2

Annual Report to

Office of Naval Research
Bioacoustic Signal Classification
Cognitive and Neural Sciences
& Mathematics Division

for the period
11/1/90 - 10/30/91

Classification of Complex Sounds
#ONR N00014-91-J-1122

Bruce G. Berg
Principle Investigator
Department of Cognitive Sciences
School of Social Sciences
University of California, Irvine
Irvine CA 92717

WORK ACCOMPLISHED

The work accomplished during the period, 11/1/90 - 10/30/91, involves the use of COSS analysis to estimate weights in various profile analysis tasks (see Berg, 1989; Berg and Green, 1990). This technique is a method for investigating how spectral information is used by listeners to discriminate complex sounds. In essence, spectral weights are estimated which provide a detailed, "microscopic" assessment of the relative influence of different spectral components on a listener's decisions. Specific studies include estimation of spectral weights using (1) broad band, "rippled" spectra, (2) narrow band spectra, and (3) simultaneous estimation of temporal and spectral weights. An overview of major findings is given below.

Discrimination of broad band, "rippled spectra

A profile-type task is essentially an empirical method for investigating how spectral information from different auditory channels is used to discriminate complex sounds. In a typical profile task, the standard consists of n-tones of equal intensity and the signal consists of an increment in the intensity of a single tone of the complex. Since overall level is varied randomly on each stimulus presentation (over a range of 20 dB or greater), absolute intensity is not a useful cue. Rather, the level of the signal component *relative* to the levels of the nonsignal components is most relevant cue (see Green, 1988). Since the tones are well-separated with respect to frequency, the task involves across-channel, level comparisons--the standard and signal-plus-standard are presumably discriminated on the basis of differences in spectral shape (e.g. flat spectrum vs. spectral "bump"). Combined with the COSS technique for estimating spectral weights, profile analysis is a useful method for investigating listeners' abilities to combine or otherwise compare information across different auditory channels in order to discriminate complex sounds.

The first set of experiments examined listeners' abilities to discriminate a flat-spectrum (eight tones of equal intensity) from a "rippled" spectra (consisting of a pattern of intensity changes across the entire spectrum). The spectral changes can be characterized as "global" (e.g. increasing the relative intensity of the low frequency components and decreasing the relative intensity of the high frequency components) or "local" (e.g. alternately increasing and decreasing the intensities of adjacent tones, thus producing a "saw tooth" shaped spectrum). Spectral weight estimates show that listeners are more efficient (relative to optimal weights) at integrating information across auditory channels when global spectral changes are made than they are at integrating information when many local changes are made. For local spectral changes, listeners tend to base their discriminations on only a limited, narrow portion of the spectrum (two or three components), largely ignoring the spectral differences occurring at other locations [see Berg and Green (1991a) for a detailed report]. This work suggests that listeners may attend to only a narrow band of frequencies in order to discriminate complex, "naturalistic" stimuli. We return to a discussion of this issue below (**Future Work** section).

Rippled spectra were also used to examine the effects of training. Spectral weights were estimated following a typical training period (about 4000 trials) and again after listeners had received an additional 10,000 trials (approximately ten hours of extended training.) Results, which are shown in Fig. 1, are mixed. The performance of two observers improved following extended training, with one showing an increase in weighting efficiency of about 50% and the other showing a ten-fold increase in weighting efficiency [see Berg (1990) for a discussion of efficiency measures.] The results provide an important demonstration that listeners are able to adjust their use of information from different auditory channels when given extended practice with feedback. Moreover, the results demonstrate that the COSS technique provides a powerful tool for monitoring an individual's ability to make optimal use of information from different auditory channels. Spectral weights also provide detailed information about individual differences. For example, "good" and "poor" profile listeners can be easily identified by the pattern of their estimated spectral weights. Berg and Green (1990) show that the estimated spectral weights for "good" profile listeners are close to optimal, whereas the weights for "poor" profile listeners are quite divergent from optimal.

Discrimination of narrow band spectra

A channels model proposed by Durlach, Braida, Ito (1986) is used to derive optimal weights in profile tasks. To test this model further, we examined performance (i.e. thresholds for an increment in the intensity of the central tone of a three tone complex) as a function of stimulus bandwidth. The bandwidth of the three-tone complex ranged from very broad (200 Hz to 5000 Hz) to very narrow (990 Hz to 1010 Hz). The signal component was 1000 Hz for all conditions. According to the channels model, thresholds should be very high when the bandwidth is very narrow (less then a critical band). First, the two nonsignal components should mask the small increment added to the signal component. Second, since all the components are contained within a single critical band, across channel comparisons cannot be made, and the model predicts that only an absolute intensity cue will remain. Reliance on absolute intensity should lead to very poor performance, because of the 20 dB range produced by the roving level procedure. A second prediction is that the weights for the three components will have roughly the same sign and magnitude, because tones are presumably unresolved within a critical band. Surprisingly, neither of these predictions is correct. Thresholds for narrow band profiles are about the same as thresholds for broad band profiles, and spectral weights for the narrowest bandwidth are essentially identical to those found for the widest bandwidth. This last finding is particularly intriguing, since it shows that the nonsignal components and the signal component have different and opposite effects on listeners' responses, *even though the tones are presumably unresolved within a critical band*.

One of the major accomplishments of this work thus far is that we have gained some understanding of the unexpected results described above. Subsequent experiments show that the perceptual cues used by listeners to discriminate complex sounds are dependent on spectral bandwidth. For broadband stimuli, across channel level comparisons are made (Berg and Green, 1990). For bandwidths spanning about two critical bands, pitch cues appear to be important (Berg, Quang, and Green, 1991). A modified version of Feth's (1974) EWAIF model (a theoretical calculation of spectral pitch) accounts for both thresholds and spectral weights in this case. For bandwidths less than a critical band, we believe that discriminations are based on envelope cues. Berg and Green (1991b) proposed a model for which the computation of a decision statistic is based on the spectrum of the envelope.

COSS analysis has been quite useful in this investigation. For one, the pattern of estimated spectral weights for the three-tone, profile task are profoundly dependent on bandwidth. When the bandwidth spans about two critical bands, the magnitude of weight for one of the nonsignal components is *greater* than the weight for the signal component, and the weight for the other nonsignal component has the same sign and roughly the same magnitude as the signal component. This contrasts with findings for broadband profiles where the two nonsignal components have weights roughly equal to -0.5 and the signal component has a weight of unity. When the bandwidth is less than a critical band, the pattern of weights are again similar to the weights for broadband stimuli. These differences in weight patterns provided the initial evidence that perceptual cues change as a function of bandwidth, making obvious the merits of COSS analysis. COSS analysis also provides a rigorous test of models. Since weight estimates are simply a means of summarizing atheoretical, behavioral data, plausible models of hearing processes should account for spectral weight estimates as well as thresholds. Taken together, the three models (channels model, modified EWAIF model, and envelope spectrum model) provide a good account of the data, predicting thresholds, spectral weights, and, in most cases, listeners' subjective reports.

This work has one far ranging implication--it illustrates that a host of perceptual cues are available to listeners for discriminating complex sounds, and that listeners are quite adaptable and able to adjust to changing listening situations. Models based on the integration of spectral energy within critical bands have been a mainstay, and to some degree, a unifying principle of signal detection theory as applied to hearing. The unexpected results of the current experiments show the need for additional computational models based on different fundamental assumptions. For one, our findings illustrate that the information obtained from a single critical band is not limited to the integration of spectral energy. Specifically, it appears that a temporal envelope can be extracted from the "output" of a single channel, and that the power spectrum of this envelope can be used to discriminate narrow band spectra. One advantage of such a model is that it is independent of source distance or absolute level. Thus, the model can account for the findings of Gilkey (1987) and Kidd, at al. (1989) that the detection of a tone in noise is largely unaffected by a roving level procedure.

Spectral-temporal weights

Preliminary work has been done in which COSS analysis is used to estimate "spectral-temporal" weights in a profile task. Whereas spectral weights quantify the relative contribution of different spectral components to a listener's decisions, temporal weights quantify the relative contribution of different temporal segments. Taken together, a spectral-temporal weight quantifies the relative contribution of a specified component during a specified temporal segment. In collaboration with Dr. Haunping Dai (a postdoc in David Green's lab), we have estimated spectral-temporal weights in a three-tone profile task. Some of the major findings from this study are listed below.

1) When the signal (e.g. intensity increment of a central, 1000 Hz tone) is presented during the full duration of the stimulus, optimal weights are the same for each temporal segment of the stimulus. Spectral-temporal weight estimates for listeners, unlike optimal weights, are not the same across different temporal segments. Typically, the largest weight is found for either the first or last temporal segment, whereas other segments have relatively small weights, often close to zero--information from these temporal segments is not used optimally.

2) When the signal is presented for only a portion of the full duration (e.g. during the last 100 ms of a 300-ms stimulus), a COSS analysis shows that listeners are able to "track" its location and give greater weight to the temporal segment containing the signal. In other words, listeners are able to isolate a 100-ms signal within the 300-ms stimulus. For a 15-ms stimulus duration, however, most listener's temporal-spectral weights remain unchanged when the temporal location of a 5-ms signal is changed. That is, listeners cannot isolate the 5-ms signal within the 15-ms stimulus, and appear to nonoptimally integrate information over the entire duration of the stimulus. The

results are consistent with current theories which posit a 100 ms integration time.

3) In theory, information across different temporal intervals is integrated in a linear fashion, that is, the information acquired from the full stimulus duration should be equivalent to a weighted sum of information acquired from different temporal segments. This assumption was confirmed empirically.

A primary motivation for estimating spectral-temporal weights in a three-tone profile task was simply to explore the viability of this procedure. By this criterion, the preliminary work was successful. We found that the temporal-spectral weight estimates are reliable (replications separated by one month yield very similar results) and that the results are, for the most part, consistent with our ideas about temporal integration.

FUTURE WORK

The general plan of research has changed little from that discussed in the initial proposal. In fact, the work completed thus far actually sharpens the goals of that initial proposal in a number of ways. We have found that spectral weight estimates provide insights about the perceptual cues that listeners use to discriminate complex stimuli, and also provide a unique method for testing computational models of hearing processes. Several new models have been proposed as a direct result of this work. COSS analysis also provides a means for monitoring performance, a method for investigating the effects of training, and an increased understanding of individual differences in performance. Finally, we have demonstrated the viability of estimating weights concurrently in the temporal and spectral domains. We believe that some degree of "closure" has been gained for a number of these issues, particularly for the discrimination of narrow band stimuli. In a very real sense, the first stage of proposed research has been completed.

Here, a brief listing is given of the work to be completed.

1) Generalize the COSS analysis to "naturalistic" stimuli. The greatest proportion of current work is directed towards this problem. The stimuli consist of digitized dolphin echolocation calls (generously provided by Dr. Whitlow Au). After scaling these calls to a range audible to humans, Au shows that human listeners can discriminate the broad band calls with roughly the same accuracy as dolphins (Au and Martin, 1989). We will use COSS procedures and analysis to investigate whether discriminations are based on a narrow frequency region(s), as suggested by the profile study of Berg and Green (1991a), or whether information is acquired across the entire spectrum. This study will also show whether or not COSS analysis is viable and useful in

listening tasks involving extremely complex spectra.

2) Generalize the COSS analysis to categorization of complex sounds. If the stimulus set consist of m different sounds, then (m-1) sets of weights are required in order to optimally categorize (or identify) the sounds. This implies that if the computational mechanisms used to *identify* a sound are similar to those used to *discriminate* sounds, then we should find evidence of more than one set of weights for listeners in a classification task.

3) Further development and refinement of COSS analysis. COSS analysis requires a large amount of data. A set of weight estimates requires between two to eight hours of listening time, depending on the specific task. Efforts to reduce this time will proceed in two directions. One is to test different experimental procedures (e.g. increasing the strength of the signal). A second is the systematic examination of a number of arbitrary parameters that are used in the 'construction' of empirical COSS functions (e.g. number of points used to define the empirical function). This can be done by both a reanalysis of existing data and the use of computer simulations. In short, we need to map out the parameter space so that judicious decisions can be made with regards to the tradeoff between the amount of data and the accuracy or reliability of weight estimates.

4) Further development of COSS theory. The theory of COSS analysis is relatively new, and we have thus far been most concerned with empirical issues. However, there may be considerably more information inherent in COSS functions than was originally perceived. Currently, I have been working on several problems with Dr. Robert Lutfi at the University of Wisconsin. We hope to gain new insights and a better understanding of the theoretical implications of COSS analysis.

5) Training of listeners. Obviously, performance is a function of how well listeners integrate spectral information. As we have shown, one difference between "good" and "poor" listeners is the manner in which they use spectral information. We have also shown that as a listener's performance improves with training, estimated weights become more similar to optimal weights. Thus far, only "passive" training in the form of feedback has been used. It seems that with knowledge about those aspects of a stimulus to which a listener attends (obtained with COSS analysis), we should be able to employ a more "active" training procedure, possibly by enhancing or "tailoring" the stimulus in order to redirect a listener's attention. Such a strategy was used successfully by Leek and Watson (1984).

GENERAL INFORMATION

I have recently transferred from the Psychoacoustic Laboratory at the University of Florida, where I was an Assistant Research Scientist working with Dr. David Green, to the University of California, Irvine, where I have a tenure-track position as an Assistant Professor in the Cognitive Science Department. Since arriving at Irvine, I have concentrated on building a new psychoacoustic laboratory. Due to a fault by the vendor, completion of the laboratory was delayed slightly because two microcomputers (IBM 486-33) had to be reordered following a six-week waiting period. Currently, the lab is fully equipped and operating. Some preliminary data is being collected, primarily to train laboratory personnel and to provide a final "debugging" check on new hardware and software. Paid listeners will be recruited by the first of the year.

An extremely competent undergraduate student, Deirdre McCarney, has been hired as a technical assistant. A second-year graduate student, Kurt Southworth, has expressed an interest in working in the laboratory. It is highly likely that he will join the laboratory by the first of the new year. Also, several of the Senior Faculty in the department, Dr. Jean-Claude Falmagne, Dr. R. Duncan Luce, and Dr. Geoff Iverson, have expressed an interest in a weekly symposium focusing on some of the issues discussed in this report.

I continue to consult with David Green about all aspects of the work. During the month of September, I returned to Florida for a three-week period to work in his laboratory and to complete several projects. In December, I will return to Florida for another three week period in order to collaborate on ongoing research.

REFERENCES

Au, W.W.L, and Martin, W.M. (1989). Insights into dolphin sonar discrimination capabilities from human listening experiments. J. Acoust. Soc. Am., 86, 1662-1670.

Berg, B.G. (1989). Analysis of weights in multiple observation tasks. J. Acoust. Soc. Am., 86, 1743-1746.

Berg, B.G. (1990). Observer efficiency and weights in a multiple observation task. J. Acoust. Soc. Am., 88, 149-158.

Berg, B.G., and Green, D.M. (1990). Spectral weights in profile listening. J. Acoust Soc Am., 88, 758-766.

Berg, B.G., and D.M. Green (1991a). Discrimination of complex spectra: spectral weights and performance efficiency. In Y. Cazals, L. Demany, and K. Horner (Eds.), *Auditory Physiology and Perception, 9th International Symposium on Hearing.* Pergamon Press (in press).

Berg, B.G. and Green, D.M. (1991b). Spectral shape discrimination of narrow-band spectra. J. Acoust. Soc. Am., 89, 1911 (abstract).

Berg, B.G., Quang, N., and Green, D.M. (1991). Discrimination of narrow band spectra: I. Spectral weights and pitch cues. In review.

Durlach, N.I., and Braida, L.D., and Ito, Y. (1986). Towards a model for discrimination of broadband signals. J. Acoust. Soc. Am., 80, 63-72.

Feth, L.L. (1974). Frequency discrimination of complex periodic tones. Percept. Psychophys., 15, 375-378.

Green, D.M. (1988). Profile Analysis: Auditory Intensity Discrimination. New York: Oxford Univ. Press

Gilkey, R.H. (1987). Spectral and temporal comparisons in auditory masking. In W.A. Yost & C.S. Watson (Eds.), *Auditory Processing of Complex Sounds*. Hillsdale, NJ: Lawrence Erlbaum. 26-36

Kidd, G., Jr., Mason, C.R., Brantley, M.A., and Owen, G.A. (1989). Roving-level tone-in-noise detection. J. Acoust. Soc. Am., 86., 1310-1317.

Leek, M.R., and Watson, C.S. (1984). Learning to detect auditory pattern components. J. Acoust. Soc. Am., 76, 1037-1044.

Publications and Presentations
1 October 1990 - 30 September 1991

Papers in referred journals: Published

Green, D.M., and Berg, B.G. (1991). Spectral weights and the profile bowl. Quarterly Journal of Experimental Psychology, 43A, 449-458.

Book Chapters: In Press

Berg, B.G., and Green, D.M. (1991). Discrimination of complex spectra: Spectral weights and performance efficiency. In Y. Cazals, L. Demany, and K. Horner (Eds.), Auditory Physiology and Perception. (In press).

Conference Presentations:

Berg, B.G., and Green, D.M. (1990). Spectral shape discrimination using three-tone complexes. Paper presented at the 120th Meeting of the Acoustical Society of America, 26-30 November, 1990, San Diego, California. (Published abstract in J. Acoust. Soc. Am., 88, S146.)

Berg, B.G., and Green, D.M. (1991). Spectral weights and discrimination of complex spectra. Paper presented at Fourteenth Midwinter Research Meeting of the Association for Research in Otolaryngology, February 3-7, 1991, St. Petersburg, Florida.

Berg, B.G., and Green. D.M. (1991). Spectral shape discrimination of narrow-band spectra. Paper presented at the 121st Meeting of the Acoustical Society of America, 29 April-3 May 1991, Baltimore, Maryland. (Published abstract in J. Acoust. Soc. Am., 89, 1991.)

Berg, B.G., and Green, D.M. (1991). Discrimination of complex spectra: Spectral weights and performance efficiency. Paper presented at the 9th International Symposium on Hearing, Carcans, France, June 9-14, 1991.

Papers in Review

Berg, B.G., Quang, N., and Green, D.M. (1991). Discrimination of narrow band spectra: I. Spectral weights and pitch cues. In review (J. Acoust. Soc. Am.)

Dai, H., and Berg, B.G. (1991). Spectral-temporal weights in a profile task. In review, (J. Acoust. Soc. Am.).

Papers in preparation.

Green, D.M., Berg, B.G., Dai, H., Eddins, D., Onsan, Z., and Nguyen, Q. (1991). Detecting spectral shape changes in very narrow frequency bands.