

The relation between auditory temporal interval processing and sequential stream segregation examined with stimulus laterality differences

SUSAN E. BOEHNKE and DENNIS P. PHILLIPS
Dalhousie University, Halifax, Nova Scotia, Canada

In this study, we examine the effects of laterality differences between noise bursts on two objective measures of temporal interval processing (gap detection and temporal asymmetry detection) and one subjective measure of temporal organization (stream segregation). Noise bursts were lateralized by presentation to different ears or dichotic presentation with oppositely signed interaural level (ILD) or time (ITD) differences. Objective thresholds were strongly affected by ear-of-entry differences, were moderately affected by ILD differences, but were unaffected by ITD differences. Subjectively, A and B streams segregated well on the basis of ear-of-entry or ILD differences but segregated poorly on the basis of ITD differences. These results suggest that perceptual segregation may be driven more effectively by differential activation of the two ears (peripheral channeling) than by differences in perceived laterality.

Our sonic environment consists of events from different sources distributed in space and time. Our perceptual system groups acoustic events that presumably emanate from a single source and segregates those from different sources into separate perceptual streams (see Bregman, 1990). Such processes are likely involved in the cocktail party phenomenon, wherein we selectively attend to a sound source of interest within a multisource environment. In this study, we are interested in how spatial location attributes might be used for segregation. To do so, we exploit an aspect of the perceptual organization of temporally interleaved sounds: that temporal judgments are impaired when they must be made between sounds that are perceived to emanate from different distal sources (Bregman & Campbell, 1971; Warren, Obusek, Farmer, & Warren, 1969). Thus, degradation in performance on temporal judgments made *between* two sounds that vary on some dimension (e.g., location or pitch) might be used as an index by which to determine the relevance of that dimension in stream segregation. Thresholds for the detection or discrimination of auditory temporal gaps (a silent interval between two sounds) may provide a simple measure of auditory temporal acuity for this purpose. In a variant of gap detection, termed *between-channel* gap detection (see Phillips, Taylor, Hall, Carr, & Mossop, 1997), the effects

on gap thresholds of varying the properties of the sounds preceding and following the silent period have been explored. The two properties that have the most striking effect on gap detection are spectral differences between the gap markers (Divenyi & Danner, 1977; Fitzgibbons, Polatsek, & Thomas, 1974; Formby & Forrest, 1991; Formby, Sherlock, & Li, 1998; Grose, Hall, Buss, & Hatch, 2001; Hall, Grose, & Joy, 1996; Kinney, 1961; Neff, Jesteadt, & Brown, 1982; Penner, 1977; Perrott & Williams, 1971; Phillips & Hall, 2000, 2002; Phillips et al., 1997; Taylor, Hall, Boehnke, & Phillips, 1999; Williams & Elfner, 1976; Williams, Elfner, & Howse, 1978) and ear-of-entry differences (Divenyi & Danner, 1977; Penner, 1977; Phillips et al., 1997; Taylor, Hall, Boehnke, & Phillips, 1999). Spectral and ear-of-entry differences between interleaved sounds are also those that have the strongest influence on auditory sequential stream segregation (Bregman, 1990; Hartmann & Johnson, 1991; van Noorden, 1975).

In the gap detection task, when the stimuli bounding the silent period are presented to different ears, gap thresholds are often 5–10 times higher than those obtained when the same gap markers are presented to a single ear, which can be as low as 2 msec (Formby et al., 1998; Penner, 1977; Phillips et al., 1997). Similarly, gap thresholds were 5–10 times longer when the markers bounding the silent period originated from free-field sources opposite each ear (i.e., at $\pm 90^\circ$ azimuth) than when both of the markers originated from the same source (Boehnke & Phillips, 1999; Phillips, Hall, Harrington, & Taylor, 1998). Two sounds originating from different free-field spatial locations result in different values of the binaural cues, which lead to lateralized percepts. These cues include interaural time difference (ITD) and interaural level difference (ILD). The two sounds will also differ monaurally in intensity

This research was supported by a Natural Sciences and Engineering Research Council (NSERC) grant to D.P.P., and by NSERC and I. W. Killam doctoral fellowships to S.E.B. It is based on part of the doctoral dissertation of S.E.B., parts of which have been previously published in abstract form. Thanks to Susan Hall for helpful comments on the manuscript. Correspondence concerning this article should be addressed to S. E. Boehnke, Centre for Neuroscience Studies, Queen's University, Kingston, ON, K7L 3N6 Canada (e-mail: susan@biomed.queensu.ca).

at each ear, and monaural intensity differences have long been known to lead to elevated gap thresholds (Penner, 1977; Plomp, 1964), most likely due to temporal masking processes (see Smiarowski & Carhart, 1975). This raises the question as to whether it is the perceptual change in “location” between the two sounds that leads to elevated thresholds in the free-field case or the difference with which the sounds activate the frequency channels of each ear. These two possibilities were dissociated in a dichotic gap detection paradigm by Oxenham (2000). Marker differences in ITD alone failed to have an effect on gap thresholds for noise stimuli. ILD differences between leading and trailing markers elevated gap thresholds, but they did so no more than did level differences of the same size as those at a single ear. This finding suggests that it is the differential activation of peripheral auditory filters, rather than the change in perceived laterality per se, that degrades performance on gap detection. Although they have been studied separately, no study has compared the absolute values of gap detection thresholds from the same subjects when the markers differ in ear of entry, ILD, or ITD alone.

Still more helpful would be knowledge of whether the effects of those spatial attributes generalize to performance in other, arguably more sophisticated, temporal-processing tasks that have been directly linked to stream segregation. Another temporal-processing paradigm, the temporal asymmetry detection task (Vliegen, Moore, & Oxenham, 1999), has recently been introduced as an objective measure of auditory sequential stream segregation (SSS; see Bregman, 1990). A simple demonstration of SSS is accomplished by presenting a sequence of two alternating sounds (A and B) differing in some attribute (usually tonal frequency), in the temporal sequence A_B_A_A_B_A . . . Note that SSS refers to the situation in which the two components do not overlap in time. A related phenomenon, *concurrent stream segregation*, refers to segregation that occurs when the two components are presented simultaneously. In the sequential case, when A and B are similar, listeners report perception of a galloping rhythm (single-stream percept or auditory fusion). As the difference between A and B is increased, listeners are more likely to report that the sequence segregates into two separate A and B streams (two-stream percept, segregation, or fission). For a range of intermediate A and B differences, it is common for listeners to experience a bistable percept (van Noorden, 1975). In the temporal asymmetry paradigm, a repeating sequence of A and B sounds, as described above, is presented, except that the temporal location of B with respect to A is systematically varied. The task of the subject is to distinguish a standard sequence in which B is centered in time between the two A sounds from a sequence in which B has been temporally shifted toward one of the A sounds. The measured threshold in this objective task represents the smallest detectable asymmetry in the duration of the silent intervals between the A_B and the B_A components. The task is easier and produces lower thresholds when a galloping rhythm is perceived than when A and B are perceived to segregate into separate perceptual streams (Vliegen et al., 1999). As such, this task is thought to provide an *objec-*

tive index of sequential stream segregation. Note that in the temporal asymmetry task, the listener is effectively biased to hear a single-stream percept (integration) in order to do the task. It is well known that mental set can modulate the emergence of a one- or two-stream percept (van Noorden, 1975).

Akin to the elevation of gap thresholds that occurs when the markers differ spectrally, temporal asymmetry detection thresholds were elevated as the A and the B components were made to differ spectrally (Vliegen et al., 1999). Furthermore, elevation of thresholds as a function of A-B spectral separation mapped onto the amount of phenomenal stream segregation reported by listeners for such A-B differences (Vliegen & Oxenham, 1999), despite differences in task demands (bias to hear two streams). Although this is suggestive of a link between performance on gap detection and temporal asymmetry detection, no study has ever directly compared thresholds for temporal asymmetry detection with those for gap detection, using similar stimuli. The present study seeks to determine whether this relation between these tasks holds true for conditions in which the components differ in spatial stimulus attributes (Oxenham, 2000; Phillips et al., 1997). That is, because a disparity between leading and trailing gap markers in ILD, but not in ITD, elevates gap thresholds, will A-B disparities in ILD, but not in ITD, elevate temporal asymmetry thresholds? If so, might A and B differences in ITD lead to less segregation than do similar differences in ILD when a more traditional subjective measure of auditory stream segregation is used?

We examined these questions in a set of three experiments. In all three experiments, we manipulated the same set of spatial cues; these were ear of entry (see Penner, 1977; Phillips et al., 1997), ILD, and ITD (see Oxenham, 2000), and differences in level between markers at one ear (to control for monaural contributions to the effects of ILD; see Oxenham, 2000; Penner, 1977). In Experiment 1, we reexamined the effects on gap detection thresholds of imposing spatial cue differences on two noise markers bounding a gap. In Experiment 2, we used the auditory temporal asymmetry task (described above), which involves recurrent sequences of two different (A, B) noise markers. Finally, Experiment 3 was a traditional *subjective* stream segregation study (see Bregman, 1990), in which listeners tracked the development of perceptual streams prompted by differentiating A and B noises of a repeating ABA_ sequence. In Experiments 2 and 3, it is reasonable to assume that the task demands differently biased listeners toward hearing one- and two-stream percepts, respectively.

EXPERIMENT 1 Gap Detection

The purpose of this experiment was to replicate and extend the findings of Oxenham (2000). Experiment 1 was thus a dichotic study of gap detection thresholds obtained using wideband noise stimuli in which the leading and the trailing markers differed in their spatial attributes. There were six experimental conditions: The leading and trailing noise markers were presented to different ears (the *Ear*

condition) or with different values of *ITD* (± 0.5 msec) or *ILD* (± 15 dB) or with different intensities presented monaurally (*loud-quiet* [LQ] or *quiet-loud* [QL], to control for the stimuli presented to each ear in the *ILD* condition), or both markers were presented diotically (the *Diotic* condition).

The chosen values of *ILD* and *ITD* were important, if somewhat arbitrary. The maximum *ITDs* generated by the human head are on the order of 600–700 μ sec, depending on head size (Middlebrooks & Green, 1991), and the maximum *ILDs* are on the order of 30 dB (at high frequencies; Middlebrooks, Makous, & Green, 1989). For tonal stimuli, the perceived intracranial lateral displacement of a source saturates after 90° of phase angle, irrespective of frequency, and it usually saturates after 15–20 dB *ILD*, irrespective of frequency (Yost, 1981). The choices of a 500- μ sec *ITD* and a 15-dB *ILD* thus guaranteed highly lateralized noise stimuli in both dichotic conditions. Whether the *ITD* and the *ILD* stimuli were *equivalently* lateralized is another matter. Harris (1960) showed that the time–intensity trading relation for their broadest band of noise (7000-Hz bandwidth) was about 0.04 msec/dB; using this trading ratio, a 15-dB *ILD* corresponds to 600 μ sec. Yost, Tanis, Nielsen, and Bergert (1975) provided psychophysical evidence suggesting that for noise stimuli, an *ITD* of 600 μ sec was judged as roughly equivalent in lateral location to a stimulus with an *ILD* near 16 dB. The present choices of a 500- μ sec *ITD* and a 15-dB *ILD* differ slightly from those used by Oxenham (2000; 640 μ sec and 12 dB), but as will be seen, this was inconsequential to the pattern of results.

Most important, it was expected that separation of the gap markers by ear of entry would result in significantly larger thresholds than would separation by *ILD* value, which should not differ from monaural level differences. There are two reasons for this. In the *Ear* (i.e., between-ears) condition, there is no common activation of peripheral channels, and thus, the task is probably accom-

plished by a relative timing operation between the events at each ear, whereas the *ILD* condition could, in theory, be performed monaurally through within-channel (i.e., within-ear) processes. Thus, the operating assumption is that within-ear processing results in lower thresholds than does between-ears processing.

Method

Subjects. The data presented here are from 2 highly practiced listeners (1 of them male) and 2 moderately practiced female listeners. All were familiar with gap detection tasks. This experiment received ethical approval from Dalhousie University's Research Ethics Board under Protocol 2000-88.

Stimuli. The stimuli (markers) were generated digitally and were presented at a sampling frequency of 44100 Hz. The markers of each trial were two newly generated independent wideband noise bursts, limited only by the sampling frequency and the transfer characteristics of the headphones. The markers were separated by an interval of silence (the gap; see Figure 1A). The leading marker was always 150 msec in duration. The duration of the trailing marker was drawn randomly on each trial from a uniform distribution ranging between 100 and 300 msec, in order to reduce the possibility that the listeners could perform the task by judging overall stimulus duration, rather than the duration of the silent interval (after Oxenham, 2000). Thus, although the leading marker was constant, the total duration of the combined stimulus was randomized. Each marker was shaped with 1-msec rise and fall times, including those used in the standard stimulus. There was no silent period (gap = 0 msec) imposed between the fall of the leading marker and the rise of the trailing marker in the comparison interval of the two-alternative forced choice (2AFC) task.

There were six stimulus conditions: *Diotic*, *ITD*, *LQ*, *QL*, *ILD*, and *Ear* (see Figure 2, bottom). In the *Diotic* condition, both the leading and the trailing noise bursts were presented diotically at 67 dB SPL. In the *ITD* condition, the noise bursts were presented dichotically at an equal level of 67 dB to each ear, but the leading marker had an *ITD* of -500 μ sec, and the trailing marker had an *ITD* of $+500$ μ sec. This was accomplished by starting the leading marker at the left ear 500 μ sec earlier than that at the right ear and starting the trailing marker's at the right ear 500 μ sec earlier than that at the left ear. This resulted in the leading marker's being lateralized to the left side of the head and the trailing marker's being lateralized to the right side of the head. In the *ILD* condition, the

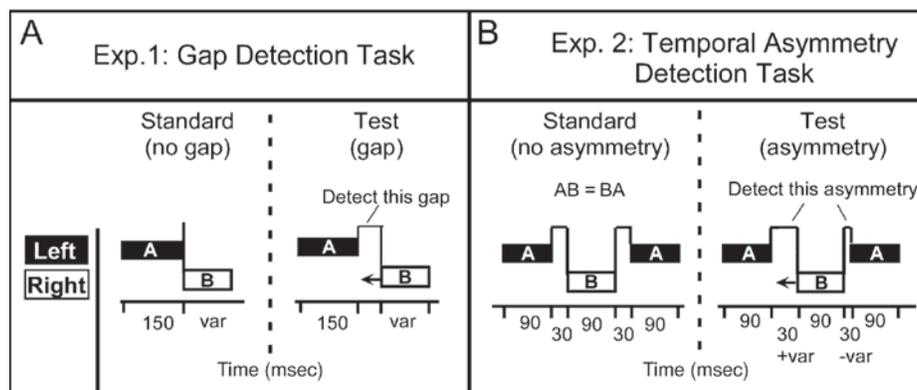


Figure 1. (A) Gap detection and (B) temporal asymmetry detection tasks for Experiments 1 and 2, respectively. Left portions of each panel indicate respective standards, which are (A) no gap and (B) temporal symmetry; right portions show the corresponding test stimuli—namely, (A) gap and (B) temporal asymmetry. In panel B, the ABA unit is part of a longer sequence that is preceded by an induction sequence of A sounds (not shown). See the text for details of respective tasks.

**EXPERIMENT 1
Gap Detection**

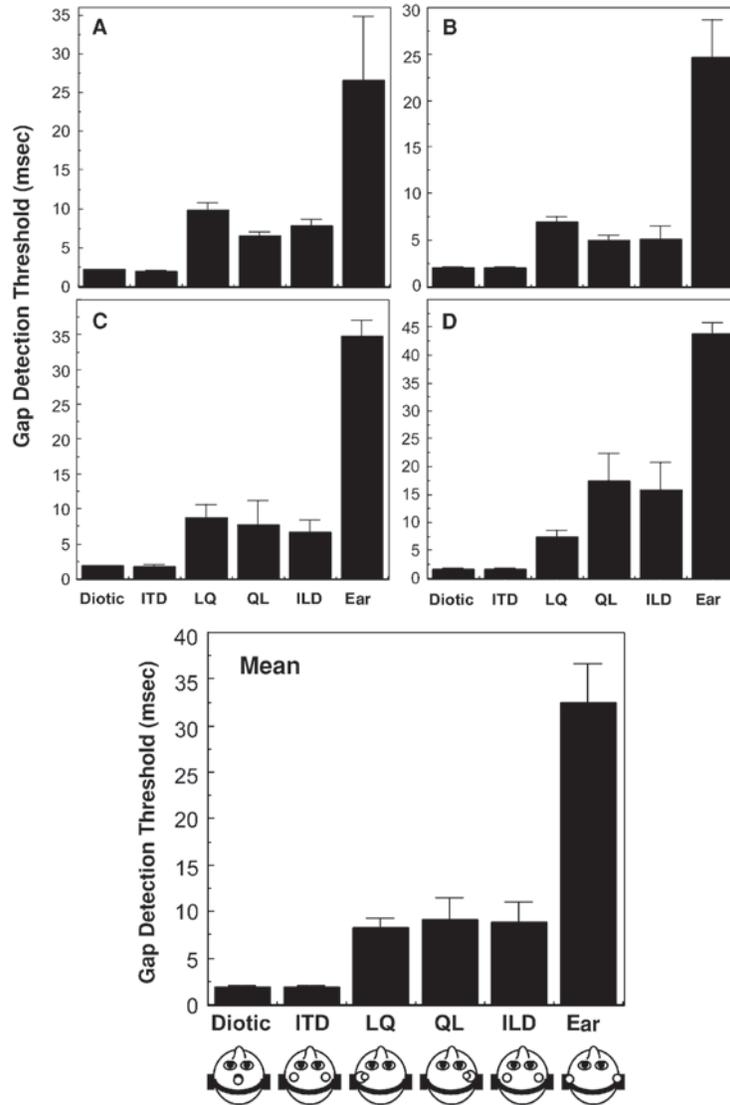


Figure 2. Individual and mean gap detection thresholds from Experiment 1 for the six conditions, as defined in the text. The schematic heads below each condition show the approximate intracranial locations of the sound images of the two markers. Error bars on the plots for individuals (A–D) represent the standard errors from the three threshold determinations contributing to the means. The error bars on the mean data (bottom panel) represent the mean of each listener’s standard errors. ITD, interaural time difference; ILD, interaural level difference; L, loud; Q, quiet.

noise bursts were presented dichotically, but the leading marker had an ILD of 15 dB, which favored the left ear (left ear, 75 dB SPL; right ear, 60 dB SPL), and the trailing marker had an ILD of 15 dB, which favored the right ear (left ear, 60 dB SPL; right ear, 75 dB SPL). This resulted in the leading marker’s being lateralized to the left side of the head and the trailing marker’s being lateralized to the right side of the head. In the monaural level difference condition in which the *leading* marker was louder (LQ), only the left channel of the stimulus in the ILD condition was presented. In the monaural level difference condition in which the *trailing* marker was louder

(QL), only the right channel of the stimulus in the ILD condition was presented. Finally, in the Ear condition, the leading and the trailing noise bursts were presented to left and right headphone channels, respectively, at 67 dB SPL.

Design and Procedure. The experiment was a fully repeated measures design; threshold estimations for each of the six conditions was obtained from each of the subjects. Each listener completed at least three threshold estimations for each condition, in no particular order. The arithmetic mean of the last three thresholds was taken as the subject’s mean threshold for that condition.

Each threshold was obtained as follows. Stimuli were presented to the subjects over AudioTechnica ATH-M40fs Precision Studio-phones or Sennheiser HD 25 headphones, with listeners seated in a sound-attenuating booth. Thresholds for detecting the silent interval were measured using a two-interval 2AFC task and a three-down, one-up adaptive method (see Levitt, 1971). The listeners were presented with a no-gap stimulus and a gap stimulus in random order and were to report whether the gap stimulus was presented first or second. The starting gap size varied among listeners but was usually 10 msec for the within-ear conditions and 30–50 msec for the between-ears conditions. After three correct detections in a row, the gap duration was reduced by a factor of 1.4, and after every incorrect response, the gap duration was increased by the same factor. This continued for eight reversals in the direction of the adaptive variable, and threshold was taken as the geometric mean of the last six reversals.

In calculating thresholds for the ITD condition, it was necessary to take into account that the gap presented to the left ear was 1 msec longer than that presented to the right ear (accounting for the 0.5-msec ITD imposed on each marker). To correct for this, we took the average gap threshold at the two ears (i.e., if the gap size at threshold was 2 msec at the right ear and 3 msec at the left, we took the real threshold to be 2.5 msec).

Results

Mean thresholds from each of the four individuals are plotted in Figure 2, with error bars representing the standard error of the three thresholds determinations that contributed to the means. Mean thresholds across listeners are plotted in the bottom panel. A one-way fully repeated measures ANOVA revealed a significant effect of condition [$F(5,15) = 32.5, p < .001$]. Simple effects indicated that mean thresholds for the Diotic (1.93 msec) and the ITD (1.87 msec) conditions did not differ [$F(1,15) < 1$], nor did those for the conditions in which the markers differed in level (ILD, 8.85 msec; LQ, 8.25 msec; QL, 9.21 msec; $F < 1$ for all comparisons). However, the thresholds for the level conditions were significantly higher than that obtained in the Diotic condition [$F(1,15) = 8.9, p < .05$]. Finally, thresholds for the Ear condition were significantly higher than those for the level conditions [$F(1,15) = 106.9, p < .001$]. Note that the Huynh–Feldt correction for sphericity violations in an ANOVA was used here and throughout this study.

Discussion

Thresholds obtained when the markers were characterized by opposite ITDs were not different from those obtained under the control condition with diotic markers (the Diotic condition). Differentiation of the gap markers by level (monaural or interaural) resulted in elevated thresholds, although they were not as elevated as when the markers were distinguished by Ear. These findings replicate those of Oxenham (2000) and extend them by demonstrating the relative ineffectiveness of marker level differences in elevating gap thresholds, as compared with marker differences in the ear of entry. The present data thus confirm the earlier study and suggest that the modest difference in the values of ILD and ITD used in the two studies was unimportant. The absolute thresholds obtained here for the ILD, LQ, and QL conditions were

lower than those obtained by Oxenham, despite the fact that we used a larger level difference. The reason for this is not clear, although it probably reflects the fact that our listeners were more practiced than his at this task. The thresholds are, however, comparable to those in previous reports from this laboratory, using different listeners (Phillips et al., 1998; Phillips et al., 1997). The important point to take from this experiment is that gap detection thresholds were most strongly affected by the similarity with which the markers activated the auditory periphery. If the markers were identical in spectrum and level at each ear (as in the Diotic and ITD conditions), thresholds were low, despite the fact that the markers perceptually switched in laterality in the ITD condition. If the markers differed in level at one or both ears, thresholds were somewhat elevated. When the markers had nonoverlapping activation patterns at the periphery, the thresholds were the most elevated.

EXPERIMENT 2

Temporal Asymmetry Detection for Sequences Differing in Spatial Attributes

In Experiment 2, we examined the effect of spatial stimulus attribute differences imposed on the components of a temporal asymmetry detection task. Listeners were presented with two interleaved sequences of sounds A and B, in the format ABA_ABA_ . . . ABA. In one sequence, the silent interval from the offset of A to the onset of B (the AB interval) was identical to the silent interval from the offset of B to the onset of the second A (the BA interval). In the comparison sequence, the AB interval was longer than the BA interval. Using a 2AFC adaptive tracking task, the smallest asymmetry that the listener could detect was determined. This task shares features of gap discrimination/detection tasks within a more complex auditory sequence (see Figure 1B). The sequence is akin to those used in auditory sequential stream segregation tasks, where listeners usually report *subjectively* whether the A and the B components are perceived to form a single auditory stream or to segregate into two separate A and B streams. Performance on the asymmetry task has been proposed as an *objective* index of sequential stream segregation, with low thresholds reflecting component differences that do not lead to segregation and high thresholds reflecting component differences that lead to segregation (Vliegen et al., 1999). The nature of this task encourages the listener to try to maintain an integrated percept of the sequences in order to achieve good performance. Thresholds in this task have been shown to be sensitive to A and B spectral differences, similar to the effect of such differences on performance in more basic temporal tasks, such as gap detection and gap discrimination. It was therefore expected that differences in A and B laterality cues would have effects on asymmetry thresholds that follow the same pattern as that observed for gap thresholds in Experiment 1: very degraded performance for A and B differences in ear of entry, equally impaired performance for A and B differ-

ences in ILD or monaural level, and normal performance for A and B differences in ITD (same as that in a diotic control condition).

Method

Subjects. Eight listeners (6 of them female, including the first author) participated. Four were practiced in auditory temporal-processing tasks, and the others were relatively naive. All had some musical training, and three (C, D, and E) performed music regularly. This experiment received ethical approval from Dalhousie University's Research Ethics Board under Protocol 2001-375.

Stimuli. The same six stimulus conditions (Diotic, ITD, LQL, QLQ, ILD, and Ear) as those in Experiment 1 were examined, except that the same frozen noise bursts were used for all A and B repetitions of a given run and the ILD and ITD values used were 12 dB and 600 μ sec. Also, in this task, the LQ and QL conditions are referred to as LQL and QLQ, to reflect that the fundamental segment of the sequence comprised three noise bursts (ABA). The ABA component of the sequences used in the temporal asymmetry task is outlined schematically in Figure 1B. The standard sequence was in the form A---A---A---A-B-A---A-B-A---...A-B-A. The three A iterations prior to the first ABA formed an induction sequence intended to bias perception toward segregation—that is, to push the As into their own stream before the first B was presented (see Vliegen et al., 1999). Five copies of the A-B-A--- stimulus were concatenated for a total sequence duration of about 3 sec ($480 \times 5 + 720$ msec induction sequence). The test stimulus was identical, except that the AB interval was greater than the BA interval, whereas the A-A interval remained constant throughout each sequence. As such, the A stream on its own remained temporally regular, as did the B stream, and the asymmetry existed only in the temporal relation between A and B. That is, the B stream started later in the test stimulus, so that each instance of B was always closer in time to the second A in each ABA sequence. The standard and the test sequences were presented 500 msec apart on each trial. Note that newly generated frozen noises were used for each run.

Design and Procedure. The experiment was a fully repeated measures design: Temporal asymmetry threshold estimations for each of the six conditions was obtained from each of the subjects. Each listener completed at least three threshold estimations for each condition, in no particular order. The arithmetic mean of the last three thresholds was taken as the subject's mean threshold for that condition.

Individual threshold estimations were obtained as follows. The sequences were presented to the subjects at a comfortable listening level (see the values in Experiment 1) over Sennheiser HD25 headphones. Thresholds for detecting the temporal asymmetry in the repeating A_B_A stream were obtained using a two-interval 2AFC and a modified two-down, one-up adaptive method (Levitt, 1971). The adaptive variable was the difference in the AB and BA intervals. The task was to indicate whether the test stimulus was in the first or the second position. Listener response was not timed, and so the intertrial interval varied. The initial asymmetry duration was above detection threshold and depended on the stimulus conditions (from 20 msec to a maximum of 30 msec). If the listener did not make a correct response at the maximum asymmetry size (30 msec), the sequence was repeated. Following two consecutive correct judgments, the size of the asymmetry was reduced by a factor of 1.4. Following a single incorrect response, the asymmetry was increased by a factor of 1.4. Each adaptive staircase continued for only eight reversals, and the geometric mean of the temporal asymmetry size for the last six reversals was defined as the detection threshold. Because of the length of each sequence, determination of a single threshold took about 6 min and required intense concentration.

Results

The mean thresholds for each listener are plotted separately in the top eight panels of Figure 3, with standard

error bars from the three threshold estimates used to generate the mean. The grand mean thresholds are plotted in the bottom panel, with standard error bars based on the 8 subjects' means. Temporal asymmetry threshold (in milliseconds) is on the y-axis.

The data were subjected to a one-way repeated measures ANOVA, which revealed a significant effect of condition [$F(7,35) = 19.59, p < .001$]. Specific hypotheses were tested, using simple effects. The mean temporal

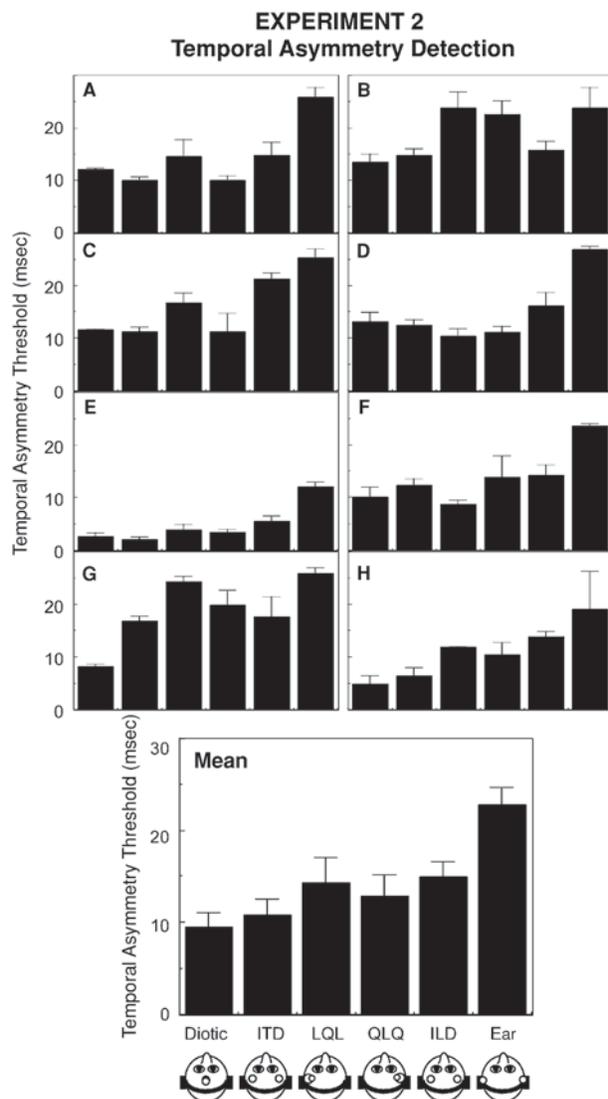


Figure 3. Individual and mean temporal asymmetry detection thresholds from Experiment 2. The schematic heads below each condition show the approximate intracranial locations of the sound images of the two markers. Error bars on the plots for individuals (A–H) represent the standard errors from the three threshold determinations contributing to the means. The error bars on the mean data (bottom panel) represent the mean of each listener's standard errors. Note that the conditions now refer to properties of the A and B noises that make up the triplet (A–B–A) that is the main unit of the sequence. Hence, loud–quiet (LQ) and quiet–loud (QL) are now recoded as LQL and QLQ. ITD, interaural time difference; ILD, interaural level difference.

asymmetry detection threshold for the Diotic condition (9.51 msec) was not significantly different from that for the ITD condition [10.76 msec; $F(1,35) < 1$]. The three conditions in which there was a level difference (QLQ, 12.8 msec; LQL, 14.3 msec; ILD, 14.9 msec) did not differ significantly ($ps > .1$) but were significantly higher than the Diotic condition ($ps < .05$). The mean threshold for the Ear condition (22.8 msec) was significantly greater than that obtained in the ILD condition [$F(1,35) = 27.5$, $p < .001$].

Thresholds were highest when A and B were presented to different ears (the Ear condition). Since the maximum possible asymmetry is 30 msec, listeners obtaining thresholds above 25 msec were not achieving any level of performance on the task, because the first adaptive step was from 30 to 25 msec. Visual inspection of the staircases indicated that the subjects could often do the task early in the staircase but “lost” whatever cue was permitting performance and that the stimulus then returned toward the maximum asymmetry at later reversals.

Discussion

The temporal asymmetry detection task used in this experiment inherently biases the listener to attempt to perceive the sequence as integrated. That is, in order to achieve good performance, it was in the listener’s best interest to try to force the A and B noises into the same perceptual stream. Thus, in terms of SSS, the task measures the segregation boundary—the point beyond which segregation must occur and integration fails. All the listeners reported the temporal asymmetry task to be cognitively and attentively demanding. The cue used to perform the task was typically a rhythmic difference between the two sequences. The standard sequence was temporally even, like the snare drum in a military march. The asymmetric sequence had a “lazy lilt” in its rhythm, akin to some jazz rhythms. This distinction was facilitated if an integrated percept could be maintained. Another strategy, reported by the listener with the lowest thresholds (E in Figure 3) involved attempting to force the first ABA iteration after the induction sequence to remain integrated. The task then became a gap discrimination task in which she determined in which sequence the first AB interval was longer. In a sense, the task was still measuring some aspect of the integration boundary, which clearly was difficult to overcome in the Ear condition, for which even the thresholds of Listener E were elevated.

The temporal asymmetry task was modeled after that used by Vliegen et al. (1999), who used pure tones differing in frequency, tone complexes that differed in fundamental frequency, or tone complexes with the same fundamental frequency but differing in the frequency range in which they were filtered. In general, the absolute values of the thresholds obtained in the present study are consistent with those in Vliegen et al. Similar to the results for the Diotic baseline condition here, Vliegen et al. found that some listeners, particularly those with musical training, achieved asymmetry thresholds near 5 msec for

the baseline condition, whereas other listeners had baseline thresholds closer to 10 msec. Their maximal thresholds approached 25 msec for the largest frequency separation, similar to those obtained in the present study in the Ear condition. Thus, in terms of absolute thresholds values, the results for the two studies are similar. Furthermore, in Vliegen et al., the greatest elevation in temporal asymmetry thresholds was for pure tones of disparate frequency or tone complexes filtered at different spectral regions—both cases in which the A and B sounds activated nonoverlapping populations of neurons at the auditory periphery. Thresholds were less affected by differences in the fundamental frequency (a centrally generated difference in pitch) of the tone complexes that activated similar spectral regions. The present data show that this rule generalizes to the spatial domain, since the stimulus condition resulting in the greatest asymmetry threshold elevation was the one in which A and B signals were presented to different ears.

An explicit comparison of the effects of spatial attributes on the two objective temporal-processing tasks is shown in Figure 4, with the asymmetry thresholds scaled to the minimum (Diotic) and maximum (Ear) gap thresholds. Although the absolute values of the thresholds in the two tasks are (unsurprisingly) different, the important point is that the relative effects of the spatial stimulus attributes were very similar in the two tasks. Thus, differences in the Ear condition always produced the highest thresholds; the ILD, LQ, and QL conditions always produced similar

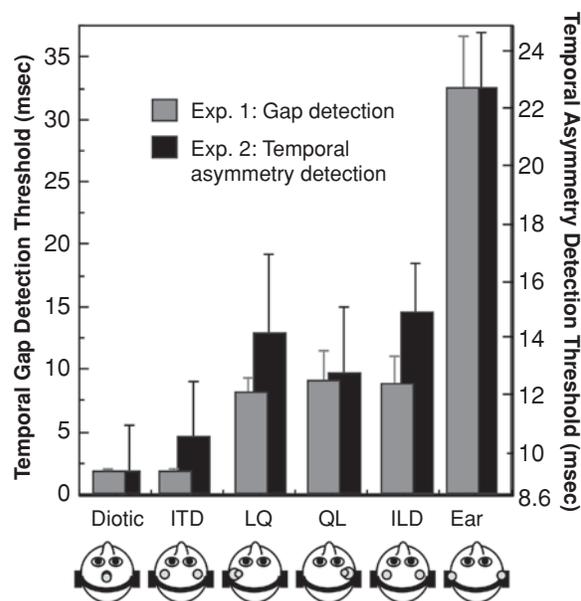


Figure 4. Direct comparison of performance on the temporal asymmetry and gap detection tasks. Stimulus conditions are defined in the text. Ordinates for the temporal asymmetry thresholds and gap detection thresholds (in milliseconds) are on the right and left axes, respectively. Thresholds for the temporal asymmetry task were normalized to gap detection performance for both the Diotic control condition and the Ear condition. ITD, interaural time difference; ILD, interaural level difference; L, loud; Q, quiet.

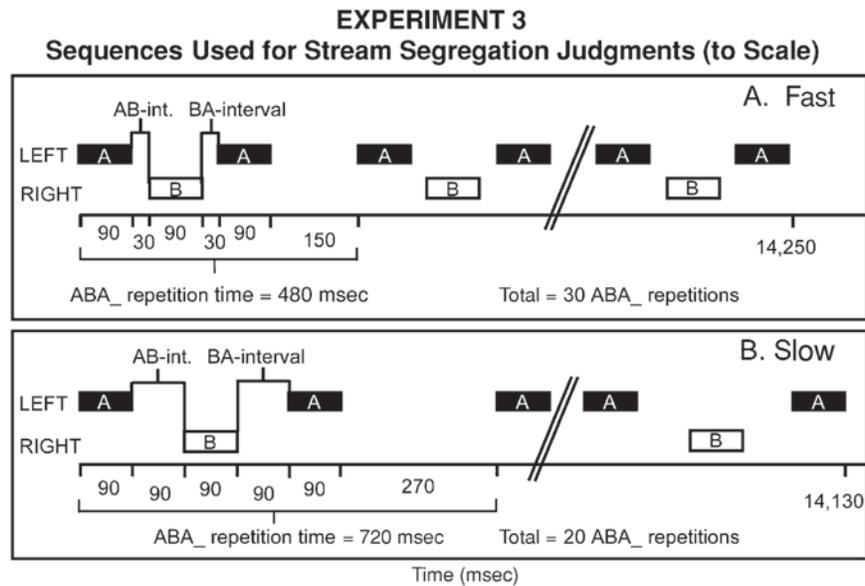


Figure 5. The temporal configuration of the A–B–A sound sequences presented to listeners in Experiment 3.

but less elevated thresholds, and the ITD and control conditions provided comparable low thresholds. There were two differences between the stimuli in the two tasks that limit this comparison. The values of ITD and ILD used differed somewhat, as did the duration of the sound components. In the gap detection case, we used a smaller ITD (500 vs. 600 μ sec) in order to prevent detection of the gap in the standard stimulus. We also used a larger ILD in the gap detection task (15 vs. 12 dB), in order to increase the effect size of imposing the level difference with our practiced listeners, whose thresholds were relatively low in the level condition, as compared with that reported by Oxenham (2000). The duration of the components (gap markers/A and B) differed due to task constraints and in order to match the values with those in the literature reported by Oxenham (2000) and Vliegen et al. (1999).

EXPERIMENT 3

Stream Segregation for Sequences Differing in Spatial Attributes

Experiments 1 and 2 confirmed that performance on a simple measure (gap detection) and a more complex measure (temporal asymmetry detection) of temporal processing were affected similarly by component differences in spatial stimulus attributes. This is important because the temporal asymmetry task is thought to index the formation of two stream percepts in classic stream segregation tasks. Thus, gap detection may be a simple task through which to index the segregation boundary. To explore the link between performance on these temporal tasks and auditory sequential stream segregation, in Experiment 3, we had listeners track their perception of a repeating sequence of A and B sounds that differed in ear of entry, ILD, ITD, and monaural level. Although measures

of streaming for such a sequence have been reported for A and B differences on a wide range of acoustic attributes (spectral and temporal pitch, timbre, and ear of entry; see Moore & Gockel, 2002, for a review), surprisingly, there are no reports on streaming with ITD and ILD, using this type of sequence.

Bregman (1990) described a number of methods that can be used to measure a subject's percept of ABA_ sequences. Typically, subjects use a scale to rate the ease with which the sequence is segregated or could be maintained as integrated (depending on task instructions). Another method requires that subjects simply state whether their percept was integrated or segregated at the end of the sequence (e.g., Beauvois, 1998; Beauvois & Meddis, 1997; Bregman, Ahad, Crum, & O'Reilly, 2000; Bregman, Ahad, & Van Loon, 2001; Rose & Moore, 2000; Vliegen & Oxenham, 1999). Bregman also referred to a method of perceptual tracking of the two states, in which listeners report changes (e.g., "one" vs. "two" streams) in their perceptual state, as they occur, with buttonpresses, for example (for recent examples, see Carlyon, Cusack, Foxton, & Robertson, 2001; Roberts, Glasberg, & Moore, 2002). In the present experiment, this perceptual tracking method was implemented for two reasons. First, pilot testing indicated that stream segregation by spatial location was often ambiguous (bistable) and was subject to perceptual state switches. Second, by using a tracking procedure, it would be possible to measure (1) the time to build up to segregation and (2) the proportion of segregation reported over time.

Method

Subjects. Perceptual tracking data were obtained from 9 listeners (7 of them female). Seven of these listeners also had participated in Experiment 2. Four were highly experienced in psychoacoustic tasks; the rest were experimentally naive. All the listeners had normal audiometric sensitivity to 8 kHz. All the procedures received ethical

approval from Dalhousie University's Research Ethics Board, under Protocol 2001-375.

Design. There were two independent variables by which sequences were characterized: the spatial stimulus attribute distinguishing the A and the B noise bursts and the rate of noise burst alternation. There were three lateralization conditions: Ear (A and B noises presented to different ears), ILD (A and B at ILDs of +12 and -12 dB, respectively), and ITD (A and B presented at ITDs of +600 and -600 μ sec, respectively). There were three control conditions: diotic (both A and B centered), LQL (A 12 dB more intense than B), and QLQ (B 12 dB more intense than A). The latter two conditions were controls for the monaural contributions to the ILD condition (see Oxenham, 2000). Each condition was presented at a fast and a slow rate. These rates were determined by different durations of silent intervals between A and B: 30 msec (as in Experiment 2) or 90 msec. The 30- and 90-msec AB (or BA) durations corresponded to an *ABA repetition time* of 480 and 720 msec, respectively, because the noise burst durations were always 90 msec. These values were chosen because the same values had been used to generate good streaming in a similar task in which stream segregation based on the pitch of tone complexes was examined (Vliegen & Oxenham, 1999). In total, this generated 12 different sequences (6 conditions \times 2 rates).

Stimulus construction. The stimuli were generated digitally and were presented at a sampling frequency of 44100 Hz at 16-bit amplitude quantization on the audio processor of an Apple Power-Macintosh. Each sequence was in the form A-B-A-_-A-B-A-_-...A-B-A, where "-" indicates a unit length of silent interval and "_" a unit the length of B (see Figure 5). Noise bursts were 90 msec in duration, including 10-msec linear rise/fall ramps. For each trial, two new independent bursts were randomly generated for the construction of the A and B components for the entire sequence (i.e., all of the As were identical in that sequence, as were all of the Bs). For the fast repetition rate, the AB and BA silent intervals (represented by - above) were 30 msec each, whereas in the slow repetition rate condition, each of the silent intervals was 90 msec in duration. The AA interval ("_-") between triplets was 2 times the AB interval plus the length of the B noise. The A and B noises were concatenated with a segment of zeros the length of the AB, BA, and AA intervals. The A-B-A-_- stimulus was replicated 30 times for the fast rate and 20 times for the slow rate and was concatenated in sequence for a total duration of 14.4 sec for any given sequence.

Procedure. The experiment was a fully repeated measures design: One- or two-stream judgments were obtained for each of the six conditions from each of the subjects.

Prior to data collection, the listeners received verbal instruction on the concept of stream segregation and were presented with several example sequences composed of A and B tones of various frequency separations (Δf). In some of these, segregation was automatic (large Δf); in others, segregation required time to "build up" (medium Δf) or did not occur at all (small Δf). After the listeners were able to label these states reliably, they performed practice sessions with each noise sequence used in the formal experiment.

The stimuli were presented at a comfortable listening level (70 dB SPL) over Sennheiser HD25 headphones, while the listeners were seated in a sound-attenuating booth. In each of 10 blocks, the sequences for each stimulus condition were presented in a random order. The listeners took breaks as needed between blocks. The subjects were told to try to "hear out" the B sounds as a separate stream but, also, to accurately report whether their percept switched back to an integrated state. A schematic diagram of the integrated and segregated states was presented via computer monitor before and during presentation of each sequence. The listeners always began rating each sequence by holding down the "1!" key on the keyboard, and as soon as they perceived the sequence as segregated into separate streams, they were to press the "2@" key. If their percept reverted to an integrated percept, they were to switch back to pressing the

"1" key. In the event that there was ambiguity as to the presence of one or two streams, they were to press the "1" key. After the sequence ended, listeners were given 5-10 sec during which they were to return to pressing the "1" key and prepare for presentation of the next sequence, which was presented automatically. Data collected for each trial consisted of a continuous record (covering the entire 14.4-sec sequence duration) logging the identity of the key depressed ("1" or "2") as a function of time.

Measures and data analysis. Two measures were extracted off-line from the perceptual tracking data: *proportion segregation*, the proportion of time spent in a segregated state, calculated as (segregation time)/(integration time + segregation time), and *buildup time*, the duration of time from sequence onset until the first report of segregation (i.e., until the first press of the "2@" key).

A post hoc analysis of the data was completed to eliminate error trials. If the listener was not pressing the "1" key within 300 msec of sequence onset, the trial was deemed an error and was removed from the analysis. If the listener made a transition to the "2" key in less than 700 msec after sequence onset (approximate time to hear one ABA iteration and press the key), the trial was classified as an anticipatory response and was removed from the analysis. Such errors were fairly rare or nonexistent in most subjects' data. After elimination of these trials, a count of the total trials left was completed. For all the listeners, at least eight error-free trials were available in each condition. To retain a consistent number of trials across the conditions, additional trials beyond eight were discarded. Thus, the results described below are based on eight repetitions of the perceptual tracking procedure for each of the 12 conditions of interest.

Results

The mean results for each condition are presented in Figure 6. Error bars are standard errors of the means across subjects. The two panels of Figure 6 show the mean proportion segregation and buildup time to first segregation. In both, values are plotted against sequence rate (interstimulus interval of 30 or 90 msec), with condition (Diotic, ITD, LQL, QLQ, ILD, and Ear) as the parameter.

Proportion of time spent in a segregated state (proportion segregation). The proportion segregation values obtained from all conditions were subjected to a two-way fully repeated measures ANOVA (6 conditions \times 2 rates). This analysis revealed a significant main effect of condition [$F(5,40) = 52.98, p < .001$], but not of rate [$F(1,8) = 3.27, n.s.$], and there was no interaction of condition with rate [$F(5,40) = 1.88, n.s.$]. A simple effects examination of the main effect of condition indicated that differentiation of A and B by Ear resulted in segregation marginally greater than that seen in the ILD condition [$F(1,40) = 3.7, p = .07$]. Imposing opposite ITDs on A and B resulted in segregation that was significantly greater than that in the Diotic condition [$F(1,40) = 15.33, p < .002$], but less than that in the ILD condition [$F(1,40) = 44.0, p < .001$]. Differentiation of A and B by level monaurally (LQL and QLQ) resulted in equally poor segregation, on average [$F(1,40) = 1, n.s.$], which was slightly greater than the Diotic baseline [$F(1,40) = 4.78, p = .05$]. However, the effect on proportion segregation when A and B differed only in monaural level was highly idiosyncratic across listeners. Four listeners reported more segregation with LQL than with QLQ; 4 other listeners reported no segregation with either, and a single listener reported more segregation with QLQ than with LQL.

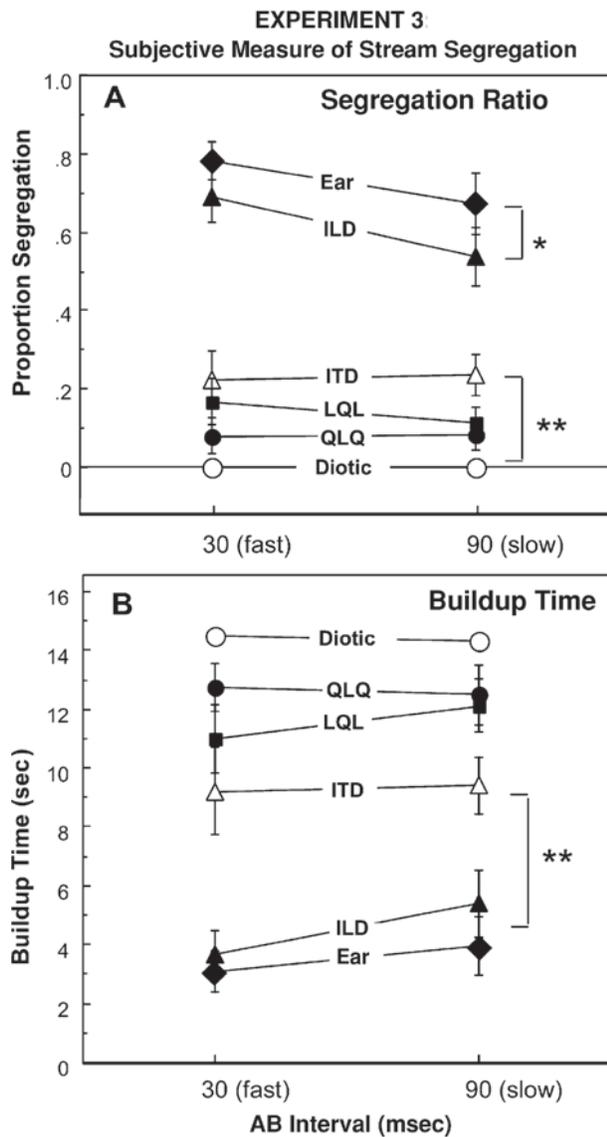


Figure 6. The mean ($n = 9$) results from Experiment 3, presented with standard error bars. Two measures from the sequence rating task data were extracted: (A) proportion segregation, the proportion of time spent in a segregated state (pressing the “2” key), and (B) buildup time, the time until the first “2” keypress, indicating segregation had occurred. ILD, interaural level difference; ITD, interaural time difference; L, loud; Q, quiet.

Buildup to segregation. The time to the first segregation (the *buildup time*) should be inversely related to the segregation ratio data, and in general, this was the case. The results of a two-way repeated measures ANOVA performed on the buildup data (Figure 6B) basically mirrored the results obtained for the proportion segregation data. As was observed with the proportion segregation data, there was a significant main effect of condition [$F(5,40) = 37.06, p < .001$] but no main effect of rate [$F(1,9) = 1.79, n.s.$] or interaction of rate with condition [$F(5,40) = 1.03, n.s.$]. Examination of the main effect of condition indicated that

separation of A and B in the Ear and ILD conditions resulted in comparably short buildup times to segregation [$F(1,40) < 1$]. Segregation on the basis of ITD was slower to buildup, as compared with that for segregation based on ILD [$F(1,40) = 21.43, p < .01$].

Comparison of objective (Experiment 2) and subjective (Experiment 3) measures of SSS. In Figure 7, the segregation ratio for the 30-msec AB interval (see above) is plotted against the temporal asymmetry detection thresholds from Experiment 2 for the 7 listeners who participated in both experiments. For conditions in which A and B differed in laterality, there was a positive linear relationship between segregation measured subjectively and the size of temporal asymmetry detection thresholds: As subjective segregation increased, so did objective thresholds. This relation can account for over 40% of the variance in the data ($R^2 = .42$). This relation did not exist when A and B differed only in intensity ($R^2 = .02$). Intensity differences elevated temporal asymmetry thresholds, but not subjective reports of segregation. Another feature is the clear separation of points for the Ear, ILD, and ITD conditions (see filled symbols), with A–B separations by Ear leading to higher values on both tasks than did AB separations by ILD, which in turn were higher than those obtained for AB separations by ITD.

Discussion

Characterization of A and B by ear of entry (the Ear condition) resulted in the highest segregation ratio and the shortest buildup time. Almost uniformly across subjects, as soon as one or two ABA repetitions had been heard, they pressed the “2@” key. The mean buildup time was 3–4 sec, but was as low as 1–2 sec in 3 listeners. Given the ABA period (480 or 720 msec), plus some time to make the decision to respond, plus the reaction time to make the response (likely, about 300 msec), segregation appeared to be nearly instantaneous. This confirms that differentiating A and B by ear of entry is a strong cue to stream segregation (Bregman, 1990). As compared with segregation by Ear, segregation by ILD resulted in somewhat lower segregation ratios but equally short buildup times. The lack of a difference between the Ear and the ILD conditions in buildup time may indicate there is a ceiling effect on the stream segregation data—that is, the variability in decision and reaction times masked any small difference in the buildup time to segregation in the ILD condition. ITD had a small but significant effect on the proportion of the sequence during which the listeners perceived segregated A and B streams (Figure 6A). This is consistent with Moore and Gockel (2002), who suggested that streaming can result whenever there are perceptual differences between two sounds; these results show that not all perceptual differences are equal.

GENERAL DISCUSSION

The goal of this study was to explore the effects of stimulus laterality differences on two objective measures of auditory temporal interval sensitivity and on judgments of stream segregation. As such, these data form a link be-

tween auditory temporal resolution and segregation of auditory streams. This is important because measures of temporal resolution are objective (“hard”), whereas judgments of perceived stream segregation have traditionally been seen as subjective (“soft”). Thus, validating the use of objective measures of stream segregation would provide a useful tool for reliable measurement of the stimulus attributes that lead to streaming.

Experiment 1 replicated and extended Oxenham’s (2000) findings in a simple gap detection paradigm. Markers with oppositely signed ITDs yielded gap thresholds at control values. Oppositely signed ILDs imposed on a gap’s markers elevated gap thresholds above control (Diotic) values, but no more so than did equivalent level differences between markers at a single ear (i.e., monaural level differences of the same size as those occurring in the ILD condition). The present study extends Oxenham’s findings by revealing that ear-of-entry differences imposed on a gap’s markers provided the greatest lengthening of gap thresholds (Figure 2). Experiment 2 showed that this *pattern* of effects of spatial stimulus attributes generalized to a more complex auditory temporal interval task: temporal asymmetry detection (Figure 3). Experiment 3 revealed the effects of the same stimulus parameters in an explicit stream segregation task (Figure 6). The temporal asymmetry task is thought to be an objective index of auditory sequential stream segregation, and in accord with this, we found similarities in the pattern of effects of

the same stimulus parameters in a subjective measure of auditory stream segregation (see Figure 7), despite different task demands: a bias to hear one stream in the objective task and a bias to hear two streams in the subjective task. There were two exceptions to the correspondence between the two tasks. The first was that ITD, although remaining relatively weak in its ability to objectively affect the auditory stream segregation judgments, nonetheless prompted some segregation, as measured by subjective judgments (Figure 6), likely due to the above-mentioned biases. The other was that monaural level differences between A and B resulted in weak stream segregation and did not affect stream segregation as much as did ILD differences. Both of these findings are compatible with the findings of Hartmann and Johnson (1991, Experiment 1, Condition 4), who showed that stimulus level differences could facilitate streaming to some degree when listeners were biased to hear separate streams in order to do an interleaved melody recognition task.

The data from the objective tasks suggest that, of the stimulus conditions examined, the strongest effect on the relative timing of auditory stimuli was exerted by Ear—that is, by differentiation of the stimuli by the peripheral neural channels through which the stimuli enter the nervous system. These observations in the spatial domain parallel previous ones in the spectral domain. Gap detection thresholds are elevated more by spectral differences between markers (in the absence of pitch differences) than by fundamental frequency differences (in the absence of spectral differences; Oxenham, 2000). The same has also been shown to be true for the temporal asymmetry detection task used in this study (Vliegen et al., 1999). The ILD, LQ, and QL conditions may have exerted effects on gap detection thresholds through nonsimultaneous masking phenomena, and it is interesting that the mean gap thresholds for these conditions were only modestly shorter than were the asymmetry detection thresholds for the same stimulus conditions (Figure 4). In neither gap detection (Figure 1; see also Oxenham, 2000) nor asymmetry detection (Figure 3) did ITD prolong thresholds, and by design, the stimuli being perceptually timed in those conditions were spectrally identical and had equal amplitudes at the two ears.

In summary, the following can be concluded about the effect of spatial stimulus attributes on temporal interval processing and stream segregation: (1) Segregation by Ear is immediate and has the strongest effect on temporal interval processing; (2) segregation by ILD is weaker than segregation by Ear; (3) segregation by ITD is weak; (4) segregation by monaural level differences is weak in subjective judgments, but can affect temporal judgments; and (5) ability to maintain an integrated percept and perform well on temporal interval judgments is related to the degree of common peripheral activation.

Segregation by Perceived Location or by Differential Activation of the Ears?

The results here exploit the fact that different cues to spatial lateralization can lead to similar location percepts

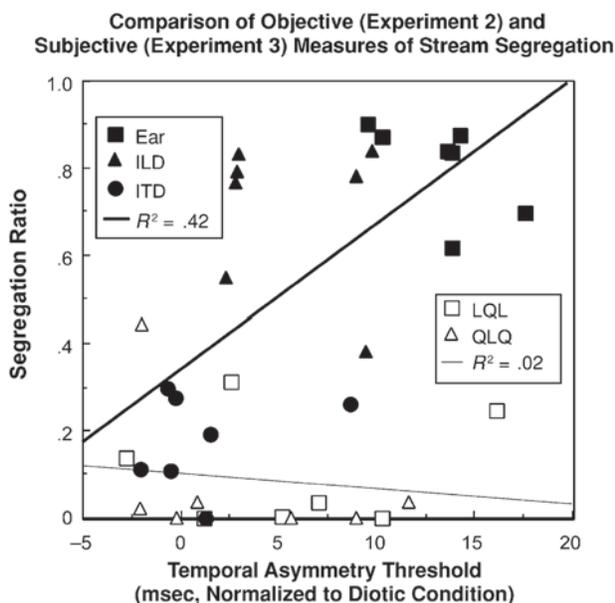


Figure 7. Segregation ratio values (fast rate: 30-msec AB interval, Experiment 3) plotted against temporal asymmetry detection thresholds (30-msec AB interval, Experiment 2) for the 7 listeners who participated in both Experiments 2 and 3. Temporal asymmetry thresholds are normalized relative to performance in the Diotic condition; that is, the value plotted is the threshold obtained for a given condition minus that obtained for the Diotic condition. A description is given in the text. ILD, interaural level difference; ITD, interaural time difference; L, loud; Q, quiet.

through different coding methods. The ITD generates lateralization while activating both ears similarly, the ILD generates a similar lateralization percept while differentially activating the two ears, and presentation of stimuli monaurally (the Ear condition) results in lateralization by activating only a single ear. These results show that although all cues lead to lateralized percepts, they are not equivalently effective in producing segregation. Thus, although perceptual differences, such as those generated by differentiating stimuli by ITD, may lead to some segregation, automatic segregation seems to require differential activation at the lowest levels of auditory processing. Although the use of these simple cues to lateralization allowed dissociation of perceived lateralization from peripheral activation, the generalization of these results to sources in the real world, where these cues act in concert, remains to be determined by future experiments.

The Role of Spatial Location in Segregation of Auditory Objects: Comparison With Other Studies

The finding that some level of sequential stream segregation by ITD differences can occur in the absence of peripheral activation differences between A and B sounds is consistent with the results of Hartmann and Johnson (1991), who found a medium benefit in the segregation of two spectrally interleaved melodies on the basis of ITD alone. This is also consistent with recent results demonstrating segregation with A and B tone complexes that were spectrally similar but differed in fundamental frequency (Vliegen & Oxenham, 1999) or component phase relations (Roberts et al., 2002) and with A and B tones with the same carrier but different amplitude modulation rates (Grimault, Bacon, & Micheyl, 2002). In all of these cases, the patterns of peripheral activation were similar for A and B sounds, but the stimuli differed in centrally generated pitch. Note, however, that in cases of *concurrent* stream segregation, fundamental frequency differences can be a powerful cue (Darwin, Brungart, & Simpson, 2003). The power of the present study was to demonstrate that ITD was *less effective* than other spatial stimulus attributes that produce similarly lateralized percepts (the ILD condition). This suggests that similarity in peripheral activation patterns provides a bottom-up starting point for the segregation process. Any bias toward an *integrated* state may, then, be overridden as other evidence for A and B differences accumulates over time. Bottom-up processes (peripheral activation pattern differences) may thus determine the immediacy of segregation, but more centrally generated perceptual differences can be used for segregation upon a “second look” by the perceptual processor. This is consistent with the fact that when segregation occurred with ITD sequences, it took somewhat longer to build up.

The pattern of results obtained in the subjective rating task, which biased the subjects to hear segregation, was similar to the pattern of benefit in recognition of individual melodies (e.g., A or B) that occurs with separation by Ear and by ITD (Hartmann & Johnson, 1991, Experi-

ment 1). Presentation of A and B to different ears (the Ear condition) resulted in the highest segregation ratings (Experiment 3 here), in accord with the greatest improvement in melody recognition (Hartmann & Johnson, 1991, Experiment 1) occurring with separation to the two ears. Giving A and B (or the two melodies) different ITD values resulted in low to intermediate ratings of segregation and an intermediate improvement in melody recognition.

Although ITD has previously been shown to be a poor segregation cue for concurrent stream segregation tasks, it is thought to be more useful for sequential tasks (Culling & Summerfield, 1995; Darwin & Hukin, 1997, 1998, 1999). For example, giving different frequency components the same ITD was not useful in successive grouping of those frequency components to solve a speech-sound recognition task (Culling & Summerfield, 1995). Neither ITD nor ILD has proved useful in segregating distracting (interfering) spectral information presented simultaneously with target spectral information on which subjects performed a fundamental frequency discrimination task, although segregation by ear of entry was very effective (Gockel & Carlyon, 1998). If, however, distracting spectral information was presented only *before and after* target spectral information, ITD was useful in segregating the distractors and improving performance on the task (Gockel, Carlyon, & Micheyl, 1999).

Differences in Laterality Effects Across Levels of Processing

It remains to discuss the reasons for the fact that ITD and ILD had greater effects in the subjective stream segregation task than they did in either of the objective detection tasks. One possibility lies in any differences in the nature of the information available to decision processors for the objective and the subjective tasks, to which we have already alluded. That is, the extent of shared peripheral activation patterns sets some starting point for a segregation process for the two sources, and a decision about the separability of streams is then modified according to information accrued during stimulus presentation. In the case of the detection tasks, high temporal acuity would result from strategies that bias the listener toward within-channel processing and toward the ignoring of stimulus information (or any perceptual attribute) that detracts from the substrate available for within-channel processing. In the stream segregation task, however, it is of advantage to exploit any stimulus feature (or perceptual attribute) that will enhance the segregation process. The extra information used in the formation of the segregated streams may take time to accrue (e.g., ITD; Figure 6B) and/or may depend on the number of frequency channels carrying information pertinent to the segregation. In this regard, note that all frequency channels carry such information in the case of the ILD condition in the present study, whereas only the low-frequency channels do so in the case of the ITD conditions. The ILD condition prompted stronger segregation and shorter buildup times to segregation than did the ITD one (Figure 6). The strength of segregation and the buildup time to segregation may thus be directly

and inversely related, respectively, to the density of evidence available to the decision processor.

REFERENCES

- BEAUVOIS, M. W. (1998). The effect of tone duration on auditory stream formation. *Perception & Psychophysics*, **60**, 852-861.
- BEAUVOIS, M. W., & MEDDIS, R. (1997). Time decay of auditory stream biasing. *Perception & Psychophysics*, **59**, 81-86.
- BOEHNKE, S. E., & PHILLIPS, D. P. (1999). Azimuthal tuning of human perceptual channels for sound location. *Journal of the Acoustical Society of America*, **106**, 1948-1955.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- BREGMAN, A. S., AHAD, P. A., CRUM, P. A. C., & O'REILLY, J. (2000). Effects of time intervals and tone durations on auditory stream segregation. *Perception & Psychophysics*, **62**, 626-636.
- BREGMAN, A. S., AHAD, P. A., & VAN LOON, C. (2001). Stream segregation of narrow-band noise bursts. *Perception & Psychophysics*, **63**, 790-797.
- 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.
- CARLYON, R. P., CUSACK, R., FOXTON, J. M., & ROBERTSON, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 115-127.
- CULLING, J. F., & SUMMERFIELD, Q. (1995). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *Journal of the Acoustical Society of America*, **98**, 785-797.
- DARWIN, C. J., BRUNGART, D. S., & SIMPSON, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *Journal of the Acoustical Society of America*, **114**, 2913-2922.
- DARWIN, C. J., & HUKIN, R. W. (1997). Perceptual segregation of a harmonic from a vowel by interaural time difference and frequency proximity. *Journal of the Acoustical Society of America*, **102**, 2316-2324.
- DARWIN, C. J., & HUKIN, R. W. (1998). Perceptual segregation of a harmonic from a vowel by interaural time difference in conjunction with mistuning and onset asynchrony. *Journal of the Acoustical Society of America*, **103**, 1080-1084.
- DARWIN, C. J., & HUKIN, R. W. (1999). Auditory objects of attention: The role of interaural time differences. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 617-629.
- DIVENYI, P. L., & DANNER, W. F. (1977). Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. *Perception & Psychophysics*, **21**, 125-142.
- FITZGIBBONS, P. J., POLLATSEK, A., & THOMAS, I. B. (1974). Detection of temporal gaps within and between perceptual tonal groups. *Perception & Psychophysics*, **16**, 522-528.
- FORMBY, C., & FORREST, T. G. (1991). Detection of silent temporal gaps in sinusoidal markers. *Journal of the Acoustical Society of America*, **89**, 830-837.
- FORMBY, C., SHERLOCK, L. P., & LI, S. (1998). Temporal gap detection measured with multiple sinusoidal markers: Effects of marker number, frequency, and temporal position. *Journal of the Acoustical Society of America*, **104**, 984-998.
- GOCKEL, H., & CARLYON, R. P. (1998). Effects of ear of entry and perceived location of synchronous and asynchronous components on mistuning detection. *Journal of the Acoustical Society of America*, **104**, 3534-3545.
- GOCKEL, H., CARLYON, R. P., & MICHEYL, C. (1999). Context dependence of fundamental-frequency discrimination: Lateralized temporal fringes. *Journal of the Acoustical Society of America*, **106**, 3553-3563.
- GRIMAULT, N., BACON, S. P., & MICHEYL, C. (2002). Auditory stream segregation on the basis of amplitude-modulation rate. *Journal of the Acoustical Society of America*, **111**, 1340-1348.
- GROSE, J. H., HALL, J. W., III, BUSS, E., & HATCH, D. (2001). Gap detection for similar and dissimilar gap markers. *Journal of the Acoustical Society of America*, **109**, 1587-1595.
- HALL, J. W., III, GROSE, J. H., & JOY, S. (1996). Gap detection for pairs of noise bands: Effects of stimulus level and frequency separation. *Journal of the Acoustical Society of America*, **99**, 1091-1095.
- HARRIS, G. G. (1960). Binaural interactions of impulsive stimuli and pure tones. *Journal of the Acoustical Society of America*, **32**, 685-692.
- HARTMANN, W. M., & JOHNSON, D. (1991). Stream segregation and peripheral channeling. *Music Perception*, **9**, 155-184.
- KINNEY, J. A. S. (1961). Discrimination in auditory and visual patterns. *American Journal of Psychology*, **74**, 529-541.
- LEVITT, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, **49** (Suppl. 2), 467-477.
- MIDDLEBROOKS, J. C., & GREEN, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, **42**, 135-159.
- MIDDLEBROOKS, J. C., MAKOUS, J. C., & GREEN, D. M. (1989). Directional sensitivity of sound-pressure levels in the human ear canal. *Journal of the Acoustical Society of America*, **86**, 89-108.
- MOORE, B. C. J., & GOCKEL, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica United With Acustica*, **88**, 320-332.
- NEFF, D. L., JESTEADT, W., & BROWN, E. L. (1982). The relation between gap discrimination and auditory stream segregation. *Perception & Psychophysics*, **31**, 493-501.
- OXENHAM, A. J. (2000). Influence of spatial and temporal coding on auditory gap detection. *Journal of the Acoustical Society of America*, **107**, 2215-2223.
- PENNER, M. J. (1977). Detection of temporal gaps in noise as a measure of the decay of auditory sensation. *Journal of the Acoustical Society of America*, **61**, 552-557.
- PERROTT, D. R., & WILLIAMS, K. N. (1971). Auditory temporal resolution: Gap detection as a function of interpulse frequency disparity. *Psychonomic Science*, **25**, 73-74.
- PHILLIPS, D. P., & HALL, S. E. (2000). Independence of frequency channels in auditory temporal gap detection. *Journal of the Acoustical Society of America*, **108**, 2957-2963.
- PHILLIPS, D. P., & HALL, S. E. (2002). Auditory temporal gap detection for noise markers with partially overlapping and non-overlapping spectra. *Hearing Research*, **174**, 133-141.
- PHILLIPS, D. P., HALL, S. E., HARRINGTON, I. A., & TAYLOR, T. L. (1998). "Central" auditory gap detection: A spatial case. *Journal of the Acoustical Society of America*, **103**, 2064-2068.
- PHILLIPS, D. P., TAYLOR, T. L., HALL, S. E., CARR, M. M., & MOSSOP, J. E. (1997). Detection of silent intervals between noises activating different perceptual channels: Some properties of "central" auditory gap detection. *Journal of the Acoustical Society of America*, **101**, 3694-3705.
- PLOMP, R. (1964). Rate of decay of auditory sensation. *Journal of the Acoustical Society of America*, **36**, 277-282.
- ROBERTS, B., GLASBERG, B. R., & MOORE, B. C. J. (2002). Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. *Journal of the Acoustical Society of America*, **112**, 2074-2085.
- ROSE, M. M., & MOORE, B. C. J. (2000). Effects of frequency and level on auditory stream segregation. *Journal of the Acoustical Society of America*, **108**, 1209-1214.
- SMIAROWSKI, R. A., & CARHART, R. (1975). Relations among temporal resolution, forward masking, and simultaneous masking. *Journal of the Acoustical Society of America*, **57**, 1169-1174.
- TAYLOR, T. L., HALL, S. E., BOEHNKE, S. E., & PHILLIPS, D. P. (1999). Additivity of perceptual channel-crossing effects in auditory gap detection. *Journal of the Acoustical Society of America*, **105**, 563-566.
- VAN NOORDEN, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation, Technische Hogeschool, Eindhoven, The Netherlands.
- VLIENEN, J., MOORE, B. C. J., & OXENHAM, A. J. (1999). The role of

- spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. *Journal of the Acoustical Society of America*, **106**, 938-945.
- VLIEGEN, J., & OXENHAM, A. J. (1999). Sequential stream segregation in the absence of spectral cues. *Journal of the Acoustical Society of America*, **105**, 339-346.
- WARREN, R. M., OBUSEK, C. J., FARMER, R. M., & WARREN, R. P. (1969). Auditory sequence: Confusion of patterns other than speech or music. *Science*, **164**, 586-587.
- WILLIAMS, K. N., & ELFNER, L. F. (1976). Gap detection with three auditory events: A single-channel process. *Journal of the Acoustical Society of America*, **60**, 423-428.
- WILLIAMS, K. N., ELFNER, L. F., & HOWSE, W. R. (1978). Auditory temporal resolution: Effects of sensation level. *Journal of Auditory Research*, **18**, 265-269.
- YOST, W. A. (1981). Lateral position of sinusoids presented with interaural intensive and temporal differences. *Journal of the Acoustical Society of America*, **70**, 397-409.
- YOST, W. A., TANIS, D. C., NIELSEN, D. W., & BERGERT, B. (1975). Interaural time vs. interaural intensity in a lateralization paradigm. *Perception & Psychophysics*, **18**, 433-440.

(Manuscript received August 19, 2003;
revision accepted for publication December 9, 2004.)