection, and auditory classification, but there is no consistent bandwidth estimate for the initial filter across these tasks. Studies by Buus (1983) and Hanna (1989, 1991) indicate that modulation rate discrimination, where the listener was asked to identify the faster of two modulation rates, show limitations due to relatively narrow (roughly one-third octave) critical-band filtering. In contrast, studies of modulation detection (e.g., Viemeister, 1979; Formby & Muir, 1988) produce bandwidth estimates broader than critical bandwidths. Studies of auditory classification by Van Tassell, Soli, Kirby, & Widin (1987) and Hanna (1990) are equivocal regarding filter bandwidth. They found that the broadband envelope contained sufficient information to convey features of complex signals, supporting the notion of broad initial filtering prior to envelope extraction by the auditory system. However, since Van Tassell, et al. and Hanna never filtered their signals prior to envelope extraction, it is unclear what role critical band filtering may have played in extracting envelope information. If the same modulation pattern was present in each band then the broadband envelope would be the same as those obtained from critical band filtering. In this case, one would expect that the broadband envelope would provide the information needed for classification even though the auditory system is using the narrowband envelopes. On the other hand, if the modulation was different in each band the broadband modulation would not necessarily preserve the perceptual information present in each band. The present study examines classification of the complex signals used by Hanna (1990), but includes conditions with stimulus bandwidths comparable to critical bands.

A second issue concerning the use of envelope cues in classification is the nature of limitations imposed by post-filtering integration. An integration time constant of 2-3 msec provides good descriptions of the data from many experiments. This value suggests that envelope rates greater than 50-80 Hz would be attenuated by post-filtering temporal integration. These effects are generally observed in studies of modulation detection and rate discrimination in noise. However, Buus (1983) found that for high-frequency two-tone complexes, modulation-rate discrimination was relatively unaffected up to rates of 640 Hz. Van Tassell, et al. (1987) suggested that modulation rates above 200 Hz are perceptually significant for classification of speech sounds. In the present study, the envelope is low-pass filtered at several frequencies to determine whether higher frequencies contribute significantly to auditory classification of complex nonspeech sounds.

METHOD

Signal conditions. Twenty-four one-second stimuli were extracted from digitized recordings of underwater sounds. Each stimulus contained an acoustic event with duration ranging from tens to hundreds of milliseconds. These twenty-four signals represented three exemplars from each of eight categories of underwater events. The recordings had been digitized at a 12.5 kHz sampling rate with 12 bits of linear encoding of amplitude.

Subjects were tested on their ability to classify the signals into the eight categories under eleven conditions. In condition 1, listeners classified the original signals. In condition 2 the stimuli were band-pass filtered from 2200 to 2800 Hz; in condition 3, the original signals were band-pass filtered from 710 to 900 Hz. The envelopes of the 2200-2800 Hz band-pass filtered signals (i.e., the envelopes from condition 2) were used to generate stimuli for conditions 4, 5, and 6. In condition 4, the envelopes modulated a 2500-Hz carrier. In condition 5, the envelopes were low-pass filtered at 100 Hz prior to modulating the 2500-Hz carrier. In condition 6, the envelope was low-pass filtered at 10 Hz prior to modulating the 2500-Hz carrier. In condition 7, an 800 Hz carrier was modulated by the envelope of the 710-900 Hz band-pass filtered signals (i.e., the envelopes from condition 3). The envelopes of the original signals (condition 1) were used to generate stimuli in conditions 8, 9 and 10. In condition 8, the envelopes were low-pass filtered at 1 kHz prior to modulating a 2500-Hz carrier. In condition 9, the envelopes were filtered at 100 Hz prior to modulating a 2500-Hz carrier. In condition 10, the envelopes were filtered at 10 Hz prior to modulating a 2500-Hz carrier. In condition 11, the envelope of the original signal was simply used as a time waveform.

Signal generation. In conditions 1-3, the original set of digitized stimuli were presented over 16-bit digital-to-analog converters with a 12.5-kHz sampling rate. Stimuli were band-passed filtered from 2200 to 2800 Hz (condition 2) or 710-900 Hz (condition 3) using a Wavetek filter (model 753A, asymptotic rejection rate of 115 dB/octave).

Envelopes of stimuli from conditions 1-3 were generated using the Interactive Lab System (ILS) from Signal Technology, Inc. The envelopes were presented over 16-bit digital-to-analog converters and low-pass filtered at 5 kHz. The carriers were generated by a Krohn-Hite oscillator (model 4180). The carrier was multiplied by the envelope and low-pass filtered at 5.0 kHz. In conditions requiring filtering of the envelope, the digitized envelopes were convolved with an exponential function (e^{-at}, $t > 0$) with decay constant chosen to produce the desired low-pass filter cutoff of 10, 100, or 1000 Hz (3-dB down points). The filter defined by the exponential function has an asymptotic rejection rate of 6 dB/oct.

In all conditions, a programmable attenuator was used to adjust the amplitude of each signal to a comfortable listening level. In addition, on each trial, stimulus levels were randomized over a 15-dB range to minimize the use of amplitude as a classification cue. An electronic switch gated the stimuli with 20-msec sine-squared ramping. Stimuli were presented to the right earphone of a Sennheiser HD430 headset.

Procedure. Initial training was conducted with reduced signal sets to facilitate the learning of the category labels. Blocks of trials using exemplars from categories one through four were run followed by blocks using exemplars from categories five through eight. On each trial, one stimulus from the original set was presented and the listener classified it into one of four categories. Feedback was given on each trial by displaying the correct response on the screen. Listeners had 720 trials by which time they were performing quite well (approximately 85% correct) with these reduced signal sets.

After initial training, data were collected over a three day period for each condition. Day one was a familiarization session and these data were not used in the analyses. Days 2 and 3 were testing days for each condition. Each of the twenty-four stimuli were presented twice within each block of forty-eight trials. Eighteen blocks of trials were run each day. Feedback was given on each trial by displaying the correct response on the screen. All listeners were tested simultaneously which prevented counterbalancing the conditions across listeners.

Listeners. Three paid volunteers served as listeners. Each had normal hearing sensitivity (less than 15 dB HL at octave frequencies from 250 to 8000 Hz). Two of the subjects had participated in a previous experiment using similar stimuli. The third subject had never heard these sounds prior to this experiment.

RESULTS & DISCUSSION

Figure 1 shows the proportion correct in each of the eleven conditions averaged across listeners. A repeated-measures analysis of variance indicated a significant effect of condition ($p. < .001$).

Critical-band information. Performance on the original signals (condition 1) was quite

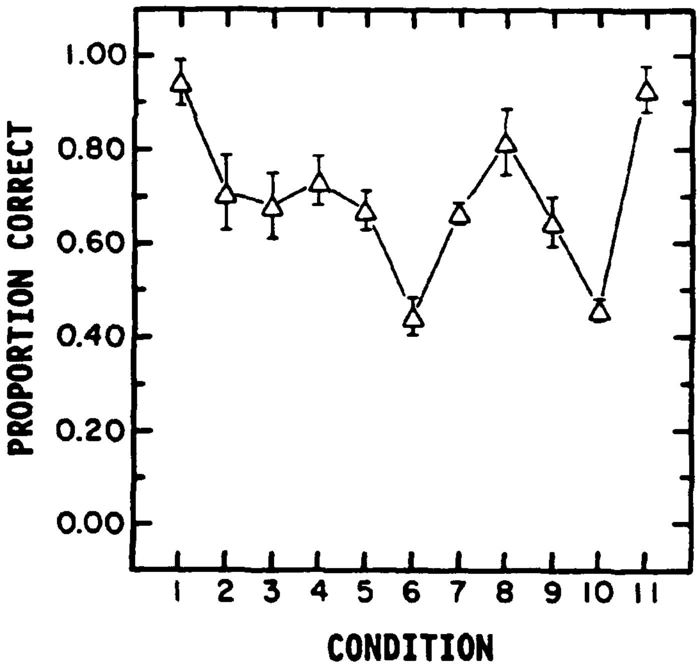

Figure 1. Proportion correct, averaged across the three listeners, for the eleven conditions. Error bars represent standard deviations across the three listeners.

good (94%). Proportions correct on the filtered signals of conditions 2 and 3 were significantly different than condition 1. Filtering the signals from 2200-2800 Hz (condition 2) decreased performance to 71% ($p < .001$); filtering from 710-900 Hz (condition 3) decreased performance to 68% ($p = .002$). The conditions that used the envelopes of the signals from these two conditions as modulators (conditions 4 and 7) did not produce a significant difference in proportions correct compared with that on the corresponding filtered signals (73% and 67%, respectively), suggesting that the envelope adequately represents the perceptual information within each critical band. Filtering the envelope of the high-frequency band at 100 Hz prior to modulation of the carrier (condition 5) also had little effect on performance (67%) compared with the unfiltered envelope condition ($p > .05$). Filtering the envelope at 10 Hz (condition 6) did produce a significant decrease to 44% ($p = .008$). *Given the shape of* the filters used to filter the envelope, these results suggest that envelope frequencies above 200 Hz were not perceptually important for the high-frequency-band filtered signals. However, perceptually important information did exist between 20 and 200 Hz and below 20 Hz.

Broadband information. Using the broadband envelope, filtered at 1 kHz (condition 8), decreased performance to 82%, which was significantly different than that with the original signals ($p = .007$). Although all three listeners showed progressively lower scores as the envelope was filtered at 100 Hz (64%) and at 10 Hz (46%), conditions 9 and 10, respectively, these successive comparisons were not statistically significant. Nonetheless, the size of these effects and their consistency across subjects suggest that modulation rates greater than 200 Hz, from 20-200 Hz, and less than 20 Hz are used for classifying these signals.

Listeners' performance in the condition with the broadband envelope filtered at 100 Hz was similar to that for conditions that used the envelopes of the 710-900 Hz and 2200-2800 Hz bands (64% vs. 67% for both conditions 5 and 7). Also, filtering the envelope at 10 Hz had a similar effect whether the envelope came from the broadband signal or the signal filtered from 2200-2800 Hz (46% vs. 44%). This similarity in performance across conditions that used different filtering suggests that the broadband envelope may simply reflect signal features that are present in narrower bands. The similarity of features present in each of these conditions will be explored further in the next section by examining the pattern of errors made in each of these conditions. A finding that similar errors are made in each condition would support the conclusion that these signals have similar modulation patterns in each frequency band and that the apparent efficacy of the broadband envelope derives from its similarity to the narrowband envelopes. The finding of different patterns of errors across conditions would indicate the use of different cues in these conditions that may or may not correspond to the cues used to classify the original unaltered signals. In this latter case, the similarity in overall performance would indicate that the broadband envelope conditions were not fully conveying the narrowband cues, since then the broadband conditions should be better than comparable narrowband conditions.

The consistent, but statistically nonsignificant, decrease in performance when broadband signal envelopes were filtered at 100 Hz, leaves open the possibility that envelope fluctuations greater than 200 Hz play a role for classification of these sounds.

The condition where the broadband envelope was played as a time waveform (condition 11) yielded 93% correct, which was not significantly different than performance with the original signals. It is curious, and perhaps suspicious, that the envelope by itself (condition 11) can be classified as well as the original signals. In fact, the envelope by itself sounds remarkably similar to the original signals. This result would suggest that the envelope contains all of the information needed to classify these signals. However, filtering the envelope at 1 kHz and using it to modulate a tonal carrier (condition 8) resulted in significantly poorer performance, implying that envelope frequencies above 1 kHz contain usable envelope information. Such an upper limit exceeds that suggested anywhere. Examination of the broadband envelope led to another possible explanation - the envelope begins to resemble a full-wave rectification of the signal. The broadband envelope thus begins to follow the fine structure of the waveform and may be providing spectral cues from the signal.

Multidimensional scaling. The data were examined in greater detail using the patterns of errors. Conditions that yield similar percents correct (e.g., conditions 2, 3, 4, 5, 7, and 9) may nonetheless have quite different signal features. This would be supported by finding different patterns of errors across conditions. Conversely, if the signal features are similar across conditions, then the patterns of errors should also be similar.

Multidimensional scaling (MDS) techniques were used to simplify the analyses of error patterns. This analysis uses a measure of similarity between signal pairs to determine a multidimensional stimulus representation. For the present study, the measure of similarity was defined as:

$$s(i,j) = \sum_{k=1}^{8} p(i,k)p(j,k) \quad i=1,24 \quad j=1,24 \quad (1)$$

where p(i,k) is the probability that stimulus i will be called category k. For a given pair of signals, this number is the probability that the two signals would be given the same category label assuming independent responses to each.

A similarity matrix (lower half, diagonal absent) was computed for each of the three listeners in each of the eleven conditions. The SINDSCAL procedure (Carroll & Chang, 1970; Kruskal & Wish, 1978) was used to scale the 33 matrices (3 subjects x 11 conditions). This technique produces a single multidimensional representation of the 24 signals. Each dimension in the signal space is taken to represent a stimulus feature that distinguishes the signals from each other. Within the SINDSCAL analysis, the similarity of two signals is predicted by Euclidean distance of signals in the multidimensional space. Differences across the 33 similarity matrices are accounted for by a weighting vector for each of the 33 matrices that gives different weights to each of the dimensions. The weighting vectors reflect the fact that these dimensions may be present to varying degree for individual listeners or in the different stimulus conditions. Conditions (or listeners) with different patterns of errors would indicate the use of different stimulus features and therefore different weighting vectors. A low weight given to a dimension would suggest the loss of a significant stimulus feature resulting from the signal manipulation used for that condition. Thus the MDS results can be used to make comparisons across conditions to determine whether common features are present in each.

SINDSCAL representations were generated for eight dimensions and less. Variance accounted for (VAF) by each solution was compared to determine the appropriate dimensionality. Starting from two dimensions, VAF increased by 7-11% with each added dimension up to six. Inclusion of additional dimensions increased VAF by only 2-3%. Thus, a six-dimensional solution, accounting for 8C% of the variance, was used for further analyses. This solution uses 342 parameters (3 subjects x 11 conditions x 6 dimension weights + 24 stimuli x 6 dimension values) to account for 6336 data values (3 subjects x 11 conditions x 24 stimuli x 8 response categories). Weight vectors were generally similar across listeners (average s.d. = .03), so results were averaged across listeners to simplify further analyses.

MDS Analysis of Narrowband Conditions. Figure 2 shows the weight vector coefficients for the original signals (condition 1, shown as triangles) and the four conditions derived from the 2200-2800 Hz filtered condition (condition 2, squares; condition 4, circles;

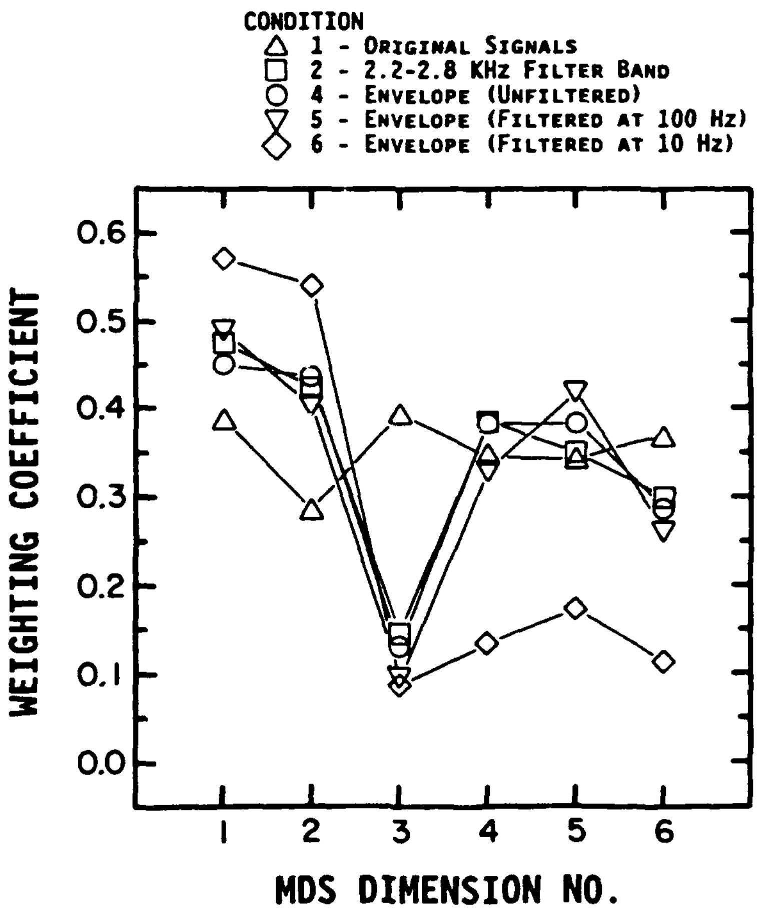

Figure 2. Weighting coefficients, averaged across listeners, for the six dimensions obtained from the SINDSCAL analysis. The parameter is the stimulis condition - condition 1 Δ, condition 2 □, condition 4 O, condition 5 ∇, and condition 6 ◊.

condition 5, downward triangles; and condition 6, diamonds). For the original signals, the weights are roughly equal across all six dimensions. In the other four conditions, two clear effects are observed. First, a very low weight is given to Dimension 3 (D3) for all, indicating that the 2200-2800 Hz band did not convey the signal feature associated with D3. Second, the weights for D4, D5, and D6 are relatively unaffected by low-pass filtering the envelope at 100 Hz (condition 5), but are greatly reduced when the envelope is filtered at 10 Hz (condition 6). This suggests that the signal features associated with these dimensions (D4, D5, and D6) are conveyed by modulation rates in the general range from 20-200 Hz. D1 and D2, which receive slightly greater emphasis as the other features are removed, are still present with only low modulation rates in the signals (less than 20 Hz). Moreover, it should be noted that the filtered band stimuli (condition 2) produce very similar weights as stimuli generated from the envelope of that band (condition 4). The differences in dimension weightings across conditions 1, 2, 4, 5, and 6 correspond well to the changes found in proportion correct shown in Fig. 1 in that conditions that are not significantly different in percent correct also show similar dimension weights and conditions that are significantly different in percent correct have different weights on at least one dimension.

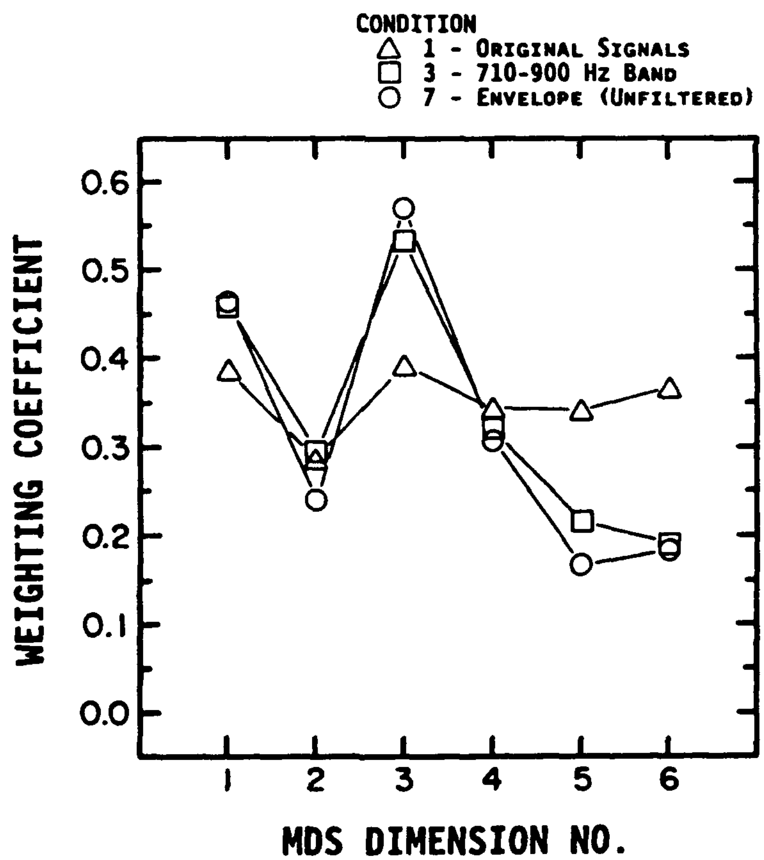

Figure 3. Weighting coefficients, averaged across listeners, for the six dimensions obtained from the SINDSCAL analysis. The parameter is the stimulus condition - condition 1 Δ, condition 3 □, and condition 7 O.

Figure 3 shows the weight vector coefficients for the original signals (condition 1, triangles) and the two conditions derived from the 710-900 Hz filtered condition (condition 3, squares; and condition 7, circles). Compared with the relatively equal weighting across dimensions for the original signals, the 710-900 Hz filtered signals show a stronger emphasis for D3 and a low weight for D5 and D6. The weights are very similar for filtered signals (condition 3) and stimuli generated from the envelope of that band (condition 7).

Figures 1, 2, and 3 taken together suggest some conclusions. Stimuli filtered to a critical bandwidth are classified in a very similar way as stimuli created by modulating a tone by the envelope of that filtered band. Thus, the envelope of critical band signals conveys much of the perceptual information of that signal. All six dimensions are present in at least one of the critical-band envelope conditions. D1, D2, and D4 are envelope features that are present in both a high-frequency and a low-frequency band and thus may be considered to be present broadband. D3 is an envelope feature that is only present in the low band and D5 and D6 are envelope features that are only present in the high band. Furthermore, percents correct (Fig. 1) and dimension weights (Fig. 2) were unaffected by low-pass filtering the envelope at 100 Hz (condition 5). It appears that classification can be accounted for by narrowband envelope cues less than 200 Hz.

<u>MDS Analyses of the Broadband Conditions</u>. Figure 4 shows the weight vector coefficients for the original signals (condition 1, triangles) and three conditions derived from the broadband envelope (condition 8, squares; condition 9, circles; and condition 10, downward triangles). Condition 8, which used the broadband envelope filtered at 1 kHz to modulate a 2500 Hz tone, contained all six dimensions, but to varying degrees. D1, D2, and D4 had slightly higher weights than for the original signals. This result is consistent with the finding that these three dimensions were envelope features that are probably present in each of several critical bands. Therefore it is expected that this information should be clearly present in the broadband envelope. On the other hand, D3, D5, and D6 were present in only one of the two bands studied. The strength of these features would be lessened in a broadband envelope. In fact the weights shown in Figure 4 support such a view because the weights for condition 8 are less than for the original signals. These dimensions, although reduced in weight, are still present to a significant degree, which may account for the relatively good performance in this condition (82% correct) - since both low-band and high-band features are available performance is expected to be better than the narrowband condition.

The reduced strength of D3, D5, and D6 in the broadband envelope condition would also explain the finding that performance found in condition 8, although high, is still significantly poorer than for the original signals. Since all six dimensions are represented in the broadband condition, one must assume that some of them are not as strong otherwise performance would be as good as for the original signals. Furthermore, the fact that performance when the broadband envelope was filtered at 100 Hz (condition 9) was no better than the narrowband envelope conditions with comparable limits on modulation rate (conditions 5 and 7) also suggests that the distinct cues present in the narrowband envelopes are not as clearly perceptible as in the broadband envelope. The results of Fig. 4 are also generally consistent with those for conditions 8, 9, and 10 of Fig. 1. The small but consistent changes in percent correct shown in Fig. 1 are accompanied by only slight changes in a single dimension in Fig. 4. These results are in contrast to those for the narrowband condition, where larger and statistically significant changes in percent correct are associated with more pronounced changes in dimension weights. Thus alterations of the broadband envelope are more weakly related to the perceptual dimensions than are changes in the narrowband envelope. For these reasons, and given the finding that the weights of D3, D5, and D6 were higher in the narrowband envelope conditions than in the broadband envelope conditions and more closely resembled the weights for the original signals, it is probable that the auditory system is normally using separate envelope information from a number of bands in spite of the good performance observed using the broadband envelope.

Figure 4 also shows that D3 is still present even with envelope filtering at 10 Hz (condition 10) suggesting that this is a low-modulation-rate (less than 20 Hz) cue present at lower spectral frequencies. D4, D5, and D6 are all relatively reduced with filtering of the envelope at 10 Hz, which agrees for the most part with the results with the 2200-2800 Hz band (Fig. 2). For D5, the filtering effect is not as pronounced for broadband envelope (Fig. 4) as for the high-frequency-band envelope

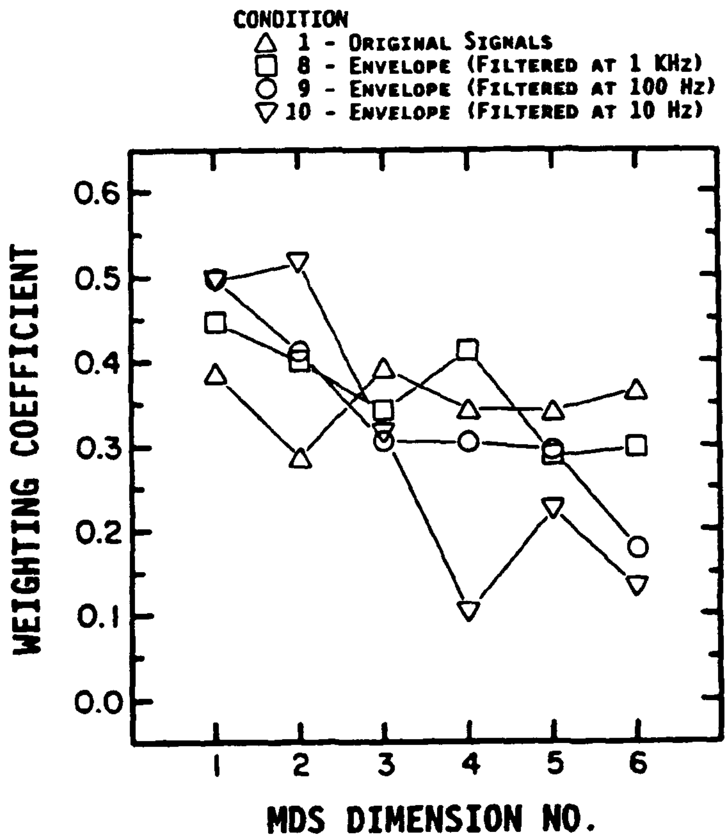

Figure 4. Weighting coefficients, averaged across listeners, for the six dimensions obtained from the SINDSCAL analysis. The parameter is the stimulus condition - condition 1 Δ, condition 8 □, condition 9 O, and condition 10 ∇.

(Fig. 2). For D6, the main reduction is observed as the envelope filter is changed from 1 kHz (condition 8) to 100 Hz (condition 9). This result may indicate a high-modulation-rate cue that was not observed in the 2200-2800 Hz conditions because this band was not sufficiently broad to support that rate of modulation. Additional conditions with somewhat broader bands or bands that more closely resemble auditory filters, which do not have the extremely sharp cutoffs used in thee present study, could determine whether these modulation rates play a significant role.

The weights for condition 11, which used the broadband envelopes as stimuli, were very similar to those for the original signals. As already stated, this result may reflect the carry-over of spectral information to the broadband envelope.

Implications for automatic classification. The finding that eight stimulus categories can be described by six perceptual dimensions or features does not seem like a parsimonious description of the data. However, one must keep in mind that these dimensions are describing 11 sets of modified stimuli. From the dimension weights we can associate specific perceptual features with the different acoustic information present in each stimulus set. Thus, each feature can be related to a range of modulation rates present in a specific frequency band (or bands). Of the six dimensions, D1, D2, and D3 depend on very low modulation rates (<20 Hz), whereas D4, D5, and D6 use intermediate rates (20-200 Hz). D3 stems from modulation patterns present at lower spectral frequencies, D5 and D6 from modulation patterns present at higher spectral frequencies, and D1, D2, and D4 from modulation patterns that are present across a broadband of frequencies. The training of automatic classification algorithms based on this specific acoustic information would be more successful than those based on the original signals for two reasons. First, algorithm input based on the specific acoustic information has a much reduced bandwidth compared with the original signals (20-200 Hz vs 5000 Hz), which improves the algorithm's performance. Second, the relevant features should be more distinct in the modified stimuli thus making it easier for the algorithm to distinguish the category features.

SUMMARY

This study provides a more detailed analysis of the importance of envelope features for aural classification than previously reported by Van Tassell, et al. (1987) and Hanna (1990). These two previous studies found evidence that the broadband envelope conveys important perceptual features. The present study found, for the brief sounds used here, that envelope features accounted for all of the features identified. In particular, classification performance, as measured by percent correct and the results of multidimensional scaling, show no difference between conditions with third-octave-band filtering and corresponding conditions using the envelope of these filtered bands to modulate a tonal carrier. Although the results are not definitive, the classification features were more clearly associated with the envelopes of critical-band filtered stimuli than of the broadband stimuli. These features were well accounted for by modulation rates less than 200 Hz, although there was some suggestion that rates greater than 200 Hz were significant for one of the features.

REFERENCES

Buus, S. (1983). Discrimination of envelope frequency, Journal of the Acoustical Society of America 74, 1709-1715.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition, Psychometrika 35, 283-319.

Formby, C., & Muir, K. (1988). Modulation and gap detection for broadband and filtered noise signals, Journal of the Acoustical Society of America 84, 545-550.

Hanna, T. E. (1989). Modulation-rate perception: Identification and discrimination of modulation rate using a noise carrier, Naval Submarine Medical Research Laboratory Technical Report No. 1128.

Hanna, T. E. (1990). Contributions of envelope information to classification of brief sounds, Naval Submarine Medical Research Laboratory Technical Report No. 1165.

Hanna, T. E. (1991). Discrimination and identification of modulation rate using a noise carrier, under review.

Kruskal, J. B., & Wish, M. (1978). Multidimensional Scaling. Eleventh in the Series: Quantitative Applications in the Social Sciences (E. M. Uslaner, Ed.), Sage Publications: Beverly Hills, CA.

Van Tassell, D. J., Soli, S. D., Kirby, V. M., & Widin, G. P. (1987). Speech waveform envelope cues for consonant recognition, Journal of the Acoustical Society of America 82, 1152-1161.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds, Journal of the Acoustical Society of America 66, 1364-1380.

OFFICE OF NAVAL RESEARCH PERCEPTUAL SCIENCE PROGRAM

Technical Reports Distribution List

Dr. Earl A. Alluisi
OUSDR (A) / R&A (E&LS)
Pentagon 3D129
Washington, DC 20301-3080

Prof. James A. Ballas
Department of Psychology
George Mason University
4400 University Drive
Fairfax, VA 22030

Mr. Charles Bates
Director, HE Division
USAFAAMRL/HE
WPAFB, OH 45433

Dr. Kenneth R. Boff
AAMRL/HE
WPAFB, OH 45433

Mr. Luiz Cabral
Naval Underwater Sys. Ctr. (Code 2212)
Bldg. 1171/1
Newport, RI 02841

Dr. Stanley C. Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Defense Tech. Info. Center (2 copies)
Cameron Station, Bldg 5
Alexandria, VA 22314

Dr. Robert A. Fleming
Naval Ocean Systems Center
Human Factors Support Grp.
1411 South Fern Street
Arlington, VA 22202-2896

Mr. Dan Greenwood
Netrologic, Inc.
5080 Shoreham Place (Suite 201)
San Diego, CA 92122

Jeffrey Grossman
Naval Ocean Systems Center
Code 4403, Bldg. 344
San Diego, CA 92152-6800

Prof. James H. Howard, Jr.
Department of Psychology
Human Performance Lab
Catholic University
Washington, DC 20064

Dr. Edgar M. Johnson
Technical Director
US Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

CAPT T. E. Jones, MSC, USN
Aviation Medicine & Human
Performance (Code 404)
Naval Medical R&D Cmd
National Capital Region
Bethesda, MD 21814-5044

Dr. Michael Kaplan
Director, Office Basic Res
US Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Richard Loutitt
National Science Foundation
Division of Behavioral & Neural Sciences
1800 G Street, N.'V.
Washington, DC 20550

CAPT William Moroney, USN
Naval Air Development Ctr. (Code 602)
Warminster, PA 18974

Dean of the Academic Dept.
U.S. Naval Academy
Annapolis, MD 21402-5018

Naval Aerospace Medical Research
Laboratory, Sensory Division Code 23
Pensacola, FL 32508

Commanding Officer
Naval Air Systems Command
Crew Station Design, NAVAIR 5313
Washington, DC 20361

Commanding Officer
Naval Biodynamics Lab
Michoud Station, Box 29407
New Orleans, LA 70189

Commanding Officer
Navy Health Research Ctr.
P.O. Box 85122
San Diego, CA 92138

Commanding Officer
Navy Personnel R&D Center
San Diego, CA 92152-6800

Director
Technical Information Div. (Code 2627)
Naval Research Laboratory
Washington, DC 20375-5000

US Naval Test Center
Aircrew Systems Branch
Systems Engineering Test Directorate
Patuxent River, MD 20670

Commanding Officer
Naval Weapons Center
Human Factors Branch (Code 3152)
China Lake, CA 93555

Dr. Michael Letsky
Office of the Chief of Naval Operations
(OP-01B7)
Washington, DC 20350

Office of the Chief of Naval Operations
OP-933D3
Washington, DC 20350-2000

Dr. Milton A. Whitcomb
NAS-National Res. Council (CHABA)
2101 Constitution Ave., N.W.
Washington, DC 20418

Office of Naval Research
Special Asst. for Marine Corps Matters
Code 00MC
800 N. Quincy Street
Arlington, VA 22217-5000

Office of Naval Research
Perceptual Science Program
Code 1142PS
800 N. Quincy Street
Arlington, VA 22217-5000
(3 copies)

LT Dennis McBride, USN
Pacific Missile Test Center
Human Factors Branch
Pt. Mugu, CA 93042

Dr. W.A. Rizzo
Head, Human Fctors Div.
Naval Training Systems Ctr
12350 Research Parkway
Orlando, FL 32826-3224

Prof. James A. Simmons
Department of Psychology
Brown University
Providence, RI 12912

Mr. Nelson F. Steele
Advanced Res Dev Corp
5457 Twin Knolls Road
Columbia, MD 21035

Dr. John Tangney
AF Office of Sci. Res.
Life Sciences Directorate, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Charles S. Weaver
MAXIM Technologies, Inc.
3000 Patrick Henry Drive
Santa Clara, CA 95054

Dr. John Weisz
Technical Director
US Army Human Engineering
Laboratory
Aberdeen Proving Ground, MD