A "looming bias" in spatial hearing? Effects of acoustic intensity and spectrum on categorical sound source localization

Lisa McCarthy[1] · Kirk N. Olsen[1,2]

Abstract Continuous increases of acoustic intensity (*up-ramps*) can indicate a looming (approaching) sound source in the environment, whereas continuous decreases of intensity (*down-ramps*) can indicate a receding sound source. From psychoacoustic experiments, an "adaptive perceptual bias" for up-ramp looming tonal stimuli has been proposed (Neuhoff, 1998). This theory postulates that (1) up-ramps are perceptually salient because of their association with looming and potentially threatening stimuli in the environment; (2) tonal stimuli are perceptually salient because of an association with single and potentially threatening biological sound sources in the environment, relative to white noise, which is more likely to arise from dispersed signals and nonthreatening/nonbiological sources (wind/ocean). In the present study, we extrapolated the "adaptive perceptual bias" theory and investigated its assumptions by measuring sound source localization in response to acoustic stimuli presented in azimuth to imply looming, stationary, and receding motion in depth. Participants ($N = 26$) heard three directions of intensity change (up-ramps, down-ramps, and steady state, associated with looming, receding, and stationary motion, respectively) and three levels of acoustic spectrum (a 1-kHz pure tone, the tonal vowel /ə/, and white noise) in a within-subjects design.

We first hypothesized that if up-ramps are "perceptually salient" and capable of eliciting adaptive responses, then they would be localized faster and more accurately than down-ramps. This hypothesis was supported. However, the results did not support the second hypothesis. Rather, the white-noise and vowel conditions were localized faster and more accurately than the pure-tone conditions. These results are discussed in the context of auditory and visual theories of motion perception, auditory attentional capture, and the spectral causes of spatial ambiguity.

Fast and accurate sound source localization is an adaptive behavior that enables organisms to efficiently locate and prioritize stimuli that pose a potential threat (Letowski & Letowski, 2012)—for example, looming (approaching) sound-emitting objects in motion. One important acoustic cue associated with auditory motion is a continuous change of intensity that perceptually is experienced as a change in loudness (Jenison, 1997; Olsen, 2014). Specifically, continuous increases (or *up-ramps*) of intensity are associated with looming sound sources in the environment, and continuous decreases (or *down-ramps*) of intensity are associated with receding sound sources (Neuhoff, 1998; Olsen, Stevens, & Tardieu, 2010). In psychoacoustic experiments, asymmetries in the perception of up-ramps and down-ramps have been reported in the context of subjective duration (DiGiovanni & Schlauch, 2007; Grassi & Darwin, 2006; Meunier, Vannier, Chatron, & Susini, 2014), global loudness (Ponsot, Meunier, Kacem, Chatron, & Susini, 2015; Ponsot, Susini, & Meunier, 2015; Stecker & Hafter, 2000), and loudness change (Neuhoff, 1998; Olsen & Herff, 2015; Olsen & Stevens,

✉ Kirk N. Olsen
k.olsen@westernsydney.edu.au

Lisa McCarthy
lisa.mccarthy.92@hotmail.com

[1] The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia

[2] Department of Psychology, Macquarie University, Sydney, NSW 2109, Australia

2010; Olsen et al., 2010; Teghtsoonian, Teghtsoonian, & Canévet, 2005). To summarize this literature, up-ramps are commonly perceived as being louder, longer, and covering a greater magnitude of loudness change than down-ramps (for a review, see Olsen, 2014).

One of the key conceptual issues underlying perceptual asymmetries in response to up-ramps and down-ramps is the theory of an "adaptive perceptual bias" for up-ramps of intensity and looming auditory motion (Neuhoff, 1998, 2001; Seifritz et al., 2002). In free-field listening conditions, a looming sound source is perceived to stop closer to a listener and travel a greater perceived distance than a receding sound source presented with an equivalent distance, duration, and stopping point (Neuhoff, 2001). Furthermore, time-to-contact studies have reported that real and implied looming stimuli are perceived to arrive at a point in space sooner than would be expected from the physical velocity of the approaching stimulus (Rosenblum, Wuestefeld, & Saldana, 1993; Schiff & Oldak, 1990).

As a result of these findings and of those from experiments measuring perceived loudness change (Neuhoff, 1998, 2001), Neuhoff concluded that up-ramps of intensity elicit adaptive behaviors that may provide a selective advantage for organisms able to perceive a looming sound source to be closer than it actually is, thus providing greater opportunity for that organism to prepare for the arrival of the source and take appropriate action, such as avoidance or retreat. Support for a potentially adaptive perceptual response to up-ramps of intensity and implied looming auditory motion has also been found in infant studies, where 4- to 6-month-old infants exhibit a significantly greater number of defensive behaviors (defined by backward movement pressure) in response to up-ramps than to down-ramps (Freiberg, Tually, & Crassini, 2001). Furthermore, up-ramps are associated with a "looming-specific" neural network that subserves auditory motion perception, space recognition, and attention (Bach, Neuhoff, Perrig, & Seifritz, 2009; Bach et al., 2008; Hall & Moore, 2003; Seifritz et al., 2002). They also elicit emotional responses that include heightened subjective and physiological arousal, unpleasantness, and perceived threat (Bach et al., 2009; Olsen & Stevens, 2013; Tajadura-Jiménez, Väljamäe, Asutay, & Västfjäll, 2010).

Therefore, if an up-ramp of intensity associated with real and implied looming auditory motion elicits potentially adaptive responses, it would be expected that in the context of sound source localization, up-ramps would be perceptually prioritized and therefore localized faster and more accurately than down-ramps, which are associated with a relatively nonthreatening receding sound source. In the present study, we investigated this hypothesis by extrapolating Neuhoff's (1998, 2001) "adaptive perceptual bias" for up-ramp intensity theory and testing it in the context of auditory spatial localization—specifically, sound source localization speed and accuracy in response to

acoustic stimuli presented in azimuth to imply looming, stationary, and receding motion in depth.

Included in the "adaptive perceptual bias" theory is an interaction between the direction of intensity change and acoustic spectrum. In the original global-loudness-change experiment (Neuhoff, 1998), 1-kHz pure-tone, tonal-vowel /ə/, (which sounds like the "a" in "about"), and white-noise stimuli were presented to participants with either 1.8-s up-ramp or 1.8-s down-ramp intensity profiles. The results indicated that the loudness change was greater in response to tonal up-ramps than to down-ramps (i.e., the pure-tone and vowel conditions), but not for white noise. Neuhoff argued that similarities in the spectral structures between the tonal stimuli in the experiment and single and potentially threatening biological sources in the environment could explain these findings (Neuhoff, 1998, 2001). In contrast, continuous broadband signals such as white noise most often reflect multiple co-occurring sound sources, such as crowd noise, the ocean, and rain (Neuhoff, 1998), all of which do not typically pose an immediate threat.

However, the assertion that single and potentially threatening biological sources in the environment are commonly characterized by tonal spectra has yet to be substantiated. Dispersed sound sources such as wind sweeping through trees or the crash of waves are spectrally similar to white noise, but a fast-approaching predatory animal roaring while in motion across the forest floor will also likely contain spectral information that is more similar to noise than to tones. Nevertheless, in the present study we investigated the interaction between the direction of intensity change associated with auditory motion and the spectrum of each ramped stimulus by measuring sound source localization speed and accuracy in response to 1-kHz pure-tone, vowel (/ə/), and white-noise stimuli. The conjecture that tonal stimuli are "perceptually salient" relative to white noise would receive support if the pure-tone and vowel stimuli were localized faster and more accurately than white noise.

Aim, design, and hypotheses

The present experiment was designed to investigate sound source localization speed and accuracy in response to variations of acoustic intensity and spectrum, with specific focus on (1) the hypothesized perceptual bias for up-ramp intensity associated with looming auditory motion; and (2) whether sound source localization is differentially affected by tonal stimuli (e.g., pure tones and vowels) relative to white noise. A secondary focus was to investigate the effects of sound source location in general. Because it has long been established that localization accuracy decreases toward a listeners' periphery (e.g., Mills, 1958; Oldfield & Parker, 1984), we expected to find similar results here. All participants were presented with three levels of acoustic spectrum, which included a 1-kHz pure tone,

a vowel (/ə/), and white noise. The intensities of these three stimuli were manipulated to continuously increase (up-ramp, looming), decrease (down-ramp, receding), or remain at a stationary, steady-state level. Therefore, the experiment was realized as a 3 × 3 within-subjects design. It was conducted under free-field listening conditions in an anechoic chamber, using a nine-loudspeaker array presented in a 180° arc on the frontal horizontal plane. The dependent variables were the speed and accuracy of participants' sound source localization responses. We hypothesized that:

1. Up-ramps of intensity would elicit faster and more accurate sound source localization responses than down-ramps of intensity;
2. Tonal stimuli (i.e., pure tones and vowels) would elicit faster and more accurate sound source localization responses than white noise.

Method

Participants

Twenty-six participants (20 females and six males) from Western Sydney University undertook the experiment. Their ages ranged from 18 to 39 ($M = 21.73$ years, $SD = 4.75$), and all reported normal hearing capabilities. Participants received course credit for their participation.

Stimuli and equipment

All stimuli were designed with intensity profiles that implied looming, receding, or stationary motion from identical starting points. This design was utilized to ensure that all stimulus onsets were controlled at an equivalent intensity of 40 decibels (dB). Up-ramps and down-ramps were then created with A-weighted intensity ranges of 40–60 dB for up-ramps (implying a looming stimulus from its 40-dB onset), and 40–20 dB for down-ramps (implying a receding stimulus from its 40-dB onset). A 40-dB steady-state intensity stimulus was also constructed that implied a stationary stimulus.

To create the intensity manipulations, steady-state versions of 1-kHz pure-tone, vowel, and white-noise stimuli were first generated. The initial vowel stimulus was a 1-s steady-state synthetic vowel (/ə/) generated from a Klatt synthesizer (Klatt, 1980) using the default sampling frequency of 8 kHz. The fundamental frequency of the vowel stimulus was 130.81 Hz. The initial 1-s white-noise and 1-s 1-kHz pure-tone steady-state stimuli were generated in Audacity (Version 2.0.2). The frequency spectra of all three steady-state stimuli are shown in Fig. 1. Up-ramps and down-ramps were then constructed from these initial steady-state exemplars in an

anechoic chamber using a custom computer program written in MAX-MSP (version 4.6.3). Minimum and maximum levels for all intensity manipulations were presented from the front speaker and measured from the point in space representing the middle of a participant's head. Intensity measurements were made with a Brüel & Kjær Hand-Held Analyser 2250 using Sound Level Meter Software BZ-7222 and recorded in the MAX-MSP program. The MAX-MSP program generated an up-ramp or down-ramp by using the minimum and maximum recorded intensity levels as the onset/offset anchors and creating a 1-s linear change of intensity between them. A stimulus duration of 1 s was chosen from the findings of a pilot experiment, where on average five participants responded faster than 1 s for all conditions (range = 614–898 ms). All stimuli were imported into Audacity and 10-ms fade-in and fade-out ramps were incorporated to remove any onset/offset clicks.

The experiment was conducted in the MARCS Institute anechoic chamber (depth = 3.88 m, width = 3.01 m, height = 2.85 m) with stimuli presented through nine Genelec 8020B speakers using an RME HDSPe RayDAT. Two Behringer Ultragain Digital ADA8200 preamplifiers were also used. The distance between the front of each speaker and the point in space signifying the center of a participant's head was 1.25 m. Nine speakers were used in a 180° arc on the frontal horizontal plane, numbered 1–9 from left to right, with Speaker 5 corresponding to the speaker directly in front of the listener. An arc of 22.5° separated the middle of one speaker from the middle of each immediately neighboring speaker. The distance from the floor to the middle of each speaker was 1.15 m. A 27-in. Diamond View LED monitor was positioned 50 cm above and 30 cm behind Speaker 5. Participants made responses using a computer mouse placed on a Logitec Comfort Lapdesk (N500) that was covered with 1-cm-thick foam to reduce unwanted acoustic reflections from the lapdesk. All stimuli and the experiment protocol were presented using MATLAB and the playrec (portaudio) function for audio playback.

Procedure

Participants were informed that the task consisted of localizing an auditory stimulus from one of the nine speakers as fast and as accurately as possible. Therefore, the task was a "categorical" localization task (Letowski & Letowski, 2012).[1] There were nine possible conditions in the experiment (3 × intensity

---

[1] Although categorical localization tasks are valid tools with which to investigate spatial hearing (Letowski & Letowski, 2012), they do have limitations. For example, one cannot know for sure the exact perceived sound source location (or location error) to the nth degree, which can otherwise be measured using an absolute localization task. Furthermore, inherent response choice constraints can influence the measurements in a categorical localization task, especially in terms of cross-plane errors and front–back errors (Perrett & Noble, 1995), both of which were not applicable to the present study.

Fig. 1 Frequency spectrograms for each steady-state (40-dB) stimulus: 1-kHz pure tone (top left panel), vowel /ə / (top right panel), and white noise (bottom panel)
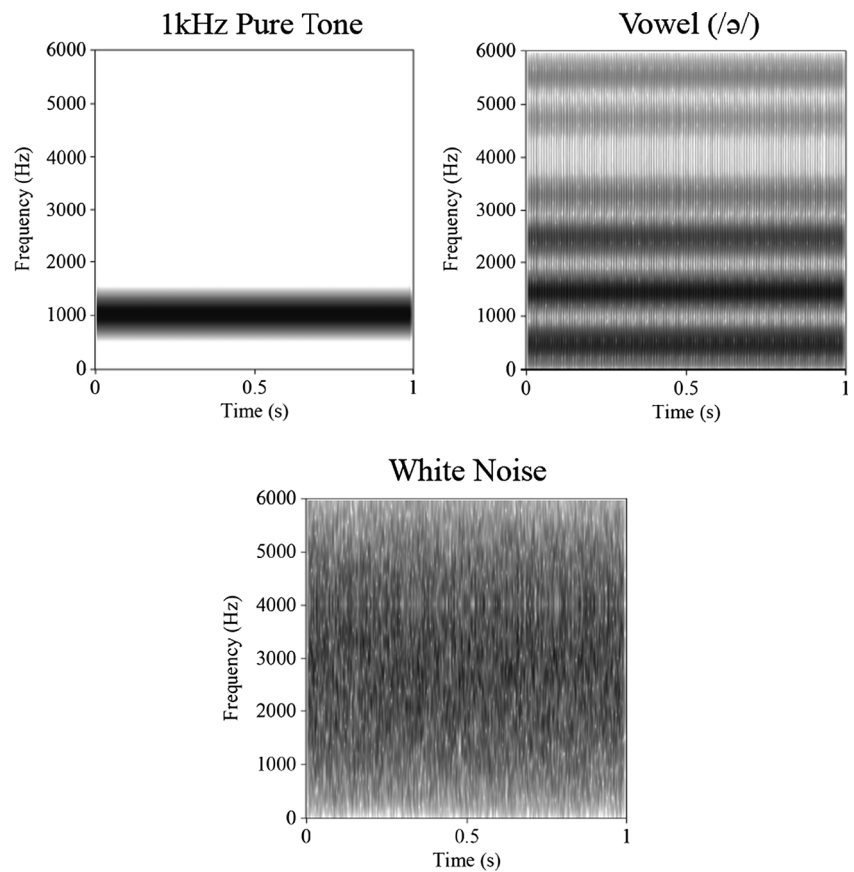
and 3 × spectrum) and 36 blocks of trials in total, split into two halves of 18 blocks with a break in-between. Each block comprised nine trials, with each trial randomly presented at each speaker location only once—hence, nine trials per block, corresponding to the nine speaker locations. Only one condition was presented within a block. Over the course of the experiment, four blocks of each condition were presented, all in a random order for each participant.

Participants were instructed to focus on a "+" symbol on the computer monitor throughout each trial and not to move their head at any time (an instruction to remain immobile during localization tasks is effective for the maintenance of posture; Noble, 1981). To begin a trial, participants were required to click the "play" button on the computer screen. Once they had localized the sound source, they were required to click the identical button (now displaying "stop" instead of "play"), and in doing so, the stimulus immediately stopped. The duration between stimulus onset and when a participant clicked "stop" constituted the localization response time measure. Participants were explicitly told not to respond merely to the onset of a sound, but rather to respond when they were confident of the sound source location. After clicking "stop," participants were required to move the mouse cursor on the monitor and to click one of nine schematically illustrated

speakers to confirm from which location they had perceived the sound to originate. This constituted the localization accuracy response. After participants chose the location of the sound source, they were required to click the "play" button again to hear the next trial. This process occurred until all nine randomly allocated trials in each block had been presented. Participants were not explicitly informed of the number of trials in each block. To ensure that participants could not anticipate the onset of a stimulus in each trial, a randomized delay between 1 and 2 s, measured from the start of the trial to the onset of the stimulus, was implemented throughout the experiment.

Two practice blocks used a 2-s, 40-dB, 1-kHz sawtooth tone stimulus. In the first practice block, each trial was presented from left to right across all speakers sequentially, so that the participant could become accustomed to each speaker location. In the second block, the presentation of each trial was randomized across all nine locations, so that the participant could become accustomed to the main experiment's randomization protocol. At the end of the first half of the main experiment (18 blocks), each participant was asked to step out of the anechoic chamber to have a break and complete a demographic questionnaire that took approximately 2–3 min. Once the questionnaire was completed, the participant reentered the

chamber to complete the second half of the experiment. Overall, 36 blocks and 324 trials were presented throughout the experiment, comprising four blocks of each of the nine conditions (not including practice trials). The experiment took approximately 50 min.

Results

Localization accuracy

The group mean localization accuracy (% correct) results indicated a significant main effect of intensity, $F(2, 50) = 6.20$, $p < .01$, $\eta_p^2 = .20$, and a significant main effect of spectrum, $F(2, 50) = 23.67$, $p < .001$, $\eta_p^2 = .49$. We found no significant interaction between the intensity and spectrum conditions, $F(4, 100) = 0.18$, $p > .05$, $\eta_p^2 = .01$. Bonferroni-corrected post-hoc comparisons were therefore conducted to investigate the specific differences within each main effect.

First, it was first hypothesized that up-ramps of intensity would elicit more-accurate localization responses than down-ramps of intensity. This hypothesis was supported. As can be seen in Fig. 2 (top panel), localization accuracy was significantly greater in response to up-ramps of intensity ($M = 87.10$ %, $SE = 1.90$) than to down-ramps ($M = 84.00$ %, $SE = 1.80$), $p < .01$, 95 % CI [.007, .054]. The results also revealed that the localization accuracy in response to steady-state intensity ($M = 86.50$ %, $SE = 1.60$) was significantly greater than that for down-ramps, $p < .05$, 95 % CI [.004, .045], but was statistically equivalent to up-ramps, $p > .05$, 95 % CI [–.033, .021].

Second, we hypothesized that tonal stimuli (pure tones and vowels) would elicit more-accurate localization responses than white noise. This hypothesis was not supported. As can be seen in Fig. 2 (bottom panel), localization accuracy was significantly greater in response to white noise ($M = 90.70$ %, $SE = 1.60$) than to the vowel ($M = 85.20$ %, $SE = 2.10$), $p < .01$, 95 % CI [.021, .089] and pure-tone ($M = 81.60$ %, $SE = 1.80$) conditions, $p < .001$, 95 % CI [.052, .131]. Finally, sound source localization accuracy was significantly greater in response to the vowel condition than to the pure-tone condition, $p < .05$, 95 % CI [.007, .066].

Localization response time

Sound source localization response times were analyzed by including (a) all response time data and (b) the response time data corresponding to correct localization responses only. Figure 3 presents descriptive statistics for each type of
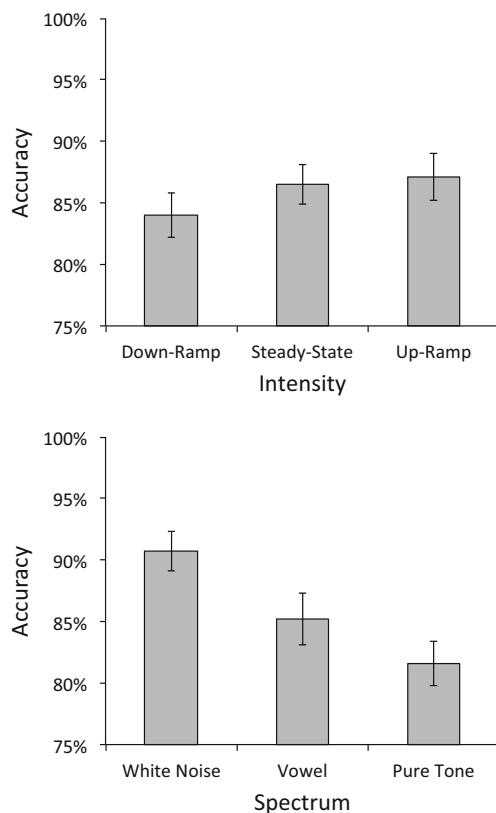

Fig. 2 Group mean sound source localization accuracy (% correct) in response to intensity conditions (top panel) and spectrum conditions (bottom panel). Error bars represent standard errors of the means
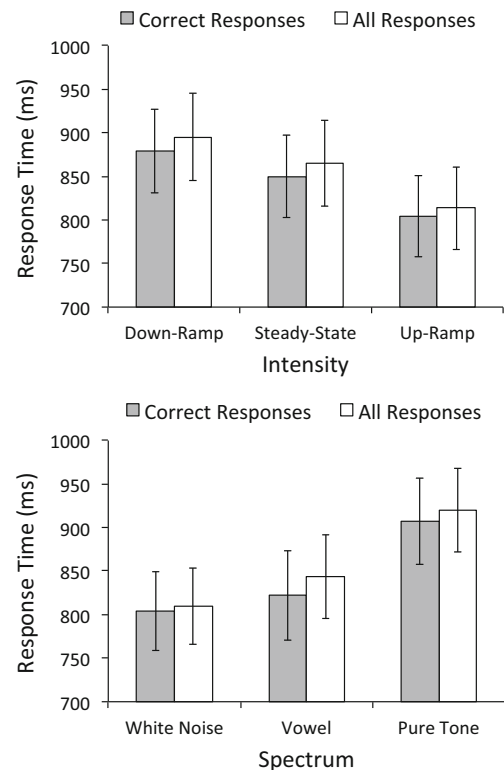

Fig. 3 Group mean sound source localization response times to intensity conditions (top panel) and spectrum conditions (bottom panel). Response times are presented for all localization responses (white bars) and for correct localization responses only (gray bars). Error bars represent standard errors of the means

analysis, and similar trends in the results can be observed. However, response times were overall faster when correct localization responses were made. This suggests no speed–accuracy trade-off. Two separate analyses of variance (ANOVAs) including Bonferroni-adjusted post-hoc comparisons were conducted on group mean response times for all responses and also for correct responses only. These analyses revealed equivalent statistical significance (or nonsignificance) for the main effects, interactions, and post-hoc comparisons. Therefore for brevity, the statistical analyses below are reported only for "all responses."

The group mean response time results indicated a significant main effect of intensity, $F(2, 50) = 10.75$, $p < .001$, $\eta_p^2 = .30$, and a significant main effect of spectrum, $F(2, 50) = 33.09$, $p < .001$, $\eta_p^2 = .57$. We found no significant interaction between the intensity and spectrum conditions, $F(4, 100) = 0.55$, $p > .05$, $\eta_p^2 = .02$. Bonferroni-corrected post-hoc comparisons were therefore conducted to investigate each main effect.

First, it was hypothesized that up-ramps of intensity would elicit faster sound source localization responses than would down-ramps of intensity. This hypothesis was supported. The sound source localization response time to up-ramps of intensity ($M = 813.47$ ms, $SE = 47.80$) was significantly faster than that to down-ramps ($M = 894.66$ ms, $SE = 50.19$), $p < .01$,

95 % CI [–130.57, –31.81]. These comparisons also showed that the localization response time to steady-state intensity ($M = 865.47$ ms, $SE = 49.55$) was significantly slower than up-ramps, $p < .05$, 95 % CI [5.00, 99.00], but was statistically equivalent to down-ramps, $p > .05$, 95 % CI [–68.80, 10.42].

Second, we hypothesized that tonal stimuli (pure tones and vowels) would elicit faster localization responses than white noise. Contrary to the hypothesis, mean response time to pure tones ($M = 920.06$ ms, $SE = 49.31$) was significantly slower than to vowels ($M = 843.41$ ms, $SE = 51.13$), $p < .001$, 95 % CI [37.52, 115.79], and white noise ($M = 810.12$ ms, $SE = 45.71$), $p < .001$, 95 % CI [76.96, 142.92]. The sound source localization response times were statistically similar between the white-noise and vowel conditions, $p > .05$, 95 % CI [–67.57, 1.01].

Location-specific accuracy and response time

Although the primary purpose of the present study was to use sound source localization as a tool to investigate intensity change and spectrum in the context of spatial hearing, here we also present descriptive statistics for the accuracy and response time to each speaker location in the top and bottom panels of Fig. 4, respectively. Inspection of Fig. 4 shows overall symmetrical response patterns, with Speaker 5 (front

Fig. 4 Group mean accuracy (top panel) and response time (bottom panel) for each speaker location. Response times are presented for all localization responses (white bars) and for correct localization responses only (gray bars). Speakers were separated by 22.5° and were presented in a 180° arc on the frontal horizontal plane. Speaker 1 was placed on the far left of the listener, and Speaker 9 was placed on the far right. Speaker 5 was placed directly in front of the listener. Error bars represent standard errors of the means
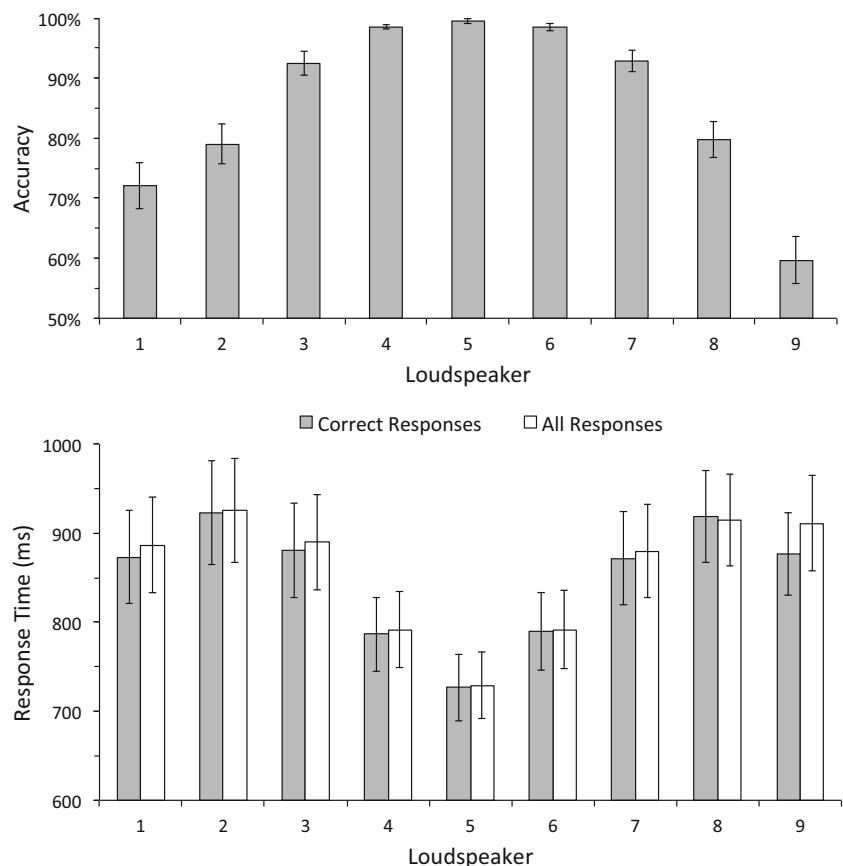
speaker) receiving the fastest and most accurate localization responses. Response times were consistently slower toward peripheral speaker locations, with a response time peak at Speakers 2 and 8. The accuracy data show a reciprocal pattern, with deteriorating sound source localization accuracy for speakers at more peripheral locations. Interestingly, there was a slight accuracy difference between Speakers 1 and 9. Nevertheless, localization accuracy consistently decreased toward the periphery on both the left and right sides of the listener.

Finally, Fig. 5 displays bubble plots indicating the distribution of localization responses to each speaker location. Larger circles indicate a greater number of localization responses to that particular speaker location, with smaller circles indicating a smaller number of responses. For example, for all stimuli presented from Speaker 1, the largest number of localization responses were made for Speaker 1, with Speakers 2 and 3 also being selected by participants, but far less frequently. This figure also illustrates the distribution of localization errors for each speaker location. For example, inaccurate responses for Speakers 1 and 9 were drawn toward the center, whereas for Speakers 2 and 8, and 3 and 7, inaccurate responses were made on either side of the speakers. Consistent with the results presented in Fig. 4, the greatest numbers of accurate responses were made at spatial locations in front of the participant.

Discussion

The aim of the present study was to investigate the possibility of a "looming bias" in spatial hearing, and more specifically, the effects of acoustic intensity and spectrum on sound source localization speed and accuracy. The results provided support for the first hypothesis, with up-ramps of intensity eliciting

significantly faster and more accurate sound source localization responses than did down-ramps of intensity. The results from the spectrum conditions did not support the second hypothesis, that tonal stimuli (i.e., pure tones and vowels) would elicit significantly faster and more accurate sound source localization responses than white noise. On the contrary, sound source localization was faster and more accurate for the spectrally complex white-noise and vowel stimuli, relative to the 1-kHz pure tone.

Effects of acoustic intensity on sound source localization

The results from the present study are consistent with prior studies that have demonstrated perceptual prioritization of stimuli on the basis of acoustic features associated with auditory motion in the environment (e.g., Bach et al., 2009; Freiberg et al., 2001; Neuhoff, 1998, 2001; Rosenblum et al., 1993; Schiff & Oldak, 1990; Seifritz et al., 2002). Perceptual prioritization of up-ramp acoustic intensity, in particular, was first demonstrated by Neuhoff (1998, 2001), who argued that up-ramps of intensity imply looming auditory motion and elicit an adaptive perceptual bias. Such a bias would provide an evolutionary advantage for organisms able to quickly and appropriately respond to looming and potentially threatening sound sources. In the context of spatial hearing, the present findings show that an up-ramp of intensity is localized faster and more accurately than a down-ramp when both dynamic stimuli originate from an equivalent point in implied space (i.e., the stimulus onsets were equivalent at 40 dB, with subsequent intensity profiles that implied looming, receding, or stationary motion over a 1-s duration and 20-dB range). Therefore, participants' tendency to localize an up-ramp stimulus faster than a down-ramp stimulus may provide some evidence of a perceptual response facilitating extra time to prepare for the arrival of a looming source.

Fig. 5 Bubble plots indicating the frequencies of speaker location responses (*y*-axis) versus the actual speaker that had presented the stimulus (*x*-axis). The sizes of bubbles are proportional to the number of localization responses to each speaker, with larger bubbles representing a greater frequency of responses to that speaker location
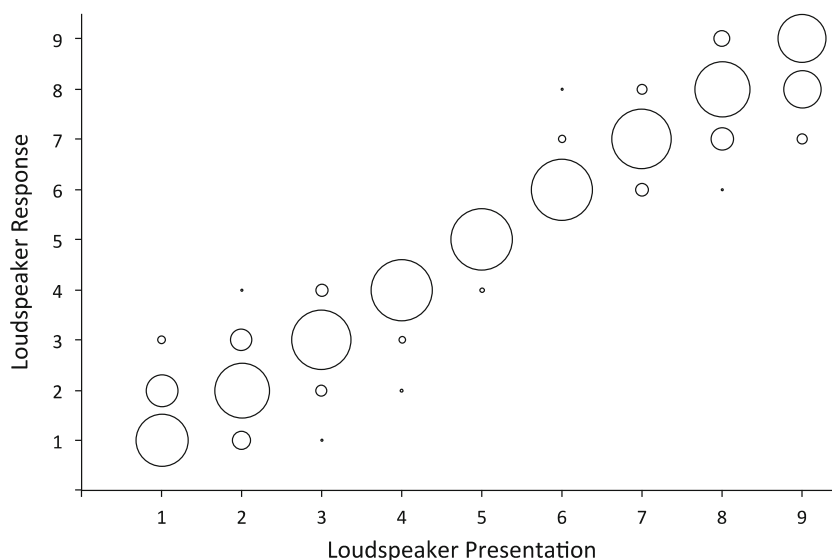
The difference in the localization response times between up-ramps and down-ramps reported here was on average 82 ms, a relatively small amount of extra time for a quick and appropriate adaptive response to a looming and potentially threatening event. However, participants made numerous repeated responses (324 trials) over 50 min, and they may have habituated to the stimuli and differences between the intensity conditions over time. A single critical response in a natural environment might show an even greater advantage for looming stimuli. Furthermore, in the context of loudness change, Neuhoff (1998) observed an interaction between the region of intensity change and the magnitude of the perceptual disparity between up-ramps and down-ramps. For example, a smaller difference in loudness change between up-ramps and down-ramps, and thus a smaller "bias for up-ramp intensity". Was observed when stimuli were presented in the region of 60–75 dB, relative to the region of 75–90 dB. In the present study, 20-dB ranges of intensity change were presented within the region of 20–60 dB. This mid-to-low region of intensity may explain why the differences in sound source localization response times between up-ramps and down-ramps were relatively small. However, it is not yet clear how the effects of intensity region in the context of loudness change measured in headphone listening conditions (e.g., Neuhoff, 1998) relate to sound source localization response times measured in free-field listening conditions.

Nevertheless, the difference in localization response times between up-ramps and down-ramps was statistically significant, and the question remains: How does the magnitude of this difference vary as a function of the range of intensity change within each ramp and the region of intensity change over which ramps are presented? This question could be investigated by presenting different magnitudes of intensity change beyond the 20-dB range and 20- to 60-dB region used here. Varying the range of intensity change within a ramped stimulus would intrinsically vary the rate of intensity change over a ramp's 1-s duration. The rate of intensity change per unit time is a key indicator of the perceived velocity of motion from a sound source (Carlile & Leung, 2016) and has been shown to influence perceptual asymmetries in both subjective duration and loudness change (Meunier et al., 2014; Olsen & Herff, 2015). Furthermore, the linear change of intensity used in the present study indicated a decelerating approaching object in the environment, rather than a constant-velocity approach (Neuhoff, 2001). Although some evidence does suggest that linear and accelerated rates of change are not discriminated well (Andreeva & Vartanyan, 1997), future work will be required in order to investigate sound source localization speed and accuracy in response to acoustic stimuli that imply accelerating, decelerating, and constant-velocity looming and receding sound sources. Finally, the up-ramps and down-ramps in the present study implied looming and receding motion from identical starting points. Future studies could also present such stimuli with identical endpoints. However, the stimulus onset intensities would vary dramatically between up-ramps and down-ramps in such a condition, and could disproportionally bias localization judgments on the basis of stimulus onset characteristics (i.e., large differences in loudness) rather than the dynamic acoustic properties of a stimulus that change over time.

Effects of acoustic spectrum on sound source localization

An influence of acoustic spectrum was also investigated here in the context of spatial hearing. Neuhoff (1998) theorized that single and potentially threatening biological sound sources in the environment share a similar harmonic structure with tonal stimuli, relative to white noise, which is more similar to non-threatening, multiple co-occurring sources in the environment (e.g., wind or the ocean). Therefore, tonal stimuli should be perceptually prioritized to a greater degree than white noise. If one extrapolates this rationale, the tonal stimuli in the present study should have elicited faster and more accurate localization responses than white noise. However, our results did not support this hypothesis. On the contrary, sound source localization was faster and more accurate for white-noise and vowel stimuli, relative to the 1-kHz pure tone. These findings provide evidence that an association between tonal stimuli/ single threatening sound sources and noise stimuli/multiple nonthreatening sound sources is not sufficient to explain the perceptual differences underlying a "perceptual bias" (Neuhoff, 1998) in the context of spatial hearing.

Rather, spectrally complex sounds have been shown to significantly improve localization performance (Morikawa & Hirahara, 2013; Morikawa, Toyoda, & Hirahara, 2013). White noise and speech are spectrally complex sounds, whereas a pure tone is not, and in the present study the relationships between spectral complexity and interaural intensity differences (IID), interaural time differences (ITD), and perceived loudness provide a likely explanation. Interaural differences arise from the spatial separation between the two ears. Specifically, the acoustic shadow of the head results in a lower intensity at the ear farthest away from the source (IID) and a time delay between the nearest ear and the ear farthest away from the source (ITD). IID is the dominant localization cue for middle-to-high frequencies, whereas ITD is the dominant localization cue for low frequencies (Dobreva, O'Neill, & Paige, 2011; Letowski & Letowski, 2012; Middlebrooks & Green, 1991). The complementary effects of these cues when used in combination can facilitate judgments of auditory spatial localization on the horizontal plane (Stevens & Newman, 1936). Therefore, in the present study, the spectrally rich and complex frequency bandwidths of white-noise and vowel stimuli most likely enhanced spatial localization performance on the horizontal plane, because interaural cues were maximized relative to the less complex 1-kHz pure tone.

Localization performance in response to the three levels of spectrum could also have been impacted by differences in perceptual loudness. For example, the large amounts of spectral information in white noise and vowels could have resulted in greater magnitudes of perceived loudness than in pure tones, even though the intensity measurements were identical across all three spectrum conditions. Evidence of such a difference in loudness has come from investigations of critical bands (Scharf, 1961; Zwicker, Flottorp, & Stevens, 1957). Critical bands are areas on the basilar membrane that comprise a specific frequency range (or bandwidth). The perceived loudness of a stimulus remains constant until the frequency bandwidth of the stimulus grows beyond the size of a critical band, after which loudness summates (Zwicker et al., 1957). Stimuli with larger frequency bandwidths, such as white noise and vowels, excite a larger number of critical bands than do stimuli with restricted bandwidths, such as pure tones, and therefore are likely to be perceived as louder than pure tones. In the present study, the influence of critical bands on loudness may explain to some extent why differences in sound source localization speed and accuracy were observed between the spectrum conditions: In everyday listening circumstances, loud sound sources are generally closer in proximity than "soft" sound sources, and thus demand a more urgent behavioral response.

Sound source localization in depth and in azimuth

The present study presented stimuli in azimuth that implied looming, receding, or stationary auditory motion in depth. Neuhoff's (1998) "adaptive perceptual bias" for up-ramp intensity was proposed in the context of looming sounds in depth, rather than in azimuth. This discrepancy may somewhat reconcile the differences in the spectrum results between those observed here and those reported in Neuhoff (2001), where real motion in depth was investigated. The finding that noise is localized more quickly and accurately in azimuth does not necessarily mean that looming tones are not "perceptually salient." In fact, accuracy may not necessarily be the best measure of perceptual salience in the context of real auditory motion in depth. For example, Neuhoff (2001) argued that a looming tonal sound source that physically moves in depth is perceptually salient because it is perceived to be closer to the listener, relative to the actual position of the sound source in space (an incorrect localization response in depth). Looming noise does not elicit such an effect, thus representing a more accurate localization response in depth. Therefore, greater localization accuracy for noise in azimuth in the present study does not necessarily fail to support the findings in Neuhoff (2001) of greater perceptual salience of looming tones over noise in depth. However, the accuracy data in the present study support the conclusion that the location in azimuth of

noise stimuli with implied motion in depth is perceived more accurately than the location of tonal stimuli.

Attentional capture and visual looming

The main findings presented here are consistent with the behavioral-urgency hypothesis from the visual-looming literature (e.g., Franconeri & Simons, 2003; von Mühlenen & Lleras, 2007). According to the behavioral-urgency hypothesis, looming stimuli capture attention quickly and efficiently because they signal events that may require a "behaviorally urgent" response. Although the present study did not explicitly measure the magnitude of attentional capture in response to up-ramps and down-ramps, indirect support for the behavioral-urgency hypothesis was observed, with up-ramps eliciting faster and more accurate localization responses than down-ramps. In this framework, it is likely that up-ramps captured attention more quickly than down-ramps, resulting in a behaviorally urgent response (fast and accurate sound source localization). This is an area for future research. Specifically, attentional-cueing paradigms commonly implemented in the visual domain could identify the magnitude of attentional capture demanded by variations of acoustic intensity and implied motion in the auditory domain. Such a design could also include audiovisual motion for greater ecological validity.

Effects of sound source spatial location

Stimuli in the present study were presented from nine speakers in a 180° arc around the participant's frontal horizontal plane. The response times and accuracy for different spatial locations displayed symmetrical response patterns, with Speaker 5 (front) receiving the fastest and most accurate localization responses. As stimuli were presented from more peripheral locations, participants' responses generally became consistently slower and less accurate. The descriptive statistics reported in Figs. 4 and 5 reveal that Speaker 2 and Speaker 8 received the largest range of inaccurate responses, in comparison to all other speakers. This suggests a greater magnitude of spatial ambiguity when stimuli were presented from these locations. The absence of head movement would have restricted participants' ability to resolve this spatial ambiguity (Letowski & Letowski, 2012), resulting in slower localization response times. While the accuracy for Speakers 2 and 8 was generally poor, the accuracy for Speakers 1 and 9 was worse. However, these speakers received faster response times than Speakers 2 and 8, which could explain their decreased accuracy in comparison. Nevertheless, the response time and accuracy findings reported in the present study are largely concordant with previous research on auditory localization, where performance was best at frontal locations and deteriorated as stimuli were presented from more peripheral locations (e.g., Mills, 1958; Oldfield & Parker, 1984).

Conclusion

The present study has contributed to existing knowledge by extrapolating the basic premise of Neuhoff's (1998, 2001) "adaptive perceptual bias" for up-ramp intensity into the domain of spatial auditory localization. Specifically, sound source localization speed and accuracy was measured in response to acoustic stimuli presented in azimuth to imply looming, stationary, and receding motion in depth. The results show that up-ramps of intensity, which are associated with looming auditory motion, are localized faster and more accurately than down-ramps of intensity, which are associated with receding auditory motion. Thus, it seems as if there may be a "looming bias" in spatial hearing when auditory looming is implied by a continuous up-ramp of intensity. However, the results from three levels of acoustic spectra did not support the hypothesized perceptual salience of tonal stimuli over white noise (Neuhoff, 1998). Rather, the richness and complexity of spectral information available to listeners is key for fast and accurate sound source localization.

References

Andreeva, I., & Vartanyan, I. (1997). The effect of the temporal integration of loudness on the auditory perception of sound source depth dynamics. *Human Physiology, 23,* 645–651.

Bach, D. R., Neuhoff, J. G., Perrig, W., & Seifritz, E. (2009). Looming sounds as warning signals: The function of motion cues. *International Journal of Psychophysiology, 74,* 28–33. doi:10.1016/j.ijpsycho.2009.06.004

Bach, D. R., Schachinger, H., Neuhoff, J. G., Esposito, F., Di Salle, F., Lehmann, C., … Seifritz, E. (2008). Rising sound intensity: An intrinsic warning cue activating the amygdala. *Cerebral Cortex, 18*, 145–150.

Carlile, S., & Leung, J. (2016). The perception of auditory motion. *Trends in Hearing, 20,* 1–19.

DiGiovanni, J. J., & Schlauch, R. S. (2007). Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds. *Ecological Psychology, 19,* 239–264.

Dobreva, M. S., O'Neill, W. E., & Paige, G. D. (2011). Influence of aging on human sound localization. *Journal of Neurophysiology, 105,* 2471–2486.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics, 65,* 999–1010. doi:10.3758/BF03194829

Freiberg, K., Tually, K., & Crassini, B. (2001). Use of an auditory looming task to test infants' sensitivity to sound pressure level as an auditory distance cue. *British Journal of Developmental Psychology, 19,* 1–10.

Grassi, M., & Darwin, C. J. (2006). The subjective duration of ramped and damped sounds. *Perception & Psychophysics, 68,* 1382–1392. doi:10.3758/BF03193737

Hall, D. A., & Moore, D. R. (2003). Auditory neuroscience: The salience of looming sounds. *Current Biology, 13,* R91–R93.

Jenison, R. L. (1997). On acoustic information for motion. *Ecological Psychology, 9,* 131–151.

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America, 67,* 971–995.

Letowski, T. R., & Letowski, S. T. (2012). *Auditory spatial perception: Auditory localization* (US Army Research Laboratory Report, pp. 1–162). Retrieved from www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA563540

Meunier, S., Vannier, M., Chatron, J., & Susini, P. (2014). Asymmetry in perceived duration between up-ramp and down-ramp sounds as a function of duration. *Journal of the Acoustical Society of America, 136,* EL166–EL172.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology, 42,* 135–159. doi:10.1146/annurev.ps.42.020191.001031

Mills, A. W. (1958). On the minimum audible angle. *Journal of the Acoustical Society of America, 30,* 237–246.

Morikawa, D., & Hirahara, T. (2013). Effect of head rotation on horizontal and median sound localization of band-limited noise. *Acoustical Science and Technology, 34,* 56–58.

Morikawa, D., Toyoda, Y., & Hirahara, T. (2013). *Head movement during horizontal and median sound localization experiments in which head-rotation is allowed.* Paper presented at the 21st International Congress on Acoustics, Acoustical Society of America, Montreal, Canada.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature, 395,* 123–124.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology, 13,* 87–110.

Noble, W. (1981). Earmuffs, exploratory head movements, and horizontal and vertical sound localization. *Journal of Auditory Research, 21,* 1–12.

Oldfield, S. R., & Parker, S. P. (1984). Acuity of sound localisation: A topography of auditory space. I: Normal hearing conditions. *Perception, 13,* 581–600.

Olsen, K. N. (2014). Intensity dynamics and loudness change: A review of methods and perceptual processes. *Acoustics Australia, 42,* 159–165.

Olsen, K. N., & Herff, S. A. (2015). The effects of acoustic intensity, spectrum, and duration on global loudness change. *Acta Acustica United With Acustica, 111,* 1094–1104.

Olsen, K. N., & Stevens, C. J. (2010). Perceptual overestimation of rising intensity: Is stimulus continuity necessary? *Perception, 39,* 695–704.

Olsen, K. N., & Stevens, C. J. (2013). Psychophysiological response to acoustic intensity change in a musical chord. *Journal of Psychophysiology, 27,* 16–26.

Olsen, K. N., Stevens, C. J., & Tardieu, J. (2010). Loudness change in response to dynamic acoustic intensity. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1631–1644.

Perrett, S., & Noble, W. (1995). Available response choices affect localization of sound. *Perception & Psychophysics, 57,* 150–158.

Ponsot, E., Meunier, S., Kacem, A., Chatron, J., & Susini, P. (2015). Are rising sounds always louder? Influences of spectral structure and intensity-region on loudness sensitivity to intensity-change direction. *Acta Acustica united with Acustica, 101,* 1083–1093.

Ponsot, E., Susini, P., & Meunier, S. (2015). A robust asymmetry in loudness between rising- and falling-intensity tones. *Attention, Perception, & Psychophysics, 77,* 907–920. doi:10.3758/s13414-014-0824-y

Rosenblum, L. D., Wuestefeld, A. P., & Saldana, H. M. (1993). Auditory looming perception: Influences on anticipatory judgments. *Perception, 22,* 1467–1482.

Scharf, B. (1961). Complex sounds and critical bands. *Psychological Bulletin, 58,* 205–217. doi:10.1037/h0049235

Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 303–316. doi:10.1037/0096-1523.16.2.303

Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schachinger, H., … Di Salle, F. (2002). Neural processing of auditory looming in the human brain. *Current Biology, 12,* 2147–2151.

Stecker, G. C., & Hafter, E. R. (2000). An effect of temporal asymmetry on loudness. *Journal of the Acoustical Society of America, 107,* 3358–3368.

Stevens, S. S., & Newman, E. B. (1936). The localization of actual sources of sound. *The American journal of psychology, 48,* 297–306.

Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., & Västfjäll, D. (2010). Embodied auditory perception: The emotional impact of approaching and receding sound sources. *Emotion, 10,* 216–229.

Teghtsoonian, R., Teghtsoonian, M., & Canévet, G. (2005). Sweep-induced acceleration in loudness change and the "bias for rising intensities.". *Perception & Psychophysics, 67,* 699–712. doi:10.3758/BF03193526

von Mühlenen, A., & Lleras, A. (2007). No-onset looming motion guides spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1297–1310. doi:10.1037/0096-1523.33.6.1297

Zwicker, E., Flottorp, G., & Stevens, S. S. (1957). Critical band width in loudness summation. *Journal of the Acoustical Society of America, 29,* 548–557.