Distractor clustering enhances detection speed and accuracy during selective listening

CLAUDE ALAIN and DAVID L. WOODS Neurosciences Center, University of California, Davis, California and VA Medical Center, Martinez, California

The effects of distractor clustering on target detection were examined in two experiments in which subjects attended to binaural tone bursts of one frequency while ignoring distracting tones of two competing frequencies. The subjects pressed a button in response to occasional target tones of longer duration (Experiment 1) or increased loudness (Experiment 2). In evenly spaced conditions, attended and distractor frequencies differed by 6 and 12 semitones, respectively (e.g., 2096-Hz targets vs. 1482- and 1048-Hz distractors). In clustered conditions, distractor frequencies were grouped; attended tones differed from the distractors by 6 and 7 semitones, respectively (e.g., 2096-Hz targets vs. 1482- and 1400-Hz distractors). The tones were presented in randomized sequences at fixed or random stimulus onset asynchronies (SOAs). In both experiments, clustering of the unattended frequencies improved the detectability of targets and speeded target reaction times. Similar effects were found at fixed and variable SOAs. Results from the analysis of stimulus sequence suggest that clustering improved performance primarily by reducing the interference caused by distractors that immediately preceded the target.

The clustering of frequencies of attended and nonattended tones has important effects on target detectability in auditory selective attention tasks. For example, Bregman and Rudnicky (1975) presented subjects with auditory sequences consisting of target tones preceded and followed by lower frequency distractors. The distractors were in turn preceded and followed by "captor" tones that varied in their similarity to the distractors. The captor tones were either lower in frequency or identical to the distractors. Similarity in frequency between the captor tones and distractors improved order judgments of the target tones, consistent with the segregation of the sequence into separate auditory streams. That is, the similarity of captors and distractors allowed them to be grouped into a separate perceptual stream from the targets, thereby facilitating target discrimination. The magnitude of the captor effect has been shown to vary with temporal predictability (Jones, Kidd, & Wetzel, 1981).

However, even without the explicit manipulation of physical clustering, the characteristics of tones preceding the target can influence the speed and accuracy of target detection (Woods, 1990). For example, when target tones are preceded by tones with similar location or frequency, reaction times (RTs) are faster than when they are preceded by tones sharing neither target feature (Woods, Alho, & Algazi, in press). Responses are further speeded following tones with both features (Woods, 1992; Woods & Alain, 1993). These priming effects, which can speed RTs by 50 msec or more, are seen for tones that occur up to 500 msec before the target. In previous clustering studies, distractors and targets were typically preceded by a short set of captor tones that varied in their similarity with the distractors and targets (e.g., Bregman & Rudnicky, 1975; Jones et al., 1981). As a result, the captor tones may have acted directly as priming stimuli, and thus were more effective in clustering conditions due to their similarity in frequency to the target.

Furthermore, because a small number of sounds presented in a predictable, repeating pattern has been used in previous studies of distractor clustering, it remains unclear whether the effect of clustering can be generalized to more complex listening situations such as random-tone sequences. In the current study, physical clustering was investigated by using unpredictable, fully randomized stimulus sequences, and both fixed and variable stimulus onset asynchronies (SOAs).

The effect of clustering on detection speed also deserves investigation. If clustering affects the sensory processing of auditory signals, it would be expected that the speed and accuracy of target detection would be similarly affected. However, without direct investigation, one cannot eliminate the possibility of tradeoffs between the accuracy and speed of sensory analysis (Luce, 1986). For example, subjects could adopt a lower response criterion in difficult conditions, sacrificing accuracy to maintain constancy in RTs.

The current study aimed to (1) investigate the interaction between physical clustering and priming; (2) explore the distractor clustering effect in fully randomized tone se-

This research was supported by grants from the NIDCD and the VA Research Service to D. L. Woods, and by a postdoctoral fellowship provided by the Quebec Government Postdoctoral Fellowship (FCAR) to the first author. Special thanks to Greg Shenaut and John Lackey for software development. Also, we would like to thank L. Nielsen-Bohlman, D. Scabini, L. Robertson, B. Yund, A. Achim, and K. Ogawa for helpful comments on the earlier drafts. Address correspondence to C. Alain, Department of Neurology, VA Medical Center, 150 Muir Road, Martinez, CA 94553 (e-mail: marva4!calain@ucdavis.edu).

quences of fixed and variable SOAs; and (3) examine the relation between RTs and accuracy in different clustering conditions. Two experiments were conducted in which subjects were asked to attend to tones of one frequency while ignoring the distracting tones of two competing frequencies. In both experiments, the tones were presented in a random order at fixed or variable SOAs. In Experiment 1, the subjects responded to infrequent, slightly longer duration targets mixed with standard tones. In Experiment 2, the targets were slightly louder than the standards. Target discriminability was adjusted in pilot studies to encourage a hierarchical selection, so that the subjects would first pay attention to the frequency (easy dimension) and then make judgments on the duration or intensity (more difficult dimension). Deviant tones were also presented at the irrelevant frequencies (e.g., same duration or intensity as the targets, but differing in frequency).

EXPERIMENT 1

In this experiment, we tested whether clustering of irrelevant frequencies would enhance the detection of occasional longer duration tones in the attended frequency. In the unclustered conditions, attended and distractor frequencies differed by 6 and 12 semitones, respectively. In the clustered conditions, distractor frequencies were grouped, and the attended tones differed from the distractors by 6 and 7 semitones, respectively. It was thought that the increased similarity between the two irrelevant frequencies would favor the segregation of the sequences into two substreams. Therefore, distractor clustering should help the monitoring of the attended frequency and thus facilitate target detection.

The tones preceding the targets varied randomly, so priming could be evaluated by examining target detection as a function of the frequency of the tone preceding the target. The interaction between priming and clustering effects could be evaluated by examining the influence of preceding stimuli on RT across conditions. If clustering mainly reduces distractibility, then we would expect larger clustering effects for targets preceded by tones of a different frequency (Bregman & Rudnicky, 1975). However, if clustering mainly enhances priming, then we might expect larger clustering effects for targets preceded by tones of identical frequency.

Method

Subjects. Sixteen paid subjects (4 females, 12 males) between 20 and 33 years of age participated. One additional subject was excluded because of poor performance during the training session.

Stimuli. The stimulus sequences, lasting approximately 10 min, consisted of 3,240 tones of three different frequencies. The three frequencies used in the different clustering conditions are shown in Figure 1. Distractor tones of 1482 Hz were presented on every block. In the evenly spaced conditions (ES), the three frequencies were equally spaced by 6 semitones (1048, 1482, and 2096 Hz, see Figure 1, top). In the clustered conditions (CL), the middle tone was paired with a distractor that was either 1 semitone lower (1482).



Figure 1. Diagram of a segment of the auditory sequences presented in different clustering conditions. Top: Evenly spaced condition, attend high pitch. Middle: Clustered condition, attend high pitch. Bottom: Clustered condition, attend low pitch. A = extreme high frequency, B = middle frequency, C = extreme low frequency. The frequency monitored in each condition is indicated with an arrow. Targets are shown with an asterisk.

and 1400 Hz, Figure 1, center) or 1 semitone higher in frequency (1482 and 1570 Hz, Figure 1, bottom).

The stimulus sequences were constructed such that each of the three frequencies occurred on 33.33% of the trials—as a short tone (40 msec) on 30.83% and as a long tone (65 msec) on 2.5%. Overall, long tones occurred in 7.5% of the trials. All the stimuli were shaped with 5-msec rise/fall times and delivered binaurally through TDH-39 headphones at 78 dB SPL. Stimulus SOAs were either fixed (onset-to-onset, 180 msec) or variable (140–240 msec with 20-msec step, M = 190 msec) in different blocks. Because auditory streaming is sensitive to the onset-to-onset interval between tones (Dannenbring & Bregman, 1976), clustering effects were expected to be more prominent in fixed SOA conditions.

Task. The subjects' task was to attend to either the low- (1048 Hz) or high- (2096 Hz) frequency tones and to respond to occasional long-duration tones at that frequency (targets). Before beginning the study, each subject practiced listening to tone sequences and detecting occasional targets. Following the practice blocks, the subjects performed in eight conditions, defined by all possible combinations of temporal regularity (fixed or variable SOAs), attended frequencies (high or low), and clustering conditions (evenly spaced or clustered). The order of conditions was counterbalanced across subjects.

The subjects were seated in a sound-attenuated room. They were instructed to maintain their attention on either the low- or highpitched tones and to ignore the other two frequencies. The subjects were told to press a button as quickly as possible when they heard occasional longer duration tones of the designated frequency (targets) while trying to make as few errors as possible. They were informed that long tones would also occur at the irrelevant frequencies.

Data analysis. Reaction times were analyzed for correct trials only, defined as buttonpresses occurring between 200 and 800 msec after target presentation. Responses outside this time window were classified as false alarms. For each subject, RTs were compared across conditions using a three-way analysis of variance (ANOVA) for repeated measures, with temporal regularity, attended frequency, and clustering types as factors. The number of misses and the number of false alarms were similarly analyzed.

Results

Accuracy. On the average, false alarms constituted 8% of all reponses and their incidence did not differ across conditions. The top panel of Figure 2 shows the effect of distractor clustering on detection accuracy. The subjects were more accurate in detecting targets in CL than in ES conditions [12.64% vs. 16.43% of misses, respectively; F(1,15) = 13.14, p < .01]. The clustering × temporal regularity interaction was not significant [F(1,15) = 2.21], suggesting that the effect of clustering on accuracy was similar at both fixed and variable SOAs. There was no significant difference between the detectability of lowand high-frequency targets.

Reaction time. The bottom panel of Figure 2 shows the effects of tonal grouping on RTs. Overall, RTs were faster in CL than in ES conditions [441 vs. 452 msec, respectively; F(1,15) = 9.19, p < .01]. Although the clustering effect appears to be reduced at fixed SOAs, the clustering × temporal regularity interaction did not reach significance [F(1,15) = 1.84]. Responses were faster for 1048-Hz than for 2096-Hz targets [438 vs. 456 msec, respectively; F(1,15) = 8.89, p < .01]. No other main effects or interactions reached significance.

Clustering/priming. To evaluate the relationship between clustering and priming, the effects of preceding stimuli on RTs were compared across conditions. This



Figure 2. Top: Proportion of misses to low- or high-frequency targets for each SOA type (fixed and variable SOAs). Bottom: Reaction time to occasional low- or high-frequency targets for each SOA type (fixed and variable SOAs) and for both conditions.

was done with a four-way ANOVA for repeated measures with preceding tone type (extreme, middle, or same), attended frequency, temporal regularity, and clustering as factors. Figure 3 shows target RTs sorted by the frequency of the immediately preceding tone (T-1; about 180 msec before the target). Overall, RTs were significantly faster when targets were preceded by nontargets of the same frequency than when they were preceded by nontargets of different frequencies [447 vs. 458 msec respectively; F(2,30) = 7.68, p < .01].

A significant clustering \times preceding tone type interaction was found. Response times were similar for targets preceded by tones of the same frequency in CL and ES conditions [448 vs. 445 msec, respectively; F(1,15) =0.31]. In contrast, RTs were faster in CL than ES conditions when targets were preceded by tones of the extreme frequency [451 vs. 466 msec; F(1,15) = 11.85, p < .01]. Responses to targets preceded by middle-frequency tones also showed a clustering effect [452 vs. 462 msec; F(1,15) =4.91, p < .05]. Further specific comparisons showed that clustering effects were significantly reduced when the target followed a tone of identical frequency [F(1,15) =11.41, p < .01].

In a previous study, Woods and Alain (1993) reported that the features of a stimulus occurring two positions before the target (T-2) influenced the speed of target detection. This raises the possibility that the stimulus at the T-2 position may have acted directly as a priming stimulus in the current experiment. The priming would have been more effective in clustering conditions because the extreme tones were closer in frequency to the targets. To examine this possibility, a post hoc analysis was performed when the target was preceded by a pair of middle tones, which were identical across clustering conditions. Responses were faster in CL than in ES conditions when the target was preceded by two successive middle tones [451 vs. 473 msec; F(1,15) = 5.02, p < .05].

Discussion

Three main findings were observed in Experiment 1. First, it was found that distractor clustering improved the accuracy of target detection (cf. Bregman & Rudnicky, 1975; Jones et al., 1981). Second, clustering speeded target detection. Third, clustering effects were similar at fixed and variable SOAs. These data show that distractor clustering may facilitate both the speed and accuracy of target detection in a variety of complex listening situations.

A sequential analysis provided some additional information on the nature of the clustering effect. When targets were preceded by tones of a different frequency, RTs were slower in ES than in CL conditions. However, clustering had only a small effect when the preceding stimulus shared the frequency of the target. Thus, in the present experiment, the clustering effect on RT appeared mediated mainly by a diminution of interference caused by the distracting tones at nonattended frequencies.

Our data also indicate that the effect of distractors on performance did not depend solely on their physical similarity with the target. Clustering speeded RTs to targets



Figure 3. Reaction times to targets immediately preceded by a tone that was extreme, middle, or of the same frequency as the target.

preceded by middle tones which were identical across conditions. Therefore, priming alone cannot account for the RT difference across conditions. In the current study, the effect of the middle tone on RT must depend on its relationship with those earlier in the sequence. One possibility is that the increase in similarity between the two irrelevant frequencies allowed them to be grouped into a separate perceptual stream from the relevant frequency, thereby facilitating target discrimination.

EXPERIMENT 2

The similar effect of clustering on performance when tones were presented at fixed or variable SOAs was an unexpected result of Experiment 1. One possible explanation is that clustering is largely determined by tonal similarity and is little influenced by the temporal regularity. However, another possibility is that the long-duration tones in Experiment 1 disrupted the temporal regularity of the stimulus sequence by randomly decreasing the interstimulus interval. To further evaluate the interaction between the temporal regularity and distractor clustering, a second experiment was performed using target and standard stimuli of identical durations.

Method

Subjects. Twelve paid subjects (6 females, 6 males) between 20 and 36 years of age participated in the second experiment, including 3 that had participated in the first. Seven additional subjects were excluded due to poor performance during the training session.

Task. Experiment 2 was identical to Experiment 1, except that the occasional long-duration tones were replaced by occasional louder stimuli. For each frequency, there were frequent soft tones (78 dB SPL) and occasional loud tones (85 dB SPL) which were presented with the same proportion as the standards and deviants in Experiment 1. The subjects received the same instructions as in the first experiment: they were told to press a button as quickly as possible when they heard the occasional louder tones of the designated frequency (targets) while trying to make as few errors as possible. The subjects were informed that loud tones might occur at the irrelevant frequencies.

Results

Accuracy. Overall, the task in Experiment 2 was more difficult than in Experiment 1, as evidenced by the number of misses (23.20% vs. 14.53%). False alarms constituted 16% of all responses and, as in Experiment 1, did not differ across conditions. The effect of clustering on detection accuracy was similar to the first experiment (see Figure 4, top). The subjects were more accurate in the CL than in the ES conditions [21.20% vs. 25.21% of misses, respectively; F(1,11) = 5.27, p < .05]. A significant clustering \times frequency interaction was found [F(1,11) = 14.49, p < .01]. Further analysis revealed that clustering-related improvements in detection accuracy were more prominent for low-frequency tones [6.82% fewer misses in the CL than the ES condition; F(1,11) =11.29, p < .01 than for high-frequency tones [1.19%] fewer misses in the CL than the ES condition; F(1,11) =0.47]. As in the first experiment, the clustering \times temporal regularity interaction was not significant [F(1,11) =0.631.



Figure 4. Top: Proportion of misses to low- or high-frequency targets for each SOA type (fixed and variable SOAs). Bottom: Reaction times to occasional low- or high-frequency targets for each SOA type (fixed and variable SOAs) and for both conditions.

Reaction time. The bottom panel of Figure 4 shows the clustering effect for target frequency at fixed and variable SOAs. Overall, RTs showed a significant effect of clustering, being faster in the CL than in the ES conditions [406 vs. 427 msec, respectively; F(1,11) = 12.41, p < .01]. As in the first experiment, the clustering × temporal regularity interaction was not significant. Responses were faster for 2096-Hz than for 1048-Hz tones [404 vs. 429 msec, respectively; F(1,11) = 12.23, p < .01]. Although the RTs tended to be faster when the tones were presented at variable versus fixed SOAs, the difference failed to reach significance [F(1,11) = 2.79].

Clustering/priming. Figure 5 shows the effect of preceding tone type on RT for both conditions. As in Experiment 1, RTs were faster when the tone in the T-1 position shared frequency with the target than when it had a different frequency [418 vs. 432 msec, respectively; F(2,22) = 9.85, p < .01]. Priming effects tended to be larger in the ES (18 msec) than in the CL conditions (11 msec), but the clustering × preceding tone type interaction failed to reach significance [F(2,22) = 1.49].

In contrast with the first experiment, clustering did not interact with priming effects. Further analysis showed that the effect of clustering was similar for targets preceded by tones of identical frequency [F(1,11) = 6.26, p < .05], or different frequencies [extreme, F(1,11) = 11.00, p < .01; middle, F(1,11) = 14.05, p < .01].

As in Experiment 1, the effect of clustering was tested when the targets were preceded by a pair of middle tones. Responses were faster in the CL than in the ES condition when the target was preceded by two successive middle tones [417 vs. 442 msec, respectively; F(1,11) = 5.12, p < .05].

Discussion

Experiment 2 replicated the major findings of Experiment 1 with intensity instead of duration targets. Both speed and accuracy were improved with distractor cluster-



Figure 5. Reaction times to targets preceded by a tone that was extreme, middle, or of the same frequency as the target.

ing. As in the first experiment, no interaction between the clustering condition and temporal regularity was found. This suggests that the variation in SOA had little influence on distractor clustering, at least in random-tone sequences.

Again, physical clustering reduced the effect of distractors on RTs. RTs were always slower in the ES than in the CL conditions when the targets were preceded by extreme or middle tones. However, in contrast to the first experiment, clustering also modulated target RTs preceded by tones of identical frequency. This suggests that distractor clustering may enhance priming, or at least may be mediated in part by mechanisms that are primingdependent.

As in the first experiment, distractor clustering reduced RTs to targets preceded by middle tones. Because these were constant, the result cannot be solely explained by a general effect of physical similarity of the preceding stimuli.

GENERAL DISCUSSION

The results from the present experiments indicate that distractor clustering improved signal detection. The analysis of the tone sequences showed that the improvement due to distractor clustering was not solely mediated by priming effects. If clustering was mediated solely by priming, then RTs to targets preceded by tones of the middle frequency (kept constant in different clustering conditions) should have been identical. However, clustering did interact to some extent with the sequence of preceding tonesits effect being larger when the target followed tones of irrelevant frequencies. Thus, in the present study, distractor clustering enhanced performance mainly by decreasing interference of irrelevant frequencies. The results from the present study are consistent with previous reports and extend the effect of physical clustering to a variety of listening conditions (Bregman & Campbell, 1971; Bregman & Rudnicky, 1975).

The effects of distractor clustering were similar in temporally predictable and unpredictable sequences. Our demonstration of the clustering effect in fully randomized sequences is consistent with previous results showing that grouping, based on tonal similarity, can occur with sequences that contain some temporal variability (Bregman, 1990; French-St. George & Bregman, 1989; Handel, Weaver, & Lawson, 1983; Tougas & Bregman, 1985). However, our results differ from what one would predict from previous reports that suggest that perceptual grouping should be more prominent at fixed SOAs (Dannenbring & Bregman, 1976; Jones, 1976; Jones et al., 1981). Unlike previous studies in which short sets of tones were presented in predictable repeating patterns, in the current study we presented tones in a random order. This had the effect of making the recognition of any particular pattern more difficult. Thus, when tones are presented in a random order, it appears that temporal regularity has less effect on performance than does the physical similarity of distracting frequencies.

As in previous studies, overall RTs showed a relative facilitation when the target tones followed nontargets of the same frequency (Woods, 1990; Woods, 1992; Woods & Alain, 1993; Woods et al., in press). This priming effect occurred whether target and nontarget tones were distinguished by duration (Experiment 1) or by intensity (Experiment 2). However, the degree to which clustering can enhance priming remains unclear because there was little additional clustering effect on the targets preceded by tones of identical frequency in Experiment 1, whereas there was a significant clustering effect in Experiment 2.

Another aspect of the current study was to examine the relation between detection speed and accuracy. In both Experiments 1 and 2, speed of target detections and accuracy were improved with clustering. This is the first demonstration of such an effect. This is consistent with the study by Alain, Achim, and Richer (in press), in which event-related brain potentials revealed enhanced attention effects in clustering conditions at latencies as short as 120 msec. This is also consistent with recent results from our laboratory showing larger attention effects in event-related brain potentials and faster RTs when distractor frequencies are clustered together (Alain & Woods, 1993).

The mechanisms whereby auditory inputs are segregated into separate streams remains unclear. In the current study, it is possible that clustering effects may be mediated by an active inhibition of the auditory processing of distracting frequencies (Woods, 1990). Clustering might make this inhibition easier to maintain, because both nearby nontarget frequencies could be inhibited simultaneously. Alternatively, physical clustering may increase the mutual cross-habituation of distractors, thereby reducing their salience (Cowan, 1988).

In summary, our results show that the effect of distractor clustering acts similarly to improve the speed and accuracy of tone identification. This is consistent with previous studies (Bregman & Rudnicky, 1975; Jones et al., 1981) and reveals that these effects of physical clustering extend to random-tone sequences. Further, our results show that physical clustering is unaffected by temporal regularity, at least when tones are presented in a random order. Finally, the analysis of stimulus sequence suggests that clustering affected performance mainly by reducing the interference caused by irrelevant frequencies. This analysis also reveals that physical clustering effects cannot solely be explained in terms of priming, but must also involve other processes such as perceptual grouping.

REFERENCES

- ALAIN, C., ACHIM, A., & RICHER, F. (in press). Perceptual context and the auditory selective attention effect on event-related brain potentials. *Psychophysiology*.
- ALAIN, C., & WOODS, D. L. (1993). Modulations in auditory cortical activity with signal clustering in humans. Manuscript in preparation.
- BREGMAN, A. S. (1990). Auditory scene analysis: The perceptual organization of sound. Cambridge, MA: MIT Press.
- BREGMAN, A. S., & CAMPBELL, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal* of Experimental Psychology, 89, 244-249.
- BREGMAN, A. S., & RUDNICKY, A. I. (1975). Auditory segregation: Stream or streams? Journal of Experimental Psychology: Human Perception & Performance, 1, 263-267.
- COWAN, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, **104**, 164-191.
- DANNENBRING, G. L., & BREGMAN, A. S. (1976). Effect of silence between tones on auditory stream segregation. *Journal of the Acousti*cal Society of America, 59, 987-989.
- FRENCH-ST. GEORGE, M., & BREGMAN, A. S. (1989). Role of predictability of sequence in auditory stream segregation. *Perception & Psychophysics*, 46, 384-386.
- HANDEL, S., WEAVER, M. S., & LAWSON, G. (1983). Effect of rhythmic grouping on stream segregation. Journal of Experimental Psychology: Human Perception & Performance, 9, 637-651.
- JONES, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323-355.
- JONES, M. R., KIDD, J. G., & WETZEL, R. (1981). Evidence for rhythmic attention. Journal of Experimental Psychology: Human Perception & Performance, 7, 1059-1073.
- LUCE, R. D. (1986). Response time: Their role in inferring elementary mental organization. New York: Clarendon.
- TOUGAS, Y., & BREGMAN, A. S. (1985). Crossing of auditory streams. Journal of Experimental Psychology: Human Perception & Performance, 11, 788-798.
- WOODS, D. L. (1990). The physiological basis of selective attention: Implications of event-related potential studies. In J. W. Rohrbaugh, R. Parasuraman, & R. Johnson (Eds.), *Event-related brain potentials: Basic issues and applications* (pp. 178-209). New York: Oxford University Press.
- WOODS, D. L. (1992). Auditory selective attention in middle-aged and elderly subjects: An event-related brain potential study. *Electroen*cephalography & Clinical Neurophysiology, 84, 456-468.
- WOODS, D. L., & ALAIN, C. (1993). Feature processing during highrate auditory selective attention. *Perception & Psychophysics*, 53, 391-402.
- WOODS, D. L., ALHO, K., & ALGAZI, A. (in press). Stages of auditory feature conjunction: An event-related brain potential study. Journal of Experimental Psychology: Human Perception & Performance.

(Manuscript received August 27, 1992; revision accepted for publication March 5, 1993.)