Separation of concurrent broadband sound sources by human listeners

Virginia Best
Department of Physiology, University of Sydney, NSW 2006, Australia

André van Schaik
Department of Physiology, University of Sydney, NSW 2006, Australia and School of Electrical and Information Engineering, University of Sydney, NSW 2006, Australia

Simon Carlile[a)]
Department of Physiology, University of Sydney, NSW 2006, Australia and Institute for Biomedical Research, University of Sydney, NSW 2006, Australia

(Received 21 August 2002; revised 14 October 2003; accepted 15 October 2003)

The effect of spatial separation on the ability of human listeners to resolve a pair of concurrent broadband sounds was examined. Stimuli were presented in a virtual auditory environment using individualized outer ear filter functions. Subjects were presented with two simultaneous noise bursts that were either spatially coincident or separated (horizontally or vertically), and responded as to whether they perceived one or two source locations. Testing was carried out at five reference locations on the audiovisual horizon (0°, 22.5°, 45°, 67.5°, and 90° azimuth). Results from experiment 1 showed that at more lateral locations, a larger horizontal separation was required for the perception of two sounds. The reverse was true for vertical separation. Furthermore, it was observed that subjects were unable to separate stimulus pairs if they delivered the same interaural differences in time (ITD) and level (ILD). These findings suggested that the auditory system exploited differences in one or both of the binaural cues to resolve the sources, and could not use monaural spectral cues effectively for the task. In experiments 2 and 3, separation of concurrent noise sources was examined upon removal of low-frequency content (and ITDs), onset/offset ITDs, both of these in conjunction, and all ITD information. While onset and offset ITDs did not appear to play a major role, differences in ongoing ITDs were robust cues for separation under these conditions, including those in the envelopes of high-frequency channels. [DOI: 10.1121/1.1632484]

PACS numbers: 43.66.Qp, 43.66.Pn [LRB] Pages: 324–336

I. INTRODUCTION

There are several acoustic cues that define the location of a sound source relative to a human listener [for review see Blauert (1983), Middlebrooks and Green (1991), and Carlile (1996)]. For a sound source located away from the midline, the difference in path length to the two ears results in a difference in the arrival time of the sounds at each ear (interaural time difference, ITD). Furthermore, a sound originating from one side of the head will be more intense in the near ear as compared to the far ear, resulting in an overall interaural level difference (ILD). As the ears are placed more or less symmetrically on the head, the binaural cues are ambiguous; a given interaural difference corresponds to a set of locations on a rough cone centered on the interaural axis. To resolve these ambiguities, the auditory system relies on the three-dimensional asymmetry of the auditory periphery. Location-dependent spectral filtering of sounds by the pinnae, head, and torso has been shown to provide monaural spectral cues to sound source location (Batteau, 1967; Gardner and Gardner, 1973). In addition, the ILD is also frequency dependent and results from differential filtering effects of the head and auditory periphery.

Importantly, sound arriving at the ears rarely comes from a single source. It is of interest to a listener to correctly attribute acoustic information to different sources to aid in the identification and localization of objects in the environment. There exists a large body of research detailing the stimulus factors that assist in auditory grouping and auditory segregation [so-called "auditory scene analysis:" see Bregman (1990)]. Auditory grouping occurs on the basis of frequency similarity and spectral continuity (Bregman and Campbell, 1971; Bregman, 1990), and segregation is aided when auditory objects differ in their spectral content or temporal structure (Perrott, 1984; Stellmack, 1994). It has also been known for several decades that the location of competing sound sources has an impact on their separability. The spatial separation of sources improves the detection of tones presented concurrently with noise [see Durlach and Colburn (1978) for review] as well as aiding complex phenomena such as speech intelligibility in noisy environments (e.g., Hirsh, 1950; Dirks and Wilson, 1969). Despite extensive work in this area, only a few studies have attempted to measure systematically the spatial resolution of the auditory sys-

[a)]Author to whom correspondence should be addressed. Electronic mail: simonc@physiol.usyd.edu.

 0001-4966/2004/115(1)/324/13/$20.00

tem for simultaneous stimuli, and to define the role of the different localization cues under these circumstances.

Perrott (1984) defined the concurrent minimum angle (CMAA) as the threshold separation angle required to distinguish two concurrent sounds. He measured the CMAA for sources distributed in a horizontal plane and reported it to be larger than the angle required to determine the direction of displacement of two *sequential* sounds (minimum audible angle, MAA). Perrott presented pairs of tones of different frequency, and asked subjects to judge the relative location of the pair by indicating whether the higher tone was to the left or right of the lower tone. Using a criterion of 75% correct, he measured CMAAs and noted a significant effect of azimuth, with CMAAs of 4°–10° at the front increasing to 30°–45° at a lateral displacement of 67°. Divenyi and Oliver (1989) extended this work to examine more complex sounds including amplitude- and frequency-modulated tones, and found a similar threshold increase for stimuli located at 80° as compared to 0° azimuth. In these studies, the concurrent stimuli differed in their frequency content and so binaural effects could not be isolated from pitch effects. In fact, Perrott (1984) reported that CMAAs increased if the frequency difference between the two sources was reduced, indicating that the two parameters are confounded. Divenyi and Oliver (1989) also noted that, for their more broadband stimuli, spectral overlap was detrimental to resolution.

These studies, and indeed most studies of multiple source perception, have only examined separation in the horizontal plane, where binaural cues are dominant and localization is relatively robust for most stimulus types. In the present study, one objective was to examine sound sources coming from many directions around the listener, displaced both horizontally and vertically. Importantly, broadband noise stimuli were chosen because they are well localized in all dimensions (Carlile *et al.*, 1997) as opposed to narrowband stimuli which are poorly localized in elevation and in the front–back dimension (Butler, 1986; Middlebrooks, 1992). Effectively, in previous experiments, the bandwidth required for accurate localization had been sacrificed for the sake of gaining an adequate frequency separation of two sources. The study reported here aimed to demonstrate how listeners perceive concurrent sounds given *only* spatial separation and *no* differences in source content, i.e., no pitch cues. It is not entirely clear what acoustical cues (for localization and/or separation) are available in the sum of two such stimuli. It is likely that interaural temporal cues related to the two independent sources are present (see Sec. VI C), but the spectrum received at each ear will be a sum of the spectra expected from the pinna filtering of each of the sources.

Most previous studies in this area have required some form of identification or localization of target sounds in the presence of distracting sounds. However, in the present study the competing sounds were indistinct, and hence subjects were asked to simply indicate whether they perceived sound arriving from a single location or from two distinct locations [this is how we define "separation;" it has also been called "detection of spatial separateness" (Noble *et al.*, 1997)]. However, our "simple" task is quite difficult for two reasons. First, the two sources presented are identical in their long-term spectrum and temporal characteristics and so cannot be segregated on the basis of their content or identity. Second, as these sounds are broadband, they both produce global activation along the basilar membrane and thus a complete sharing of receptors must take place. Separation cannot commence at the periphery as is the case with differing tones, which are processed in distinct frequency channels, but must rely on more central processing of spatial cues.

The first experiment aimed to examine the ability of listeners to resolve the simultaneous pair and compared this ability in different spatial regions around the listener. The following experiments examined in some detail the contribution of different localization cues to this perception, with a focus on the possible role of ITD. The ITD cue has two main components: a transient onset (and offset) ITD and an ongoing ITD. For pure tones the ongoing ITD refers to cycle-by-cycle phase differences, and is only useful at low frequencies (below about 1.5 kHz) where periodicity information is preserved by phase-locking neurons in the auditory pathway (see Moore, 1997). For complex waveforms, however, it is known that interaural delays in the slowly varying envelope of high-frequency channels (above 1.5 kHz) can also be processed by the auditory system (Henning, 1974, 1980; McFadden and Pasanen, 1976). Lateralization studies have shown that ongoing ITD is a more potent indicator of lateral position than is onset/offset ITD, with the latter only having a strong influence when the signal is brief [shorter than about 10 ms (Tobias and Schubert, 1959)] or the ongoing ITD is ambiguous (Kunov and Abel, 1981; Buell *et al.*, 1991). In the following experiments, separation of concurrent noise sources was examined upon removal of low-frequency content (experiment 2a), onset/offset ITDs (experiment 2b), both of these in conjunction (experiment 3a), and all ITD information (experiment 3b). These manipulations were made possible by exploiting a virtual auditory space stimulation technique.

II. METHODS

A. Subjects and environment

Four subjects (two males and two females, aged 23–33 years) participated in the experiments. All had normal hearing by audiometric testing and some experience in auditory psychophysical testing. Experiments consisted of a series of short tests, during which subjects were seated in a darkened, sound-attenuating chamber. Stimuli were presented in virtual auditory space (see Sec. II B) over in-ear headphones at approximately 50 dB sensation level. A small amount of training was conducted prior to testing, and this is described in Sec. II D.

B. Stimulus generation

Virtual auditory space (VAS) refers to a spatial auditory percept that is created using headphone presentation. For the presentation of multiple concurrent sounds at finely spaced locations this was an essential tool, as it avoided the need for multiple speakers and an extremely sensitive placement system. In addition, some of the stimulus manipulations used would simply not be achievable using stimuli presented from loudspeakers in ‘‘real’’ space. Individualized VAS is generated by accurately simulating the wave pattern at the eardrum occurring after free-field stimulation with an external sound source (Carlile, 1996). For different sound source locations in space relative to a listener's head, a different pattern of filtering is imposed on the sound by the head and outer ears (head-related transfer function, HRTF). These functions are recorded routinely in our laboratory by the placement of microphones in the ear canal to record the impulse response to a specific broadband stimulus [blocked-ear canal recording (Middlebrooks *et al.*, 1989; Moller *et al.*, 1995]. The recording stimulus was a 1024-bit Golay code pair presented 12 times, with the resultant input averaged to increase the signal-to-noise ratio (Zhou *et al.*, 1992). Subjects were seated in an anechoic chamber, within which a hoop system carrying a loudspeaker could be moved to position the loudspeaker at any location on an imaginary sphere (radius 1 m). The subject's head was positioned in the center of this sphere by fixing a laser from the front (0°,0°) on the nose, and a laser from the left side (−90°,0°) on the ear canal entrance. Subjects were instructed to keep their heads in this position for the duration of the recording process (about 30 min) with the assistance of a chin-rest. To minimize artifacts that may arise as a result of small head movements during the Golay code pair repetitions (Zahorik, 2000), head position was continuously monitored using an electromagnetic head-tracker (Polhemus) and recording stopped immediately when deviations occurred. Subjects were instructed on how to readjust their heads using a small LED display that signaled the direction of deviation relative to the calibrated position. Impulse responses were obtained for 393 positions evenly spaced on the sphere. This procedure has been described in more detail elsewhere (Pralong and Carlile, 1996).

Microphone (Sennheiser KE 4-211-2) and system transfer functions were removed from the impulse response functions by deconvolving the microphone transfer function, and HRTFs were extracted. Location-independent components were removed to leave only the directional transfer functions (DTFs) for each ear (see Middlebrooks *et al.*, 1989). The two DTF filters corresponding to a particular location could then be convolved with any sound stimulus and delivered via in-ear headphones (Etymotic Research ER-2) to the subject to give rise to a virtual externalized stimulus at that location. Analog-to-digital and digital-to-analog conversion for recording and playback occurred at a sampling rate of 80 kHz [Tucker-Davis Technologies (TDT) System II] and stimuli were delivered to an amplifier via a programmable attenuator (TDT: PA4). Stimulus synthesis and delivery as well as data recording and visualization made use of MATLAB 5.3 software (Mathworks Inc.).

Two tests were undertaken to ensure that listeners were valid subjects for VAS experimentation. First, subjects were required to confirm subjectively that their individualized VAS stimuli were realistic and externalized. Second, a standard localization test was performed to ensure that listeners could localize with a suitable level of accuracy in both the free-field and in VAS [see Carlile *et al.* (1997) for details]. Briefly, this test required a subject to stand in darkness in the center of the anechoic chamber and indicate his/her perceived location of a series of 150-ms broadband noise bursts (300 Hz to 16 kHz). These were presented randomly from a subset of 76 recording positions, and responses were indicated using a nose-pointing technique and recorded via the head-tracker. All subjects used in this study were found to be accurate localizers in both the free-field and in VAS, and a summary of their localization data can be found in the Appendix.

Importantly in this study, it was required that stimuli be presented at often closely spaced locations on the virtual sphere of space. However, in the recording process, only a discrete set of 393 locations were obtained distributed on the sphere at approximately 10°–13° intervals. Thus a spatial interpolation procedure, based on spherical thin-plate splines and principal components analysis, was employed to allow locations *within* this grid to be simulated (Carlile and Leung, 2001). It has been confirmed in the laboratory that this procedure produces DTFs that are statistically identical to equivalent measured DTFs (in terms of the structure of the filters), and that virtual sound stimuli based on these DTFs are localized as well as those based on measured ones (Carlile *et al.*, 2000).

To generate stimuli imitating concurrently activated sound source pairs, two independent random noises were filtered with DTFs corresponding to the two desired locations and then added together. For ‘‘zero separation’’ stimuli, the same procedure was followed but the two DTF pairs used were identical.

C. Psychophysical procedures

The experiment took the form of a single-interval forced choice procedure (see Sec. II D for a short discussion of this approach). Subjects were presented with a stimulus pair and were asked to indicate, by pressing one of two buttons, whether they perceived sound to be coming from one or two source locations.

Stimulus locations are described throughout the study using azimuth and elevation with respect to a single-pole coordinate system. For experiments 1, 2a, and 2b (see below) testing took place at five reference locations on the 0° elevation horizontal plane: 0°, 22.5°, 45°, 67.5°, and 90° azimuth (Fig. 1). In each trial, one stimulus in the concurrent pair was presented from a reference location and the other from a test location displaced in either azimuth (horizontal separation) or elevation (vertical separation) or from the same location. Note that binaural cues change with horizontal separation (maximum rate-of-change at the front, minimum rate-of-change at the side) and also with vertical separation using the single-pole coordinate system (maximum rate-of-change at the side and *no* change at the front). For each reference location, 15 test locations were chosen on the basis of preliminary testing. For vertical separation, the testing range was the same for each of the five locations, and spanned the entire available range of elevations: from −45° to 90° (directly overhead). Ranges for horizontal separation varied with location as required to cover a suitable range: ±21° for 0° azimuth; ±32° for 22.5° azimuth; ±42° for 45° azimuth; ±53° for 67.5° azimuth; and ±63° for 90° azimuth.

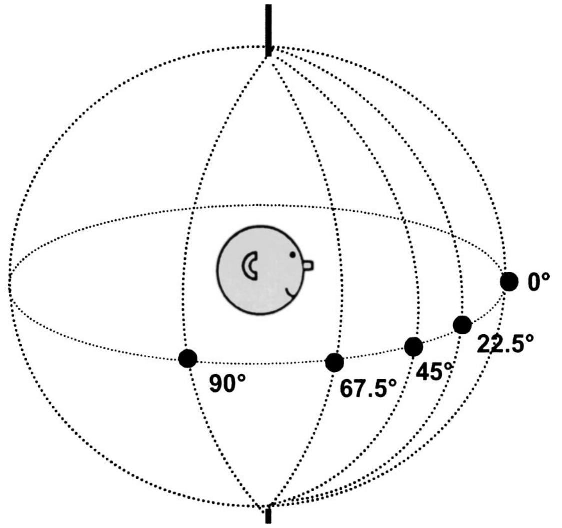

FIG. 1. The single pole coordinate sphere used to define locations in space. The black dots illustrate the five reference locations examined: 0°, 22.5°, 45°, 67.5°, and 90° azimuth, all on the 0° elevation plane at the level of the ears.

All reference location/separation combinations were presented ten times each in a random order, giving 1500 trials in an experimental block (5 reference locations, 2 directions, 15 test separations, 10 repetitions). These were broken down into ten tests of 150 trials, each lasting about 15 min. For experiments 3a and 3b testing only took place at the most frontal reference location (0° azimuth). The same test separations were used for this location as in the other experiments (vertical: 15 values between −45° and 90° elevation; horizontal: 15 values between −21° and 21° azimuth). These shorter experiments each comprised 300 trials (two tests of 150 trials).

Results from each block of experiments were analyzed in turn. Results from a particular reference location were pooled and sorted and psychophysical curves for horizontal and vertical separation were plotted. These illustrate, for each test separation, the percentage of times (out of ten repetitions) that the subject responded that he/she perceived two sources.

D. Response criterion, training, and controls

In establishing the experimental protocol for examining the spatial perception of concurrent stimuli, the widely used two-interval discrimination task was trialed; it required subjects to choose which interval contained two sources and which contained only one. However, this approach was quickly rejected because it was clear that comparisons could be made on the basis of a number of available cues (such as timbral changes), and not necessarily the ones of interest in this study (spatial cues). The single-interval paradigm was thus adopted in order to best measure the effects of spatial separation on the perception of concurrent sources. When presented with a particular stimulus, subjects were required to respond as to whether they perceived one or two source locations. Although this approach did not *ensure* that responses were related to a clear percept of two sources, it did encourage listeners to use cues that were (subjectively) spatial in nature.

One difficulty associated with the single-interval subjective task is that response biases (such as pressing one button more often than the other) cannot be effectively removed as they can in a discrimination task. Only if this bias is consistent within a subject can comparisons across conditions be made with confidence. Another potential problem with the one-interval task is its subjectivity: the experimenter must rely on subjects to adopt a criterion for responding and to maintain this criterion throughout testing. These potential difficulties were dealt with in two ways. First, subjects underwent a small amount of training to stabilize their performance before commencement of data collection. Each subject was run through two tests of the format described in Sec. II C, allowing him/her to become familiar with the stimulus and task and to establish a comfortable criterion. Second, throughout an experimental block, a specific repeated control was included in each test to ensure that a particular subject's bias and criteria did not fluctuate. This control consisted of a small set of 30 trials, which were interleaved with the 150 experimental trials. These were five repeated trials at six horizontal separation values from the 0° reference location set (separations of 0°, 3°, 6°, 9°, 12°, 15°). Responses to these trials provided a measure of response bias and criterion for each test. At the end of an experimental block, the control sets were examined to confirm that they were stable across testing. If a set deviated greatly from the rest, as determined by the experimenters on the basis of informal monitoring, the test from which it came was repeated. Only two tests had to be repeated on this basis (in one subject).

Despite the good within-subject consistency, the fact that response bias and criteria varied across individuals meant that it was difficult to pool responses across the population. Thus for this study the emphasis is not on the absolute value of responses, but rather the pattern of performance of each subject across the various stimuli and test locations.

III. EXPERIMENT 1: BROADBAND STIMULI

A. Stimuli

Experiment 1 was conducted to examine the ability of subjects to separate concurrent broadband sounds. Random noises containing frequencies from 300 Hz to 16 kHz were used (as for localization testing, see Sec. II B). For a given stimulus, two independent noises were filtered with appropriate DTFs and ramped by applying a raised cosine to the first and last 10 ms. The two binaural signals were then added to create concurrent stimuli as depicted in Fig. 2(a).

B. Results

Psychophysical curves for the first experiment are shown in Fig. 3. Data are displayed for individual subjects [Figs. 3(a)–(d)] and mean data are shown in Fig. 3(e). The left-hand panels show data for concurrent sources that were separated in azimuth (horizontally) about the reference location and the five subplots in a panel show data for each of the

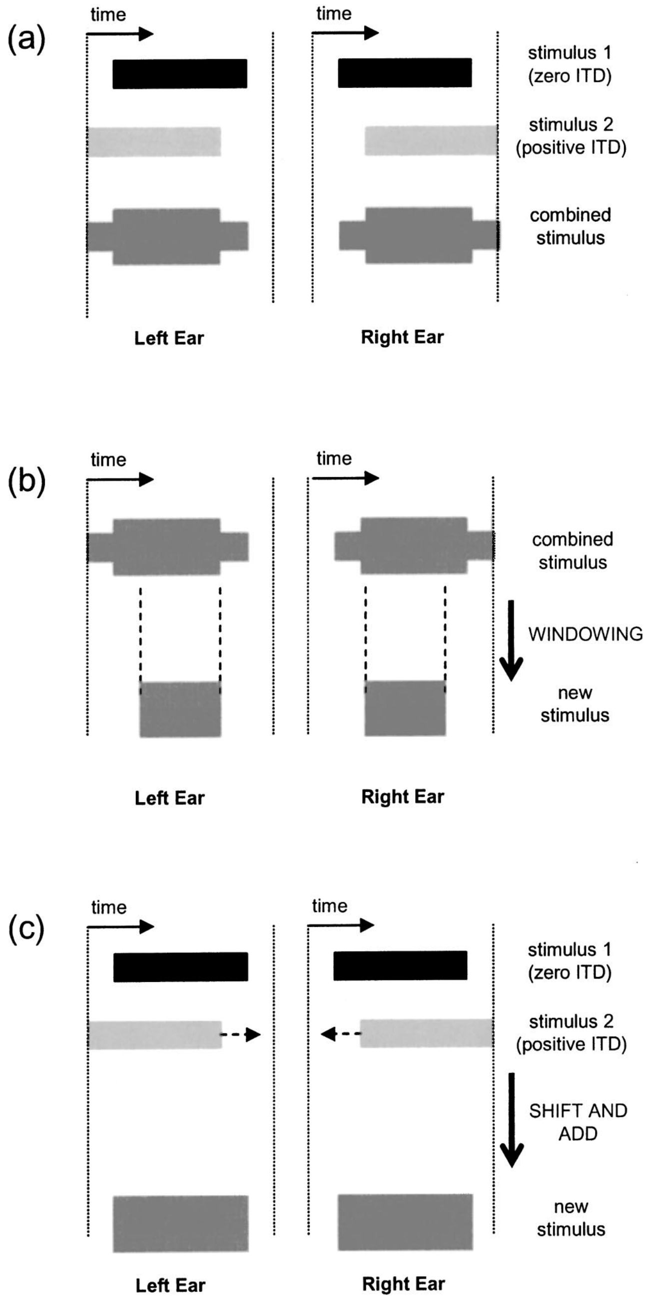

FIG. 2. Temporal features of stimulus manipulations used in experiments 1, 2b, and 3b. The examples given are for a reference stimulus located frontally and a test stimulus at a single lateral location. Left and right ear stimuli are shown in left and right columns. (a) Experiment 1: To create a concurrent pair, the reference stimulus (zero ITD—black bars) is added to a laterally displaced sound (positive ITD—light gray bars) to create a new stimulus. The ITD difference creates two onsets and offsets in each ear, as well as a phase difference in the ongoing portion of the signals. (b) Experiment 2b: The combined stimulus just described is windowed identically in each ear to remove the onset and offset cues. The ongoing phase differences remain in the signal. (c) Experiment 3b: To remove all ITD differences between the two stimuli, the laterally displaced stimulus (light gray bars) is shifted in time in each ear before adding to align it to the zero ITD stimulus. This removes both onset/offset and ongoing ITD differences, but individual spectral characteristics are preserved.

five reference locations. The abscissa shows the separation of the test stimulus with respect to the reference, where negative values indicate leftward separation and positive values indicate rightward separation. Plotted on the ordinate is the percentage of trials (out of 10) where the subject reported that he/she perceived two sources in the presentation.

A feature that is evident for all subjects (and the mean data) is that the psychophysical curves for horizontal separation broaden as the reference position is displaced more laterally, indicating a decrease in spatial sensitivity with increasing laterality of the sources (left panels, Fig. 3). For the most frontal reference location (0°,0°), 100% response rate was reached with much smaller separation angles than for the most lateral location (90°,0°). All subjects showed a similar pattern of results in this experiment, although there were marked differences in overall response levels. For example, subject 4 responded readily to the detection of both sources, as indicated by the relatively narrow troughs [Fig. 3(d)], while subject 1 tended to give more conservative responses [Fig. 3(a)]. This individuality is examined in Sec. VI A. A feature of the data for all subjects is an asymmetry in the psychophysical curves for 67.5° reference location with a drop in response rate as test stimuli were moved towards the back (positive separation). An explanation for this is given in Sec. III C below.

The right-hand panels of Fig. 3 show data plotted in a similar way, but for sources separated vertically about the reference location. Again each of the five reference locations are plotted in separate subplots, and the axes are as described above, except that the separation was in elevation (vertical), with negative values indicating downward separation and positive values indicating upward separation. In these data, the opposite trend can be seen, in that the curves appear to get slightly sharper at the more lateral locations. By far the most striking feature of these data is the frontal midline location. Subjects 1 and 3 show essentially flat psychometric functions indicating that they never (subject 1) or rarely (subject 3) perceived both sources in the pair. Subject 2 gave more positive responses to the stimulus pairs overall but the curve is still relatively flat, indicating that responses were not related to separation. Responses from subject 4 were inconsistent, but he seemed to be able to resolve the sources at large positive separation values. An examination of this subject's HRTFs revealed a marked asymmetry for vertical midline positions, resulting in an elevation-dependent ITD. Specifically, the difference in ITD between elevations of 0° and 90° was measured to be 0.08 ms (compared to the mean value of 0.01 ms for the other three subjects) and this difference is indeed large enough to be a useful cue for separation (see Sec. VI B and Fig. 6).

C. Discussion

The trends seen in the data as the spatial location of testing was varied point strongly to a role for binaural cues in this particular task. The data obtained for horizontal separation on the 0° elevation plane are certainly consistent with this idea. As a result of the position of the two ears on this plane, binaural cues change approximately as a sine function of azimuth, with the rate-of-change being maximal at 0° azimuth and decreasing towards 90° azimuth. As a result, performance in spatial discrimination tasks (Mills, 1972) and absolute localization (Carlile *et al.*, 1997) become poorer in

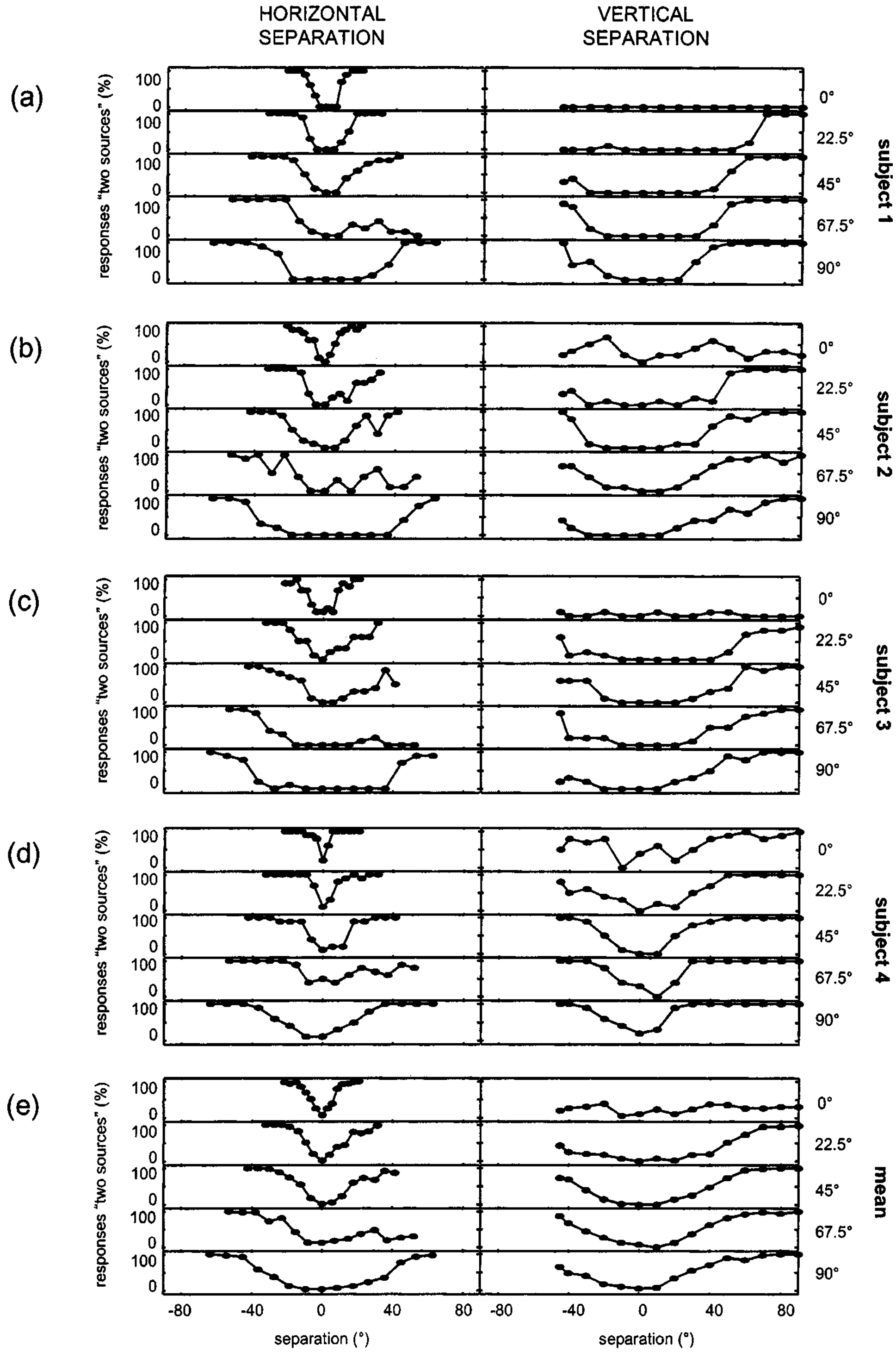

FIG. 3. Psychophysical curves for experiment 1. (a)–(d) show data for subjects 1–4, respectively, and (e) is mean data. The left-hand column shows horizontal separation data and the right-hand column shows vertical separation data. The five subplots in each panel show data for each of the five reference locations (top to bottom: 0°, 22.5°, 45°, 67.5°, and 90° azimuth). For the given test separations (abscissa), the curves show the percentage of trials in which the subject perceived two sources (ordinate).

the horizontal dimension as azimuth increases. In addition, it follows that if separation is dependent on differences in the binaural cues between a pair of stimuli, then the *angle* required to achieve adequate binaural separation would increase with laterality. This is seen in the present data, where the angle of separation required to reach a given response rate increased at the more lateral positions. Consistent with this, Perrott (1984) reported that the concurrent minimum audible angle of tones increased at lateral positions, and Divenyi and Oliver (1989) found a similar effect for amplitude- and frequency-modulated tones.

A special case can be seen at the 67.5° reference position when the test stimuli were separated towards the back. The value of the ITD and ILD for the test stimulus increases as the test location is moved towards 90° azimuth (separations of 9° and 18°) but then *decreases* back towards the reference value when the test location passes 90° (separations of 27° and 36°). When the test location is 112.5° azimuth (separation of 45°), this is the reflection of 67.5° azimuth about the interaural axis and so the binaural cues for the test and reference location are near to identical. Thus these data support the notion that the task is dependent on binaural differences between the two sources.

The pattern of responses seen with vertical separation also suggests that the rate-of-change of the binaural cues is, at least in part, responsible for performance. At the most lateral reference location (90° azimuth), vertical separation of a concurrent stimulus causes a greater change in binaural cues than at more frontal locations, as the displacement is towards the far ear and away from the near ear. Furthermore, at this location, a given separation in the vertical direction changes the binaural cues by approximately the same amount as the same separation in the horizontal direction. Consistent with this, the psychophysical curves at 90° azimuth are similar for horizontal and vertical separation.

Perhaps the clearest demonstration that binaural cues are important for this task is the fact that when stimuli were separated along the vertical midline, subjects generally perceived only one source (or were confused). In other words, given only one set of binaural cues, the auditory system assumes one source. The implication here is that spectral cues are not sufficient under these circumstances to indicate the presence of the distinct sources. This may, at first, seem surprising because spectral cues are responsible for an accurate ability to localize single sources on the vertical midline: average errors are approximately 5° at (0°,0°) for broadband noise (Carlile *et al.*, 1997; Butler *et al.*, 1990). In addition, Perrott and Saberi (1990) reported that subjects could detect a change in location between sequentially presented click trains when they were separated by 3.5° along the vertical midline. Clearly, however, when a pair of broadband sounds is presented concurrently, as in the current experiment, the individual spectral cues become less useful as they sum at each ear. This may not be the case for more band-limited sounds, where there is less spectral overlap.

As discussed in the Introduction, it is known that ITDs are useful for separating concurrent sources under many circumstances. When two sounds are presented concurrently from different locations, there is a different arrival time at each ear for each source. This results in two different ITDs, and it is feasible that these may be recovered by the auditory system and used to separate the two sounds (see Sec. VI C). It is less clear, on the other hand, how useful relative ILD could be for indicating the presence of distinct sources, as signal energy would be added at each ear, producing a *single* new ILD in any frequency channel. As ITD was the most likely candidate for the trends seen in the data from experiment 1, experiments 2 and 3 were undertaken to confirm and characterize its role.

IV. EXPERIMENT 2: EFFECT OF REMOVING SINGLE COMPONENTS OF THE ITD

A. Stimuli

Experiment 2 was conducted in the same way as experiment 1, but with two stimulus manipulations to remove a component of the ITD cue. In experiment 2a, stimuli were high-pass filtered in order to remove low-frequency ongoing timing information. Filtering was performed in the Fourier domain by zeroing the magnitude of all components below 2 kHz and taking the inverse transform to return to a time domain waveform. In experiment 2b, the aim was to remove onset and offset ITD information, and this was achieved by adding the two DTF-filtered noises together and then windowing the resultant left and right ear stimuli such that they ramped on and off at exactly the same points in time. Specifically, 2 ms was trimmed from each end of the stimuli, which was enough to ensure the removal of the two onset and offset cues that were previously available [Fig. 2(b)]. Because 4 ms of signal was lost in the windowing process, the original sounds were generated with duration 154 ms to give rise to a final 150-ms signal. Technically, this manipulation sets the onset and offset ITDs to zero, as opposed to removing them. Nonetheless, it provides the opportunity to observe the effects of removing the location-dependent variation in this cue.

B. Results

Figure 4 shows data for high-passed stimuli (solid lines) and stimuli with onsets and offsets removed (dashed lines) as compared to the broadband data described in experiment 1 (dotted lines). The same set of separations was tested in this experiment as in experiment 1; however, the symbols have been omitted from this figure for clarity. The curves represent the percentage of responses in which the subjects responded that they heard both sources in the concurrent pair. Results for subjects 1–4 are shown [Figs. 4(a)–(d)] as well as the mean [Fig. 4(e)].

It can be seen for each individual that there was no substantial effect of removing the onset and offset ITD cues on performance; the curves (dashed lines) are generally in agreement with the broadband condition (dotted lines). This is true for both horizontal (left panels) and vertical (right panels) separation and is confirmed in the mean data.

It can also be seen that for subject 4, behavior did not change substantially as a result of the high-pass filtering. The curves (solid lines) for this subject generally agree with the broadband condition (dotted lines), with minor deviations. For subject 2 this is generally true also, although some larger deviations can be seen particularly for horizontal separation at the more lateral locations. For subjects 1 and 3, the high-pass filtering had a more obvious impact on performance, but again only at the more lateral reference locations. Here the psychophysical curves became flatter, and 100% separation rate was reached far less often than in the broadband condition. In an examination of the mean data [Fig. 4(e)] it appears that the front-most reference location (0° azimuth) was not affected by the high-pass filtering. However, some flattening of the curves is apparent at more lateral locations (especially 67.5° and 90° azimuth), representing a decrease in sensitivity to the task.

C. Discussion

It was interesting to find that the separation of concurrent sounds at the frontal location was not impaired by high-pass filtering the signals. It seems that in this situation, the ongoing low-frequency ITDs are not useful, or are redundant, and the perception is maintained by high-frequency acoustical cues in this case. It has been shown previously that high-pass filtering does not significantly affect the accuracy of localization of *single-source* stimuli (Butler and Hu-

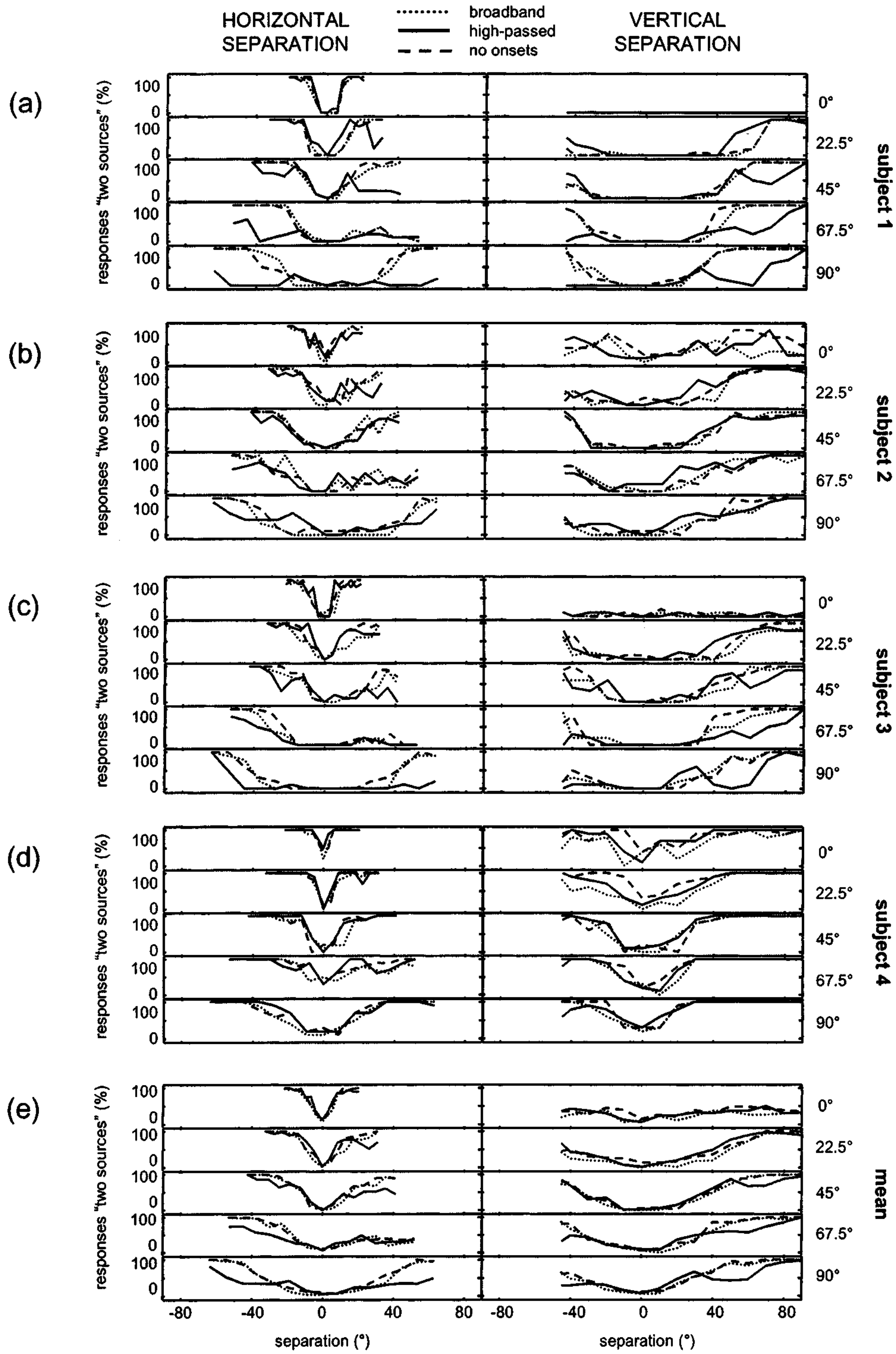

FIG. 4. Psychophysical curves for experiment 2. (a)–(d) show data for subjects 1–4, respectively, and (e) shows mean data. Results are shown for broadband stimuli (dotted lines; same as Fig. 3), stimuli high-pass filtered at 2 kHz (solid lines), and stimuli with onset/offset cues removed (dashed lines). All other details as for Fig. 3.

manski, 1992; Carlile and Delaney, 1999), and the present findings suggest that high-frequency cues can also be sufficient for identifying that more than one location is present in a *multiple-source* stimulus. In some subjects, high-pass filtering was seen to disrupt performance at the more lateral locations. Indeed the task was performed more poorly at these locations in the broadband condition, but with high-passed stimuli performance in some subjects dropped severely. In fact, for one subject the response rate dropped to around zero for *any* stimulus pair presented at the most lateral location [subject 1, Fig. 4(a)]. It is possible that this is related to ITD discrimination thresholds in this region, and this is discussed further in Sec. VI B.

Another finding from experiment 2 was that the presence of ITDs at the onset and offset of the stimulus was not necessary to match the performance of subjects seen in experiment 1. This is somewhat surprising considering that many neurons in the auditory pathway respond preferentially to the onset of a sound stimulus (Pickles, 1988). It is also curious because the onset cue is available and consistent across *all* frequency channels. Indeed for very short impulsive sounds such as clicks, it is the only cue available, and is crucial for lateralization (Tobias and Schubert, 1959). For longer duration stimuli under anechoic conditions, the ongoing phase cue is likely to be a more reliable cue because it can be integrated over time to improve the estimate. It seems that the transient offset and onset disparities are redundant in the presence of this very robust cue. A similar redundancy of onset ITD information has been shown for the lateralization of ongoing noise bursts presented over headphones (Tobias and Schubert, 1959).

V. EXPERIMENT 3: EFFECT OF REMOVING ITD COMPONENTS IN COMBINATION

A. Stimuli

Experiment 3a was a combination of the two conditions described in experiment 2, where stimuli were high-pass filtered at 2 kHz *and* left and right ear signals were each windowed as just described. In this way, onset and offset ITDs as well as ongoing low-frequency ITD information was removed. This allowed an examination of the usefulness of *high-frequency* information for this separation task, preserving only high-frequency ongoing ITD as well as level and spectral cues.

In the final experiment (experiment 3b) the aim was to remove *all* timing information. To achieve this, the two DTF-filtered stimuli were first created as for condition 1. Following this, the left and right ear signals representing the test stimulus were shifted in time (if required) such that their onsets matched those of the left and right ear signals of a stimulus corresponding to location (0°,0°). The signals were then ramped and added [see Fig. 2(c)]. The result of this manipulation was that the onsets and offsets of the stimulus pair coincided, and the ongoing time differences were matched too. Thus the timing information in isolation corresponded to only one location in space (0°,0°), but the spectral and interaural level information from each individual source was still contained in the composite signal.

These conditions were tested only at the frontal location (0°,0°), as results from the previous conditions had indicated that this location was the only one where no degradation was seen using the ITD manipulations individually.

B. Results

Each panel in Fig. 5 contains data from the previous experiments, as well as data from experiment 3, for horizontal separation at the frontal location only. Individual data for subjects 1–4 are shown [Figs. 5(a)–(d)] as well as the mean data [Fig. 5(e)]. Experiment 1 (dotted lines), experiment 2a (thin solid lines), and experiment 2b (thin dashed lines) have already been discussed, and again the relatively narrow troughs can be seen, indicating a good level of performance in this testing range. The thicker lines show data from experiment 3: high-pass filtered stimuli with onset/offset cues removed (thick solid lines) and stimuli with all ITD cues removed (thick dashed lines).

It can be seen that the curves from experiment 3a (thick solid lines) are very similar to the previous data, indicating that the combination of stimulus manipulations from experiment 2 did not degrade performance at this location. In fact, it seems that in this condition subjects tended to report *more often* that they perceived two sources.

Psychophysical curves for experiment 3b (thick dashed lines) show large individual differences, but it can clearly be

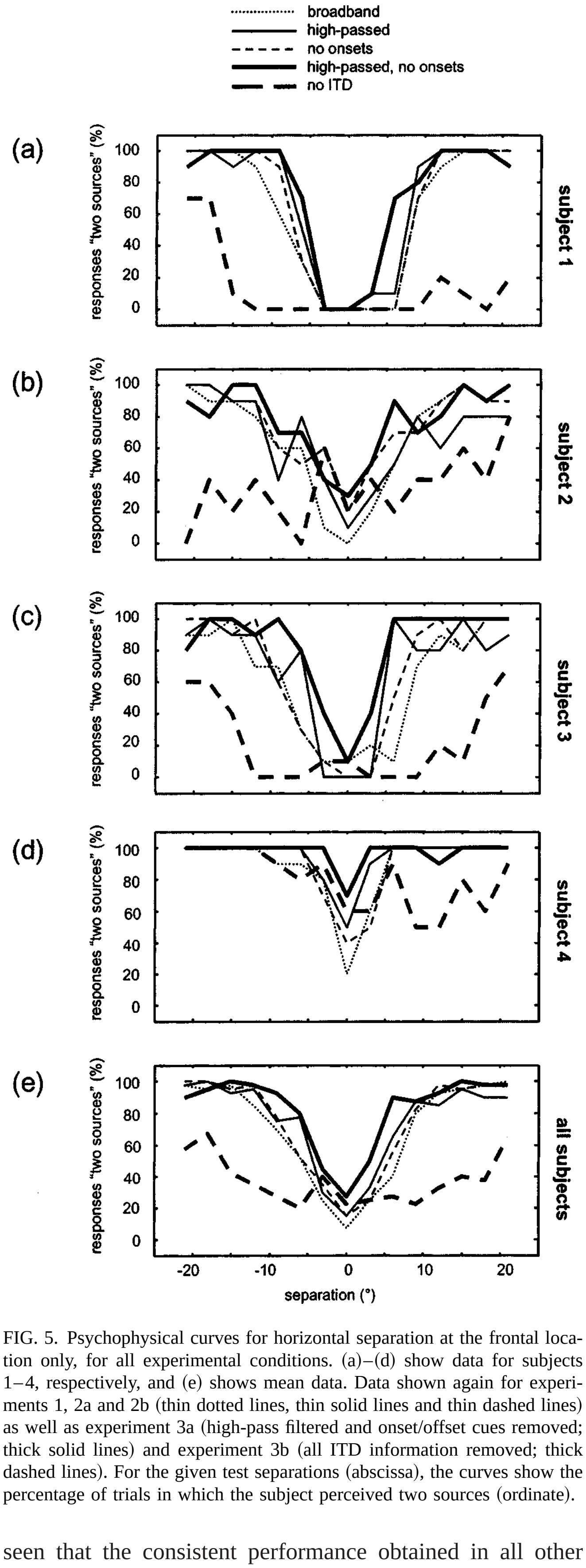

FIG. 5. Psychophysical curves for horizontal separation at the frontal location only, for all experimental conditions. (a)–(d) show data for subjects 1–4, respectively, and (e) shows mean data. Data shown again for experiments 1, 2a and 2b (thin dotted lines, thin solid lines and thin dashed lines) as well as experiment 3a (high-pass filtered and onset/offset cues removed; thick solid lines) and experiment 3b (all ITD information removed; thick dashed lines). For the given test separations (abscissa), the curves show the percentage of trials in which the subject perceived two sources (ordinate).

seen that the consistent performance obtained in all other conditions was disrupted in this condition. Across all spatial configurations, the number of trials where both stimuli in a concurrent pair were reported was lower than in the other conditions. This is true for the individual subjects [Figs.

5(a)–(d)] and for the mean [Fig. 5(e)]. Even for the largest separations, response rate never reached 100% for subjects 1–3. Subject 4 showed a marked performance asymmetry, displaying a reasonable ability to distinguish the two sources with leftward separation, but a disrupted performance with rightward separation.

C. Discussion

Results from experiment 1 suggested that binaural differences drove the separation of concurrent broadband noise sources in the four subjects examined. In experiment 2a, in which the stimuli were high-pass filtered, it was found that low-frequency phase information was not crucial to perform this task, although it aided performance at the more lateral locations. In experiment 2b, in which onset and offset cues were eliminated, it was seen that onset and offset ITD information was not critical and did not even appear to contribute to the performance seen in experiment 1. As there was a possibility of redundancy between these two sets of cues (low-frequency ITDs and onset/offset ITDs), they were both eliminated in experiment 3a to assess their combined importance. At the frontal location, performance was not disrupted under these conditions. Remarkably, subjects could determine the number of sources in a frontally presented signal with this highly impoverished set of localization cues. In the final experiment (experiment 3b), all ITD information was effectively removed by aligning the left and right ear signals in time for each member of the concurrent stimulus pair. The dramatic change in performance in this condition demonstrated that despite the redundancy in the cues, *some* ITD information is crucial for this task. It seems that interaural level differences (ILDs) are ruled out as major cues for this task, as they were available but did not maintain performance in this experiment. This suggests that in experiment 3a, subjects were reliant on ITDs in the envelope of the ongoing high-frequency signal. This is consistent with previous findings that envelope ITDs in high-frequency channels can be useful for sound lateralization (Henning, 1974, 1980; McFadden and Pasanen, 1976). The surprising point is that this cue *alone* produced performance levels equal to those seen when robust low-frequency ITD information was available. This is unexpected, especially because several studies have demonstrated the high-frequency ITD to be a much weaker cue than the low-frequency ITD for lateralization (Yost, 1976; Bernstein and Trahiotis, 1982).

An examination of Fig. 5 reveals that for experiment 3b, there were some effects at larger separations, with subjects reporting the perception of two sources despite the absence of ITD cues. It is possible that this is an artifact of the conflicting ITD and ILD cues at these extremes, where the ITD was zero but the overall ILD was nonzero as one of the source locations was displaced from the midline. Across subjects the mean overall ILDs at the extremes were measured to be 4 dB (test azimuth $-21°$) and -4.5 dB (test azimuth $21°$). Indeed it was reported by subjects that some stimuli in this experiment produced an ‘‘unnatural’’ percept, and this phenomenon has been reported previously for stimuli where interaural parameters are in conflict (e.g., Gaik, 1993).

VI. GENERAL DISCUSSION

A. Subject performance

In an overview of this collection of data, it is apparent that there are strong individual differences between subjects. As discussed in Sec. II D, efforts were made to ensure that each subject behaved consistently across trials, but as expected the response criterion adopted by a particular subject was quite individualized. This individuality showed up mainly in the overall tendency of subjects to respond that they perceived two source locations in a stimulus presentation. In order to quantify this and inspect the individual differences, a ‘‘false alarm’’ rate was calculated from the trials in which both noise sources emanated from the same location (‘‘zero separation’’ trials). Across the entire testing set there were 250 of these trials, and for each subject the total percentage of false alarms was calculated. For subjects 1–3 this rate was very low (0%, 6%, and 1.5%, respectively) but for subject 4 the occurrence was higher (26%). This indicates that subject 4 had a greater tendency to respond positively to two source locations, i.e., a less strict response criterion. However, despite these intersubject differences, the overall patterns in the data were the same across subjects.

Furthermore, it is clear from the data that subjects could resolve the pair of noises with a high degree of certainty under the right conditions (determined by both spatial configuration and individual criterion). This is in contrast to a recent study by Braasch (2002), who presented similar stimuli to subjects in order to examine and model localization in the presence of a distracter. Braasch presented concurrent 200-ms broadband noises separated in azimuth in the frontal region (analogous to our 0° reference location), and reported that in the majority of presentations, subjects perceived only a *single* auditory image. This is surprising because in that study the separation angles ranged from 15° to 90°, and in the present study subjects could fully resolve stimuli at the maximum separation of 21°. As the stimuli were identical apart from a small difference in their duration, the different responses of subjects are most likely related to the different tasks which they were asked to perform. Whereas our subjects were asked to report the number of source locations they perceived, the primary task of Braasch's subjects was to give an estimate of the location of the auditory image (if there was only one) or the most lateral location (if there was more than one image). Perhaps the fact that our subjects were asked to focus exclusively on the number of sources allowed them to adopt a more sensitive criterion to this parameter.

It is important to consider exactly what cues the subjects in the present study may have used in their assessment of the number of sources in a stimulus. It is clear that the perception of more than one source was a binaural effect; noise pairs separated along the vertical midline remained perceptually fused. Taken together, the results suggested that the presence of two different ongoing ITDs in the presented signal was responsible for separation. However, there are at least two possible ways in which this ITD cue might be used by the auditory system: (1) the two ITDs might be explicitly extracted from the composite signal, or (2) the presence of

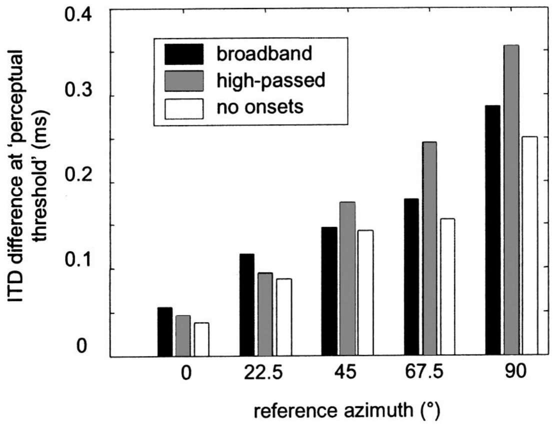

FIG. 6. Average ITD difference at perceptual threshold for experiments 1, 2a and 2b as a function of reference location. Values are pooled across subjects and direction of separation (see text for further details).

two ITDs in the mixed signal may be inferred from the neural representation. These two possibilities are discussed below.

B. A consideration of ITD sensitivity

If subject responses were driven by an ability to clearly perceive the two component ITDs, then it can be reasoned that the pattern of responses observed should bear some relation to ITD sensitivity with a single source. This idea is examined by considering the pattern of responses from experiments 1, 2a, and 2b which spanned the five reference locations.

Perceptual thresholds were estimated from the psychophysical curves for all subjects in these three conditions. Perceptual threshold was defined as the separation value whereby two source locations were reported to be perceived in 75% of cases. As "positive" and "negative" separation values were tested (corresponding to left and right, or up and down) an upper and lower threshold was obtained separately from the two halves of each curve. It was not possible to obtain a value in some cases, where performance did not reach the 75% level within the range of testing. For each case where a threshold value could be obtained, it was possible to express the angle in terms of the *ITD difference* between the two sources. Figure 6 shows the average ITD difference at perceptual threshold for each of the five reference locations. This value is pooled across the four subjects and the four directions of separation (left, right, down, up) for cases where threshold could be obtained.

It can be seen that, on average, listeners in the present experiment required about 50-μs difference in ITD to consistently perceive two sources when they were presented concurrently around the frontal reference position (0° azimuth). Klumpp and Eady (1956) made measurements of ITD thresholds for listeners using headphones. They found that for a broadband noise with a reference ITD of 0 μs, the threshold ITD for determining whether a successive stimulus was to the left or right was 10 μs. This value is substantially smaller than that for the present experiment, suggesting that, if ITD is the dominant cue in our task, it is more difficult to extract from a *concurrent* presentation of stimuli than from a sequential one. According to Fig. 6, the ITD difference required for separation increased as the reference location became more lateral. A similar effect has been reported for sequentially presented stimuli presented over headphones, where just-noticeable changes in ITD increase with baseline ITD. Klumpp and Eady (1956) used a band-pass filtered noise stimulus (150–1700 Hz) and found that the threshold for a 0-μs ITD stimulus was 9 μs but increased to 29 and 50 μs for ITDs of 430 and 790 μs, respectively. Other studies have shown similar effects for high- and low-frequency transients (Hafter and DeMaio, 1975) and 500-Hz tones (Hershkowitz and Durlach, 1969; Domnitz and Colburn, 1977).

The data in Fig. 6 also indicate that ITD thresholds are higher on average in the high-pass condition for the more lateral locations (45°, 67.5°, and 90° azimuth), confirming the observations made in experiment 2a. This finding is consistent with the study of Hafter and DeMaio (1975) where it was found that just-noticeable ITD differences were larger for high-frequency (3–4 kHz) clicks compared to low-frequency (0.1–2 kHz) clicks. These authors also reported that this disparity increased with increasing baseline ITD, and for the largest ITD examined (500 μs) ITD discrimination performance using high-frequency clicks was unmeasurable in one subject.

Thus, although there are substantial magnitude differences, it seems that there is some correspondence between the concurrent ITD thresholds estimated from the present data and reported sequential ITD discrimination thresholds. In both situations there is an increase in threshold with increasing reference ITD. In addition, while ITD in high-frequency regions is a useable cue, a low-frequency component improves performance at large values of ITD.

C. A consideration of ITD extraction from a mixed signal

Although the analysis presented above is consistent with the idea that ITDs are useful cues for the separation of broadband noise sources, one must consider by what mechanisms the auditory system might recover these cues from the summed stimuli it receives. Preliminary computational modeling using biologically plausible elements indicates that a running cross-correlation of the time-domain inputs to the left and right ears may provide an adequate explanation of these data (e.g., Best *et al.*, 2002). The emergence of two peaks in the cross-correlation function suggests that the auditory system may be able to extract individual ITD estimates for each of the sources.

However, an important effect of two independent sources mixing is decorrelation of the signals at the two ears. It is possible that this decorrelation is detected and used by subjects as an indication that there is more than one source. In terms of a cross-correlation mechanism, it may be that performance related less to the presence of two ITD peaks and more to the *lack of a strong peak*. Certainly it is clear from the literature that listeners are sensitive to interaural correlation (Pollack and Trittipoe, 1959a,b; Gabriel and Colburn, 1981) and that decorrelation is related to the perception of image width (e.g., Blauert and Lindemann, 1986). It is hoped that further experimentation and more sophisticated

TABLE I. Summary of localization data for the four subjects in the free-field and in virtual auditory space. Listed are spherical correlation coefficients and the percentage of trials in which a cone-of-confusion error occurred (in five tests each of 76 locations) under the two conditions.

Subject	Condition	SCC	% COC errors
1	Free-field	0.92	1.8
	VAS	0.85	3.7
2	Free-field	0.92	1.3
	VAS	0.92	2.1
3	Free-field	0.93	1.1
	VAS	0.90	2.9
4	Free-field	0.91	0.3
	VAS	0.91	1.8

modeling will illuminate this important issue of how well the ITDs of concurrent stimuli are preserved in the auditory system.

It is important also to note that the data in this study were collected under anechoic conditions, and the acoustics of real-world listening conditions may have some effect on the ability of subjects to discriminate one source from two. Certainly the presence of echoes could be expected to confuse the estimate of source numbers (by increasing the number of peaks and/or the interaural decorrelation). On the other hand, it has been shown that the impact of reverberation on the localization of a single source is minimal in some cases (Hartmann, 1983; Shinn-Cunningham, 2000).

VII. CONCLUSIONS

Spatial resolution with concurrent sources was examined systematically with a focus on the role of the different localization cues. The study was unique in that it (a) employed broadband stimuli that were distinguishable *only* on the basis of location, (b) focused on the detection of multiple sources in an acoustic stimulus rather than on issues of identification and/or localization, and (c) examined stimuli distributed both horizontally *and* vertically. It was shown that a concurrent pair of broadband white noises could be resolved if a sufficiently large binaural difference was present. In particular, differences in ongoing ITDs were shown to be robust cues for separation, including those in the envelopes of high-frequency channels (above 2 kHz).

APPENDIX: LOCALIZATION PERFORMANCE IN THE FREE-FIELD AND VIRTUAL AUDITORY SPACE

Prior to experimentation, the ability of the four subjects to localize broadband noise bursts in the free-field and in virtual auditory space (VAS) was assessed. Each subject underwent five localization tests (described in Sec. II B) in each condition, and the values in Table I summarize the data. Target and response locations were analyzed in terms of the lateral-polar angle coordinate system, and cone-of-confusion (COC) errors were defined as trials in which the target and response polar angle differed by more than 90°. The occurrence of this type of error is expressed as a percentage of the total number of trials. The remaining data were assessed statistically using the spherical correlation coefficient (SCC), which provides a measure of the correspondence between the perceived and actual directions on the sphere (1=perfect correlation; 0=no correlation). See Carlile *et al.* (1997) for more detail.

Batteau, D. W. (**1967**). ''The role of the pinna in human localization,'' Proc. R. Soc. London, Ser. B **158**, 158–180.

Bernstein, L. R., and Trahiotis, C. (**1982**). ''Detection of interaural delay in high-frequency noise,'' J. Acoust. Soc. Am. **71**(1), 147–152.

Best, V., van Schaik, A., and Carlile, S. (**2002**). ''The perception of multiple broadband noise sources presented concurrently in virtual auditory space,'' in *Proc. Audio Engineering Society 112th Convention*, Munich, Germany, paper 5549.

Blauert, J. (**1983**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge).

Blauert, J., and Lindemann, W. (**1986**). ''Spatial mapping of intracranial auditory events for various degrees of interaural coherence,'' J. Acoust. Soc. Am. **79**, 806–813.

Braasch, J. (**2002**). ''Localization in the presence of a distracter and reverberation in the frontal horizontal plane: II. Model algorithms,'' Acust. Acta Acust. **88**, 956–969.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge).

Bregman, A. S., and Campbell, J. (**1971**). ''Primary auditory stream segregation and perception of order in rapid sequences of tones,'' J. Exp. Psychol. **89**, 244–249.

Buell, T. N., Trahiotis, C., and Bernstein, L. R. (**1991**). ''Lateralization of low-frequency tones: Relative potency of gating and ongoing interaural delays,'' J. Acoust. Soc. Am. **90**, 3077–3085.

Butler, R. A. (**1986**). ''The bandwidth effect on monaural and binaural localization,'' Hear. Res. **21**, 67–73.

Butler, R. L., and Humanski, R. A. (**1992**). ''Localization of sound in the vertical plane with and without high-frequency spectral cues,'' Percept. Psychophys. **51**(2), 182–186.

Butler, R. A., Humanski, R. A., and Musicant, A. D. (**1990**). ''Binaural and monaural localization of sound in two-dimensional space,'' Perception **19**, 241–256.

Carlile, S. (ed.) (**1996**). *Virtual Auditory Space: Generation and Applications* (Landes, Austin).

Carlile, S., and Delaney, S. (**1999**). ''The localization of spectrally restricted sounds by human listeners,'' Hear. Res. **128**, 175–189.

Carlile, S., and Leung, J. (**2001**). ''Rendering sound sources in high fidelity virtual auditory space: Some spatial sampling and psychophysical factors,'' in *Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents and Virtual Reality*, edited by M. Smith, G. Salvendy, D. Harris, and R. Koubek (Erlbaum, Hillsdale, NJ), pp. 599–603.

Carlile, S., Jin, C., and van Raad, V. (**2000**). ''Continuous virtual auditory space using HRTF interpolation: Acoustic and psychophysical errors,'' *Proc. First IEEE Pacific-Rim Conf. Multimedia*, Sydney, Australia, pp. 220–223.

Carlile, S., Leong, P., and Hyams, S. (**1997**). ''The nature and distribution of errors in sound localization by human listeners,'' Hear. Res. **114**, 179–196.

Dirks, D. D., and Wilson, R. H. (**1969**). ''The effect of spatially separated sound sources on speech intelligibility,'' J. Speech Hear. Res. **12**, 5–38.

Divenyi, P. L., and Oliver, S. K. (**1989**). ''Resolution of steady-state sounds in simulated auditory space,'' J. Acoust. Soc. Am. **85**, 2042–2052.

Domnitz, R. H., and Colburn, H. S. (**1977**). ''Lateral position and interaural discrimination,'' J. Acoust. Soc. Am. **61**, 1586–1598.

Durlach, N. I., and Colburn, H. S. (**1978**). ''Binaural Phenomena,'' in *The Handbook of Perception*, edited by E. C. Carterette and M. P. Friedman (Academic, New York).

Gabriel, K. J., and Colburn, H. S. (**1981**). ''Interaural correlation discrimination: I. Bandwidth and level dependence,'' J. Acoust. Soc. Am. **69**, 1394–1401.

Gaik, W. (**1993**). ''Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling,'' J. Acoust. Soc. Am. **94**, 98–110.

Gardner, M. B., and Gardner, R. S. (**1973**). ''Problem of localization in the median plane: Effect of pinnae cavity occlusion,'' J. Acoust. Soc. Am. **53**, 400–408.

Hafter, E. R., and De Maio, J. (**1975**). ''Difference thresholds for interaural delay,'' J. Acoust. Soc. Am. **57**, 181–187.

Hartmann, W. M. (**1983**). ''Localization of sound in rooms,'' J. Acoust. Soc. Am. **74**, 1380–1391.

Henning, G. B. (**1974**). ''Detectability of interaural delay in high-frequency complex waveforms,'' J. Acoust. Soc. Am. **55**, 84–90.

Henning, G. B. (**1980**). ''Some observations on the lateralization of complex waveforms,'' J. Acoust. Soc. Am. **68**, 446–454.

Hershkowitz, R. M., and Durlach, N. I. (**1969**). ''Interaural time and amplitude jnds for a 500-Hz tone,'' J. Acoust. Soc. Am. **46**, 1464–1467.

Hirsh, I. J. (**1950**). ''The relation between localization and intelligibility,'' J. Acoust. Soc. Am. **22**, 196–200.

Klumpp, R. G., and Eady, H. R. (**1956**). ''Some measurements of interaural time difference thresholds,'' J. Acoust. Soc. Am. **28**, 859–860.

Kunov, H., and Abel, S. M. (**1981**). ''Effects of rise/decay time on the lateralization of interaurally delayed 1-kHz tones,'' J. Acoust. Soc. Am. **69**, 769–773.

McFadden, D., and Pasanen, E. G. (**1976**). ''Lateralization at high frequencies based on interaural time differences,'' J. Acoust. Soc. Am. **59**, 634–639.

Middlebrooks, J. C. (**1992**). ''Narrow-band sound localization related to external ear acoustics,'' J. Acoust. Soc. Am. **92**, 2607–2624.

Middlebrooks, J. C., and Green, D. M. (**1991**). ''Sound localization by human listeners,'' Annu. Rev. Psychol. **42**, 135–159.

Middlebrooks, J. C., Makous, J. C., and Green, D. M. (**1989**). ''Directional sensitivity of sound pressure levels in the human ear canal,'' J. Acoust. Soc. Am. **86**, 89–108.

Mills, A. W. (**1972**). ''Auditory Localization,'' in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), pp. 303–348.

Moller, H., Sorensen, M. F., Hammershoi, D., and Jensen, C. B. (**1995**). ''Head-related transfer functions of human subjects,'' J. Audio Eng. Soc. **43**(5), 300–321.

Moore, B. C. J. (**1997**). *An Introduction to the Psychology of Hearing* (Academic, London).

Noble, W., Byrne, D., and Ter-Host, K. (**1997**). ''Auditory localization, detection of spatial separateness, and speech hearing in noise by hearing impaired listeners,'' J. Acoust. Soc. Am. **102**, 2343–2352.

Perrott, D. R. (**1984**). ''Concurrent minimum audible angle: A reexamination of the concept of auditory spatial acuity,'' J. Acoust. Soc. Am. **75**, 1201–1206.

Perrott, D. R., and Saberi, K. (**1990**). ''Minimum audible angle thresholds for sources varying in both elevation and azimuth,'' J. Acoust. Soc. Am. **87**, 1728–1731.

Pickles, J. O. (**1988**). *An Introduction to the Physiology of Hearing* (Academic, London).

Pollack, I., and Trittipoe, W. J. (**1959a**). ''Binaural listening and interaural noise cross correlation,'' J. Acoust. Soc. Am. **31**, 1250–1252.

Pollack, I., and Trittipoe, W. J. (**1959b**). ''Interaural noise correlations: Examination of variables,'' J. Acoust. Soc. Am. **31**, 1616–1618.

Pralong, D., and Carlile, S. (**1996**). ''The role of individualized headphone calibration for the generation of high fidelity virtual auditory space,'' J. Acoust. Soc. Am. **100**, 3785–3793.

Shinn-Cunningham, B. G. (**2000**). ''Learning reverberation: Considerations for spatial auditory displays,'' in *Proc. Int. Conf. Auditory Display*, Atlanta, GA, pp. 126–134.

Stellmack, M. A. (**1994**). ''The reduction of binaural interference by the temporal nonoverlap of components,'' J. Acoust. Soc. Am. **96**, 1465–1470.

Tobias, J. V., and Schubert, E. D. (**1959**). ''Effective onset duration of auditory stimuli,'' J. Acoust. Soc. Am. **31**, 1595–1605.

Yost, W. A. (**1976**). ''Lateralization of repeated filtered transients,'' J. Acoust. Soc. Am. **60**, 178–181.

Zahorik, P. (**2000**). ''Limitations in using Golay codes for head-related transfer function measurement,'' J. Acoust. Soc. Am. **107**, 1793–1796.

Zhou, B., Green, D. M., and Middlebrooks, J. C. (**1992**). ''Characterization of external ear impulse responses using Golay codes,'' J. Acoust. Soc. Am. **92**, 1169–1171.