Effect of silence between tones on auditory stream segregation*

Gary L. Dannenbring and Albert S. Bregman

McGill University, Montreal, Canada H3C 3G1
(Received 22 September 1975; revised 22 December 1975)

The present study investigates the effect on auditory stream segregation of the duration of silent gaps between items of a repeating sequence of sine tone stimuli. While previous investigators had shown that gaps reduced the tendency for items to segregate, the present experiment showed that this was not the case; rather, in adjusting tone durations for stream segregation thresholds, subjects simply compensated for increased gap duration, keeping the total time (tone duration plus gap duration) constant.

Subject Classification: [43]65.75, [43]65.54, [43]65.68.

INTRODUCTION

A phenomenon which has recently been investigated in several studies is that of auditory stream segregation (Bregman and Campbell, 1971; Bregman and Dannenbring, 1973). Basically, this is a phenomenon in which a rapid sequence of high and low tones splits into two separate perceptual streams, one consisting of the high tones and the other of the low tones. Bregman and Campbell showed that judgments of the correct order of tones was possible only within a stream and not across streams. A similar result was reported by Warren, Obusek, Farmer, and Warren (1969), who found that subjects had a great deal of difficulty in determining the order of four different sounds (tone, hiss, buzz, vowel "ee") when each sound lasted 200 msec. They found that the duration of each of the sounds had to be increased to 700 msec for naive listeners to be able to correctly determine the order of the sounds.

In a second paper, Warren and Warren (1970) stated that the perception of the temporal order of a repeating sequence of vowel sounds was enhanced when there were brief spaces between the sounds, keeping the total time between vowel onsets constant. Essentially the same results were also reported by Thomas, Cetti, and Chase (1971), again using vowel sounds. More recently, van Noorden (1975) has presented data that seem to show the importance of the duration of the gap between sine tone stimuli in reducing segregation effects.

Nickerson and Freeman (1974), however, were unable to demonstrate any clear effect of introducing longer gaps between elements of a sine tone pattern. This result, together with the fact that van Noorden (1975) reported data on only one subject (himself), indicates that further research was needed to investigate the role played by the gap duration in sequences of sine tone stimuli. The present experiment was designed to explore this effect.

I. METHOD

A. Subjects

The subjects were 14 students taking an experimental psychology course at Vanier College in Montreal who volunteered their services.

B. Apparatus

The sine tone stimuli were generated by a PDP-11 computer (Digital Equipment Corp.) operating a Wavetek model 136 VCA/VCG tone generator through a D/A converter. The durations of the tones were adjustable by a knob connected to the computer through an A/D converter. Trials were terminated by a button press, and the next trial then started automatically. The stimuli were amplified using a Sony TA-1055 stereo amplifier and presented to the subject seated in a small room through Sennheiser HD-414 stereo headphones at 82 dB SPL, measured using a General Radio type 1551-C sound-level meter with a flat plate coupler. This should only be considered an approximation of the intensity of the sounds at the ear, since Sennheiser HD-414 headphones do not form a seal around the ear, but rather have a thin (~5 mm) foam rubber pad between the headphones and the ear. Background noise in the room was ~54 dB SPL.

C. Stimuli and procedure

The basic stimulus pattern for each trial was a repeating sequence of four tones, 2000, 614, 1600, and 400 Hz, always presented in that order. The tones were turned on and off instantaneously at zero-voltage crossing of the sine wave. There were six different gap durations (G) between the tones: 1, 10, 25, 50, 75, and 100 msec. The duration of the tones (D) was variable, with subjects controlling this duration by adjusting a knob. The time intervals involved in the experiment are diagrammed in Fig. 1.

The subject was seated in a small room and the

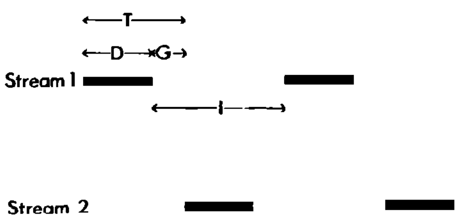

FIG. 1. A sequence of four tones which illustrates the time intevals in the experiment. D is the duration of each tone. G is the gap between successive tones. T is the onset-to-onset interval. $I(=D+2G)$ is the interval between tones in the same stream.

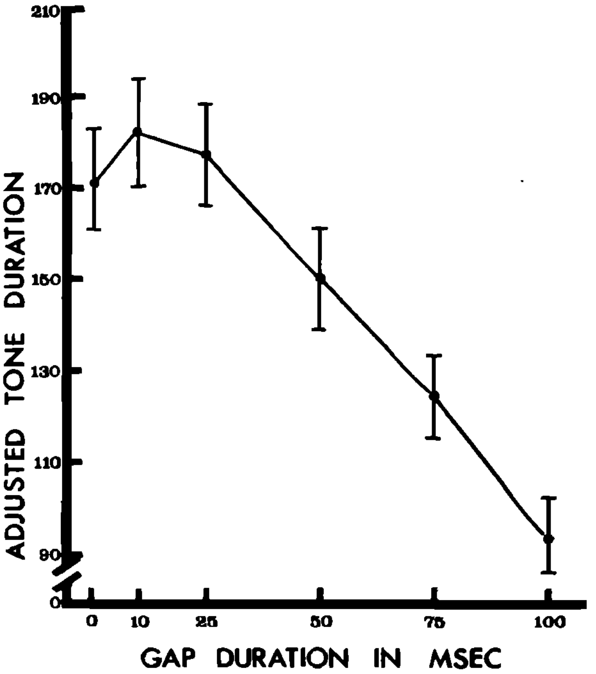

FIG. 2. Mean thresholds between one and two perceptual streams as a function of gap duration. Brackets indicate ±1 standard error of the mean.

procedure was described to him. Each trial began with a short (500 msec) 1000-Hz warning tone, followed by a 500-msec pause and then the trial itself. Subjects were instructed to adjust the duration of the tones in the repeating sequence until their perception of the pattern changed from one stream to two (or from two streams to one). Each trial began with the knob set to produce the shortest duration of the tones (resulting in two perceptual streams); subjects were encouraged to use the full range of the knob by turning it first until they perceived one stream, then slowly back until the tones split again into two streams. They were free to turn the knob back and forth until they were satisfied that they had found a threshold value; at this point they pushed a button which caused the computer to record the knob setting. They then turned the knob back to the shortest tone duration, which started the next trial.

Each subject received all six trials in four different blocks, in a different random order for each block, for a total of 24 trials. In addition, there was a practice trial at the beginning consisting of stimuli with 1-msec gaps between the tones. All of the subjects seemed to understand the procedure and felt that they could find a point between one and two perceptual streams during the practice trial. Most subjects took ~20 min to complete the experiment.

II. RESULTS AND DISCUSSION

The mean thresholds between one and two perceptual streams, together with the standard error of the mean for each point, are presented in Fig. 2. The values (in milliseconds) are the tonal durations as adjusted by the subjects. An analysis of variance revealed a significant effect of increasing gap duration, $F(5,65)=21.55$ $p<0.001$. However, as can be seen from Fig. 2, subjects were basically adjusting the tone duration (D) to compensate for the gap duration (G). This is implied by the fact that the slope of the function, at least for $G\geq 25$ msec, is close to -1 (-1.107). Thus, the total time $(T)=G+D$ remained fairly constant. Actually, it appears from Fig. 2 that there may be two components to this function, since for $G=1$, 10, and 25 msec the slope is fairly flat (0.187). Fitting a straight line by the least squares method through all of the points results in a slope of -0.865.

These results differ greatly from what one would predict from other statements in the literature. Van Noorden (1975) asserts that the interval I in Fig. 1, namely, the gap between the offset of one tone and the onset of the next one *in the same stream*, is what determines stream segregation. If this were true, subjects should have held $I=2G+D$ constant; i.e., should have subtracted 2 msec from D as the experimenter added 1 msec to G. This predicts a slope of -2.0 in Figure 2, which is clearly not the case. Furthermore, any contribution of the silent gap G, over and above its contribution to the onset-to-onset interval T should create a slope whose absolute value is greater than 1. This was not obtained. It seems, therefore, that the onset-to-onset interval T determines stream segregation independently of whether that interval is filled by sound or by silence (given that at least 25 msec of silence have occurred). Below $D=25$ msec it seems that the addition of a silent gap has no effect, not even in contributing to T. This suggests that the effects of very short silences do not fall on the same continuum as those of longer ones.

Gaps in sequences of tones have been shown to be important in improving discrimination of an element of a pattern occurring immediately before the gap. This is true when the gap is the space following a sequence, causing an improvement in the discrimination of the final element (Ortmann, 1926; Divenyi and Hirsh, 1974), or when the gap is a shorter one located within a sequence of tones (Watson, Wroton, Kelly, and Benbassat, 1975). If a listener's inability to properly identify and discriminate elements of the tonal sequence was contributing to stream segregation, introducing longer gaps between tones should affect stream segregation; yet this was clearly not the case in our data.

It might be supposed, alternately, that the effect of gaps between tones could be affected by the rather slow decay of auditory sensation (Plomp, 1964). If the sensation from one tonal element had not decayed below threshold level before the onset of a second element in the same frequency stream (assuming that this decaying sensation is not masked by the tones in the other stream), one might expect these items to be "linked" perceptually. However, this effect should again decrease as gap duration increased, and this did not occur.

The present results are more in agreement with those of Nickerson and Freeman (1974), who did not find a clear effect of gap duration. Both the present experiment and that of Nickerson and Freeman used repeating sequences of four tones, while van Noorden (1975) used two-tone sequences; it is possible that this is contributing to the differences between results. The other two studies which reported an effect of the gap duration (Warren and Warren, 1970; Thomas *et al.*, 1971) used repeating sequences of vowel sounds as stimuli. Warren and Warren stated that the introduction of silences caused the sequence of vowels to sound more like normal speech. Thus, it is possible that a different mechanism is involved in the perceptual ordering of these sounds; this may be the reason for the different effect of gaps in vowel sequences. Also, Thomas *et al.* (1971) simply compared their results with an earlier experiment in which no gaps were present between the stimuli, and the different results might simply reflect other differences between the experiments rather than an effect of gap duration.

A final point which should be noted concerns the difference between the present study, in which subjects are finding thresholds for stream segregation, and the studies involving the correct ordering of vowel sounds. We believe that auditory stream segregation is a basic, fairly low level process, similar to Neisser's (1967) notion of "preattentive processes." The present experiment was designed as a direct, sensitive measure of stream segregation. The inability of subjects to correctly name the order of vowel sounds in a repeating sequence may be due to factors other than stream segregation effects. For example, when listeners are required to report the order of sounds, those sounds must first be accurately discriminated and identified; this is probably not necessary when subjects are finding stream segregation thresholds. Since studies such as Watson *et al.* (1975) have shown that discrimination of an element is improved when a gap follows the element, this may be the reason for improved performance in ordering vowel sounds with gaps between the sounds, rather than any reduction of segregation of items into separate streams. Clearly, further research is needed to investigate the role played by these factors.

*This research was supported by grants from the National Research Council of Canada, the Quebec Ministry of Education, and the McGill University Graduate Faculty.

Bregman, A. S., and Campbell, J. (1971). "Primary auditory stream segregation and perception of order in rapid sequences of tones," J. Exp. Psychol. 89, 244–249.

Bregman, A. S., and Dannenbring, G. L. (1973). "The effect of continuity on auditory stream segregation," Percept. Psychophys. 13, 308–312.

Divenyi, P. L., and Hirsh, I. J. (1974). "Identification of temporal order in three-tone sequences," J. Acoust. Soc. Am. 56, 144–151.

Neisser, U. (1967). *Cognitive Psychology* (Appleton–Century–Crofts, New York).

Nickerson, R. S., and Freeman, B. (1974). "Discrimination of the order of the components of repeating tone sequences: Effects of frequency separation and extensive practice," Percep. Psychophys. 16, 471–477.

Noorden, L. P. A. S. van (1975). "Temporal coherence in the perception of tone sequences," Ph.D. thesis (Institute for Perception Research, Eindhoven, Holland).

Ortmann, O. (1926). "On the melodic relativity of tones," Psychological Monographs, 35, No. 162.

Plomp, R. (1964). "Rate of decay of auditory sensation," J. Acoust. Soc. Am. 36, 277–282.

Thomas, I. B., Cetti, R. P., and Chase, P. W. (1971). "Effect of silent intevals on the perception of temporal order for vowels," J. Acoust. Soc. Am. 49, 85 (A).

Warren, R. M., Obusek, C. J., Farmer, R. M., and Warren, R. P. (1969). "Auditory sequence: Confusion of patterns other than speech or music," Science 164, 586–587.

Warren, R. M., and Warren, R. P. (1970). "Auditory illusions and confusions," Sci. Am. 233, 30–36.

Watson, C. S., Wroton, H. W., Kelly, W. J., and Benbassat, C. A. (1975). "Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals," J. Acoust. Soc. Am. 57, 1175–1185.