Journal of Experimental Psychology:
Human Perception and Performance
1976, Vol. 2, No. 3, 337–346

Perceiving and Counting Sounds

Dominic W. Massaro
University of Wisconsin—Madison

Observers counted sequences of 20-msec tones that were presented to the same spatial location or alternated between spatial locations. Counting accuracy increased with increases in the silent interval between tones and increased at a faster rate when the tones were presented to the same location. In the second study, subjects monitored the same sequences of tones for a probe tone that was either higher or lower than the other tones. Probe recognition improved with increases in the silent interval between tones, and there was no significant decrement in monitoring tones alternated between spatial locations. In the next two studies, observers processed a sequence of tones presented at the same spatial location, but the tones could alternate in frequency by slightly more than an octave. Counting accuracy was poorer when the tones alternated in frequency. When subjects monitored these sequences of tones for the duration of a probe tone, recognition leveled off at a lower asymptote for sequences of tones alternating in frequency. The deficit in counting sounds alternating between spatial locations or frequencies appears to be due to a process that has difficulty integrating successive tones that are perceived at different spatial locations or pitch levels.

Observers have difficulty counting the number of brief events such as sounds, taps, or lights presented in a relatively rapid sequence (Cheatham & White, 1954; Garner, 1951; Lechelt, 1975; White, 1963). Although this task has been assumed to measure a perceptual deficit, it is worthwhile to distinguish between the perception of an event and the output required by the initial perception of that event. An observer might perceive that an event has occurred but fail to act upon it. In this task, acting would involve incrementing the appropriate counter. If the subjects are attempting to count events presented at 10 events/sec, they could easily fall behind since the fastest rate of subvocal speech is about 6 syllables/sec. A limitation of subvocal speech or some other process might account for the underestimation rather than some failure of initial perception of events.

Recently, observers have been required to count sounds presented to the same ear or alternated between the ears (Guzy & Axelrod, 1972; Harvey & Treisman, 1973). Subjects in these experiments have tended to underestimate sounds alternated between the ears relative to a single ear presentation. In terms of the above analysis of the counting task, this result could be due to either the initial perception of each of the sounds, the later process of incrementing the appropriate counter, or both. When sounds are alternated between the ears, subjects may fail in the initial perception of some of the sounds because of a problem of attending to more than one spatial location. The perception of each sound may be dependent on attention, and we may be able to attend to only one location at a time (Broadbent, 1958, 1971). In this case, sounds alternating between the ears will exceed the capacity of the perceptual mechanism if the rate of presentation exceeds the time required to perceive the sounds and to switch between the respective locations. Even if sounds could be perceived without attention to the spatial location of their arrival, attention might be necessary

---

This research was supported by a grant from the National Institute of Mental Health, Public Health Service. Portions of the present research were reported at the Psychonomic Society, Boston, 1974.

Terry Benzschawel, James Bryant, Wendy Idson, Michael Cohen, and Michael Olson assisted in the experimental work.

Requests for reprints should be sent to Dominic W. Massaro, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

for incrementing the counter in memory. In this case, although sounds alternating between the two ears might be perceived as well as sounds presented to just one ear, it may be more difficult to act on sounds (i.e., updating the appropriate counter) perceived at different locations in space. This analysis shows that the observed decrement might occur at either the perceptual or counting stage of processing or at both stages of processing in the counting task.

The central purpose of this series of experiments is to determine the stage of processing that is responsible for relatively poor counting performance when sounds are alternated across two points on some auditory dimension. The experiments to be reported utilize highly practiced observers in order to measure the degree of counting resolution under optimal conditions. Subjects are presented with a fixed set of stimulus alternatives, one per trial, and are required to make a response from the same fixed set of alternatives. Furthermore, feedback is provided on each trial to optimize the subject's performance under all conditions. The dependent measure is the average percentage correct under each experimental condition. Given a fixed set of stimulus and response alternatives, this measure corresponds to the *d'* measure of signal detectability theory, which specifies what the subject knows under each of the experimental conditions. The percentage correct measure indexes how countable tones are under each experimental condition. A second measure, the mean counting response, indexes the degree to which the subject was likely to underestimate or overestimate the number of tones presented in a sequence.

EXPERIMENT 1

Method

Subjects. Seven subjects were tested for 5 days. The students, recruited from introductory psychology classes, volunteered to fulfill a course requirement.

Procedure. Subjects heard a sequence of short tones on each trial, indicated the total number of tones in the sequence, and then were informed of the number of tones actually presented. The test tones were 20-msec sine waves of 800 Hz and 86 dB SPL. The total number of tones in the sequence was 5, 6, 7, or 8. The rate of presentation was varied by varying the duration of the silent interval between successive tones in a sequence. For a given sequence, the silent intervals between tones were 25, 45, 75, 105, 135, 165, 205, or 255 msec. The tones in a given sequence were presented to the same ear or alternated between the ears. A sequence of tones presented to the same ear could be presented to either the left or right ear and a sequence of tones alternated between the ears could begin on either the left or right ear. The four types of trials (left-right . . . , right-left . . . , right-right . . . , left-left . . .) were equally likely to occur on any trial.

Each trial began with the presentation of the sequence of tones followed by a 2-sec response interval. The subject made the appropriate response by hitting one of four pushbuttons labeled 5, 6, 7, and 8, respectively. Subjects were required to make one of these four responses on each trial. Feedback was given by a visual presentation of the digit 5, 6, 7, or 8 for 250 msec. The intertrial interval was 500 msec. Within a given session, all experimental conditions were chosen randomly with equal probability. Subjects were instructed to count the total number of tones in the sequence and were informed that the rate of presentation could vary, and that sequences might alternate from ear to ear or be presented to the same ear. They were explicitly told to count the total number of tones regardless of the rate of presentation and whether or not the tones alternated between the ears.

The procedure was aimed at preventing the successful utilization of any strategy other than actually counting the tones. Four lengths were used so that knowing that the sequence had an odd or even number of tones would not be sufficient for a correct answer. All 64 possible trial types (alternating or same ear × 8 rates × 4 test lengths) were equally likely to occur on any trial. This prevented subjects from learning to use the duration of a sequence as a reliable cue to the number of tones it contained. If the rate of presentation did not vary within the session, subjects could make reliable judgments on the basis of duration alone. The completely random presentation also worked against the utilization of different strategies under the different experimental conditions. Subjects could not know the rate of presentation or whether the tones were alternating until they heard at least two tones.

Two sessions of 305 trials each were given each day for 5 consecutive days. There was a short break between the two sessions. The first day and the first 5 trials of each experimental session were eliminated from the data analysis. The results are, therefore, based on an average of 37 observations per subject at each of the 64 experimental conditions. The percentage of correct judgments was averaged across the four different lengths of the test sequence, giving about 150 observations for each subject at each of the remaining 16 conditions (alternating or same ear × 8 rates of presentation).

Apparatus. All experimental events were controlled by a PDP-8L computer. The tones were generated by a digitally controlled oscillator

(Wavetek Model 155) and were presented over matched headphones (Grason Stadler TDH-49). The feedback was given visually over displays made of light-emitting diodes (Monsanto Model MDA III). (These displays are described in Nealis, Engelke, and Massaro, 1973.) Four subjects were tested simultaneously in separate sound-attenuated rooms.

Results

The accuracy of the counting task can be measured in terms of the percentage of correct identifications averaged across the four lengths of the test sequence. The rate of presentation of the test sequences is expressed in terms of processing time between the onsets of successive tones. The top panel of Figure 1 plots performance as a function of processing time under the alternating and single ear presentations for six of the seven observers. One observer averaged only 62% correct at the two slowest rates of presentation. Given that the other six observers averaged over 90% at these rates, this observer was eliminated from the data analysis. The performance of the six observers in Figure 1 shows large effects of the rate of presentation and the alternating versus the single ear presentation. Performance improved with increases in processing time from roughly 38% to 98%, $F(7, 35) = 70.47$, $p < .001$. All six observers showed a decrement in performance in the alternating ear relative to the single ear condition, $F(1, 5) = 10.56$, $p < .0025$. The advantage of the single ear presentation also interacted with processing time, as might be expected from the different levels of performance across the processing time variable, $F(7, 35) = 3.42$, $p < .001$. A 10% advantage of presenting the tones to the same ear emerged at a processing time of 95 msec/tone, and this advantage actually increased until the alternating ear caught up to the perfect performance of the single ear presentations at the slower rates of presentation.

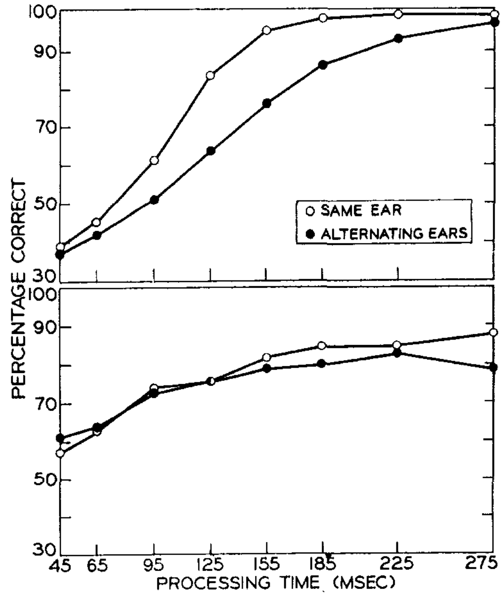

FIGURE 1. Percentage of correct counts of the number of tones in the test sequence (top panel) and correct probe frequency recognitions (bottom panel) as a function of processing time for each tone and whether the tones were presented to the same ear or alternated between the ears.

The percentage of correct identifications measures accuracy in the counting task. We can also ask whether the observers were biased to over- or underestimate the number of sounds in the sequence. Table 1 presents mean counting response as a function of processing time for each tone and whether the tones were presented to the same ear or alternated between ears. Given the fixed set of alternatives of 5, 6, 7, and 8, the mean response for an unbiased subject should be 6.5. In contrast to the accuracy measure, processing time and the attention conditions had very little effect on response bias. Subjects overestimated the number of tones coming along the same ear at the shortest processing interval and underestimated the

TABLE 1

MEAN COUNTING RESPONSE AS A FUNCTION OF PROCESSING TIME FOR EACH TONE AND WHETHER THE TONES WERE PRESENTED TO THE SAME EAR (EXPERIMENT 1)

Processing time (msec)	Attention condition	
	Same ear	Alternating ear
45	6.73	5.68
65	6.48	5.89
95	6.45	6.44
125	6.47	6.51
155	6.51	6.57
185	6.51	6.57
225	6.51	6.53
275	6.50	6.51

number of alternating tones at the two shortest processing intervals. Subjects showed no tendency to over- or underestimate the number of tones at the other 13 experimental conditions. These results, in conjunction with those in Figure 1, show that alternating tones between the ears decreased counting performance without inducing the observers to over- or underestimate the number of tones in the sequence.

Discussion

The results show a real processing deficit when tones are alternated between the ears relative to being presented to the same ear. The deficit interacts in a very orderly way with the processing time available for each tone. When the tones are presented at a very fast rate, so that performance is near chance, the presentation condition has no effect. Slowing the tones down to about 10 tones/sec is sufficient to pull performance well above chance with a significant advantage of the same ear relative to the alternating ear presentations. Slowing the rate down to about 4 tones/sec gives the alternating condition time to catch up to the single ear condition at essentially perfect performance.

The decrement in the counting task under the alternating ear condition could be due to either the perceptual or counting processes or both, as described in the introduction. The tone-counting experiment requires the tones to be both perceived and counted. A task that requires perception but not counting would test whether alternating tones between the ears interferes with their initial perception relative to the single ear situation. Harvey and Treisman (1973) took this approach in a similar experiment. Their observers underestimated and showed more variability in counting tones switched between the ears than tones presented to the same ear. In another study observers monitored a sequence of six 80-msec tones, presented at the same two rates as the first study, and reported whether all of the tones were presented at the same background frequency (1,000 Hz) or whether one of the tones was the target frequency (1,050 Hz). Subjects made a two-alternative forced-choice response at the end of each trial. The six tones were presented to the same ear or alternated between ears. The results showed that subjects missed about 2% more targets under the alternating than the single ear presentations, with about a 1% difference in false alarm rates. Reaction times were also 40 msec slower with the alternating than with single ear presentations. The small differences between the alternating and single ear presentations would not be able to account for the very large differences in the counting task. Unfortunately, the almost perfect accuracy in the target recognition task in the Harvey and Treisman experiment puts a ceiling on performance that would preclude observing any difference between the alternating and single ear presentations. The idea of whether the deficit observed in counting alternating tones is due to a decrement in the initial perception of each tone needs to be tested without the problem of a ceiling effect.

The next experiment was designed to replicate the first study in all aspects except the processing task required by the subject. Instead of counting the tones on each trial, subjects monitored the sequence of tones for a probe tone that was either higher or lower in frequency than the other tones in the sequence. A probe tone was presented on each trial and the task was to indicate whether the probe tone was higher or lower. We can expect that the accuracy of identifying the pitch of the probe tone would increase with increases in processing time for each tone. Pitch perception improves with increases in the silent interval after a short tone before presentation of a second tone (Massaro, 1970). However, if subjects can perceive tones alternating between the ears as well as tones coming to the same ear, there should be no difference between these two conditions. Experiment 2 attempts to locate the poor performance with alternating tones in Experiment 1 at either the initial perception or counting stage of processing.

EXPERIMENT 2

Method

Subjects. Eight subjects volunteered to participate in the experiment for 5 consecutive days. Two subjects were eliminated after the third day because they failed to perform at better than 65% in the task. Six subjects were paid for participation and the other two received experimental points in an introductory psychology course.

Procedure. The procedure was identical to Experiment 1 with the following exceptions. One of the tones in the sequence of tones was equally likely to be Δf (delta frequency) Hz higher or lower than the other 800-Hz tones. The first and last tones in the sequence were always 800 Hz. The value of Δf was maintained at a level at which subjects' performance averaged 75% correct. The Δf values leveled off at 45 Hz for the first group and 25 Hz for the second group across the last 4 days of the experiment. Subjects were required to indicate whether the probe tone was higher or lower on each trial regardless of the number of tones, the rate of presentation, or whether the tones were alternated between the ears or came in the same ear. The subject made the appropriate response by depressing one of two pushbuttons labeled *high* and *low,* respectively. The visual feedback *HI* or *LO* indicated whether the probe tone was higher or lower.

The results were pooled over the number of tones in the sequence in the data collection. The first day and the first 5 trials of each session were eliminated from the data analysis. The 2,400 observations across the four experimental days give roughly 150 observations at each of the 16 experimental conditions (8 processing times × alternating vs. same ear). The dependent measure is the average percentage of correct high and low judgments.

Results

The bottom panel of Figure 1 plots the percentage of correct recognitions for the six observers as a function of processing time for each of the attention conditions. Recognition performance improved about 25% with increases in the processing time available for each tone in the sequence, $F(7, 35) = 21.8$, $p < .001$. Performance was only 2% better when the tones came to the same ear than when they were alternated between the ears, $F(1, 5) = 2.95$, $p > .12$, and this variable did not interact with processing time, $F(7, 35) = 2.03$, $p > .05$. In contrast to the large decrement in counting performance in Experiment 1, observers show no significant decrement in recognizing the pitch of a probe tone when the tones are alternated between the ears.

Discussion

One might argue that the tones alternating between the ears produce less backward recognition masking than tones presented to the same ear. If this were the case, the advantage of selective attention in the same ear condition could have been offset by the disadvantage of greater backward recognition masking. However, there is no evidence that a contralateral masker produces less backward recognition masking than an ipsilateral masker (Massaro, 1970; Massaro & Cohen, 1975; Pisoni, 1972). Differences in backward recognition masking cannot account for the failure to find an advantage in monitoring tones presented to the same ear relative to monitoring tones alternated between the ears.

The results of Experiments 1 and 2 support the idea of two stages of processing in counting a sequence of tones. The tones might be resolved perceptually (their pitch) but not conceptually (how many). The results are consistent with the hypothesis that the counting deficit must have occurred at a later stage than the initial perceptual resolution of each of the tones. The tones in the single ear presentation are perceived and localized in the same spatial location, whereas the tones in the alternating ear presentation are perceived at different locations in space. The counting task requires the subject to increment a counter for each sound. In the alternating presentation, the subject must count sounds which have been perceived at different locations in space. In the single ear presentation, the sounds are perceived in the same location, making them much easier to count. The next two experiments ask whether a counting deficit without a perceptual deficit also occurs when tones are alternated in frequency relative to being presented at the same frequency.

EXPERIMENT 3

Method

Subjects. Eight students recruited from an introductory psychology class served 1 hour a day for 5 consecutive days.

Procedure. On each trial, subjects were presented with a sequence of short tones, estimated the total number of tones presented, and were then given feedback on the exact number of tones presented on that trial. The tones were 20-msec sine waves presented binaurally. Three independent variables were manipulated in the present study. The total number of tones in the sequence was 5, 6, 7, or 8. The silent interval between successive tones was 25, 45, 75, 105, 135, 165, 205, or 255 msec. The tones in a given sequence were presented at the same frequency or alternated between two frequencies. When all of the tones were presented at the same frequency, they were equally likely to be presented at A_4(440 Hz) or B_5(988 Hz). When the tones alternated between frequencies they alternated between A_4 and B_5 with an equal likelihood of either of these notes starting the sequence. Therefore, the four kinds of trials (A_4A_4 . . . , B_5B_5 . . . , A_4B_5 . . . , B_5A_4 . . .) were equally likely to occur on any trial. The 440 Hz and 988 Hz tones were equated for subjective loudness. Both tones presented at 86 dB SPL at each ear appeared to be equally loud. All other procedural details were the same as those in Experiment 1.

On each day, subjects received two sessions of 300 trials each with about a 10-min. break between the 15-min. sessions. Before the experiment proper, the first four subjects received two sessions on Day 1 and one session on Day 2 of sequences of 6 or 7 items. The second four subjects had only two sessions of practice on the 6- or 7-tone sequences on Day 1. Therefore, the data of the first four subjects are taken from the last 3 days of the experiment, whereas the data of the second four subjects are taken from the last 4 days of the experiment. The first five trials of each session were eliminated from the data analysis.

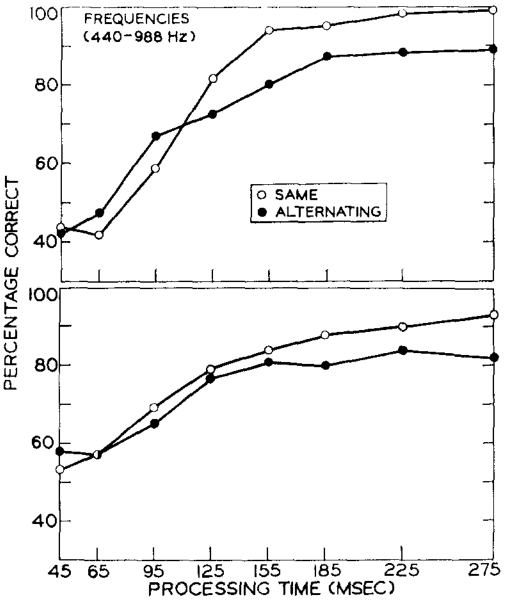

FIGURE 2. Percentage of correct counts of the number of tones in the test sequence (top panel) and correct probe duration recognitions (bottom panel) as a function of processing time for each tone and whether the tones were presented at the same frequency or alternated between two frequencies.

Results

The dependent measure is the percentage of correct identifications averaged over the four tone-sequence lengths. The rate of the presentation of the tones is given in terms of processing time expressed as the time between the onsets of successive tones. The top panel of Figure 2 plots the percentage or correct counts for the 8 observers as a function of processing time and whether the tones were presented at the same or alternating frequencies. The figure shows that average performance improved about 50% with increases in the processing time for each tone under both the same and alternating frequency conditions, $F(7, 49) = 97$, $p < .001$. Overall performance was better in the same than in alternating frequency conditions, $F(1, 7) = 6.17$, $p < .05$. The significant interaction between processing time and the same or alternating frequency conditions, $F(7, 49) = 8.35$, $p < .001$, reflects the consistent 10% advantage of the same frequency condition at the five slowest rates in contrast to no difference at the fastest rate and a 7% advantage of the alternating condition at processing times of 65 and 95 msec. The latter differences between the same and alternating condition were statistically significant, $F(1, 7) = 8.73$, $p < .025$, and $F(1, 7) = 10.02$, $p < .025$, respectively.

Although all eight observers showed some decrement in the alternating condition at the intermediate rates of presentation, there were some differences in counting performance among the eight observers. Four of them counted at essentially perfect accuracy under both the same and alternating conditions at the longest processing interval. The other four observers reached perfect performance under the same condition but showed a significant decrement in counting alternating tones even at the longest processing interval. The poor performance of these four subjects accounts for the overall 10% decrement at the 275-msec processing interval.

Table 2 presents the mean counting responses across the experimental conditions.

The results show that the subjects did not significantly underestimate the number of tones in the alternating frequency conditions. Given the extended practice of the observers with information feedback, it is not surprising that they were capable of maintaining an unbiased criterion and, thus, optimizing the percentage of correct identifications.

TABLE 2

Mean Counting Response as a Function of Processing Time for Each Tone and Whether the Tones Were Presented at the Same Frequency (Experiment 3)

Processing time (msec)	Attention condition	
	Same frequency	Alternating frequency
45	6.76	6.39
65	6.46	6.45
95	6.30	6.39
125	6.40	6.42
155	6.51	6.48
185	6.52	6.50
225	6.50	6.50
275	6.50	6.50

Discussion

The results show that the accuracy of counting a sequence of tones depends on whether the tones are presented at the same or alternating frequencies. Further, the difference between these two conditions appears to be dependent on the rate of presentation of the tones. At rates of 10–15 tones/sec, there is an advantage in counting tones alternating between frequencies. Counting performance is substantially better when the tones are presented at the same frequency at rates of 8 tones/sec or slower. The crossover in performance at processing times of 95 and 125 msec shows that this interaction is not due simply to the lower level of performance at the fast rates relative to the slow rates of presentation (cf. Figure 2, top panel).

The effect of processing time appears to function differently depending on whether the tones alternate between two different locations or two different frequencies. It appears that the differences between these two dimensions could be due to the particular values that were used rather than the dimensions themselves. Tones presented to the right and left ears are maximally separated, whereas the 440- and 988-Hz tones differed only by slightly more than an octave. In another series of experiments, the advantage of the alternating frequencies at the faster rates of presentation was replicated with 440- and 988-Hz tones but not when the alternating frequencies were 440 and 3,950 Hz (Massaro, Note 1). In the latter condition, the decrement in counting tones alternating between two frequencies appeared to emerge at a faster rate of presentation, analogous to counting tones alternating between the ears. This result suggests that the psychophysical difference between the alternating frequencies will influence the size of the decrement as a function of processing time. If this is the case, we might expect that alternating tones between two relatively close locations would produce results more in line with those seen in the top panel of Figure 2 rather than the top panel of Figure 1. Some evidence for this comes from Axelrod and Powazek's (1972) finding that difficulty of judging the rate of alternating clicks decreased with decreases in the spatial separation between the clicks. Given that the latter study employed only one rate of presentation, the interaction of spatial separation and processing time remains to be determined. The most reasonable conclusion appears to be that counting sounds alternating between different frequencies is very similar to counting sounds alternating between different spatial locations.

Analogous to counting tones alternating between the ears, subjects have more difficulty counting tones alternating between frequencies than tones presented at the same frequency. In order to locate the poor performance at either the perceptual or counting stage of processing, the next experiment required subjects to perceive the tones without counting them. Subjects monitored the sequence of tones for a longer or a shorter probe tone. The task was to indicate whether the probe tone was longer or shorter. Increasing the processing time for each tone should lead to better recognition of the probe tone. If subjects have more difficulty perceiving tones alternating in frequency relative to tones presented at the same frequency, then probe recognition should be poorer in the alternating than single frequency conditions. Experiment 4 provides a test of whether the deficit in counting tones alternating in frequency is due to a perceptual or a counting stage of processing.

EXPERIMENT 4

Method

Subjects. Five subjects participated 1 hour a day for 5 days. Four additional subjects participated for four days. Three other subjects were eliminated after one or two days because of extremely poor performance in the task. The subjects were paid for participating in the study or volunteered for extra credit in an introductory psychology course.

Procedure. The procedure was identical to Experiment 3 with the following exceptions. One of the tones in the sequence of tones was equally likely to be Δt (delta time) msec longer or shorter than the 20-msec test tones. The first and last tones in the sequence were never changed in duration. The value of Δt was adjusted and maintained at a level at which subjects performed an average of 75% correct. The Δt values were usually within the range of 6–10 msec on the four experimental days. The task was to indicate whether the probe tone was longer or shorter than the other tones in the sequence regardless of the number of tones in the sequence, the rate of presentation, or whether the tones were presented at the same frequency or alternated between the two frequencies. Subjects depressed one of two pushbuttons, labeled L and S, indicating whether the probe tone was longer or shorter. The visual feedback L or S indicated whether the probe tone was longer or shorter on that trial.

The results were taken from the last 4 days of the experiment for the first five subjects and from the last 3 days for the last four subjects. This gives roughly 38 observations per day at each of the 16 experimental conditions (8 processing times × alternating vs. same frequency). The dependent measure is the average percentage of correct longer and shorter judgments.

Results

The bottom panel of Figure 2 presents the percentage of correct recognitions of the probe tone for the nine observers. Recognition performance improved about 32% with increases in the processing time available for each tone, $F(7, 56) = 62$, $p < .001$. Overall performance was about 3% better when the tones came in at the same frequency than when they alternated in frequency, but this effect was mainly due to performance at the three longest processing intervals. Although the main effect of attention condition was not significant, the interaction between processing time and presentation conditions was statistically significant, $F(7, 56) = 5.71$, $p < .001$. Figure 2 shows some decrement in the probe monitoring task at the rates of presentation that gave a significant decrement in counting performance.

Discussion

Subjects had more difficulty monitoring tones for a duration change when the tones were alternated in frequency than when they were presented at the same frequency. The decrement in monitoring performance may be due to less stimulus information in the alternating frequency than in the single frequency case. Subjects may have made successive comparison judgments during the tone sequences in order to recognize the probe tone as longer or shorter than the other tones. The change in duration of the tone produces a timbre difference and the timbre of the longer probe tone is purer than that of the shorter probe tone. However, the timbre also depends on the frequency of the tone. When the tones come in at the same frequency and, therefore, the same timbre, it would be optimal to make successive comparisons between the timbre of the successive notes. In contrast, the tones alternating in frequency will be heard as having different timbres. In this case, it will be difficult to make successive comparison judgments in order to determine the duration of the probe tone. Therefore, the advantage of monitoring tones in the single frequency case might be due to the advantage of making successive comparisons of tones that have the same frequency. In this case, the difference in timbre of the probe tone is easier to recognize than in the case of alternating frequencies. It seems likely that the decrement of monitoring tones in the alternating frequency condition is not dependent on an attention mechanism but is simply due to a more difficult stimulus situation than the single frequency task. A similar argument can be made for monitoring tones alternating between the ears since the same tone may give different pitch perceptions to the two ears. This might account for the slight decrement in monitoring tones alternating between the ears in Experiment 2.

General Discussion

The results of the second experiment showed no facilitative effect of monitoring for a probe tone that could occur at a single spatial location relative to occurring at either of two spatial locations. The results support the idea that initial (primary) recognition of a tone cannot be enhanced by selectively attending to some auditory dimension. In contrast, subjects show a huge decrement in counting tones presented at different spatial locations relative to just one spatial location (Experiment 1). These results are compatible with the idea that deriving meaning from an auditory sequence can be enhanced by selectively attending to some auditory location (Massaro, 1976).

The second two experiments with tones alternating in frequency were not completely successful at locating the counting deficit at the initial perceptual or counting stage of information processing. The major problem was that the monitoring task does not provide a direct measure of the perceptual resolution of each test tone. Observers are able to utilize a successive comparison judgment in the single frequency condition but not in the alternating frequency condition. Even with this stimulus bias, however, subjects appear to show a larger decrement in counting sounds alternating between different frequencies relative to monitoring these tones for the probe tone.

The experiments can be interpreted in terms of Harvey and Treisman's (1973) processing load account of the tone monitoring and tone counting experiments. The authors argue that the two tasks require different degrees of perceptual processing and that the deficit in performance on the counting relative to the monitoring task can be explained by a simple processing load idea. In the monitoring task, the response can be triggered by the output of a single frequency analyzer (Treisman, 1969). The counting task, however, requires a conscious perception of the tones, which is dependent on a full analysis of the separate dimensions for each tone. Alternating the tones between the ears makes it more difficult to give a full analysis to each of the tones but does not disrupt the output of a single frequency analyzer.

Recent findings by Moray (1975) and Gilliom and Sorkin (1974) are compatible with the results of this series of experiments. Their subjects were able to monitor two different channels, such as spatial location or frequency, for a signal without a performance decrement relative to the single channel condition. Subjects in Moray's and Gilliom and Sorkin's tasks show a performance decrement on one channel in the two-channel task, however, when a signal is detected on the other channel. This means that detecting a signal on one channel causes an interrupt that reduces monitoring efficiency on the other channel. A performance decrement is found only when two signals can occur along both channels on the same trial. When two signals cannot occur simultaneously, subjects can monitor two spatial locations and two frequencies without a performance decrement relative to the single channel condition (Sorkin, Pohlmann, & Gilliom, 1973). Our tone-counting task is more analogous to those trials on which a signal is detected on one channel and further processing is required on the other channel. Subjects can terminate processing when they recognize the probe tone in the monitoring task but must process (count) all of the tones in the tone-counting task.

The present experiments can also be interpreted in terms of Posner's cost-benefit analysis (e.g., Posner & Snyder, 1975). In the tone-monitoring experiment, attention to a given location in space or to a particular tone frequency may not benefit perceptual processing. In the tone-counting experiment, however, subjects must attend to one spatial location or frequency to count the tone and then switch to another spatial location or frequency to count the following tone. The switching produces a significant cost (performance decrement) relative to the single channel condition. There is very little cost in the probe monitoring experiment, since it may not be necessary to switch attention to resolve the pitch or duration of the probe tone. One must switch attention, however, to count the sequence of tones that are perceived at different spatial locations or at different pitches.

We have attempted to argue that the decrement in counting sounds alternated across different frequencies or spatial locations occurs because the sounds are perceived at different pitches or spatial locations and not because of an initial failure to perceive each sound. The sounds are resolved in both the single and alternating presentations, since selective attention may not be possible within the auditory system at this level (Massaro, 1975a, 1975b). The decrement must be due to a process that has difficulty integrating and counting tones that are perceived at different spatial locations or frequencies.

REFERENCE NOTE

1. Massaro, D. W. *The effects of rate of presentation and frequency alternations on counting pure tones.* Unpublished manuscript, University of Wisconsin, 1976.

REFERENCES

Axelrod, S., & Powazek, M. Dependence of apparent rate of alternating clicks on azimuthal separation between the sources. *Psychonomic Science,* 1972, *26,* 217–218.

Broadbent, D. E. *Perception and communication.* New York: Pergamon Press, 1958.

Broadbent, D. E. *Decision and stress.* London: Academic Press, 1971.

Cheatham, P. G., & White, C. T. Temporal numerosity: III. Auditory perception of number. *Journal of Experimental Psychology,* 1954, *47,* 425–428.

Garner, W. R. The accuracy of counting repeated short tones. *Journal of Experimental Psychology,* 1951, *41,* 310–316.

Gilliom, J. D., & Sorkin, R. D. Sequential vs. simultaneous two-channel signal detection: More evidence for a high-level interrupt theory. *Journal of the Acoustical Society of America,* 1974, *56,* 157–164.

Guzy, L. T., & Axelrod, S. Interaural attention shifting as response. *Journal of Experimental Psychology,* 1972, *95,* 290–294.

Harvey, N., & Treisman, A. M. Switching attention between the ears to monitor tones. *Perception & Psychophysics,* 1973, *14,* 51–59.

Lechelt, E. C. Temporal numerosity discrimination: Intermodal comparisons revisited. *British Journal of Psychology,* 1975, *66,* 101–108.

Massaro, D. W. Preperceptual auditory images. *Journal of Experimental Psychology,* 1970, *85,* 411–417.

Massaro, D. W. Backward recognition masking. *Journal of the Acoustical Society of America,* 1975, *58,* 1059–1065. (a)

Massaro, D. W. *Experimental psychology and information processing.* Chicago: Rand McNally, 1975. (b)

Massaro, D. W. Perceptual processing in dichotic listening. *Journal of Experimental Psychology: Human Learning and Memory,* 1976, *2,* 331–339.

Massaro, D. W., & Cohen, M. M. Preperceptual auditory storage in speech recognition. In A. Cohen & S. G. Nooteboom (Eds.), *Structure and process in speech perception.* New York: Springer-Verlag, 1975.

Moray, N. A data base for theories of selective listening. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance, V.* New York: Academic Press, 1975.

Nealis, P. M., Engelke, R. M., & Massaro, D. W. A description and evaluation of light-emitting diode displays for generation of visual stimuli. *Behavior Research Methods and Instrumentation,* 1973, *5,* 37–40.

Posner, M. I., & Snyder, C. R. R. Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance, V.* New York: Academic Press, 1975.

Pisoni, D. B. Perceptual processing time for consonants and vowels. *Haskins laboratories status report on speech research, SR 31/32,* 1972, 83–92.

Sorkin, R. D., Pohlmann, L. D., & Gilliom, J. D. Simultaneous two-channel signal detection. III. 630- and 1400-Hz signals. *Journal of the Acoustical Society of America,* 1973, *53,* 1045–1050.

Treisman, A. M. Strategies and models of selective attention. *Psychological Review,* 1969, *76,* 282–299.

White, C. T. Temporal numerosity and the psychological unit of duration. *Psychological Monographs,* 1963, *77*(12, Whole No. 575).

(Received November 21, 1975)