

# Dissociation of perceptual judgments of what and where in an ambiguous auditory scene

Andrew H. Schwartz and Barbara G. Shinn-Cunningham<sup>a</sup>

*Hearing Research, Boston University, Boston, Massachusetts 02215*

(Received 6 January 2010; revised 8 September 2010; accepted 10 September 2010)

Whenever an acoustic scene contains a mixture of sources, listeners must segregate the mixture in order to compute source content and/or location. Some past studies have explored whether perceived location depends on which sound elements are perceived within a source. However, no direct comparisons have been made of “what” and “where” judgments for the same sound mixtures using the same listeners. The current study tested if the sound elements making up an auditory object predict that object’s perceived location. Listeners were presented with an auditory scene containing competing “target” and “captor” sources, each of which could logically contain a “promiscuous” tone complex. In separate blocks, the same listeners matched the perceived spectro-temporal content

logically be part of either of two competing objects, can show an apparent disconnect between the degree to which an interfering tone contributes to an auditory object’s perceived spectro-temporal content (Shinn-Cunningham *et al.*, 2007) versus the object’s perceived spatial location (Lee *et al.*, 2009). These experiments employed a synthetic vowel and a stream of “captor” tones designed to promote segregation of the third harmonic from the vowel (the “ambiguous tone”). The vowel, captors, and ambiguous tone were played with various simulated source locations, and subjects were asked to judge either the vowel’s spectrum or location. In the absence of the captors, the ambiguous tone contributed strongly to the perceived spectrum of the vowel, even when the ambiguous tone and the other vowel components had different spatial cues. Estimates of the vowel location in this situation generally fell in between the reported locations of the ambiguous tone alone and the vowel alone (without the ambiguous third harmonic). This result is consistent with the idea that localization is determined by integrating cues across the sound components that are perceived as making up the target object. In trials with the captor stream, listeners heard the ambiguous tone as contributing only weakly, if at all, to the spectrum of the vowel. However, despite this large change in the spectral judgment of the vowel due to the presence of the captors, the captors had a very small influence on the perceived location, which was always strongly affected by the spatial cues in the ambiguous tone.

Although these results appear to contradict the consistent-object hypothesis, it is still possible that the hypothesis holds in these conditions. For instance, a sound element that only contributes a small amount to the perceived spectrum of an object could still greatly influence the perceived location of the object. Thus, the consistent-object hypothesis could explain these results if the perceived location of the vowel is influenced strongly by the presence of a component sound element that has a relatively low intensity.

Here we explicitly test the consistent-object hypothesis in two-object mixtures with ambiguous grouping cues by quantifying the level of contribution of a “promiscuous” tone complex (which could logically be grouped into either of two competing auditory objects) to both the perceived spectro-temporal content and the perceived location of a target tone complex (one of the two competing objects). We used a repeating captor tone complex, similar to the captor stream in previous experiments (Lee *et al.*, 2009; Shinn-Cunningham *et al.*, 2007) to promote segregation of the promiscuous complex from the target complex. We then manipulated the level of the captor stream and promiscuous complex to change the degree to which the promiscuous complex contributed to the perception of the target’s spectral content and perceived location.

First, to understand how the level of the promiscuous complex influenced the perceived spectral content or location of the target complex, we asked listeners, in separate sessions, to match either the spectral content or location of the target complex for different intensities of the promiscuous complex in trials where the captors were absent (“no-captor trials”). We then had subjects perform the same perceptual judgments in trials with the captors present (“ambiguous-

mixture trials). We used data from the no-captor trials to estimate the “effective level” of the promiscuous complex contributing to subjects’ matches in the ambiguous mixtures. By combining the results from the no-captor trials with the results from the ambiguous mixtures, we directly tested the consistent-object hypothesis. We found that the contribution of the promiscuous complex to the target’s spectral content did not quantitatively predict the level of contribution to its perceived location, contradicting the consistent-object hypothesis.

## II. METHODS

In separate tasks, subjects were trained to match either the spectral content/timbre (“what” task) or the location (“where” task) of a repeating target tone complex. Each trial repeatedly alternated back and forth between the stimulus to be matched (the target) and a “match” stimulus. Both the target and match stimuli lasted three seconds each time they were played; they alternated repeatedly, allowing subjects to compare them back to back. During each presentation of the match stimulus, listeners could adjust either the match stimulus’ spectral content (in the “what” task) or its laterality (in the “where” task) to perceptually match the corresponding attribute of the target, as described below. There was no limit on the number of target/match presentations per trial; when satisfied with the match, the subject ended the trial by pressing a button.

All tone complexes were 75 ms in duration with 6 ms linear on/off ramps. The target had a fundamental frequency of 110 Hz and had harmonics as described in Fig. 1(A) (solid

bars). The target was repeated every 300 ms during the target presentation portion of a trial (corresponding to 10 repetitions). A second “promiscuous tone complex [Fig. 1(A), open bars] had a fundamental frequency of 330 Hz (the third harmonic of the target) and was presented simultaneously with the target. In ambiguous-mixture trials, two “captor complexes, identical to the promiscuous complex, were presented sequentially prior to each presentation of the promiscuous complex and target. These captor complexes were separated by 100 ms, creating an isochronous sequence of the 330-Hz-fundamental complex [Fig. 1(B)]. Thus, the promiscuous complex could be perceptually grouped with either the target or the captors, or possibly with both or with neither (Shinn-Cunningham *et al.*, 2007). In no-captor conditions, only the target and simultaneous promiscuous complex were present, forming a single harmonic complex with a fundamental frequency of 110 Hz.

Stimuli were played on commercially available hardware (Tucker Davis Technologies, Alachua, FL) and delivered to subjects via Etymonic ER-1 insert earphones. The level of the combined target+promiscuous tone complex was roughly 73 dB SPL prior to any attenuation of the promiscuous tone complex. Thirteen subjects participated in both the main “what task and the “where task. Six of these thirteen subjects (selected based on availability, rather than any other criterion) participating in a follow-up control task to ensure that “what matches were based on perceived spectral content rather than loudness. All subjects, ranging in age from 18–30, had clinically normal hearing. Subjects gave written consent, as overseen by the Boston University Charles River Campus IRB, and were compensated \$10/h for their participation.

## A. What task

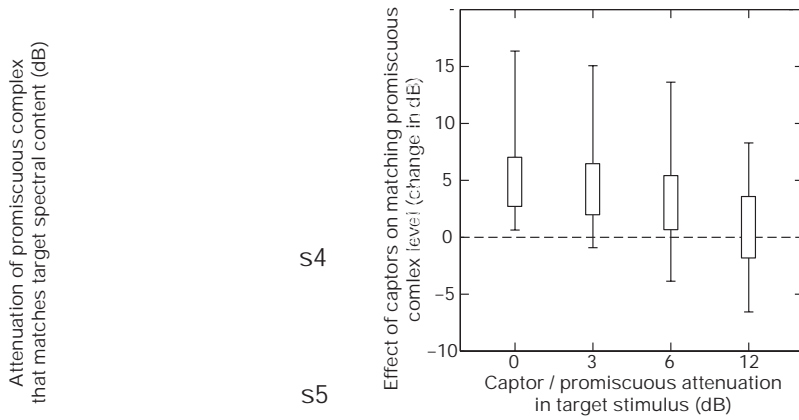
In the “what task, the match stimulus was a diotic tone complex consisting of the target and promiscuous tone complexes. To reset any buildup of streaming between the test and match tone complexes, both the match and the presentation period were followed by a white noise burst at roughly 60 dB SPL whose length varied randomly from 600–1000 ms. This noise burst was preceded by a pause whose length varied randomly from 400 to 700 ms. The noise burst presented after the match stimulus had a 50 ms silent gap starting at 200 ms to help subjects avoid confusing the test and match stimuli. Subjects could control-

cuous tone attenuations (four values), for a total of 64 trials per session. The 64 trials in each session were randomly ordered, separately for each subject.

## B. Where task

For the “where” task, the match stimulus was an acoustic ILD pointer (Bernstein and Trahiotis, 1985; Trahiotis and Stern, 1989; Buell and Hafter, 1991; Best *et al.*, 2007) made up of a 200-Hz-wide band of noise centered at 2 kHz, played at a level of roughly 55 dB SPL. Subjects could change the ILD of the pointer during presentation of the match stimulus by pressing two buttons (the same as those used in the “what” task) to increase or decrease the ILD between  $\pm 30$  dB. As in the “what” task, ILD changes were performed in real time with a sufficiently small step size that the ILD of the pointer changed continuously and smoothly as long as a button was depressed.

Subjects were trained to use the ILD pointer before the main data collection began. During training, the test stimulus was a 200-Hz-wide band of noise centered at 2 kHz with a 2 Hz, 50%-depth sinusoidal modulation envelope that distinguished it from the match stimulus. We used this noise stimulus rather than the tone complexes used in the main experiment (which contained ITDs) so that we could objectively verify that subjects could reliably match the stimulus ILD with the pointer ILD. The test stimulus was given a fixed ILD selected from a uniform distribution between  $\pm 20$  dB. Subjects were asked to match the lateral location of the test noise stimulus by adjusting the ILD of the match stimulus. Training was conducted in 40-trial-long runs, each of which typically lasted 20–30 min. At the start of each experimental session, subjects trained until they completed one training run in which the RMS difference between the ILDs of the test and match stimuli was less than 4 dB. Subjects typically met this criterion after 2–3 training runs on the initial day of “where” testing. Subjects who completed train-



possible that they instead matched some other target attribute. Specifically, listeners may have adjusted the level of the promiscuous complex in the match stimulus to equate the perceived loudness of the target and match stimuli, rather than spectral shape.

To test for this possibility, six of the original subjects participated in a control task. This experiment was similar to the main “what” task, except that (1) for brevity, we presented target stimuli with only two attenuations (0 and 3 dB) of the captors and promiscuous complex, and (2) in half of the trials, the subject controlled the attenuation of the target complex (within the match stimulus), rather than the attenuation of the promiscuous complex (leaving the promiscuous complex level in the match stimulus unchanged). To encourage listeners to use the same strategy, regardless of whether they adjusted the promiscuous complex or the target complex, we intermingled trials randomly. A total of eight different trial types were used: all combinations of two target-stimulus attenuations (either 0 or 3 dB), two stimulus elements adjusted by subjects during the matching task (either target complex or promiscuous complex), and two types of target stimuli (captor either present or absent). Subjects performed six matches for each trial type, for a total of 48 trials per subject. Each subject completed these trials in a single, brief (less than 1 h long) session, following completion of both the main “what” task and the “where” task. They were instructed to perform this task just as they had performed the main “what” task.

If listeners were matching target loudness and if the presence of the captors decreased the target loudness (for instance), then the presence of the captors would cause listeners to attenuate whichever components they controlled in the match stimulus, whether they could adjust the promiscuous-complex level or the target-complex level. However, if they were matching perceived target timbre,

no adjustments would be needed in the two matching paradigms. Specifically, if captors decreased the perceptual contribution of the promiscuous complex to the target and listeners matched the target spectral content as instructed,

they would attenuate the match stimulus’ promiscuous complex if they controlled its level, but they would increase the match stimulus’ target complex if they controlled its level. Moreover, such a pattern of adjustment would change the overall loudness of the match stimulus in opposite directions in the adjust-promiscuous-complex and adjust-target-complex trial. Therefore, if this pattern is observed, it rules out the possibility that listeners matched loudness.

### III. RESULTS

#### A. WHAT TASK

Subjects generally set the intensity of the promiscuous complex in the match stimulus in the no-captor trials close to the physical intensity of the promiscuous complex in the test stimulus [Fig. 2(A), circles], even though the promiscuous complex and target had different ITDs. The mean difference between subjects’ responses and the true, physical attenuation of the promiscuous complex was 1.4 dB, with a standard deviation of 2.6 dB. These results confirm that subjects were able to reliably match the spectral content of the target-plus-promiscuous-complex object using our procedures. Moreover, these results are consistent with results from past studies suggesting that listeners do not segregate sound elements on the basis of ITD alone (Culling and Summerfield, 1995; Darwin and Hukin, 1999).

These data were fitted by a least-squares linear regression [Fig. 2(A), dotted lines] representing the relation between the true physical level of the promiscuous complex (at 0  $\mu$ s ITD) and the perceived level of the promiscuous complex grouped with the target complex (at 600  $\mu$ s ITD). To ensure response reliability, subjects whose least-squares fit yielded an RMS error of over 7 dB (twice the training criterion) were excluded from analysis (two out of 13 subjects were excluded for not meeting this criterion). Note that this RMS error refers only to the prediction error between the linear fit and the underlying data, and does not make the explicit assumption that subjects will “correctly” match the true promiscuous complex intensity, which differed in ITD from

tones in the target complex. However, it was generally true that the perceived contribution of the (straight ahead) promiscuous complex to the composite target-plus-promiscuous-complex object was very close to the full intensity of the promiscuous complex.

Subjects generally set the level of the promiscuous complex in the match stimulus to be lower in the with-captor trials than in the no-captor trials [in Fig. 2(A), triangles lie above circles]. This result is consistent with the captors reducing the contribution of the promiscuous complex to the target. However, the degree of this reduction was inconsistent across subjects, demonstrated in the three panels of Fig. 2(A). For s3, the captors had little effect on the perceived spectral content of the target (circles and triangles fall almost on top of each other in the top panel). For s5, the presence of the captors reduced the promiscuous complex contribution by roughly 15 dB (triangles fall well above the circles in the bottom panel). For s4, the captors had a modest effect (triangles fall above the circles in the middle panel).

To summarize results across subjects, we quantified the effect of the captors on the perceived spectral content of the target by computing the difference between the response in the presence of the captors [triangles in Fig. 2(A)] and in their absence [circles in Fig. 2(A)]. Full and inter-quartile ranges of this difference are plotted in Fig. 2(B) for the subjects who responded consistently (i.e., who passed the RMS criterion described above). Overall, the captors tended to reduce the contribution of the promiscuous complex to the target [values in Fig. 2(B) generally tend to be positive or near zero]. The effect tended to decrease as the level of the captors and promiscuous complex decreased. It is worth noting that this decrease may in part be due to the limited response range available (subjects could set the promiscuous complex level in the match stimuli between +2 and -20 dB).

For the 6 and 12 dB conditions, the upper range of the response attenuations reaches the maximum possible value. In all conditions there was large inter-subject variability in the effect of the captors. Yet despite this variability the individual subjects were very reliable in their responses, with some subjects showing large, consistent shifts in the perceived intensity of the promiscuous complex contributing to the target's spectral content due to the captors [e.g., s5 in Fig. 2(A)].

## B. WHAT TASK

To account for possible shifts between sessions in subjects' maps from ITD and ILD to perceptual space and to enable more direct comparisons across subjects (who may not use the ILD scale identically; e.g., [Bernstein and Trahitis, 1985](#); [Best et al., 2007](#)), responses in the "where" task were shifted and normalized within each session so that zero represents the mean ILD response to the diotic control trials and one represents the mean ILD match to the 600  $\mu$ s ITD target alone (the largest ITD used).

Figure 3(A) plots the normalized laterality of the target as a function of the promiscuous-complex attenuation for the same three subjects whose "what" task results are shown in Fig. 2(A). The results for the no-captor trials (the largest ITD used

target location data are well summarized by a straight line [e.g., see the correspondence of dotted lines and circles in Fig. 3(A)], justifying the use of a least-squares linear regression to fit these responses. The resulting linear relationship maps the physical attenuation of the promiscuous complex to a perceived location of the target auditory object.

The presence of the captors reduced the influence of the promiscuous complex on the perceived location of the target (see also Best *et al.*, 2007; Lee *et al.*, 2009); the perceived locations tend to be more lateral with the captors than without [triangles in Fig. 3(A) tend to fall above the circles]. However, the size of this effect differs across subjects. For s5 (bottom panel), the captors have a negligible influence on the perceived target location, while this effect is moderate for s3 and s4 (top two panels). Moreover, as discussed at greater length below, these individual differences are not predicted by the individual differences in the “what” task.

As in the “what” task, we quantified the effect of the captors by computing the difference between the response in the presence of the captors [Fig. 3(A), triangles] and in their absence [Fig. 3(A), circles]. The across-subject full and interquartile ranges of this difference are plotted in Fig. 3(B). Values above zero indicate that the addition of the captors caused the target to be perceived more laterally than in the no-captor condition (i.e., to be less influenced by the 0  $\mu$ s ITD of the promiscuous complex). Despite the fact that inter-subject differences were large, Fig. 3(B) shows that most subjects perceived the target as more lateral when the captors were present compared to when the captors were absent (there was typically a positive shift in perceived location due to the captors).

In addition to affecting the influence of the promiscuous tones on the target, the captors also had some influence on the perceived location of the target itself. To address this influence, we included a session in which subjects matched the target location when the target and captors were present, but the promiscuous complex was absent. In these trials, subjects tended to perceive the target further to the side when the captors were present than when the target was played alone (see also Best *et al.*, 2005; Braasch and Hartung, 2002; Lee *et al.*, 2009; Lorenzi *et al.*, 1999). This resulted in nor-

malized localization responses greater than one, demonstrating “repulsion” between the captors and the target (see, for example, Lee *et al.*, 2009).

Figure 4(A) shows data for these no-promiscuous-complex trials for the same three subjects as in Figs. 2(A) and 3(A), as well as for one additional subject who showed an even stronger repulsion effect than any of the other three example subjects. As with the results for “what” and “where,” these measures reveal large subject differences. Repulsion was negligible for s3, s4, and s5 [triangles fall near one in the top three panels of Fig. 4(A)]; oshows data for these no-complex trials for the same three subjects as in

spatial judgments of the target in ambiguous mixtures. If, in computing the target's location, binaural cues in different frequency components are weighted according to their contribution to the target's spectral content (the consistent-object hypothesis), then the effective levels computed from the "what" and "where" tasks for a given subject should be equal.

The effect of the captors was often large enough that, for stimuli in which the promiscuous complex was attenuated by 6 and 12 dB, responses were outside the range of responses observed in corresponding no-captor trials. For these responses, computing an effective level of the promiscuous complex would require extrapolation of the linear fits to the no-captor data [e.g., see Fig. 3(A), top panel; triangles for the 6 and 12 dB attenuations lie well above the range of  $\alpha$ -axis values described by the dotted line]. Moreover, the relationship between physical attenuation of the promiscuous complex and the normalized perceived location cannot be linear over an infinite range (e.g., once the promiscuous complex is attenuated enough that it has no measurable impact on the target, further attenuation will not change the perception of the target). Therefore, we restricted all statistical analyses that used effective levels of the ambiguous matches to the 0 and 3 dB attenuations, where such extrapolation was not generally needed.

We can quantify how well our data fit the consistent-object hypothesis by plotting, for each subject and stimulus condition, the effective attenuation of the promiscuous complex in the "where" task (hereafter, "effective spatial attenuation") against the effective attenuation of the same stimulus in the "what" task ("effective spectral attenuation"). The consistent-object hypothesis predicts that these quantities should be equal, so values plotted this way should lie along the identity line.

Figure 5(A) shows the means and 95% confidence intervals of the effective attenuations in the two tasks for the same three subjects from previous figures (filled symbols).

For subject s3, the effective spatial attenuation is larger than the effective spectral attenuation (all points fall above the diagonal in the top panel). The opposite is true for subject s5, where the effective spectral attenuation is larger than the effective spatial attenuation (all points fall below the diagonal in the bottom panel). Other subjects' results lie in between these extremes; for instance, data for s4 fall on the diagonal, in accordance with the consistent-object hypothesis [middle panel of Fig. 5(A)]. Although the inter-subject differences are large, intra-subject differences are small, demonstrating response reliability. Thus, though the consistent-object hypothesis may describe results for some subjects in our population [e.g., s4, in the middle panel of Fig. 5(A)], it does not generally hold for all subjects (see below for a more thorough statistical analysis).

Figure 5(B) summarizes the group data by plotting, for Average displacement from identity line (dB)



target), which may be incorrect; however, this analysis gives a first-order correction for repulsion effects. We used the resulting corrected lateralization values to calculate a corrected effective promiscuous complex level [open symbols in Fig. 5(A)].

As seen in Fig. 5(A), correcting for repulsion does not improve the fit of the effective level points to the identity line. Moreover, given that such a correction will generally tend to reduce the effective spatial attenuation of the promiscuous complex, and given that some subjects already demonstrate less effective spatial attenuation than effective spectral attenuation [e.g., s5 in Fig. 5(A)], we conclude that repulsion cannot account for the observed departure from the consistent-object hypothesis.

Although inter-subject variability is large, individual subjects are relatively consistent in how they respond. We therefore analyzed individual results to see if we could reject the consistent-object hypothesis for individual subjects. Assuming response variations for a given subject and given stimulus condition can be accurately modeled as Gaussian-distributed noise, we can test for significant differences in the effective spectral and spatial attenuations (which, being affine transformations of subjects' responses, are also normally distributed) using a paired t-test against the null hypothesis (the consistent-object hypothesis). Specifically, the null hypothesis posits that the effective spatial attenuation and effective spectral attenuation have the same distribution. As discussed above, effective attenuations computed from extrapolated values of the linear fits are unreliable; therefore, we only analyzed results for the 0 and 3 dB conditions.

Figure 6 plots histograms of the resulting p-value distributions for the individual subjects' data. The small p values for some subjects in our population suggest that the consistent-object hypothesis can be rejected for several, but not all, of our subjects (four out of eleven subjects in the 0 dB condition and six out of eleven subjects in the 3 dB condition). However, in such a population analysis, a small fraction of trials (here, individual subject's t-test results) will turn out to be significantly different simply by chance. To assess the significance of our t-test p-value distribution, we can treat the outcome of each t-test as a Bernoulli trial, where the event " $p < 0.05$ " defines a success. The probability of obtaining four or more such successes (the result of the 0 dB condition) in eleven independent trials is approximately  $1.6 \times 10^{-3}$ . For six or more successes (the result of the 3 dB

condition), this probability is approximately  $5.8 \times 10^{-6}$ . Note that for the 0 dB condition, all four significant p-values were below 0.01; defining " $p < 0.01$ " as a success gives a stricter probability of approximately  $3.1 \times 10^{-6}$ . Hence, we conclude that these failures of the null hypothesis across the population of tested subjects were not observed by chance. Thus, although results for some subjects are well fit by the consistent-object hypothesis, we can reject the consistent-object hypothesis as describing a general property that holds for all subjects in the population at large.

#### D. C. $\Psi$ $\Psi$

In the main "what task, the captors' presence caused listeners to attenuate the level of the promiscuous tones to match the target stimuli. However, these results could arise if the captors reduced the perceived loudness of the target and listeners matched loudness rather than spectral shape. Our control experiment tested for this possibility by asking listeners to perform the same matching task as in the main "what task, both when controlling the level of the match stimulus' promiscuous complex-285.8Nistimulus'-.2i-249etimub50-285.s tos2.887ausedbothg4.3c1-665.9bc1-665ulus'-455.-cuous-52-onclmatng-40

matched the spectral shape of the target stimulus, the mean response attenuations of the promiscuous complex should be negatively correlated with the mean response attenuations of the target complex.

For each subject, we computed the Pearson's correlation coefficient relating the attenuations of the promiscuous complex to attenuations of the target complex for the same target stimulus. These Correlation coefficients ranged from  $-0.67$  to  $-0.98$ , with a mean of  $-0.87$ . From this, we conclude that subjects were matching the timbre of the tone complexes rather than overall stimulus loudness.

#### **IV. SUMMARY AND DISCUSSION**

Recent experiments showed that a reduction in the contribution of a tone to a target vowel's spectrum did not appear to result in an equivalent reduction in its contribution to

## V. CONCLUSIONS

1. Two harmonically related tone complexes (target and promiscuous complexes) were grouped together as a single object despite having different ITDs. The perceived location of the composite, grouped object varied with the level of the promiscuous complex, consistent with object location being determined by a weighted integration of binaural cues across perceptually grouped frequency components.
2. Adding a captor stream reduced the effective level of the promiscuous complex for both judgments of the target's spectrum and of the target's location.
3. The reduction in the contribution of the promiscuous complex to the perceived location of the target differed from the reduction in the contribution to the perceived spectral content of the target. Although some individual subjects' spatial and spectral judgments were influenced in the same way by the captors, effects varied markedly between subjects. These results contradict the consistent-object hypothesis, and show a dissociation between how the brain computes what an object is and where it is located.
4. Some subjects showed "repulsion" between the captors and the target. However, this cannot account for the discrepancies observed between the effective spectral and spatial contributions of the promiscuous complex to the target.

## ACKNOWLEDGMENTS

This work was supported by Grants NIDCD ROI DC009477 (to BGSC) and by training Grant NIDCD T32 DC00038 (supporting AS). We would like to thank Timothy Streeter for assistance with the TDT system, Adrian K.C. Lee for his help in setting up the experimental paradigm, and Lorraine Delhorne for assistance in recruiting subjects and running experiments. Two anonymous reviewers provided helpful, insightful comments on an earlier version of this manuscript.

- Bernstein, L. R., and Trahiotis, C. (1985). "Lateralization of low-frequency, complex waveforms: The use of envelope-based temporal disparities, J. Acoust. Soc. Am. **77**, 1868-1880.
- Best, V., Gallun, F. J., Carlile, S., and Shinn-Cunningham, B. G. (2007). "Binaural interference and auditory grouping, J. Acoust. Soc. Am. **121**, 1070-1076.
- Best, V., van Schaik, A., Jin, C., and Carlile, S. (2005). "Auditory spatial perception with sources overlapping in frequency and time, Acta Acust. Acust. **91**, 421-428.
- Braasch, J., and Hartung, K. (2002). "Localization in the presence of a distracter and reverberation in the frontal horizontal plane. I. Psychoacoustical data, Acta Acust. Acust. **88**, 942-955.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA), pp. 1-790.
- Buell, T. N., and Hafter, E. R. (1991). "Combination of binaural information across frequency bands, J. Acoust. Soc. Am. **90**, 1894-1900.
- Culling, J. F., and Summerfield, Q. (1995). "Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay, J. Acoust. Soc. Am. **98**, 785-797.
- Darwin, C. J. (1995). "Perceiving vowels in the presence of another sound: A quantitative test of the 'old-plus-new' heuristic, in *Learning and Cognition: Reactions and Interactions: Theoretical and Methodological*, Elsevier, Amsterdam, pp. 1-12.
- Darwin, C. J., and Hukin, R. W. (1999). "Auditory objects of attention: The role of interaural time differences, J. Exp. Psychol. Hum. Percept. Perform. **25**, 617-629.
- Dye, J. (1990). "The combination of interaural information across frequencies, J. Acoust. Soc. Am. **88**, 483-493.