

Pattern identification of pure tones and frequency glides by untrained listeners

JOHN E. KERIVAN

*Naval Submarine Medical Research Laboratory,
Naval Submarine Base New London, Groton, Connecticut 06340*

and

BERNARD J. CAREY

Department of Electrical Engineering, University of Connecticut, Storrs, Connecticut 06268

Six subjects identified the order of four-event sequences. Contiguous pure tones (713, 1,031, 1,209, and 1,514 Hz in permuted orders) were presented by earphones at 40 dB SL, with individual events (tones) from 20, to 40, 60, and 300 msec in duration. Again, silent intervals of 20 or 60 msec were inserted among tones of 20 or 40 msec duration. Finally, the pure tones of 713 and 1,209 Hz were combined, in any four-event sequence, with two glissandi chosen from 466 to 714 Hz, from 714 to 1,208 Hz, and their mirror reversals. The temporal and frequency continuity both of tonal and of glissando-plus-tonal sequences affected the identification of sequential order. Degraded performance in the glissando-plus-tonal condition was attributed partially to a subjective experience of pitch blurring. The inclusion of silent intervals in the sequences of the shorter pure-tone durations improved identification performance to that of contiguous sequences of equal overall duration, i.e., adding silent processing time was as efficacious as increasing by the same amount the duration of the individual frequency event.

Experimental evidence suggests that several variables significantly affect the identification of temporal order of auditory sequences. Some of these pertain to the acoustic stimulus properties, while others relate to listener training and experimental procedures (Divenyi & Hirsh, 1974; Neisser & Hirst, 1974; Nickerson & Freeman, 1974).

A few investigators have looked at the role of frequency glides (glissandi) in connecting tones within a sequence (see Bregman & Dannenbring, 1973). Others have reported on pitch perception of tone bursts in the presence of continuous vs. discontinuous glissandi (Nabelek, Nabelek, & Hirsh, 1973). However, the use of a glide as an event in the sequence has not been investigated.

Interevent silent intervals resulted in the improvement of discrimination performance in pure-tone sequences (Watson, Wroton, Kelly, & Benbassat,

1975). However, gap durations were not adequately sampled to determine the extent of pre- and post-event masking.

We therefore studied the frequency and temporal continuity of sequences of pure tones alone, and of pure tones plus glissandi on pattern identification. The main objective of our first experiment was to determine listener identification performance for permuted sequential orders of four pure tones. In a second experiment, glissandi were substituted for two of the tones. In a third experiment, silent intervals (gaps) were inserted among pure tones of very brief event durations.

METHOD

Experiment 1 used four contiguous pure tones: 713, 1,031, 1,209, and 1,514 Hz. Each tone (event) had a rise and decay duration equal to 10% of the stated steady state. Event durations within each sequence were identical; they ranged from 20 to 300 msec in 20-msec steps. In Experiment 3, interevent gaps of 20 and 60 msec were placed between the tones at the 20- and 40-msec event durations. All 24-sequence permutations appeared once for all event durations.

Experiment 2 was exactly as Experiment 1, except that two linear glissandi chosen from 466 to 714 Hz, from 714 to 1,208 Hz, and their mirror reversals replaced the tones of 1,031 and 1,514 Hz.

Three classes of frequency continuity were established: continuous (C), when one glissando connected both tones; intermediate (I), when only one tone and one glissando were connected; and discontinuous (D), when there were no frequency connections between tones and glissandi (see Figure 1).

Stimulus materials were generated by a PDP-9 computer and

This research is based on a Master's thesis submitted to the University of Connecticut by J. E. Kerivan. We wish to thank J. Donald Harris for his advice and assistance in all phases of the study, particularly in the development of the second experiment. We also want to thank Michael McClean and Kenneth Randolph, who provided valuable help with parts of this work. Portions of this work were presented at the 47th Annual Meeting of the Eastern Psychological Association, 1976. The opinions expressed in this paper are those of the authors and are not necessarily the official views of the U.S. Navy. Requests for reprints should be sent to John E. Kerivan, Naval Submarine Medical Research Laboratory, Code 31, Box 900, Naval Submarine Base New London, Groton, Connecticut 06340.

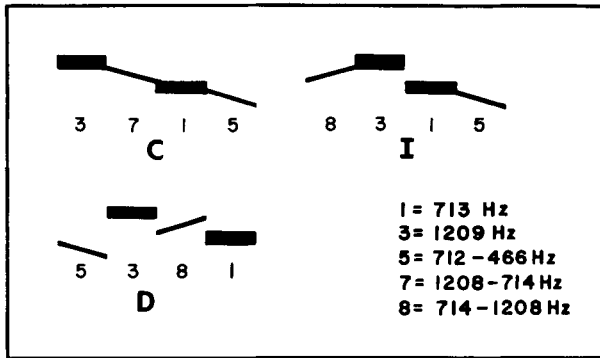


Figure 1. Examples of the three categories of frequency continuity for the stimuli in Experiment 2. In the experiment, additional sequence orders representing these categories were given.

were recorded and played back on a Tandberg 15 SL tape recorder at a speed of 7½ ips. A Krohn-Hite Model 310-C filter was set to bandpass 380-1,800 Hz; voltage levels were monitored with a VTVM.

Six adult subjects were asked to identify a four-event sequence by ordering appropriately labeled cards from left to right representing time (see Warren, Obusek, & Charles, 1972). All subjects were untrained on the task; however, two had training in music, two were experienced in other psychoacoustic listening tasks, and two were naive. The subjects were given up to three chances to judge each sequence. This number was recorded on a separate sheet by the experimenter and later related to performance. It was thought that this approach would reflect the degree of sequence identification difficulty as a function of event duration.

Sequences were presented binaurally in an IAC booth through a pair of TDH-39 earphones mounted in Tracor Otopups at a 40-dB sensation level (based on audiometric data from 500 to 1,500 Hz). The 300-msec presentation rate was the practice rate for all experiments; thus, relatively slight training was given. Event durations and sequence orders within experiments, and the experiments themselves, were counterbalanced. Feedback regarding the number of correct responses was given at the end of each event-duration session. The experiments were run on different days.

RESULTS AND DISCUSSION

Identification performance across sequences in all experiments is shown Figure 2. The $d' = 1.0$ level correspond to an average of approximately 20% correct decisions and was defined as the identification threshold (see Elliott, 1964). Figure 2 shows that identification performance in Experiment 1 was above threshold for the four experienced subjects at the 20-msec event duration (see Points 1, 4). The other two subjects (see Point 7) did not reach threshold until event duration was 100 msec.

Experiment 2 (glissandi) subjects with musical training, (Point 2) did not reach threshold until event duration was 100 msec. Those with other listening experience (Point 5) reached threshold at 80 msec. However, the naive listeners (Point 8) did not attain threshold until event duration reached 140 msec.

Performance in Experiment 3 (silent intervals) is plotted at the four event durations from 40 to 100 msec. (Each event duration here is taken as the total event plus gap time; thus, the 20-msec-event/20-msec-gap data were plotted at the 40-msec-event duration.) Once again, the performance of the experienced subjects (Points 3, 6) was superior to that of the naive (Point 9). Experienced subjects were above threshold at the 40-msec event duration, while the naive subjects reached threshold only at 80 msec (20-msec tone, 60-msec silence).

A product-moment correlation was performed between performance vs. the number of required judgments at each event duration: $r = -.76$ in Experiment 1, $-.72$ in Experiment 2, and $-.74$ in Experiment 3.

Comparisons within Figure 2 reveal event duration and temporal continuity effects. (1) All identification thresholds were better in Experiment 1 than

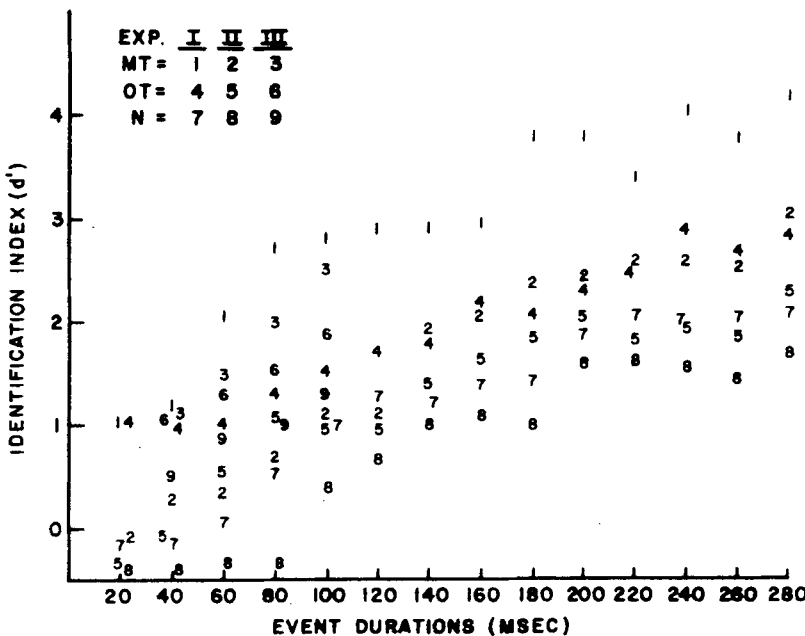


Figure 2. Identification performance for all experiments across sequence orders as a function of event duration. Numbers represent averaged data for three subject categories: MT = musically trained, OT = other training, and N = naive. Two subjects per category.

in Experiment 2; glissandi clearly produced difficulties. (2) Interevent gaps for the shorter tones improved performance over sequences even of longer tonal duration (Experiments 1 vs. 3). For example, compare Points 4 vs. 6, 7 vs. 9, and 1 vs. 3.

These results are in agreement with findings of Massaro (1970) and Watson et al. (1975). Our data show that the auditory system is capable of processing very brief tones, e.g., 20 msec/event, 60 msec/gap, with the same temporal resolution as 80-msec contiguous tones. This processing appears to be based on the total tone-gap period. The "consolidation time," elucidated by Massaro and reiterated by Watson et al., serves as one possible explanation of this effect. As the total tone/gap duration approaches the 120-160 msec outlined by Watson et al. (100 msec total duration in our study), identification performance for the brief-event durations improves.

All of the four experienced subjects were above threshold identification performance in Experiment 1 even at the 20-msec-event duration. The amount of training was minimal (one block of 24 trials) and as such cannot qualify the subjects as highly experienced in this particular situation. In a similar task, Divenyi and Hirsh (1974) noted that naive subjects required 75 msec per event to reach identification threshold. In our study, 100 msec per event were needed by the two naive subjects to reach threshold for pure tones. This difference could be the result of one additional sequential event (Divenyi and Hirsh used three-event sequences) or of a difference in individual abilities.

Glissandi considerably degraded identification performance for all subjects. Bregman and Dannenbring

(1973) also used glissandi, finding an improvement to connect tones; however, it is clear that the use of a glissando as an event, even when it connects events in frequency, does not necessarily (in our data, ever) facilitate sequential order judgments (cf. d' for Experiment 1 vs. for Experiment 2 in Table 1). For example, comparison of the pure tone vs. glissandi plus tonal sequences both exhibiting frequency and temporal continuity (category C) shows that those in Experiment 1 were identified correctly ($d' = 3.4, 3.4$) and those in Experiment 2 were identified less well ($d' = 2.0$ and 1.83 , respectively). One possible explanation for the discrepancy between our data vs. those of Bregman and Dannenbring is that our identification tasks were obviously more difficult than the discrimination tasks used by Bregman and Dannenbring.

Using the same criterion for thresholds as that presented earlier ($d' = 1.0$ level), one sees that those sequences in both Experiment 1 and Experiment 2 that are unidirectional in frequency are most often identified correctly. In addition, those which approximate this frequency continuity are also identified more often than others. For example, sequence order 3214 and 2341 in Experiment 1 had performance levels of $d' = 2.5$ and 2.28 , respectively, while the similar sequence orders 3751 and 3716 in Experiment 2 had thresholds at $d' = 1.77$ and 1.71 , respectively. A rank order correlation ($r = .78$) for sequence orders between Experiment 1 and Experiment 3 shows that sequence identifiability remained relatively constant despite the inclusion of silent gaps.

Event order effects are clearly illustrated in Table 2,

Table 1
Sequence Order Identifications Given in d' Scores Where $d' = 1.0$ is Threshold of Identifiability

EXPERIMENT 1				EXPERIMENT 2			
Sequence	(d')	Sequence	(d')	Sequence	(d')	Sequence	(d')
1234	3.4	3124	1.97	3715 C	2.00	6371 C	1.38
4321	3.4	4132	1.97	6183 C	1.83	3761 I	1.34
4312	2.5	3412	1.97	8315 I	1.80	8136 D	1.34
3214	2.5	1243	1.86	3751 I	1.77	3618 C	1.34
2341	2.28	2314	1.77	3716 C	1.71	1835 C	1.34
3421	2.28	2143	1.77	1583 I	1.69	8361 I	1.32
3241	2.28	1423	1.69	7163 I	1.69	1836 C	1.15
1432	2.23	3142	1.65	3185 I	1.63	1736 D	1.04
4231	2.14	1342	1.65	7153 C	1.55	1573 I	1.04
2134	2.12	2431	1.63	1358 D	1.42	6137 I	1.04
4123	2.02	4213	1.48	8135 D	1.38	3517 D	1.00
1324	2.00	2413	1.26	6318 D	1.38	5381 D	.80

NOTE—Sequence numbers are frequencies: 1 = 713; 2 = 1,031; 3 = 1,209; 4 = 1,514; 5 = 712-466; 6 = 466-712; 7 = 1,208-714; and 8 = 714-1,208 Hz. (C) Continuity, (I) intermediate, and (D) discontinuity represent frequency continuity categories.

Table 2
Identification Performance For All Tones in Each Event Position
Across Subjects in Experiments 1 and 3

Frequency	EVENT POSITION				TOTAL
	1	2	3	4	
713Hz	1.83	1.49	1.21	2.32	1.75
1031Hz	1.15	1.55	1.42	1.28	1.35
1209Hz	1.45	1.49	1.49	.97	1.35
1514Hz	2.36	1.75	1.83	2.93	2.12
TOTAL	1.67	1.55	1.49	1.75	

NOTE— d' scores are given; threshold of identifiability is $d' = 1.0$.

giving the d' values for four alternative decisions where the 1.0 level is the threshold of identification. Inspection of this table shows that the highest tone was the easiest to identify in any event position. The lowest tone was the next best identified, and the intermediate tones were equally identifiable. These results are consistent with the findings of Watson et al. (1975) of a "tonal component effect" and Divenyi and Hirsh's (1974) "salience of the final tone." Based on our findings, a clear hierarchy of sequence identifiability appears to be a function, among other things, of ordinal position [the first and last tones (events) serving as landmarks]; frequency continuity is also an important factor. The inclusion of silent gaps within the sequence greatly changes absolute levels of performance, but does not significantly affect this clear hierarchy.

Subjective impressions on the part of subjects indicated that a pitch blurring of the glissandi might have been the primary cause for the overall reduced performance in Experiment 2. They observed this effect only when the glissando was in the fourth or last position, and they stated that it seemed to extend perceptually beyond its range, making comparisons

between the two glissandi in the sequence practically impossible. A frequency analysis of the acoustic stimuli from the earphones with a 6-cc coupler and CRO insured that the phenomenon was indeed subjective and was not a "ringing" effect of the transducer.

It is clear that much more needs to be done before the effects of glissandi as events in temporal sequences can be adequately quantified.

REFERENCES

- BREGMAN, A. S., & DANNENBRING, G. The effect of continuity on auditory stream segregation. *Perception & Psychophysics*, 1973, 13, 308-312.
- DIVENYI, P. L., & HIRSH, I. J. Identification of temporal order in three-tone sequences. *Journal of the Acoustical Society of America*, 1974, 56, 144-151.
- ELLIOTT, P. B. Tables of d' . In J. A. Swets (Ed.), *Signal detection and recognition by human observers*. New York: Wiley, 1964.
- MASSARO, D. W. Preperceptual auditory images. *Journal of Experimental Psychology*, 1970, 84, 411-417.
- NABELEK, I., NABELEK, A., & HIRSH, I. J. Pitch of sound bursts with continuous or discontinuous change of frequency. *Journal of the Acoustical Society of America*, 1973, 53, 1305-1312.
- NEISSER, U., & HIRST, W. Effect of practice on the identification of auditory sequences. *Perception & Psychophysics*, 1974, 15, 391-398.
- NICKERSON, R. S., & FREEMAN, B. Discrimination of the order of the components of repeating tone sequences: Effects of frequency separation and extensive practice. *Perception & Psychophysics*, 1974, 16, 471-477.
- WARREN, R. M., & OBUSEK, C. J. Identification of temporal order within auditory sequences. *Perception & Psychophysics*, 1972, 12, 86-90.
- WATSON, C. S., WROTON, H. W., KELLY, W. J., & BENBASSET, C. A. Factors in the discrimination of tonal patterns. 1. Component frequency, temporal position, and silent intervals. *Journal of the Acoustical Society of America*, 1975, 57, 1175-1185.

(Received for publication March 5, 1976;
revision accepted September 15, 1976.)