ELSEVIER

Hearing Research 174 (2002) 149–157

HEARING RESEARCH

www.elsevier.com/locate/heares

Concurrent motion detection based on dynamic changes in interaural delay

Kourosh Saberi [a,*], Prisilia Tirtabudi [b], Agavni Petrosyan [a], David R. Perrott [b], Thomas Z. Strybel [c]

[a] *Department of Cognitive Sciences, University of California, Irvine, CA 92697, USA*
[b] *Department of Psychology, California State University, Los Angeles, CA 90032, USA*
[c] *Department of Psychology, California State University, Long Beach, CA 90840, USA*

Received 11 January 2002; accepted 28 August 2002

Abstract

The ability to detect a dynamic change in the interaural delay of a pure tone in the presence of a distracter tone of a different frequency was investigated in four conditions: (1) a control condition in which no distracter tone was present, (2) the distracter tone was stationary (fixed interaural delay), (3) the distracter had an interaural delay that changed in the same direction as that of the target tone, i.e., concurrent auditory motion in the same direction, and (4) the distracter had an interaural delay that changed in a direction opposite to that of the target tone, i.e., concurrent auditory motion in opposite directions. In a cued single-interval two-alternative forced-choice design, the observer had to determine if the target tone had a constant or dynamic interaural delay. The target was a 500-Hz tone and the distracter was a tone with a frequency of 300, 510, 550, 600, 800, or 1000 Hz. Detection was also examined for a range of stimulus durations, rates of change in interaural delay (i.e., velocity), and extent of change in interaural time difference (i.e., 'distance'). Results showed that the best performance (highest d') was associated with the no-distracter condition, followed by the stationary-distracter, opposite-direction, and same-direction conditions, respectively. Detection improved with increasing frequency difference between distracter and target tones, but was nonetheless lower than that associated with the no-distracter condition, even when the distracter frequency was several critical bands removed from the target frequency.

Key words: Binaural; Dynamic interaural delay; Auditory motion

1. Introduction

Several studies in recent years have investigated signal detection and identification in multisource environments. These studies have examined the recognition of auditory temporal patterns presented through separated sources (Arbogast and Kidd, 2000; Kidd et al., 1998), localization in virtual environments (Bolia et al., 1999), localization of concurrently presented tones from different free-field sources (Perrott, 1984), and speech intelligibility in the presence of simultaneously active, but spatially segregated interfering speech, i.e., the so-called cocktail-party effect (Sayers and Cherry, 1957; Yost et al., 1996; Yost, 1997). As a rule, signal detection, identification, or localization improves with increasing spatial or spectral separation between signal and distracters and declines with increasing number of interfering sources.

Often, however, a natural multisource auditory environment requires signal detection under conditions of changing interaural information, either as a result of a target source in motion or due to a listener's head rotation. A number of previous studies have investigated

* Corresponding author. Tel.: +1 (949) 824 6310; Fax: +1 (949) 824 2307.
E-mail address: kourosh@uci.edu (K. Saberi).

Abbreviations: ITD, interaural time difference; ILD, interaural level difference; MDI, modulation detection interference

0378-5955/02/$ – see front matter
PII: S0378-5955(02)00652-4

the roles of velocity (Perrott and Musicant, 1977b; Waugh et al., 1979; Grantham, 1986), movement trajectory (Saberi and Perrott, 1990), stimulus spectrum (Perrott and Tucker, 1988; Chandler and Grantham, 1992; Saberi, 1996; Saberi and Hafter, 1995) as well as other dynamic stimulus features in motion detection (Perrott and Nelson, 1969; Perrott and Musicant, 1977a; Grantham, 1986; Strybel et al., 1989; Saberi and Hafter, 1997). These studies have shown that motion detection is optimum for mid-range velocities, and better for azimuthal or oblique trajectories than for vertical trajectories. Studies that have isolated the role of dynamic interaural cues have revealed that at high velocities, motion detection is less salient when based on a dynamic interaural time difference (ITD) compared to a dynamic interaural level-difference cue (ILD). This finding has been referred to as 'lag of lateralization' (Blauert, 1972) or 'binaural sluggishness' (Grantham, 1984; Grantham and Wightman, 1978) and implies a low-pass filtering of the rate of changing interaural delay.

An area of research that has yet to be investigated is the auditory processing of motion in the presence of interfering sound sources or in multisource environments. Previous studies have shown that dynamic covariation of a non-target sound may interfere with the detection of a target signal, even when the target and interfering waveforms are spectrally separated by more than a critical band such as that observed for modulation detection interference (MDI; Yost and Sheft, 1989; Yost et al., 1989). Studies of binaural interference have also shown interference with detection or lateralization of spatially stationary waveforms (Saberi, 1995; Trahiotis and Bernstein, 1990; Heller and Trahiotis, 1994).

The current study investigated the ability of human observers to detect the motion (i.e., dynamic change in ITD) of a target tone of one frequency in the presence of a distracter tone of a different frequency. Motion was simulated using a binaural-beat paradigm. To emphasize motion cues we randomized both the initial ITD of motion, as well as the ITD of stationary control tones to eliminate absolute position cues, and restrict the range of peak ITD to within less than half the period of the highest frequency employed which eliminates phase wrapping and loudness fluctuations at extreme lateral positions. We examined four primary distracter conditions: (1) no distracter, (2) the distracter was stationary (no motion), (3) the distracter had an ITD changing in the same direction as that of the target (motion in the same direction), and (4) the distracter had an ITD changing in the opposite direction to that of the target (motion in opposite direction). When the distracter ITD was dynamic, it had the same rate of change as the ITD of the target tone (constant velocity). For conditions 2–4 we examined the ability to detect motion in the presence of a distracter frequency of $f_d = f_t + \Delta$ Hz where the magnitude of Δ ranges from 10 to 500 Hz. This range of frequency differences ensured conditions in which the target and distracter frequencies fell within as well as outside a critical band. Finally, we examined the effects of stimulus duration, distance[1] traveled (in ITD units of μs), and motion velocity on concurrent motion discrimination.

2. Materials and methods

Four subjects, two male and two females, participated in this experiment. All had normal hearing based on self-report, and were paid an hourly wage for their participation. Experiments were conducted in an acoustically isolated steel chamber (Industrial Acoustics Company). The target stimulus was a dichotic tone of frequency f_t Hz presented to the left ear and $f_t + \Delta f_t$ Hz presented to the right ear. The distracter frequency was $f_d = f_t + \Delta$ where $\Delta = -200$, 10, 50, 100, 300, or 500 Hz presented to the left ear and $f_d + \Delta f_d$ presented to the right ear (see below). Stimuli were generated in a Dell OptiPlex-Gx1 computer and presented binaurally through digital-to-analog converters (Sound Blaster Live, −120 dB noise floor) and Sony Headphones (MDR-V1) at a sampling rate of 10 kHz and a level of 65 dB SPL. They were low-pass filtered at 5 kHz and had linear rise–decay times of 10 ms.

2.1. Stimulus conditions

There were four primary stimulus conditions. In all conditions, motion was simulated by increasing the frequency of the waveform to the right ear. The waveforms to the two channels for the four conditions were:

2.1.1. Condition 1: no distracter

$$X_l(t) = \sin(2\pi f_t t + \phi_t)$$
$$X_r(t) = \sin\left(2\pi f_t \left[1 + \frac{V}{10^6}\right] t\right) \quad (1)$$

where $X_l(t)$ and $X_r(t)$ are the time waveforms to the left and right ears respectively, $f_t = 500 + \varepsilon$ Hz is the frequency to the left ear, ε is a 25-Hz random perturbation on each trial to eliminate monaural pitch cues, t is in seconds, V is the rate of change in interaural delay (i.e.,

[1] For convenience we will refer to a change in ITD as 'distance' traveled, partly because a change in azimuthal displacement in the free-field produces a corresponding monotonic change in ITD, and partly because a change in ITD is usually associated with a lateral position change of an intracranial image along the interaural axis.

velocity) ranging from 200 to 1600 µs/s as described below, and

$$\phi_t = 2\pi f_t \frac{\mathrm{ITD_i}}{10^6}$$
$$\text{where} \begin{cases} \mathrm{ITD_i} \sim U(0, \mathrm{ITD_m}) \text{ if motion} \\ \mathrm{ITD_i} \sim U(-\mathrm{ITD_m}, \mathrm{ITD_m}) \text{ if stationary} \end{cases} \quad (2)$$

$\mathrm{ITD_m}$ in µs is the maximum allowable ITD for that run, equal to distance traveled. This value is specified by the experimenter as described in the next section, and is one of the following values depending on the selected stimulus sets: 100, 200, 400, or 800 µs. The initial interaural delay $\mathrm{ITD_i}$ is randomly selected from a uniform distribution with a range from 0 to $\mathrm{ITD_m}$ on motion trials. Note that the positive phase advance is added to the left channel, producing an initial ITD bias favoring the left ear. If the trial was a no-motion trial, then $V=0$, the frequencies to the two ears were the same, and $\mathrm{ITD_i} \sim U(-\mathrm{ITD_m}, \mathrm{ITD_m})$.

Although when the interaural delay was dynamic, the right ear always received the higher frequency (i.e., motion toward the right ear), and although the range of dynamic ITDs was restricted to the region bound by ±'distance' µs, the start point (i.e., initial ITD) was always randomized by ϕ_t to reduce potential cueing on absolute position. For example, if the distance traveled was 200 µs, then the start point could be anywhere from −200 to 0 µs. Negative and positive signs of ITD, by convention, correspond to the waveforms leading to the left and right ears, respectively. Such a randomization severely degrades fixed-location cues, forcing observers to rely on a *change* in ITD. For this example, if the interaural delay was constant (i.e., no-motion trial), then the ITD was randomly selected from −200 to +200 µs. The duration of the sound in the no-motion trials was always matched to that of the motion trials.

2.1.2. Condition 2: stationary distracter

$$\begin{aligned} X_l(t) &= \sin(2\pi f_t t + \phi_t) + \sin(2\pi f_d + \phi_d) \\ X_r(t) &= \sin\left(2\pi f_t \left[1 + \frac{V}{10^6}\right] t\right) + \sin(2\pi f_d t) \end{aligned} \quad (3)$$

where $f_d = f_t + \Delta$ is the distracter frequency, and ϕ_d is the distracter starting phase, selected according to the same rules as, but independent of ϕ_t, the starting interaural phase difference of the target tone (see Eq. 2). Again, on no-motion trials, $V=0$ and the target tones had the same frequency, but different starting phases at two ears. The duration of the distracter sound was always matched to that of the target sound.

2.1.3. Condition 3: Distracter and target tones having the same direction of change in interaural delay (i.e., same direction of motion)

$$\begin{aligned} X_l(t) &= \sin(2\pi f_t t + \phi_t) + \sin(2\pi f_d t + \phi_d) \\ X_r(t) &= \sin\left(2\pi f_t \left[1 + \frac{V}{10^6}\right] t\right) + \sin\left(2\pi f_d \left[1 + \frac{V}{10^6}\right] t\right) \end{aligned} \quad (4)$$

where ϕ_t and ϕ_d are independent uniform deviates as described earlier.

2.1.4. Condition 4: Distracter and target tones having opposite directions of change in interaural delay (i.e., opposite directions of motion)

$$\begin{aligned} X_l(t) &= \sin(2\pi f_t t + \phi_t) + \sin\left(2\pi f_d \left[1 + \frac{V}{10^6}\right] t\right) \\ X_r(t) &= \sin\left(2\pi f_t \left[1 + \frac{V}{10^6}\right] t\right) + \sin(2\pi f_d t + \phi_d) \end{aligned} \quad (5)$$

where ϕ_t and ϕ_d are independent uniform deviates as described earlier.

2.2. Procedure

The experiment was conducted using a block design. On a given run, a combination of distance by velocity was randomly selected from one of five sets of predetermined values. The five distance (ITD)/velocity (V) sets were: (1) 400 µs, 200 µs/s, (2) 800 µs, 800 µs/s, (3) 400 µs, 1600 µ/s, (4) 200 µs, 400 µs/s, and (5) 100 µs, 400 µs/s. Thus, the stimulus durations for the five sets, respectively, were 2, 1, 0.25, 0.5, and 0.25 s. We selected these five sets to cover a wide range of velocities, distances, and durations, and to allow some cross-set comparisons. For example, sets 1 and 3 have the same distance traveled, but different velocities and durations. Sets 4 and 5 have the same velocity, but different distances and durations, and sets 3 and 5 have the same durations, but different distances and velocities. As an aside, we should note that for the longest distance conditions, 800 µs, we did not use distracter frequencies greater than 600 Hz, because the distance traveled would have exceeds the half period of the higher-frequency distracters of 800 and 1000 Hz (cf. Fig. 2).

The run included 50 trials and usually lasted less than 10 min. A minimum of four runs were completed by each subject for each case (i.e., each frequency difference of each of four distracter conditions of each of five stimulus sets). On each run, on a random basis, one of the five distance-by-velocity sets, one of four distracter conditions (i.e., same direction, opposite direction, stationary, or none), and a frequency difference between target and distracter tones were selected when applicable. The experimental design was a cued single-interval two-alternative forced-choice paradigm. On each trial, a 0.5-s monaural cue that signalled the frequency of the target tone was followed by a 0.5-s silent period, followed, with equal *a priori* probability, by either a stationary target tone or a target tone that simulated movement toward the right at the specified velocity and distance. The cue was provided only when a distracter tone was simultaneously presented with the target tone to reduce cognitive errors based on a decay of memory for the target frequency, which was particularly useful when the frequency difference between the target and distracter was small. The observer's task was to respond either 'no-motion' or 'motion' by pressing one of two keys on the computer keypad. If the frequencies of the target tones to the two ears were the same and the subject responded no-motion, a correct-response feedback was displayed on the monitor. If the frequencies were different and the subject responded motion, again, a correct feedback was displayed. Otherwise, an incorrect-response feedback was displayed. The protocol for experiments on human subjects were approved at the University of California by Irvine's Institutional Review Board.

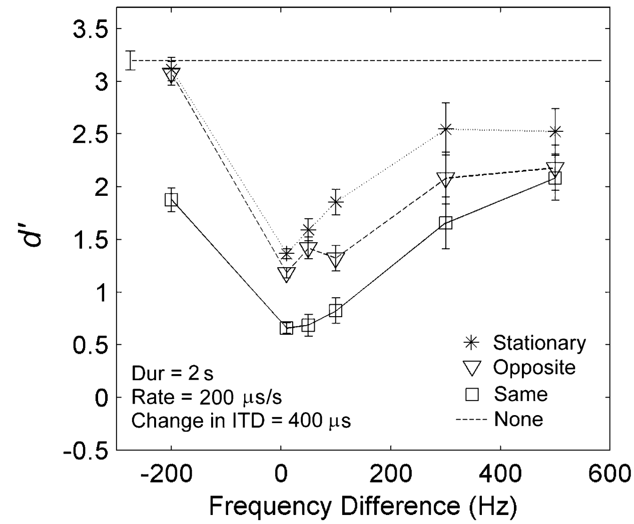

Fig. 1. Index of detectability, d', for detection of motion, averaged across four subjects as a function of the frequency difference between distracter and target tones. The parameter is distracter-movement condition. The dashed horizontal line shows averaged d' in the absence of a distracter tone. The asterisks show d' for a constant-ITD distracter. The inverted triangles show averaged d's for the condition in which the direction of change in distracter ITD is opposite to that of the target tone. The squares show d's for the condition in which the distracter and target tones had the same direction of change in ITD. Error bars represent one standard error of the mean.

3. Results

Fig. 1 shows averaged data from four subjects for stimulus set 1, i.e., distance of 400 μs, velocity of 200 μs/s, and duration of 2 s. The abscissa represents the frequency difference between the target and distracter, Δ. The ordinate represents the index of detectability d' (Green and Swets, 1966) which was estimated by transforming hit and false-alarm rates to *z*-scores and finding their difference (Macmillan and Creelman, 1991). Because some observers on some conditions did not make response errors in the 200 trials (i.e., a d' of infinity), as is customary in such cases we have assumed a small inattention rate of about 2% (Macmillan and Creelman, 1991; Green, 1995; Saberi and Green, 1997) and have imposed a ceiling value of 3.5 on d' estimates. Threshold for detection is customarily assumed to be equal to $d' = 1$.

The dashed horizontal line represents the no-distracter condition. Although this latter condition produces a single d' value, it is plotted as a dashed line across the range of frequency differences to facilitate visual comparison. The error bars represent one standard error. The asterisks, squares, and inverted triangles represent data for distracter conditions 2–4 respectively (see Section 2). Several trends are evident in these data. First, detection is lower for all distracter conditions compared to the no-distracter case. Second, as the distracter frequency approaches the target frequency, detection of the motion of the target sound declines, with poorest performance usually observed for the smallest frequency difference of 10 (i.e., within a critical band). As the frequency difference between the target and the distracter increases, so does the detection index, however, even at a frequency difference of 500 Hz (a separation of about four critical bands), performance is still not as good at that for the no-distracter condition. Motion detection also improves when the distracter frequency is lowered to values below the target frequency, although it appears that at least for two conditions (stationary distracter and opposite-direction of movement), detection improves to d' values nearly as high as that for the no-distracter condition.

An important observation concerns the relative performance for the four stimulus conditions. The no-distracter condition produces the highest d', followed by the stationary-distracter condition, followed by the condition in which the distracter is moving in the direction opposite to that of the target tone, and finally, followed by the condition in which the distracter is moving in the same direction as that of the target tone. This ordering of performances is maintained at all frequency differences between the target and distracter. A repeatedmeasures analysis of variance (ANOVA) test on the data of Fig. 1 showed a significant effect of distracter condition, $F_{2,6} = 10.15$, $P < 0.05$, frequency difference between distracter and target tones, $F_{5,15} = 9.13$, $P < 0.01$, and no interaction effect between these two factors, $F_{10,30} = 1.82$, ns.

Figs. 2–5 show d' as a function of the frequency difference between the target and distracter tones for the remaining four distance/velocity stimulus sets, plotted in the same manner as Fig. 1. The pattern of performance as a function of frequency difference and distracter condition for Figs. 2 and 3 are similar to that of Fig. 1 in that the same-direction condition generally produces the lowest d's and the stationary distracter condition produces the best performance for the three conditions that include distracters. It is interesting to note, however, that the lowest detection index for the 'same-direction' condition generally occurs at a frequency difference of 100 Hz, not 10 Hz as was the case for Fig. 1. Also note that the three conditions that produced the pattern of performances described thus far include the three 'longest'-distance conditions, i.e., 400 and 800 µs. The points at the right side of Fig. 2 have been excluded to avoid phase wrapping (see Section 2). An ANOVA test on the data of Fig. 2 showed a significant effect of distracter condition, $F_{2,6} = 8.94$, $P < 0.05$, a significant effect of frequency difference between distracter and target tones, $F_{3,9} = 14.98$, $P < 0.01$, and no interaction effect between these two factors, $F_{6,18} = 1.55$ ns. An ANOVA test on the data of Fig. 3 showed a significant effect of distracter condition, $F_{2,6} = 26.69$, $P < 0.01$, a significant effect of frequency difference between distracter and target tones, $F_{5,15} = 12.15$, $P < 0.01$, and significant interaction between these two factors, $F_{10,30} = 2.55$, $P < 0.05$.

The data of the two 'shorter'-distance conditions (Figs. 4 and 5) show a different pattern of performance as a function of distracter condition. Fig. 4 shows that although the performance for the stationary distracter condition is still the best of the three distracter conditions, there appears to be little difference between the same- and opposite-direction conditions at the larger frequency differences, and an advantage of the same- over opposite-direction condition when the frequency difference is small. An inspection of the data of individual subjects for this figure, however, shows that some

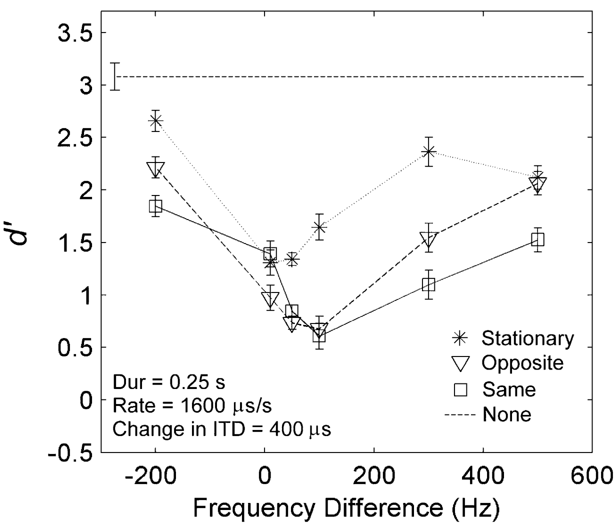

Fig. 3. Same as Fig. 1, except that the stimulus duration is 0.25 s, the extent of change in ITD is 400 µs, and the rate of change in ITD is 1600 µs/s.

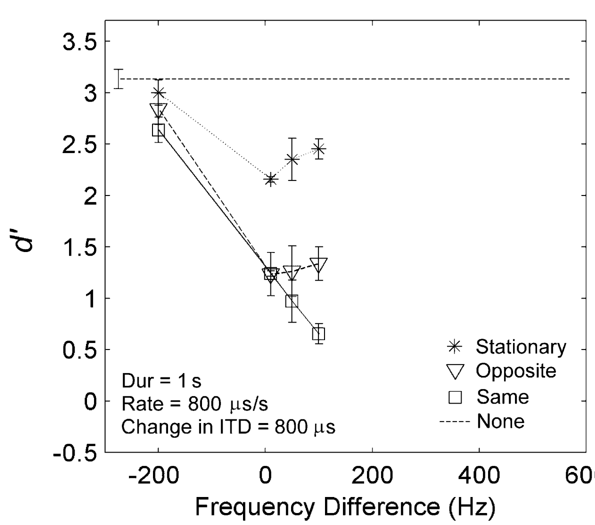

Fig. 2. Same as Fig. 1, except that the stimulus duration is 1 s, the extent of change in ITD is 800 µs, and the rate of change in ITD is 800 µs/s.

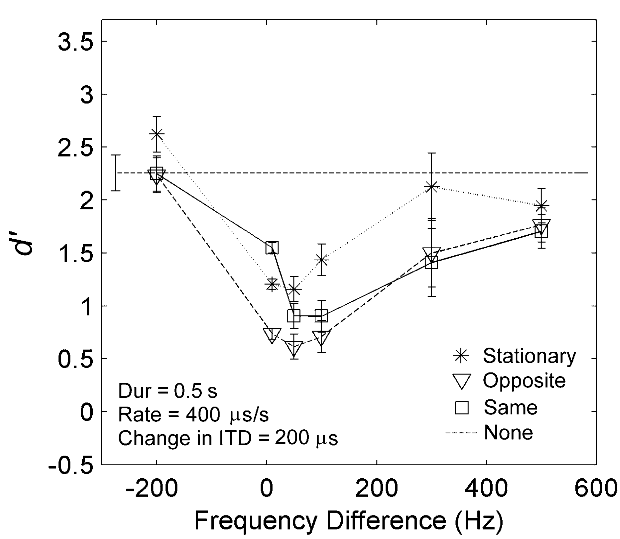

Fig. 4. Same as Fig. 1, except that the stimulus duration is 0.5 s, the extent of change in ITD is 200 µs, and the rate of change in ITD is 400 µs/s.

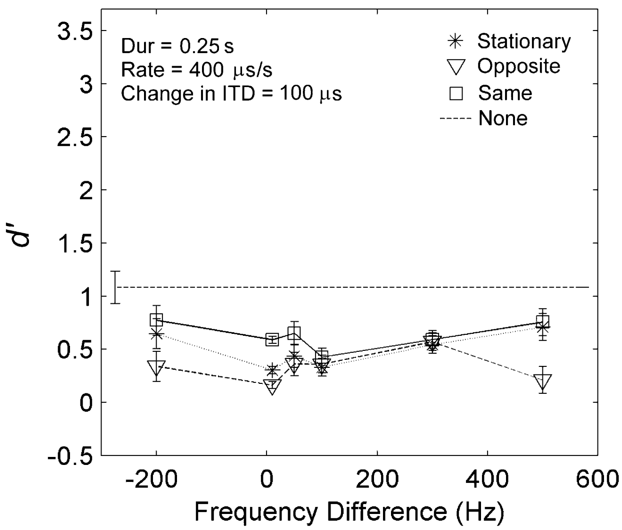

Fig. 5. Same as Fig. 1, except that the stimulus duration is 0.25 s, the extent of change in ITD is 100 µs, and the rate of change in ITD is 400 µs/s.

subjects perform better in the same-direction and some in the opposite-direction condition, and thus, the difference between the two conditions is not significant ($F_{2,6} = 4.38$, ns). However, there were significant effects of frequency difference between target and distracter ($F_{5,15} = 5.07$, $P < 0.01$) and a significant interaction effect ($F_{10,30} = 2.22$, $P < 0.05$). Fig. 5 shows that when the distance traveled is relatively small (100 µs), there is little apparent effect of frequency difference and of distracter condition, possibly because the d' values are quite small and below the conventionally accepted value of threshold at $d' = 1$. An ANOVA test on the data of Fig. 5 showed no significant effect of distracter condition, $F_{2,6} = 4.79$ ns, no significant effect of frequency difference between distracter and target tones, $F_{5,15} = 0.85$ ns, and no interaction between these two factors, $F_{10,30} = 1.47$ ns.

4. Discussion

The difference in detection between distracter conditions may be better understood by considering the percepts generated under each condition. Figs. 1–3, which represent the longer 'distances' traveled (400 and 800 µs), show that the same-direction condition generally produces the poorest detection performance, especially when the frequency difference between target and distracter is within a critical band. The critical band for frequencies below 500 Hz is approximately 100 Hz, and 15–20% of the frequency above 500 Hz (Zwicker, 1961; Scharf, 1970; Moore, 1997). We should, however, note that estimates of binaural critical band are larger than monaural estimates (Kollmeier and Holube, 1992). Consider, for example, the square symbol at a frequency difference of 10 Hz in these figures. When the target ITD and distracter are dynamic and have the same-direction of change, a single image with a coherent percept of smooth motion is perceived. When the target is stationary (no-motion trials) and the distracter is dynamic, then an image whose extensity varies within a trial is reported. By extensity, we are referring to a perceptual spatial extension or increase in perceived size of an auditory image or object. This variation in extensity results from one static ITD and one changing ITD in tones that are very close in frequency. Thus, observers may base their judgments on coherent motion versus changing extensity. The averaged performance for this case is not very high, partially because on both 'motion' and 'no motion' trials, a 10-Hz amplitude fluctuation due to beating makes distinctions between the two types of trials somewhat more difficult.

When the frequency difference is larger but still within a critical band (i.e., 50 or 100 Hz), performance remains low, although common loudness beats are replaced with a percept of sound quality best described as 'roughness'. At still larger frequency differences (e.g., 500 Hz; square symbol to the right of Fig. 1), where the target and distracter frequencies are more than one critical band apart, detection improves, but remains lower than the no-distracter condition. Thus, interference from a distracter tone, either when it is spatially stationary or dynamic, degrades the detection of motion of a target image even when target and distracter frequencies are separated by several critical bands. This may possibly be due to the fact that the skirts of auditory filters at these frequencies extend beyond a critical band, partly due to the wider critical bands for binaural compared to monaural conditions (Kollmeier and Holube, 1992), and partially due to perceptual grouping as described below.

For the 'opposite'-direction condition, where the change in target ITD is opposite in direction to that of the distracter, the percepts are quite different than those described above for the same-direction condition. When the trial is a 'no motion' trial (i.e., target tone stationary), a moderate change in extensity is perceived. However, on 'motion' trials, the change in extensity is considerably more pronounced. Performance on the opposite-direction condition could be superior to the same-direction case because in the opposite-direction condition, the leftward motion of the distracter coupled with the rightward motion of the target produced greater change in extensity. The difference in performance for the same- and opposite-direction conditions, however, although significant, is not as large as may be expected from the very different stimuli and their corresponding percepts. This may be due to the fact that in both conditions, subjects are making comparison of changes in extensity, with the opposite direction providing a moderately greater change. In the opposite-direction condition, motion and no-motion trials both produce percepts of a change in extensity, with a greater change corresponding to motion trials. Whereas for the same-direction condition, the comparison is between a fixed-extensity image ('motion' trials) versus a image with varying extensity ('no-motion' trials).

For the opposite-direction condition and the larger frequency differences, two distinct images are perceived that clearly appear to approach and cross by each other. When the frequency difference is small, a single image is perceived whose extensity modulates, becoming smaller, and then broadening again. For the 'stationary'-distracter condition, the task is generally not as difficult as the task for the same- and opposite-direction conditions. The observer on 'no motion' trials perceives a completely fixed-position image, and on 'motion' trials, perceives either two images when the frequency difference is large, or a single image with reduced velocity of motion, or fluctuating extensity.

An interesting finding is the asymmetry in performance for distracter frequencies above and below the target frequency. Excluding Fig. 5, where performance for all distracter conditions are below threshold, every averaged d' (nine out of nine) at a frequency difference of −200 Hz is higher than the corresponding d' at +300 Hz, in spite of the fact that the former is closer in frequency to the target than the latter and thus expected to produce greater interference. Such asymmetry may be related to the logarithmic spacing of auditory filter resonant frequencies (Moore, 1997). In logarithmic coordinates, the higher-frequency distracter is effectively closer than the lower-frequency distracter to the target tone. It should be also noted that in terms of critical bands, the distracter at −200 Hz and +300 Hz are nearly equally far (about two critical bands) from the 500-Hz target. Although evaluating frequency separation in terms of critical bands reduces the asymmetry, some asymmetric effects remain. This is particularly clear if one compares −200 and +500 Hz. For these, eight or nine comparisons yield better performance with the low-frequency distracter than with the high-frequency distracter, even though the low-frequency distracter is closer both in terms of critical bands and in terms of logarithmic frequency.

When the distance traveled is small, i.e., 100 or 200 μs, then a different pattern of performance is observed (Figs. 4 and 5). For these cases, the 'same'-direction condition does not generally produce the lowest d's. Performance is either nearly equal to or slightly better than that for the 'opposite'-direction condition. As noted in Section 3, an inspection of the individual-subject data shows that this difference between the 'same' and 'opposite' conditions is not significant since some observers show the reverse pattern. The small distance traveled is apparently not sufficient to produce a difference in performance based on a change in extensity ('opposite' condition) compared to a change in location ('same' condition). The decline of performance for the conditions shown in Fig. 5 may be attributed to the short distance traveled, not to the short stimulus duration. Note from this figure, that the distance traveled is 100 μs and the stimulus duration is 0.25 s. Compare these data to the higher d's shown in Fig. 3 which are also based on a duration of 0.25 s, but a distance of 400 μs.

The across-frequency channel effects observed here are similar to effects reported for monaural processing of complex multiband waveforms. Among these effects, in particular, is MDI in which the detection of modulation information in one frequency band is made more difficult by the presence of a comodulated band at a remote frequency (Yost and Sheft, 1989, 1990; Yost et al., 1989; Mendoza et al., 1995). This interfering effect is observed even if the flanking non-target waveform is several critical bands removed from the center frequency of the target band. MDI exhibits parallels to interference with concurrent motion detection in that the dynamic co-variation of one stimulus interferes with the detection of another. For instance, concurrent motion in the same direction generally produces lower thresholds than other conditions, including concurrent changes in the opposite direction to the target stimulus. In addition, both MDI and motion interference are maintained when target and interfering stimuli are spectrally separated.

Studies of MDI, as well as several other cross-frequency effects such as profile analysis (Green, 1988), comodulation masking release (Hall et al., 1984; Grose and Hall, 1992), and comodulation detection differences (McFadden, 1987; Cohen and Schubert, 1987; Fantini and Moore, 1994) have shown that stimulus features that contribute to auditory object formation can result in interaction among remote frequency channels. These features include the shape of the spectrum, onset synchrony, or dynamic co-variation in either intensity or spectrum of sounds, all of which are important cues to perceptual grouping (Bregman, 1990). Studies of cross-spectral effects have often implicated the involvement of different cues for within versus between frequency-channel interaction. That is, depending on whether the target and distracter frequencies are within or outside a critical band, different cues are available to an observer even though interference (or signal enhancement) may appear to be affected similarly in both cases (Green et al., 1995; Bacon and Konrad, 1993). In our observations, we have found that although interference with motion detection occurs at all frequency differences tested, the perceptual descriptions by subjects are quite different for within- compared to across-channel conditions. In the former case, subject report that interference is stronger because of the presence of perceptual beats and loudness fluctuations which are not perceived when the target and interfering tones are more than a critical band apart. At these larger frequency differences, beat fluctuations are not heard. However, other types of cues to perceptual grouping, such as synchronous onsets, coherent movement, as well as similar movement velocities (rates of change in ITD) are the dominant features that interfere with motion detection. These latter cues, of course, also contribute to motion detection interference when the target and distracter are within a critical band.

Other than the monaural cross-channel effects noted above, several studies have investigated across-frequency effects with lateralization of binaural waveforms. Saberi (1996) has shown that the lateralization of one band of noise may be facilitated when a second spectrally remote band with the same interaural delay is comodulated with the target band, compared to independependantly modulated bands. Further evidence of across-frequency-channel effects in binaural hearing, for stationary waveforms, has been provided in studies of masking level differences, judgments of laterality, and the discrimination of ITDs or ILDs (Trahiotis and Bernstein, 1990; Bernstein and Trahiotis, 1992, 1995; Heller and Trahiotis, 1994; Dye et al., 1996; Dye, 1997; Zurek, 1985). The current findings, therefore, we believe fit well within the context of a broader range of auditory phenomena concerned with binaural interference, dynamic co-variation of spectrally remote events, and stimulus features that contribute to perceptual grouping and auditory object formation.

Acknowledgements

Work supported by NIH. We thank two anonymous reviewers for their helpful comments.

References

Arbogast, T.L., Kidd, G., 2000. Evidence for spatial tuning in informational masking using the probe-signal method. J. Acoust. Soc. Am. 108, 1803–1810.

Bacon, S.D., Konrad, D.L., 1993. Modulation detection interference under conditions favoring within- or across-channel processing. J. Acoust. Soc. Am. 93, 1012–1022.

Bernstein, L.R., Trahiotis, C., 1992. Spectral interference in a binaural detection task: Effects of temporal fringe and masking bandwidth. J. Acoust. Soc. Am. 94, 735–742.

Bernstein, L.R., Trahiotis, C., 1995. Binaural interference effects measured with masking-level difference and with ITD- and IID-discrimination paradigms. J. Acoust. Soc. Am. 98, 155–163.

Blauert, J., 1972. On the lag of lateralization caused by interaural time and intensity differences. Audiology 11, 265–270.

Bolia, R.S., D'Angelo, W.R., McKinley, R.L., 1999. Aurally aided visual search in three-dimensional space. Hum. Factors 41, 664–669.

Bregman, A.S., 1990. Auditory Scene Analysis. MIT Press, Cambridge.

Chandler, D.W., Grantham, D.W., 1992. Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. J. Acoust. Soc. Am. 91, 1624–1636.

Cohen, M.F., Schubert, E.D., 1987. The effects of cross-spectrum correlation on the detectability of a noise band. J. Acoust. Soc. Am. 81, 721–723.

Fantini, D.A., Moore, B.C.J., 1994. Profile analysis and comodulation detection differences using narrow bands of noise and their relation to comodulation masking release. J. Acoust. Soc. Am. 95, 2180–2191.

Dye, R.H., 1997. The relative contributions of targets and distractors in judgments of laterality based on interaural differences of level. In: Gilkey, R.H., Anderson, T.R. (Eds.), Binaural and Spatial Hearing in Real and Virtual Environments. Lawrence Erlbaum, New Jersey Associates.

Dye, R.H., Stellmack, M.A., Grange, A.N., Yost, W.A., 1996. The effect of distractor frequency on judgments of target laterality based on interaural delays. J. Acoust. Soc. Am. 99, 1096–1107.

Grantham, D.W., 1984. Discrimination and dynamic interaural intensity differences. J. Acoust. Soc. Am. 76, 71–76.

Grantham, D.W., 1986. Detection and discrimination of simulated motion of auditory targets in the horizontal plane. J. Acoust. Soc. Am. 79, 1939–1949.

Grantham, D.W., Wightman, F.L., 1978. Detectability of varying interaural temporal differences. J. Acoust. Soc. Am. 63, 511–523.

Green, D.M., 1988. Profile Analysis: Auditory Intensity Discrimination. Oxford University Press, New York.

Green, D.M., 1995. Maximum-likelihood procedures and the inattentive observer. J. Acoust. Soc. Am. 97, 3749–3760.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Wiley, New York.

Green, D.M., Dai, H., Saberi, K., 1995. Shape discrimination of spectra with various bandwidths. In: Manley, G.A., Klump, G.M., Koppl, C., Fastl, H., Oeckinghaus, H. (Eds.), Advances in Hearing Research. World Scientific Publishers, Singapore, pp. 501–511.

Grose, J.H., Hall, J.W., III, 1992. Comodulation masking release for speech stimuli. J. Acoust. Soc. Am. 91, 1042–1050.

Hall, J.W., III, Haggard, M.P., Fernandez, M.A., 1984. Detection in noise by spectro-temporal pattern analysis. J. Acoust. Soc. Am. 76, 50–56.

Heller, L.M., Trahiotis, C., 1994. Interference in detection of interaural delay in a SAM tone produced by a second, spectrally remote SAM tone. J. Acoust. Soc. Am. 97, 1808–1816.

Kidd, G., Mason, C.R., Rohtla, T.L., Deliwala, P.S., 1998. Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. J. Acoust. Soc. Am. 104, 422–431.

Kollmeier, B., Holube, I., 1992. Auditory filter bandwidths in binaural and monaural listening conditions. J. Acoust. Soc. Am. 92, 1889–1901.

Macmillan, N.A., Creelman, C.D., 1991. Detection Theory: A User's Guide. Cambridge University Press, New York.

McFadden, D., 1987. Comodulation detection differences using noise-band signals. J. Acoust. Soc. Am. 81, 1519–1527.

Mendoza, L., Hall, J.W., Grose, J.H., 1995. Modulation detection interference using random and sinusoidal amplitude-modulation. J. Acoust. Soc. Am. 97, 2487–2492.

Moore, B.C.J., 1997. An Introduction To The Psychology Of Hearing. Academic Press, New York.

Perrott, D.R., Tucker, J., 1988. Minimum audible movement angle as a function of signal frequency and the velocity of the source. J. Acoust. Soc. Am. 83, 1522–1527.

Perrott, D.R., 1984. Concurrent minimum audible angle: A re-examination of the concept of auditory spatial acuity. J. Acoust. Soc. Am. 75, 1201–1206.

Perrott, D.R., Musicant, A.D., 1977a. Rotating tones and binaural beats. J. Acoust. Soc. Am. 61, 1288–1292.

Perrott, D.R., Musicant, A.D., 1977b. Minimum auditory movement angle: Binaural localization of moving sound sources. J. Acoust. Soc. Am. 62, 1463–1466.

Perrott, D.R., Nelson, M.A., 1969. Limits for the detection of binaural beats. J. Acoust. Soc. Am. 46, 1477–1481.

Saberi, K., 1995. Lateralization of comodulated complex waveforms. J. Acoust. Soc. Am. 98, 3146–3156.

Saberi, K., 1996. An auditory illusion predicted from a weighted cross-correlation model of binaural interaction. Psych. Rev. 103, 137–142.

Saberi, K., Green, D.M., 1997. Evaluation of maximum-likelihood estimators for non-intensive auditory psychophysics. Percept. Psychophys. 59, 867–876.

Saberi, K., Hafter, E.R., 1995. A common neural code for frequency- and amplitude-modulated sounds. Nature 374, 537–539.

Saberi, K., Hafter, E.R., 1997. Experiments on auditory motion discrimination. In: Gilkey, R.H., Anderson, T.R. (Eds.), Binaural and Spatial Hearing in Real and Virtual Environments. Earlbaum, New Jersey, pp. 315–327.

Saberi, K., Perrott, D.R., 1990. Minimum audible movement angles as a function of sound source trajectory. J. Acoust. Soc. Am. 88, 2639–2644.

Sayers, B.M., Cherry, E.C., 1957. Mechanism of binaural fusion in hearing of speech. J. Acoust. Soc. Am. 29, 973–987.

Scharf, B., 1970. Critical bands. In: Tobias, J.V. (Ed.), Foundations of Modern Auditory Theory, Vol. I. Academic, New York, pp. 157–202.

Strybel, T.Z., Manligas, C., Perrott, D.R., 1989. Auditory apparent motion under binaural and monaural listening conditions. Percept. Psychophys. 45, 371–377.

Trahiotis, C., Bernstein, L.R., 1990. Detectability of interaural delays over select spectral regions: Effects of flanking noise. J. Acoust. Soc. Am. 87, 810–813.

Waugh, W., Strybel, T.Z., Perrott, D.R., 1979. Perception of moving sounds: Velocity discrimination. J. Audit. Res. 19, 103–110.

Yost, W.A., 1997. The cocktail party effect: 40 years later. In: Gilkey, R., Anderson, T. (Eds.), Binaural and Spatial Hearing in Real and Virtual Environments. Erlbaum Press, New Jersey.

Yost, W.A., Sheft, S.A., 1989. Across-critical-band processing of amplitude-modulated tones. J. Acoust. Soc. Am. 85, 848–857.

Yost, W.A., Sheft, S.A., 1990. A comparison among three measures of cross-spectral processing of amplitude modulation with tonal signals. J. Acoust. Soc. Am. 87, 897–900.

Yost, W.A., Sheft, S.A., Opie, J., 1989. Modulation interference in detection and discrimination of amplitude modulation. J. Acoust. Soc. Am. 86, 2138–2147.

Yost, W.A., Dye, R.H., Sheft, S.A., 1996. Simulated 'cocktail party' with up to three sound sources. Percept. Psychophys. 58, 1026–1036.

Zurek, P.M., 1985. Spectral dominance in sensitivity to interaural delay for broadband stimuli. J. Acoust. Soc. Am. 78 (Suppl.), S18.

Zwicker, E., 1961. Subdivision of the audible frequency range into critical bands (Frequenzgruppen). J. Acoust. Soc. Am. 33, 248.