Influence of spatial and temporal coding on auditory gap detection

Andrew J. Oxenham[a)]
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 23 July 1999; revised 18 October 1999; accepted 4 January 2000)

This study investigated the effect on gap detection of perceptual channels, hypothesized to be tuned to spatial location or fundamental frequency (f_0). Thresholds were measured for the detection of a silent temporal gap between two markers. In the first experiment, the markers were broadband noise, presented either binaurally or monaurally. In the binaural conditions, the markers were either diotic, or had a 640-μs interaural time difference (ITD) or a 12-dB interaural level difference (ILD). Reversing the ITD across the two markers had no effect on gap detection relative to the diotic condition. Reversing the ILD across the two markers produced a marked deterioration in performance. However, the same deterioration was observed in the monaural conditions when a 12-dB level difference was introduced between the two markers. The results provide no evidence for the role of spatially tuned neural channels in gap detection. In the second experiment, the markers were harmonic tone complexes, filtered to contain only high, unresolved harmonics. Using complexes with a fixed spectral envelope, where the f_0 (of 140 or 350 Hz) was different for the two markers, produced a deterioration in performance, relative to conditions where the f_0 remained the same. A larger deterioration was observed when the two markers occupied different spectral regions but had the same f_0. This supports the idea that peripheral coding is dominant in determining gap-detection thresholds when the two markers differ along any physical dimension. Higher-order neural coding mechanisms of f_0 and spatial location seem to play a smaller role and no role, respectively. [S0001-4966(00)02204-9]

PACS numbers: 43.66.Mk, 43.66.Pn, 43.66.Hg [DWG]

INTRODUCTION

Gap detection has long been used as a measure of temporal resolution in the auditory system (Plomp, 1964b; Penner, 1977; Buus and Florentine, 1985; Moore *et al.*, 1993). Often, for sinusoids and broadband noise, silent gaps of 5 ms or less can be detected. This minimum detectable gap duration has been interpreted as revealing a fundamental ''sluggishness'' in the auditory system's response to very rapid changes in sound level. In another class of gap-detection experiments, the sounds before and after the gap, known as ''markers,'' differ along a certain physical dimension. For instance, using sinusoids, the two markers may have different frequencies. Previous studies have found that gap-detection thresholds increase as the frequency difference between the two markers increases, and that at large frequency differences, the minimum detectable gap increases to 20 ms or more, i.e., approximately an order of magnitude greater than when the two markers are at the same frequency (e.g., Williams and Perrott, 1972; Formby and Forrest, 1991; Formby *et al.*, 1998). A similar deterioration in performance, both in traditional gap detection (Phillips *et al.*, 1997) and in modulation gap detection (Grose *et al.*, 1999), is found when the two markers are presented to separate ears.

The effect of a difference in marker frequency has generally been attributed to the frequency selectivity established in the auditory periphery, and has also been modeled in this way (Forrest and Formby, 1996; Heinz *et al.*, 1996). When the two marker frequencies are the same or very similar, the markers stimulate the same region of the cochlear partition, which in turn leads to responses from the same population of auditory nerve fibers. Thus any perceived interruption (fluctuation, onset, or offset) in the stimulus is a reliable cue for detecting the gap. When the two markers have different frequencies, they are separated in the cochlea such that they maximally stimulate different places along the cochlear partition, which in turn leads to different populations of auditory nerve fibers responding to each frequency. In this case, the offset of the first tone and the onset of the second are always perceived, whether the gap is present or not. Thus the perceived onset or offset is no longer a reliable cue, and the gap can only be detected by a timing comparison across different neural channels (e.g., Hanekom and Shannon, 1998). These two cases are therefore often referred to as ''within-channel'' and ''between-channel'' gap detection, respectively.[1]

This explanation relies on the fact that stimulus frequency is a neurally and perceptually relevant dimension: If the auditory system were not frequency selective, then no discontinuity between the two markers would be perceived, regardless of the frequency difference. Similarly, gap detection might be used to probe other, higher-level, organizational principles in the auditory system. Two dimensions that are known to have neural representations established at a level higher than the cochlea are spatial location (Moore, 1991; Brainard, 1994) and periodicity (Langner, 1992). Recently, Phillips *et al.* (1998) reported finding an influence of spatial perceptual channels on gap detection. In their experiment, listeners were seated individually in a room with two

[a)]Electronic mail: oxenham@mit.edu

 0001-4966/2000/107(4)/2215/9/$17.00

loudspeakers, positioned to the left and right of the listener. In one condition, listeners were required to detect a gap in a broadband noise emitted from one of the loudspeakers. As in studies using headphones, thresholds were generally less than 5 ms. In the second condition, the sound preceding the gap (marker 1) was played from one loudspeaker, while the sound following the gap (marker 2) was played from the other loudspeaker. Here, thresholds for gap detection were much worse, often by an order of magnitude. Phillips *et al.* suggested that this was due to the auditory system representing the two stimuli in different spatially tuned neural channels and that, as with the markers of different frequencies, the auditory system was not able to compare timing efficiently across these different channels.

While the study of Phillips *et al.* (1998) suggests that gap detection may indeed be a straightforward way of probing higher-order neural representations of sound, some questions remain. First, the localization of sound is generally achieved by combining information from the two ears in the form of interaural time differences (ITDs) and interaural level differences (ILDs) (Blauert, 1997). From the study of Phillips *et al.*, it is not clear which of these cues was dominant in producing the deterioration in gap detection. Second, presenting the markers from different locations results in monaural level differences between the two markers, at least for frequencies above about 500 Hz (Blauert, 1997). It has long been known that a monaural level difference between the two markers produces a deterioration in gap detection (Plomp, 1964b; Plack and Moore, 1991). Thus it is possible that the results of Phillips *et al.* (1998) do not represent an effect of spatial hearing at all, but rather reflect the monaural effect of introducing a level difference across the gap. These questions can be addressed by presenting the sounds over headphones. In this way, ITDs and ILDs can be manipulated independently to assess the relative contributions of each to changes in gap detection. Furthermore, the relative contribution of the monaural system can be assessed by comparing binaural and monaural performance using the same level differences.

Another acoustic dimension that seems to be represented in the higher levels of the auditory system is periodicity, or repetition rate (Langner, 1992). For complex sounds, this representation is believed to be orthogonal to the tonotopic representation of frequency, already established in the cochlea. It may be especially relevant for complex harmonic sounds, such as speech, where higher-order harmonics are no longer spectrally resolved in the auditory system, but where they combine within individual auditory filters to form periodic waveforms with a repetition rate equal to that of the fundamental frequency (f_0). Complexes comprising only high-order, unresolved harmonics can be used to dissociate the effects of spectral and temporal cues (Vliegen and Oxenham, 1999). Although the pitch of complexes containing only unresolved harmonics is weaker than that of resolved harmonic complexes, it remains reasonably clear and can be used to make judgments of musical intervals (Houtsma and Smurzynski, 1990).

If periodicity represents an important neural coding principle, and if gap detection reflects such coding, then gap detection should be worse for conditions where the f_0 of the two markers is different, even if the spectral envelope and overall level remain constant. This seems not to have been tested before, although some evidence suggests that gap *discrimination* deteriorates somewhat as the f_0 difference between the markers is increased (Vliegen *et al.*, 1999). Further evidence for the effect of periodicity on gap detection comes from a study of cochlear implant patients (Chatterjee *et al.*, 1998). In their experiment, Chatterjee *et al.* found that when stimulating a single electrode, performance was generally worst when the pulse rates of the two markers were very different. This is somewhat analogous to changing the f_0 while keeping the spectral envelope constant.

The experiments described here explore the effects of spatial and temporal cues on auditory gap detection. In the first experiment, the roles of ITDs, ILDs, and monaural level differences on gap detection in broadband noise were investigated. In the second experiment, unresolved harmonic complexes were used to separate the roles of spectral envelope and periodicity in gap detection.

I. EXPERIMENT 1. INFLUENCE OF SPATIAL PERCEPTUAL CHANNELS ON GAP DETECTION

A. Stimuli

Gap-detection thresholds were measured using broadband noise as markers. In the binaural conditions, the markers were either identical in the two ears (diotic), or had an ILD of 12 dB or an ITD of 640 μs. These differences were sufficient to fully lateralize the stimuli to one or the other side (Blauert, 1997). The following ITD conditions were tested: (1) Both markers lateralized to the left; (2) both markers lateralized to the right; (3) the first and second markers lateralized to the left and right, respectively; and (4) the first and second markers lateralized to the right and left, respectively. The ILD condition comprised marker 1 being lateralized to the left and marker 2 to the right. In two of the four monaural conditions, the ILD condition was repeated with either the left or the right headphone disconnected, such that a monaural level difference occurred between the two markers, with either the first marker (left ear only) or the second marker (right ear only) 12 dB higher. In the other two monaural conditions, the markers were at the same level and were both presented to either the left or the right ear. A schematic diagram of some of the stimulus configurations is shown in Fig. 1.

The spectrum level of the markers was 36 dB SPL, except in conditions involving a level difference. In such cases, the spectrum levels were 36 and 24 dB SPL (nominal overall levels of 79 and 67 dB SPL). The duration of the first marker was fixed at either 10 or 150 ms. Two durations were tested because Phillips *et al.* (1997, 1998) have reported that a short initial marker increases the difference in performance between within-and between-channel gap detection. The duration of the second marker was varied randomly in each interval from 100 to 300 ms, with uniform distribution. This was intended to render a possible cue of overall duration unreliable (Formby and Forrest, 1991). The markers were gated abruptly; because they were broadband, no audible

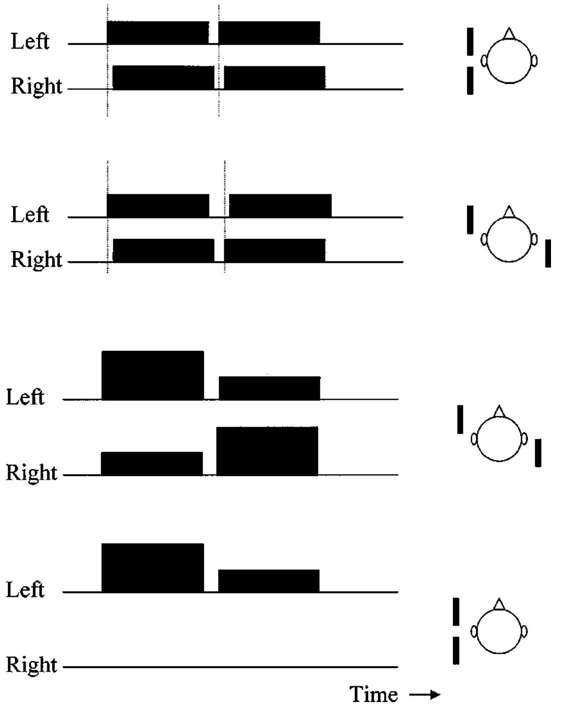

FIG. 1. Schematic diagram of the stimulus configurations for the ITD conditions and the binaural and monaural level differences (left column), and the perceived lateralization of the stimuli (right column). In the first panel, the left side leads by 640 μs for both the first and second marker; both markers are lateralized to the left. In the second panel, left leads right for the first marker, but right leads left for the second marker. This results in a shift in lateralization between the two markers. In the third panel, left is 12 dB higher in level than right for the first marker, and vice versa for the second marker, again leading to a shift in lateralization between the two markers. In the fourth panel, the stimulus is presented only to the left ear, but with the same level difference (12 dB) as used in the third panel.

spectral splatter was generated by this procedure. Before each trial, a new 655-ms sample of Gaussian noise was generated. The four noise bursts (markers 1 and 2 for both intervals) were cut independently and randomly with replacement from this long buffer.

All stimuli were generated digitally at a sampling rate of 50 kHz, and played out via a digital-to-analog converter (TDT DA2), The stimuli were lowpass filtered at 20 kHz (TDT FT5) and passed through a programmable attenuator (TDT PA4) before being presented to the listener via a headphone buffer (TDT HB6) and a Sony MDR-V6 headset.

B. Procedure

Thresholds for the minimum detectable gap were measured using a two-interval, two-alternative forced-choice method with a 3-down 1-up interleaved adaptive tracking procedure, which tracks the 79%-correct point on the psychometric function. Listeners were presented with two intervals, separated by an interstimulus interval of 400 ms, and were required to select the interval containing the gap. Correct-answer feedback was provided after each trial. In the ''no gap'' interval, the two markers generally abutted each other, such that there was no gap between the offset of marker 1 and the onset of marker 2. The exception to this was in the ITD conditions where the lateralization changed between markers. In these cases, even in the ''no gap'' interval, a gap of 1.28 ms ($2\times 640\,\mu$s) was present in one of the two ears. This gap was not reported as being heard by any of the listeners, which is not surprising given that the minimum detectable gap is usually found to be between 2 and 3 ms. Thresholds in this condition are given as the duration of the longer of the two gaps.

The initial gap duration was set to 80 ms, or 20 ms in cases where it became clear in pilot experiments that thresholds were less than 10 ms. Three independent tracks were interleaved in each run, so that each trial within a run was selected randomly from one of the three tracks. The initial step size was an increase or decrease by a factor of 1.78. After three reversals, the step size in that track was reduced to an increase or decrease by a factor of 1.26. Each track was terminated after five reversals, and the (geometric) mean of the gap duration at the last two reversals from each of the three tracks was defined as threshold. Thus each track threshold was the mean of six values. Reported thresholds are the (geometric) mean of at least three such runs for each listener. In conditions involving a level difference, thresholds were found to be more variable and so a total of six runs were measured. The conditions were not run in any particular order and listeners were often presented with a number of different condition in one session.

C. Subjects

Six female listeners participated as subjects in this experiment. All had pure-tone thresholds of 15 dB HL or less at octave frequencies between 250 and 8000 Hz. All but one were college students and all were paid for their participation. The ages of the listeners ranged from 21 to 45 years (median age: 28.5). All listeners were given at least 2-h practice, spread across the different conditions, before data were collected. Two of the listeners had previous experience in psychoacoustic tasks.

D. Results

The pattern of results was similar for all listeners, and so only mean data are shown in Fig. 2. The error bars denote ± 1 standard error of the mean across listeners. The left panel shows results from conditions in which there was no interaural or level difference between the markers. The filled squares represent the diotic condition and the circles and triangles represent results from the left and right ear only, respectively. The results are broadly in line with previous studies of gap detection in noise (Plomp, 1964b; Penner, 1977; Forrest and Green, 1987; Green and Forrest, 1989); thresholds for the 150-ms marker 1 are generally around 3 ms. Thresholds for the 10-ms marker are somewhat higher, with mean thresholds of between 6 and 7 ms. No general ear advantage is apparent, and there seems to be little if any benefit gained from having the stimuli presented to both ears. By pooling the monaural and diotic conditions for each listener and using a repeated-measures analysis of variance (ANOVA), the difference in thresholds between the two marker 1 durations was found to be significant ($F_{1,5}=11.05$; $p<0.05$).

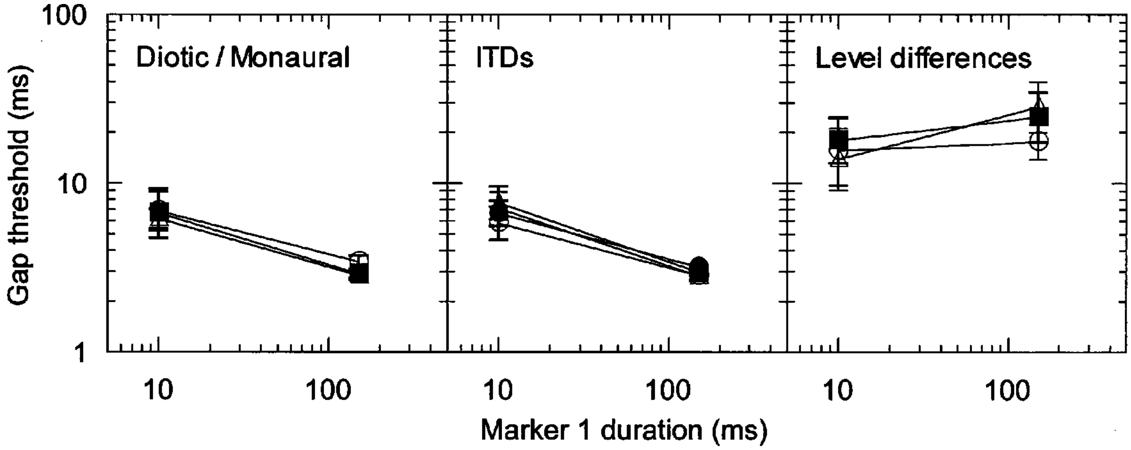

FIG. 2. Mean thresholds for detecting a gap between two broadband noise markers that were presented monaurally or diotically (left panel), with 640-μs interaural time differences (middle panel), or with 12-dB interaural or intermarker level differences (right panel). Error bars represent ± 1 standard error of the mean across six listeners. In the left panel, filled squares represent thresholds in the diotic condition, while the triangles and circles represent thresholds from stimuli presented monaurally, to the right and left ear, respectively. In the middle panel, circles denote conditions where left led right for marker 1; triangles denote conditions where right led left for marker 1. Open symbols represent conditions where the ITD was the same for both markers, and filled symbols represent conditions where the ITD was reversed for the second marker. In the right panel, circles denote the conditions presented to the left ear only, where marker 1 was 12 dB higher in level than marker 2. Triangles denote the conditions presented to the right ear only, where marker 1 was 12 dB lower in level than marker 2. Filled squares represent the binaural conditions, where the two previous conditions were presented together, producing an interaural level difference of 12 dB which was reversed between the two markers.

The middle panel of Fig. 2 shows data from binaural conditions with ITDs. Circles show conditions where marker 1 was lateralized to the left; triangles show where marker 1 was lateralized to the right. Open symbols represent conditions where both markers were lateralized to the same side; filled symbols represent conditions where the lateralization was reversed from marker 1 to marker 2. If perceived position had a large effect on gap thresholds, as proposed by Phillips *et al.* (1998), one might expect thresholds for conditions where the lateralization reverses across the gap to be higher than for conditions where the lateralization remains constant. In fact, the middle panel shows that changing the lateralization has essentially no effect on thresholds. This is despite the fact that all listeners reported hearing a large shift in the perceived location of the noise in these conditions. As with the monaural and diotic conditions, the difference in thresholds between the two marker 1 durations was found to be significant ($F_{1,5}=16.87$; $p<0.01$).

The right panel shows the effect of introducing monaural and binaural level differences. In many cases, performance was worse than in the diotic condition (squares in the left panel) by up to an order of magnitude. The mean threshold, pooled across all level-difference conditions, was 19 ms. The differences in thresholds between the monaural/diotic conditions and the level-difference conditions were highly significant for both the 10-ms marker 1 ($F_{1,5}=26.37$; $p<0.005$) and the 150-ms marker 1 ($F_{1,5}=50.38$; $p<0.001$). Listeners generally found these conditions more difficult and reported that the cues were different from those used in the other conditions. The idea that a different process may have been involved is supported by the somewhat higher variability associated with these thresholds, even when plotted on a log scale. There was, however, no noticeable difference in thresholds between the binaural and monaural level differences. In contrast to the previous two conditions, an ANOVA, again using individual mean data pooled across the three level-difference conditions, showed no significant effect of marker 1 duration ($F_{1,5}=3.13$; $p>0.1$).

E. Discussion

The data from the diotic and ITD conditions (left and middle panels of Fig. 2, respectively) show that thresholds for the 10-ms marker are generally higher than for the 150-ms marker. This is not consistent with the findings of Phillips *et al.* (1997, 1998), who found little or no effect of marker 1 duration in conditions where the markers had the same spectrum or location. The present results are also inconsistent with the results of most previous headphone studies, which have also reported little or no effect of marker duration (Abel, 1972; Penner, 1977; Forrest and Green, 1987). However, the results do seem to be consistent with a more recent study (Snell and Hu, 1999), in which gap detection was found to be poorest for a gap placed near the onset or the offset of a 150-ms broadband noise. It is not clear what accounts for these discrepancies. One consideration is that the amount of practice given to listeners in the present study was generally less than that reported in previous studies. It is possible that thresholds would have decreased further in the 10-ms conditions, if more extensive practice had been given. In support of this conjecture, He *et al.* (1999) also found an increase in gap thresholds as they decreased the overall marker duration from 400 to 100 ms. They used groups of young and aged listeners, all of whom had received only about 30 min practice on each condition. Also, Snell and Hu (1999) found that the effect of gap placement was greatest for their inexperienced listeners.

Unlike the monaural/diotic and ITD conditions, there was no consistent effect of marker 1 duration in the level difference conditions (Fig. 2, right panel). In contrast, Phillips *et al.* (1997, 1998) found that, for most of their listeners, thresholds increased with decreasing marker 1 duration. Again, the reason for this discrepancy is not clear. However, some listeners reported that they were comparing the intervals between the *onsets* of the two markers to determine which interval contained the gap. If used at all, such a strategy would be more effective with the 10-ms marker, as the interval between the marker onsets is smaller for the shorter marker 1 than for the longer one.

Despite these relatively small differences, the main results of experiment 1 are clear: Gap detection is not affected by changing the ITD between marker 1 and marker 2 (Fig. 2, middle panel), suggesting that a change in perceived location does not necessarily elevate gap-detection thresholds. Although the ILD across the gap did have a large effect on gap-detection thresholds, essentially the same deterioration was produced by a monaural level difference between the two markers. Introducing binaural information produced no deterioration beyond that observed for the monaural level differences. This suggests that the binaural system plays little or no role in these tasks and that thresholds in all conditions can be understood solely by considering the signals presented to each ear separately. This in turn suggests that the spatial effect on gap detection reported by Phillips *et al.* (1998) may have been due to the monaural level differences introduced at each ear by presenting the stimuli from different sides of the head. Thus the effect of marker location reported by Phillips *et al.* is probably not a reflection of spatially tuned neural channels, but instead reflects the well-known consequence of presenting the two markers at different levels (Phomp, 1964b; Penner, 1977; Plack and Moore, 1991).

The finding that lateralization *per se* has no effect on gap thresholds is consistent with one condition tested by Phillips *et al.* (1997). In their experiment 4, they found that presenting marker 1 to the left ear only and marker 2 to both ears had only a small effect on gap thresholds, relative to the monaural condition. It therefore seems that if monaural information is available, listeners' performance is not greatly affected by binaural information.

The present results may also be useful in interpreting the results of Boehnke and Phillips (1999). As in Phillips *et al.* (1998), they measured gap detection using two loudspeakers. However, instead of fixing the speakers to the left and right of the listener, they measured gap-detection thresholds for a number of relative speaker locations. The results showed that performance depended primarily on whether the speakers were located ipsilaterally or contralaterally, and that the exact positions of the speakers on each side did not significantly affect thresholds. Boehnke and Phillips (1999) interpreted these results as providing evidence for two broadly tuned neural channels, roughly representing the left and right hemispheres. They related their findings to animal physiological and behavioral data showing fairly broad azimuthal tuning in auditory cortical regions, with each hemisphere best represented in contralateral cortical regions. In light of the present results, however, an alternative explanation can be made in terms of monaural level differences between the two markers. Large overall differences in ILD only occur when two stimuli are located in opposite hemispheres. Therefore, the pattern of results observed by Boehnke and Phillips (1999) is exactly what would also be predicted simply on the basis of monaural level differences.

The question remains as to why a level difference between the two markers has such a large effect on gap detection. It is interesting that threshold values and variability are more akin to a between-channel task than a within-channel one, despite the fact that the two markers are identical in everything but level. Plomp suggested that in the case of marker 1 being higher in level than marker 2, the onset of marker 2 was masked by marker 1. Thus the function relating gap threshold with the level of marker 2 may be treated as a measure of forward masking, or the ''decay of auditory sensation'' (Plomp, 1964b). Similarly, backward masking may govern conditions where marker 1 is lower in level than marker 2 (Plack and Moore, 1991). However, in this case it is difficult to reconcile the present results with our current understanding of backward masking, as it is generally accepted that little or no backward masking occurs for gaps longer than about 10 ms (e.g., Oxenham and Moore, 1994), and thresholds in the present experiment were often greater than 20 ms. It is possible that gap detection is generally performed by detecting an onset transient that is otherwise not detectable. In the case of marker 2 being higher in level than marker 1, the onset of marker 2 is always detectable, whether a gap is present or not. This makes the onset cue unusable for detecting a gap, and listeners may be forced to adopt another, less efficient, strategy involving temporal discrimination, rather than detection.

In summary, for gap detection in broadband noise, perceived changes in lateralization are neither necessary nor sufficient to produce elevated thresholds. At this stage, therefore, there seems to be little reason for postulating any additional influence of spatial coding on gap detection. It is more parsimonious to assume that the results of Phillips *et al.* (1998) and Boehnke and Phillips (1999) are due to monaural level differences between the two markers, rather than to spatially tuned neural channels.

II. EXPERIMENT 2. EFFECT OF F_0 AND SPECTRAL ENVELOPE ON GAP DETECTION

Experiment 1 showed that the effect of spatial separation on gap detection is probably due to monaural level differences, rather than a reflection of spatially tuned neural channels. The role of neural channels tuned to different repetition rates (or f_0s) remains open, although a study in cochlear implant patients (Chatterjee *et al.*, 1998) and a study of temporal discrimination in normal-hearing listeners (Vliegen *et al.*, 1999) both suggest that f_0 differences might have an effect on gap detection. This experiment investigated the separate influences of f_0 and spectral envelope on gap detection by using harmonic tone complexes consisting only of high harmonics, which are generally considered to remain unresolved in the auditory system.

A. Effect of fundamental frequency

1. Method

Equal-amplitude harmonic tone complexes with f_0s of 140 and 350 Hz were used as markers. The components were

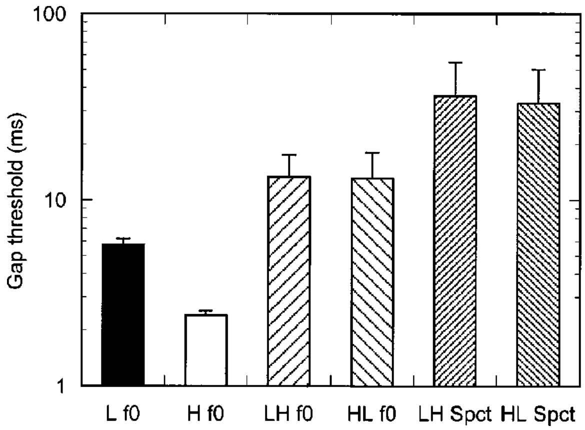

FIG. 3. Mean thresholds for detecting a gap between two harmonic complexes. Error bars represent ±1 standard error of the mean across six listeners. The first four bars (labeled ''f_0'') represent conditions in which the spectral envelope was held constant by bandpass filtering the complexes between 4 and 8 kHz. The f_0 of each marker was either 140 Hz (''*L*'') or 350 Hz (''*H*''). The f_0 s of the two markers were either the same (''*L*'' or ''*H*'') or different (''*LH*'' or ''*HL*''). The last two bars (labeled ''Spct'') represent conditions in which the f_0 was held constant at 140 Hz. One marker was bandpass filtered between 2 and 4 kHz (''*L*''), while the other marker was bandpass filtered between 4 and 6 kHz (''*H*''). The order of these two markers was interchanged to produce the two conditions ''*LH*'' and ''*HL*.''

added in sine phase. The duration of marker 1 was 150 ms; the duration of marker 2 was randomized between 100 and 300 ms. Each marker was gated abruptly and had a random starting phase. The random starting phase ensured that there was usually a phase discontinuity between the two markers in both the ''gap'' and ''no-gap'' intervals of a trial, even when the f_0 was the same. After gating, the stimuli were bandpass filtered with cutoff frequencies of 4000 and 8000 Hz (TDT PF1; attenuation slopes of 72 dB/octave). This eliminated any audible splatter due to gating, and also ensured that only components with harmonic numbers greater than 11 were at full amplitude. Plomp (1964a) has shown that generally only harmonics with numbers up to between 5 and 8 can be resolved by the auditory system.

The overall level of the complex tones after filtering was approximately 72 dB SPL. By keeping the overall level the same, both f_0s should have evoked the same loudness, as the overall bandwidth was also held constant. Four conditions were tested: (1) Both markers had an f_0 of 140 Hz, (2) both markers had an f_0 of 350 Hz; (3) the first and second markers had f_0s of 140 and 350 Hz, respectively; and (4) the first and second markers had f_0s of 350 and 140 Hz, respectively.

Stimuli were presented to the left ear only. Four repetitions of each conditions were measured, and the conditions were presented in random order. Otherwise, the method of stimulus generation, presentation, and measurement was the same as that in experiment 1. The same six listeners participated.

2. Results

Again, the individual patterns of results were fairly similar, and so the (geometric) mean data are plotted as the first four bars of Fig. 3. The first two bars show thresholds when both markers had the same f_0 of either 140 Hz (''Lf_0'') or 350 Hz (''Hf_0''). The next two bars show results for conditions where the f_0 changed across the markers from 140 to 350 Hz (''LHf_0'') or from 350 to 140 Hz (''HLf_0''). All six listeners showed an effect of f_0 for the two same-f_0 conditions (filled and open bars); the average gap threshold at 140 Hz was 5.7 ms compared with 2.4 ms at 350 Hz. Using a paired *t*-test, this difference was found to be highly significant [$t(5)=15.53$; $p<0.001$].

Four of the six listeners and the mean data showed a deterioration in performance when the f_0 was altered across the gap. For the other two listeners, performance was similar to that of the lower f_0. The mean threshold, pooled across listeners and the two conditions, was 13.2 ms. This was significantly higher than the threshold for the 140-Hz same-f_0 condition [$t(13)=3.95$; $p<0.005$]. Thus it appears that changes in f_0 may have an effect on gap detection, even if the spectral envelope is held constant.

B. Effect of spectral envelope

The previous section showed an effect of f_0 on gap-detection thresholds, although the effect was generally not as large as has been reported for sinusoids of different frequencies (Formby *et al.*, 1998). This section examines whether a change in spectral region is sufficient to produce a deterioration in gap detection, even if the f_0 (and hence the pitch) stays constant.

1. Method

Complex tones with an f_0 of 140 Hz, generated in sine phase, were used as markers. Again markers were gated on and off abruptly, and the starting phase of each marker was randomized. The markers were filtered after gating with a 2-kHz-wide bandpass filter (TDT PF1; attenuation slopes of 72 dB/octave). The cutoff frequencies were either 2000 and 4000 Hz, or 4000 and 6000 Hz. The overall level of the complex after filtering was approximately 74 dB SPL. Markers 1 and 2 were always filtered into different spectral regions; both low-high and high-low conditions were tested. The two markers were generated and filtered separately and then combined (TDT SM3) before being passed to the headphone buffer. Stimuli were presented to the left ear only. Four repetitions of each condition were measured. All six listeners from experiment 1 participated.

2. Results

The mean results are shown as the last two bars of Fig. 3. It should be noted that the bandwidth of these conditions was 2 kHz instead of 4 kHz. The narrower bandwidth would be expected to produce somewhat higher thresholds even in the absence of a spectral difference between the markers. However, using noise markers, Eddins *et al.* (1992) found that thresholds increased by a factor of $\sqrt{2}$ for a halving in bandwidth. The mean increases observed here are closer to an order of magnitude. Individual thresholds range from around 10 ms for two listeners to over 100 ms for another listener. The mean threshold (pooled across low-high and high-low conditions) was 35 ms. This is significantly higher than both the same-f_0 and different-f_0 conditions with the fixed spectral envelopes ($p<0.005$). It is interesting to note that the two listeners who showed virtually no effect of f_0, and the smallest effect of spectral region, were also the most experienced listeners, although by this stage all listeners had been exposed to at least 20-h gap detection.

C. Discussion

All listeners showed a larger gap threshold for the 140-Hz f_0 than for the 350-Hz f_0. This difference of more than a factor of two may be due to differences in the periodic fluctuations within each stimulated auditory filter. First, the fluctuations are slower for the 140-Hz complex, with a period of 7.14 ms (i.e., greater than the mean threshold) compared with 2.86 ms for the 350-Hz complex. The gaps may therefore be less discriminable from the inherent periodic dips in the filtered envelope. Second, the lower f_0 means that the components are spaced more closely together. Therefore, the waveform at the output of each auditory filter will be a combination of more components, giving a more modulated envelope than for the higher f_0 complex. A similar explanation in terms of random fluctuations (as opposed to the deterministic stimuli used here) has been proposed by Glasberg and Moore (1992) to account for poor gap detection in narrow-band noise.

The results in Fig. 3 show that the effect of changing spectral region was generally much greater than that of changing f_0. This supports the idea that the deterioration in gap detection observed with sinusoidal markers of different frequencies is primarily due to the effects of the frequency-to-place mapping in the cochlea. Changes in f_0 also have an effect, in line with the cochlear-implant study of Chatterjee *et al.* (1998). However, in line with the findings of Vliegen *et al.* (1999), spectral cues seem to have a stronger influence than f_0 cues in temporal discrimination tasks.

While the stimuli were designed to significantly reduce spectral cues in the different f_0 conditions, they cannot be totally ruled out of these experiments. In particular, it is possible that resolved combination tones (Goldstein, 1967; Smoorenburg, 1972a, b), spaced at harmonic integers below the stimulus passband, might have been audible and may have contributed to an overall spectral difference. It seems unlikely, however, that these low-level components had a strong influence on the results. Vliegen and Oxenham (1999) examined the role of resolved combination tones in a sequential streaming task, and found no change in the results when the combination tones were masked by low-pass noise. If combination tones did play a role, then the effect of f_0 on gap detection observed in this experiment may be greater than the ''true'' effect of f_0 alone. This is turn would strengthen the conclusion that spectral and/or level differences are of paramount importance in determining gap thresholds.

III. GENERAL DISCUSSION

Experiment 1 showed that spatial cues have no effect on gap detection, if monaural level differences between the markers are not present. This conflicts with a growing body of literature in which it is claimed that gap detection can be used to examine perceptual organization at levels higher than the cochlea (Phillips *et al.*, 1998; Boehnke and Phillips, 1999; Taylor *et al.*, 1999), a view that is also expounded in a recent review article (Phillips, 1999). Based on the results of experiment 1, it seems likely that all the results ascribed to effects of spatially tuned neural channels can in fact be explained in terms of the resulting monaural level differences between the two markers. In general, it seems that monaural information can be used to detect gaps, even if binaural information encourages perceptual segregation. Because of this, it is probably not possible to use gap detection to investigate the role of spatially tuned neural channels.

Based on data currently available, it seems that differences between the two markers in level and frequency spectrum can both have a strong effect on gap detection. Differences in periodicity (and hence pitch), independent of spectral envelope (or electrode for cochlear implants) also have an effect, but it seems to be weaker than that of spectrum, at least for normal-hearing listeners. This is consistent with recent results from a temporal discrimination task (Vilegen *et al.*, 1999).

A. The role of spatial coding in auditory streaming

At first sight, it may seem surprising that lateralization based on ITD's has no effect on gap detection, and that lateralization based on ILDs has no effect over and above that produced by the monaural level differences. On the other hand, it has been shown that ITDs play essentially no role in the perceptual grouping of concurrent sounds (Culling and Summerfield, 1995). Also, for sequential sounds, Deutsch (1974, 1975) found that frequency cues for pure tones overrode lateralization cues in forming ''auditory streams.'' This was true even though individual tones were presented to one ear at a time, producing an infinite ILD. Thus spatial cues may generally play a secondary role in the formation of auditory objects and streams (Darwin and Carlyon, 1995). If poor performance in gap detection is a reflection of stream segregation, then it follows that spatial cues should have a minimal effect.

B. Explaining within-and between-channel differences in gap detection

Gap detection generally seems poorest when the markers stimulate different peripheral frequency channels, regardless of the perceived pitch or location. There seems to be little consensus as to the mechanisms underlying this effect. Some previous studies have ascribed it to the cognitive load of attending to more than one perceptual channel. For instance, Fitgibbons *et al.* (1974) suggested that the shift of attention from one channel to another during the gap may impair performance. Phillips *et al.* (1997; see also Phillips, 1999) proposed that both channels may be attended to concurrently, but that attending to one channel may reduce the resources available for monitoring events in the other channel. Both these explanations seem unlikely for the following reasons. First, it has been shown for the detection of tones in noise that performance is only slightly degraded when the tone is presented randomly at one of four possible frequencies (Schlauch and Hafter, 1991), suggesting that it is possible to attend to at least four frequency channels without significant attentional cost, at least in a detection paradigm. Second, detection of a multitone complex improves as the number of tones is increased from one to ten, consistent with the predictions of multiple observations in independent frequency channels, with no additional attentional load assumed (Buus *et al.*, 1998). Third, and most relevantly, gaps in multiple spectral bands are more detectable than a gap in a single band (Grose and Hall, 1988; Green and Forrest, 1989; Grose, 1991; Hall *et al.*, 1996), suggesting that listeners are able to efficiently combine information about gaps across frequency. Taken together, these results suggest that it is possible to attend to more than one frequency at a time, and that the sharing of attentional resources cannot explain the deterioration in gap thresholds by an order of magnitude or more when the markers are at different frequencies. Note that in all the studies cited above, the signals (or gaps) occurred at expected spectral locations. The situation is different when a signal has unexpected spectral (Greenberg and Larkin, 1968; Scharf *et al.*, 1987) or temporal (Wright and Dai, 1994) characteristics. However, in the case of all between-channel gap-detection studies so far undertaken, the spectral (or spatial) characteristics of both markers have not been randomized, and so can be treated as being expected by the listener.

In the Introduction, it was suggested that within-and between-channel gap detection may reflect very different detection strategies. For within-channel gap detection, with the markers at the same level, the detection of any transient (onset or offset) is sufficient to signal the presence of a gap, and no judgment of timing is necessary. Furthermore, the task can be done using a one-interval paradigm, where listeners are asked whether or not they detected a gap (He *et al.*, 1999). If the two markers are at different frequencies, transients are detectable whether a gap is present or not. Thus listeners are forced to make a judgment of timing, comparing either the two onsets or the offset of marker 1 with the onset of marker 2. This task is one of discrimination, rather than detection, and would be very difficult to perform in a one-interval paradigm without extensive prior training. This may also explain why performance deteriorates so dramatically in ''within-channel'' gap detection, when a level difference is introduced between the two otherwise identical markers.

In a study of increment and decrement detection, Oxenham (1997) suggested that onsets may provide a more salient cue than do offsets. This suggestion was made to account for the asymmetry between increment and decrement detection at very short durations, but it has some physiological support in the abundance of onset cells in the auditory system (Pickles, 1988). Onset cells are often broadly tuned, with bandwidths sometimes extending over several critical bands. This characteristic may help to explain some interesting results from the study of Formby *et al.* (1998). They found that the detection of a gap between two sinusoidal markers of equal frequency and level was impaired when an additional tone at a different frequency was gated with the second marker. In contrast, performance was hardly affected by gating an additional tone with the first marker. This suggests that the onset of the additional second marker tone made the gap more difficult to detect. It is therefore possible that the additional onset ''masked'' the onset of the second marker by stimulating neural units that would otherwise have provided information as to the presence of the gap. Similar neural mechanisms may be responsible for modulation detection interference, or MDI (Yost *et al.*, 1989).

IV. SUMMARY

In experiment 1, thresholds for detecting a gap between two broadband markers were measured for conditions in which interaural time, interaural level, or monaural across-marker level differences were introduced. Reversing the perceived lateralization of marker 2 with respect to marker 1 by varying the interaural time difference did not affect gap thresholds. Reversing the perceived lateralization by varying the interaural level difference did result in a deterioration in performance. However, the same deterioration was found in monaural conditions, when a level difference was introduced between marker 1 and marker 2. Thus perceived lateralization is neither necessary nor sufficient to produce a deterioration in gap-detection thresholds, while monaural level differences (Plomp, 1964b; Penner, 1977; Plack and Moore, 1991) are sufficient to produce as large an effect as is observed when the markers are presented from loudspeakers located on different sides of the head (Phillips *et al.*, 1998). The results do not support the idea that gap detection can be used to investigate the role of spatially tuned neural channels in temporal processing (Phillips *et al.*, 1998; Boehnke and Phillips, 1999; Phillips, 1999).

In experiment 2, the effect on gap detection of using harmonic-complex markers with different spectral envelopes or fundamental frequencies was investigated using complexes consisting of only high, unresolved harmonics. Changing the fundamental frequency across the gap significantly increased thresholds, but changes in the spectral envelope produced a larger deterioration in thresholds. This suggests that the initial frequency-to-place mapping in the cochlea is dominant in determining gap-detection thresholds for dissimilar markers, while temporal coding plays a secondary role (Vliegen *et al.*, 1999).

ACKNOWLEDGMENTS

This work was supported by NIH/NIDCD Grants Nos. R01 DC 00187 and R01 DC 03909 and was carried out while the author was at the Institute of Hearing, Speech, and Language at Northeastern University, Boston, MA. Karen Stafford assisted in collecting the data. Mary Florentine, Craig Formby, Dennis Phillips, and Chris Plack provided helpful comments on previous versions of this manuscript.

[1]Dichotic presentation, where the markers are presented to separate ears, may be regarded as ''between-channel,'' even if the two markers have the same spectral content, because separate peripheral pathways are being activated by the two markers.

Abel, S. M. (**1972**). ''Discrimination of temporal gaps,'' J. Acoust. Soc. Am. **52**, 519–524.

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge).

Boehnke, S. E., and Phillips, D. P. (**1999**). ''Azimuthal tuning of human perceptual channels for sound location,'' J. Acoust. Soc. Am. **106**, 1948–1955.

Brainard, M. S. (**1994**). ''Neural substrates of sound localization,'' Curr. Opin. Neurobiol. **4**, 557–562.
Buus, S., and Florentine, M. (**1985**). ''Gap detection in normal and impaired listeners: The effect of level and frequency,'' in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer, New York).
Buus, S., Muesch, H., and Florentine, M. (**1998**). ''On loudness at threshold,'' J. Acoust. Soc. Am. **104**, 399–410.
Chatterjee, M., Fu, Q. J., and Shannon, R. V. (**1998**). ''Within-channel gap detection using dissimilar markers in cochlear implant listeners,'' J. Acoust. Soc. Am. **103**, 2515–2519.
Culling, J. F., and Summerfield, Q. (**1995**). ''Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay,'' J. Acoust. Soc. Am. **98**, 785–797.
Darwin, C. J., and Carlyon, R. P. (**1995**). ''Auditory organization and the formation of perceptual streams,'' in *Handbook of Perception and Cognition*, edited by B. C. J. Moore (Academic, San Diego), Vol. 6.
Deutsch, D. (**1974**). ''An auditory illusion,'' Nature (London) **251**, 307–309.
Deutsch, D. (**1975**). ''Two-channel listening to musical scales,'' J. Acoust. Soc. Am. **57**, 1156–1160.
Eddins, D. A., Hall, J. W., and Grose, J. H. (**1992**). ''Detection of temporal gaps as a function of frequency region and absolute noise bandwidth,'' J. Acoust. Soc. Am. **91**, 1069–1077.
Fitzgibbons, P. J., Pollatsek, A., and Thomas, I. B. (**1974**). ''Detection of temporal gaps within and between perceptual tonal groups,'' Percept. Psychophys. **16**, 522–528.
Formby, C., and Forrest, T. G. (**1991**). ''Detection of silent temporal gaps in sinusoidal markers,'' J. Acoust. Soc. Am. **89**, 830–837.
Formby, C., Sherlock, L. P., and Li, S. (**1998**). ''Temporal gap detection measured with multiple sinusoidal markers: Effects of marker number, frequency, and temporal position,'' J. Acoust. Soc. Am. **104**, 984–98.
Forrest, T. G., and Formby, C. (**1996**). ''Temporal gap detection thresholds in sinusoidal markers simulated with a single-channel envelope detector,'' Aud. Neurosci. **3**, 21–33.
Forrest, T. G., and Green, D. M. (**1987**). ''Detection of partially filled gaps in noise and the temporal modulation transfer function,'' J. Acoust. Soc. Am. **82**, 1933–1943.
Glasberg, B. R., and Moore, B. C. J. (**1992**). ''Effects of envelope fluctuations on gap detection,'' Hear. Res. **64**, 81–92.
Goldstein, J. L. (**1967**). ''Auditory nonlinearity,'' J. Acoust. Soc. Am. **41**, 676–689.
Green, D. M., and Forrest, T. G. (**1989**). ''Temporal gaps in noise and sinusoids,'' J. Acoust. Soc. Am. **86**, 961–970.
Greenberg, G. S., and Larkin, W. D. (**1968**). ''Frequency-response characteristics of auditory observers detecting signals of a single frequency in noise: The probe-signal method,'' J. Acoust. Soc. Am. **44**, 1513–1523.
Grose, J. H. (**1991**). ''Gap detection in multiple narrow bands of noise as a function of spectral configuration,'' J. Acoust. Soc. Am. **90**, 3061–3068.
Grose, J. H., and Hall, J. W. (**1988**). ''Across-frequency processing in temporal gap detection,'' in *Basic Issues in Hearing*, edited by H. Duifhuis, H. P. Wit, and J. P. Horst (Academic, New York).
Grose, J. H., Hall, J. W., and Buss, E. (**1999**). ''Modulation gap detection: Effects of modulation rate, carrier separation, and mode of presentation,'' J. Acoust. Soc. Am. **106**, 946–953.
Hall, J. W., Grose, J. H., and Joy, S. (**1996**). ''Gap detection for pairs of noise bands: Effects of stimulus level and frequency separation,'' J. Acoust. Soc. Am. **99**, 1091–1095.
Hanekom, J., and Shannon, R. V. (**1998**). ''Gap detection as a measure of electrode interaction in cochlear implants,'' J. Acoust. Soc. Am. **104**, 2372–2384.
He, N., Horwitz, A. R., Dubno, J. R., and Mills, J. H. (**1999**). ''Psychometric functions for gap detection in noise measured from young and aged subjects,'' J. Acoust. Soc. Am. **106**, 966–978.
Heinz, M. G., Goldstein, M. H., and Formby, C. (**1996**). ''Temporal gap detection thresholds in sinusoidal markers simulated with a multi-channel, multi-resolution cochlear model,'' Aud. Neurosci. **3**, 35–56.
Houtsma, A. J. M., and Smurzynski, J. (**1990**). ''Pitch identification and discrimination for complex tones with many harmonics,'' J. Acoust. Soc. Am. **87**, 304–310.
Langner, G. (**1992**). ''Periodicity coding in the auditory system,'' Hear. Res. **60**, 115–142.
Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1993**). ''Detection of temporal gaps in sinusoids: Effects of frequency and level,'' J. Acoust. Soc. Am. **93**, 1563–1570.
Moore, D. R. (**1991**). ''Anatomy and physiology of binaural hearing,'' Audiology **30**, 125–134.
Oxenham, A. J. (**1997**). ''Increment and decrement detection in sinusoids as a measure of temporal resolution,'' J. Acoust. Soc. Am. **102**, 1779–1790.
Oxenham, A. J., and Moore, B. C. J. (**1994**). ''Modeling the additivity of nonsimultaneous masking,'' Hear. Res. **80**, 105–118.
Penner, M. J. (**1977**). ''Detection of temporal gaps in noise as a measure of the decay of auditory sensation,'' J. Acoust. Soc. Am. **61**, 552–557.
Phillips, D. P. (**1999**). ''Auditory gap detection, perceptual channels, and temporal resolution in speech perception,'' J. Am. Acad. Audiol. **10**, 343–354.
Phillips, D. P., Hall, S. E., Harrington, I. A., and Taylor, T. L. (**1998**). '' 'Central' auditory gap detection: A spatial case,'' J. Acoust. Soc. Am. **103**, 2064–2068.
Phillips, D. P., Taylor, T. L., Hall, S. E., Carr, M. M., and Mossop, J. E. (**1997**). ''Detection of silent intervals between noise activating different perceptual channels: Some properties of central auditory gap detection,'' J. Acoust. Soc. Am. **101**, 3694–3705.
Pickles, J. O. (**1988**). *An Introduction to the Physiology of Hearing* (Academic, London).
Plack, C. J., and Moore, B. C. J. (**1991**). ''Decrement detection in normal and impaired ears,'' J. Acoust. Soc. Am. **90**, 3069–3076.
Plomp, R. (**1964a**). ''The ear as a frequency analyzer,'' J. Acoust. Soc. Am. **36**, 1628–1636.
Plomp, R. (**1964b**). ''The rate of decay of auditory sensation,'' J. Acoust. Soc. Am. **36**, 277–282.
Scharf, B., Quigley, S., Aoki, C., Peachy, N., and Reeves, A. (**1987**). ''Focused auditory attention and frequency selectivity,'' Percept. Psychophys. **42**, 215–223.
Schluach, R. S., and Hafter, E. R. (**1991**). ''Listening bandwidths and frequency uncertainty in pure-tone signal detection,'' J. Acoust. Soc. Am. **90**, 1332–1339.
Smoorenburg, G. F. (**1972a**). ''Audibility region of combination tones,'' J. Acoust. Soc. Am. **52**, 603–614.
Smoorenburg, G. F. (**1972b**). ''Combination tones and their origin,'' J. Acoust. Soc. Am. **52**, 615–632.
Snell, K. B., and Hu, H. (**1999**). ''The effect of temporal placement on gap detectability,'' J. Acoust. Soc. Am. **106**, 3571–3577.
Taylor, T. L., Hall, S. E., Boehnke, S. E., and Phillips, D. P. (**1999**). ''Additivity of perceptual channel-crossing effects in auditory gap detection,'' J. Acoust. Soc. Am. **105**, 563–566.
Vliegen, J., Moore, B. C. J., and Oxenham, A. J. (**1999**). ''The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task,'' J. Acoust. Soc. Am. **106**, 938–945.
Vliegen, J., and Oxenham, A. J. (**1999**). ''Sequential stream segregation in the absence of spectral cues,'' J. Acoust. Soc. Am. **105**, 339–346.
Williams, K. N., and Perrott, D. T. (**1972**). ''Temporal resolution of tonal pulses,'' J. Acoust. Soc. Am. **51**, 644–647.
Wright, B. A., and Dai, H. (**1994**). ''Detection of unexpected tones with short and long durations,'' J. Acoust. Soc. Am. **95**, 931–938.
Yost, W. A., Sheft, S., and Opie, J. (**1989**). ''Modulation interference in detection and discrimination of amplitude modulation,'' J. Acoust. Soc. Am. **86**, 2138–2147.