Auditory Objects of Attention

C. J. Darwin and R. W. Hukin

Laboratory of Experimental Psychology, University of Sussex, Brighton, BN1 9QG England

Abstract: Although interaural time differences (ITDs) are the dominant cue for the lateralisation of complex sounds, they are remarkably weak at grouping together simultaneous sounds. Experiments are described which explore the relationship between this observation and auditory attention. A small difference in ITD between two sentences spoken in the same voice is sufficient to allow a listener to say which of two target words occurred in the attended sentence, even when the two sentences are resynthesised to have the same fundamental frequency. By contrast, a large difference in ITD does not allow a listener to exclude a single harmonic from a target steady-state vowel within an attended sentence. The results can be explained by assuming that listeners are not able to attend directly to frequency components that share a common ITD. Rather, listeners attend to the direction of auditory objects, whose frequency composition is determined by grouping cues such as onset time and harmonicity, and whose lateral position is determined from the lateralisation cues of the component frequencies.

A PARADOX

On the one hand, both our everyday experience and experimental evidence (1, 2) show that auditory attention can be directed towards sounds that come from a particular spatial direction. On the other hand, although interaural time difference (ITD) is the dominant natural cue for determining the direction of a complex sound (3, 4), it is remarkably ineffective at helping listeners to group together the simultaneous frequency components that make up a particular sound source (5, 6, 7). We have proposed (8, 9) a theoretical framework, building on a previous suggestion of Woods & Colburn (10), that allows a resolution to this paradox.

A PROPOSED RESOLUTION

A proposed resolution to this paradox (on the right side of Figure 1) distinguishes between grouping mechanisms responsible for the formation of auditory objects (which make very little use of ITD) and the determination of the subjective location of a grouped auditory object which may be based on the pooled ITDs of the grouped frequency components (11, 12). Attention may then be directed to the position of the grouped auditory object, but cannot be directed (as in the left-hand part of Figure 1) to frequency components that simply share a common ITD.
We have shown (9) that listeners can attend across time to one of two simultaneous spoken sentences distinguished by very small differences in ITD. Listeners heard two simultaneous sentences played on a monotone (with a difference in Fo of 0, 1, 2 or 4 semitones between the sentences). Each sentence contained a target word, and the two target words ('dog', 'bird') were simultaneous and bounded by silence to eliminate co-articulatory cues. Their task was to attend to one of the sentences and say which target word occurred in it. When the sentences were presented without reverberation and differed by only ±45μs in ITD (corresponding to about 10° angular separation), listeners reported above 80% correctly which target word came from the attended sentence. We have subsequently shown that addition of reverberation (in the right panel of Fig. 1) significantly impairs listeners' performance but they were still well above chance at all but the smallest ITD.

![Figure 2](image)

**FIGURE 2.** Effect of ITD, Fo and reverberation on attention to one of two sentences.

By contrast, in a different experiment, listeners did not use such continuity of ITD to determine which individual frequency components should form part of the sentence to which they are attending (±600 μs had very little effect in segregating a single harmonic from a vowel embedded in a sentence). The conceptually simpler framework that listeners can simply attend to an ensemble of frequencies that share a common ITD (on the left side of Figure 1) is difficult to reconcile with these, and other, experimental results.

**ACKNOWLEDGMENTS**

Rob Hukin is supported by Medical Research Council (UK) grant G9505738N.

**REFERENCES**