Perception & Psychophysics
1989, 46 (4), 395-396

Notes and Comment

Propagation of constraints in auditory organization

ALBERT S. BREGMAN and YVES TOUGAS
McGill University, Montreal, Quebec, Canada

In normal listening situations, sounds can occur in mixtures. For correct pattern recognition to occur, the auditory system must decompose such mixtures, assigning the resolvable components of the sound to the separate perceptual structures that represent individual sounds. A way of studying the heuristics that are used by the auditory system is to do so through the capturing of a component out of a mixture by the use of a preceding "captor" tone. Figure 1, pattern 1, shows a pattern of tones presented repetitively in a cycle. One part of it is a mixture of two pure tones, B and C, which can either fuse to form a complex tone with a rich timbre or segregate and be heard as two separate pure tones. Many factors have previously been shown to affect the fusion of such tones (reviewed by Bregman, in press). One factor is the presence of a preceding pure tone, A, that can capture one of the tones (say B) into a sequential stream. In Figure 1 there is also a second captor (tone D), which is positioned in frequency to capture C into a lower frequency stream (remember that the stimulus is a repeating cycle).

In the present paper, we are interested in the following question: When two pure-tone streams are formed, one high in pitch and one low, are the two decisions—the one that extracts B from the mixture, and the one that extracts C—made independently? It is possible that instead of finding independent decisions, we could observe an interaction. The extraction of B might facilitate the extraction of C. Why is this interesting? Another way of stating our question is to ask whether constraints can "propagate" across the auditory "field" in the manner that was conceptualized by Gestalt psychology. This interest relates to work in computer vision. When local evidence is ambiguous, its interpretation can be forced by decisions made at other visual locations (see review by Witkin & Tenenbaum, 1983). A generalized theory based on this method has been called "relaxation labeling" (e.g., by Zucker, 1977).

In the present experiment, we looked to see whether perceiving B as a pure tone and not just as a partial of the mixture BC depended on the interpretation given to C. The experiment is interesting, because an attention-based theory would make a different prediction. It would say that we hear B as a separate tone because tone A, heard in isolation, has provided the selective attention with a model of what to listen for. If attention is conceptualized as a filter, tone A "tunes" the filter so that it can "strip out" B from the BC mixture. There could also be a separate filter tuned by D that could strip C out of the mixture, but there would be no reason to imagine that these filters could affect one another. Later pattern recognition could simply select the output of whichever filter proved more useful.

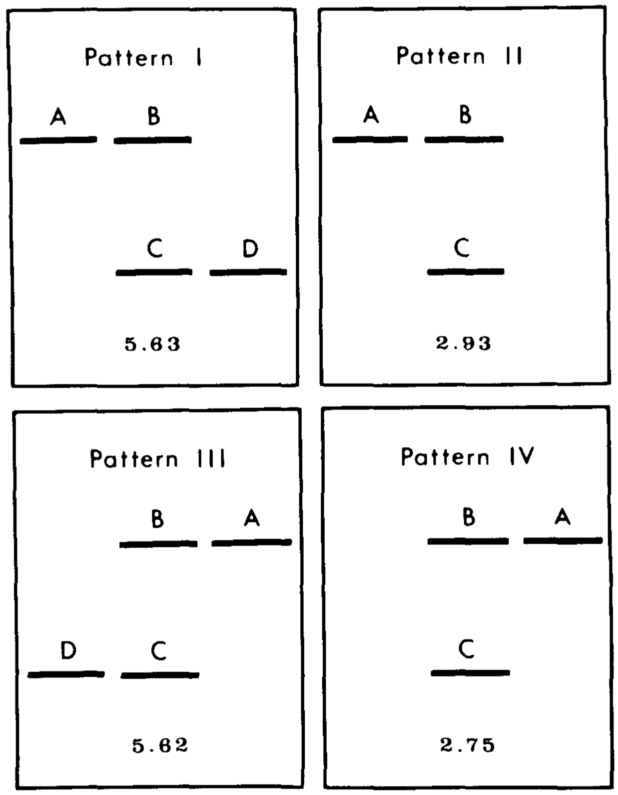

Figure 1. The four cyclical patterns judged by listeners. Horizontal bars represent sinusoidal components. The numbers are the means of ratings for the "clarity of the two high tones as a pair."

We used the four patterns shown in Figure 1. The lower tones were at 400 Hz and the higher ones at 848 Hz. Tones were 140 msec in duration, with 30 msec between consecutive tones. Patterns II and IV had a silence replacing the missing D tone, with the cycle having the same overall repetition rate as in Patterns I and III. We tested 22 young adults in a sound-isolated chamber and asked them to judge the "clarity of the two high tones as a pair" on a 7-point scale. They heard 20 cycles of a pattern and were given 10 sec to make their ratings.

The numbers in Figure 1 give the means for these ratings. A and B were heard more clearly as a pair in the

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. Reprints are available from A. S. Bregman, Psychology Department, McGill University, 1205 Dr. Penfield Ave., Montreal, Quebec H3A 1B1, Canada.

two conditions in which tone D was present. Tukey tests showed a significant difference between either of the patterns that contained tone D and either of the ones that did not. No other differences were significant.

This result is not consistent with a notion of grouping as the result of a filter-like extraction of tones. Some sort of theory in which decisions interact is implied. Our theory is that the co-occurrence of tones B and C is viewed by the auditory system as either a complex tone or two pure tones. It is as if these two interpretations were in conflict and any evidence that strengthened one weakened the other. There is a heuristic of auditory interpretation that says that if a subset of partials from a spectrum resembles a set that has been heard in isolation recently, then that subset in the more complex spectrum is a reoccurrence of the earlier set and is not merely a part of a more complex sound (see Bregman, in press). For this reason, the presence of D strengthens the interpretation of C as a separate tone. Apparently, this interpretation helps to "force" the decision that BC is not a complex tone, and this favors the release of B to group with A.

REFERENCES

BREGMAN, A. S. (in press). *Auditory scene analysis*. Cambridge, MA: MIT Press.

WITKIN, A. P., & TENENBAUM, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld, *Human and machine vision* (pp. 481-543). New York: Academic Press.

ZUCKER, S. W. (1977). Relaxation labeling and the reduction of local ambiguity. In C. H. Chen (Ed.), *Pattern recognition and artificial intelligence*. New York: Academic Press.

(Manuscript received April 10, 1989;
revision accepted for publication April 24, 1989.)