Effects of forward masking on sound localization in cats: basic findings with broadband maskers

Yan Gai, Janet L. Ruhland, and Tom C. T. Yin

Department of Neuroscience, University of Wisconsin, Madison, Wisconsin

Submitted 10 April 2013; accepted in final form 9 July 2013

Gai Y, Ruhland JL, Yin TC. Effects of forward masking on sound localization in cats: basic findings with broadband maskers. J Neurophysiol 110: 1600–1610, 2013. First published July 10, 2013; doi:10.1152/jn.00255.2013.—Forward masking is traditionally measured with a detection task in which the addition of a preceding masking sound results in an increased signal-detection threshold. Little is known about the influence of forward masking on localization of free-field sound for human or animal subjects. Here we recorded gaze shifts of two head-unrestrained cats during localization using a search-coil technique. A broadband (BB) noise maskeder was presented straight ahead. A brief signal could come from 1 of the 17 speaker locations in the frontal hemifield. The signal was either a BB or a band-limited (BL) noise. For BB targets, the presence of the forward masker reduced localization accuracy at almost all target levels (20 to 80 dB SPL) along both horizontal and vertical dimensions. Temporal decay of masking was observed when a 15-ms interstimulus gap was added between the end of the masker and the beginning of the target. A large effect of forward masking was also observed for BL targets with low (0.2–2 kHz) and mid (2–7 kHz) frequencies, indicating that the interaural timing cue is susceptible to forward masking. Except at low sound levels, a small or little effect was observed for high-frequency (7–15 kHz) targets, indicating that the interaural level and the spectral cues in that frequency range remained relatively robust. Our findings suggest that different localization mechanisms can operate independently in a complex listening environment. Forward masking; free-field; ILD; ITD; localization

In addition, masking refers to the ability of an additional sound to interfere with the perception of a target sound. With regard to the timing of the masker and the target, there are three general classes of masking: simultaneous, forward, or backward. In a complex environment, listening in fluctuating noise can sometimes benefit from “dip listening.” That is, when the masker’s amplitude is momentarily small, a large signal-to-noise ratio (SNR) compared with the average can be achieved (Cooke 2006). However, we cannot always take advantage of this dip listening, because even when the noise is at its minimum, there can be temporal masking (forward or backward masking) coming from adjacent high amplitudes of noise. This type of masking is especially problematic for hearing-impaired listeners who typically show poor temporal resolution and a strong effect of forward masking (see review, Reed et al. 2009), which could be a reason for their degraded speech recognition in noise (Lutman 1991).

Forward masking has usually been studied by measuring the increased signal-detection threshold resulting from the addition of a preceding masking sound (Deatherage and Evans 1969; Kollmeier and Gilkey 1990; Oxenham and Plack 2000; Smiarowski and Carhart 1975; Waltzman and Levitt 1978). The masking effect declines with increased temporal gap between the masker and the signal. For example, adding a 10-ms gap to the end of a broadband masker can decrease the detection threshold of a subsequent click by 15 dB (Smiarowski and Carhart 1975). In real life, signals and noise often come from different locations. The ability to localize a signal from background noise enables the listener to attend to the sound source of interest, resulting in improved detection or perception of the sound (Garadat and Litovsky 2007; Ison and Agrawal 1998; Kidd et al. 1998; Nityananda and Bee 2012; Saberi et al. 1991). Human and animals’ localization performance under simultaneous masking has been measured in a number of studies (Good and Gilkey 1996; Lingner et al. 2012; Lorenzi et al. 1999). Physiological studies have also demonstrated masking effect of a preceding sound on neural responses to a following sound (e.g., Brosch and Schreiner 1997; Nelson et al. 2009; Relkin and Turner 1988; Singheiser et al. 2012; Smith et al. 1977). Comparatively, little is known about forward masking of free-field localization.

There are many studies on binaural release from forward masking under headphone conditions (Bekhterev et al. 2002; Deatherage and Evans 1969; Kollmeier and Gilkey 1990; Zwicker and Zwicker 1984). The detection threshold decreases when the signal changes from diotic (same for both ears; NoSo) to dichotic (anti-phase across ears; NoS π). This threshold decrease is usually called a binaural masking level difference (BMLD). Unfortunately, this finding cannot easily be generalized to the free-field condition. First, in most of those forward masking studies the signals are low-frequency tones with only the interaural time difference (ITD) cue included. Depending on the signal frequency, the value (Sπ: half of the tone cycle) often exceeds the subject’s physiological ITD range. In contrast, sound locations in free field are encoded by both interaural time and level (ILD) cues, as well as spectral cues created by head-related transfer functions (HRTFs) (see reviews, Tollin and Yin 2009; Yin 2002). The advantage of delivering sound through headphones is that individual cues can be strictly controlled. However, under free-field condition, “the combination of binaural cues and monaural, envelope-based detection cues may act synergistically to produce better performance than expected from the combination of independent cues” (Hall et al. 2006). In addition, attentional factors are more likely to play a role in the free-field condition acting against spatial release from forward masking. On the one hand, when the signal originates from a location far from the masker, the two sounds can be easily separated, resulting in a lower detection threshold. On the other hand, when the masker is presented first, listeners’ attention may be drawn to the masker’s
vicinity and thereby have a higher chance of missing a remote and brief signal. This issue may be less prominent when the masker and signal are presented concurrently.

The present study examined sound localization under forward masking in a free-field setup. The search-coil technique enables accurate measurement of gaze movements as an indication of perceived spatial locations of the sound source, rather than providing the subjects with a limited number of choices. A high spatial resolution is critical given the possibility that the effect of masking on localizing a loud target may be subtle. We aimed to answer two major questions in this study.

First, does forward masking affect localization mechanisms? The binaural release studies mentioned above suggest that, at near-threshold levels, localization mechanisms can be involved to facilitate signal detection. Yet when the signal is well detected, it is unclear whether or not localization alone can be affected.

Second, if forward masking disturbs localization mechanisms, are all the cues equally affected? According to the classical duplex theory (Mills 1972; Rayleigh 1907), the ITD cue is important for localization of low-frequency sound, whereas the ILD is important for localization of high-frequency sound. For a headphone study to examine the masking effect on spectral cues, HRTFs need to be measured and incorporated to the generation of stimuli. This may be one reason why binaural detection studies have been heavily focused on the ITD cue. Another reason is probably because ITD cues seem to be dominant for real-world broadband stimuli (Macpherson and Middlebrooks 2002; Wightman and Kistler 1992). We conducted this experiment with free-field stimuli where all three cues were combined using broadband stimuli or uli where all three cues were combined using broadband stimuli or.

The behavioral experiment was conducted in a dimly lit soundproof chamber (2.2 × 2.5 × 2.5 m; Fig. 1A). The eye coil was connected to a magnetic search coil system that enabled the vertical and horizontal components of eye movements to be monitored with a resolution of <1°. Speakers (model 40-1310B; RadioShack, Fort Worth, TX) and light-emitting diodes (LEDs; 2.0-mm diameter, red light, λmax = 635 nm) were placed in the frontal hemifield between ±50° in azimuth and ±30° in elevation. The 17 speakers included a central speaker (0°, 0°), six speakers in the horizontal plane, six speakers in the sagittal plane, and four speakers along the diagonals (±20°, ±20°). Speakers were hidden from view by a black translucent cloth. An LED was suspended over the center of each speaker and, when illuminated, could be easily seen through the cloth. A Tucker-Davis Technologies III system was used for stimulus generation and data collection.

**Behavioral paradigm and acoustic stimuli.** Cats were trained by operant conditioning with food reward. Eye coils were calibrated with a behavioral procedure (Populin and Yin 1998) that relied on the natural instinct of the cat to look at a light source that suddenly appears in the visual field. The output of the coil system was recorded when the eye assumed a stationary position at the end of movements to known locations. The vertical and horizontal components of eye positions were fit separately with linear equations. The coefficients obtained from the fitting procedure were used to convert the voltage output of the coil system to degrees of visual angle.

Both cats had participated in previous auditory and visual tasks. The only additional training required for the present study was a quick adaptation with softer maskers (30 dB SPL) and longer targets (i.e., 100 or 50 ms) for a few days (data not included). Cats performed the experiment with their heads unrestrained and a feeding tube attached to the head post. In each trial, the cat was required to fixate an LED presented from straight ahead (0°, 0°) and maintain gaze fixation within the acceptance window for a variable period of time (600–1,000 ms). For control trials, no sound was played along with the fixation LED. During a forward masking trial, a broadband (BB) noise masker (0.1–30 kHz) was presented at (0°, 0°) the same time period as the LED (Fig. 1B). For both control and masking trials, when the LED and/or the sound masker was extinguished, a short acoustic signal (25 ms) was presented immediately (called FM0) or 15 ms (called FM15) after the termination of the fixation light and the masker sound (Fig. 1B). The cat was required to make a gaze saccade to the apparent location of the signal. If during the 600–1,000 ms following the offset of the LED the gaze remained within a specified acceptance window around the target location, the cat was given a food reward. Data were analyzed regardless of whether a reward was received. Forward masking trials were randomly mixed with control trials and other types of trials (e.g., visual-target trials).

**METHODS**

**Subjects and experimental setup.** All surgical and experimental procedures were approved by the University of Wisconsin Animal Care and Use Committee and complied with the guidelines of the National Institutes of Health. Two adult female cats were implanted with stainless-steel coils (S1700127-FEP; Alan Baird Industries, Hohokus, NJ) for the eyes under aseptic surgical conditions. The coils were threaded subcutaneously and connected to connectors that were embedded in the dental acrylic cap on the skull.

The behavioral experiment was conducted in a dimly lit sound-attenuated chamber (2.2 × 2.5 × 2.5 m; Fig. 1A). The eye coil was connected to a magnetic search coil system that enabled the vertical and horizontal components of the eye movements to be monitored with a resolution of <1°. Speakers (model 40-1310B; RadioShack, Fort Worth, TX) and light-emitting diodes (LEDs; 2.0-mm diameter, red light, λmax = 635 nm) were placed in the frontal hemifield between ±50° in azimuth and ±30° in elevation. The 17 speakers included a central speaker (0°, 0°), six speakers in the horizontal plane, six speakers in the sagittal plane, and four speakers along the diagonals (±20°, ±20°). Speakers were hidden from view by a black translucent cloth. An LED was suspended over the center of each speaker and, when illuminated, could be easily seen through the cloth. A Tucker-Davis Technologies III system was used for stimulus generation and data collection.

**Behavioral paradigm and acoustic stimuli.** Cats were trained by operant conditioning with food reward. Eye coils were calibrated with a behavioral procedure (Populin and Yin 1998) that relied on the natural instinct of the cat to look at a light source that suddenly...
The 25-ms target was either a BB noise (0.1–30 kHz) or a band-limited (BL) noise. Three frequency ranges were tested for the BL condition: low (0.2–2 kHz), mid (2–7 kHz), and high frequencies (7–15 kHz). Note that the bandwidths of the BL noise used here were broader than what are normally used for narrow-band sound (e.g., wider than the critical band). According to the duplex theory, this arrangement ensured that the low frequency covered only the ITD cue, the mid frequency had some ITD and spectral cues (see Fig. 11 in Sabin et al. 2005; see Discussion for the limitation of the duplex theory and the possible role of envelope ITD cue). The overall sound level of the BB noise masker was fixed at 50 dB SPL. The overall sound level of the BB or BL target varied between 20 and 80 dB SPL. All the noise tokens (the masker or the signal, BB or BL) were independent samples but were kept frozen for each condition. To verify that the major results were not associated with special tokens of frozen noise, several conditions were repeated using different frozen noise samples.

Data analysis. Horizontal and vertical gaze positions were determined separately by a velocity criterion (Populin and Yin 1998; Tollin et al. 2005). The beginning of the gaze movement was marked as the point in time when steady fixation ended, i.e., when the magnitude of the velocity exceeded 2 standard deviations (SDs) of the mean velocity computed during the initial steady fixation. The latency was computed as the time when the gaze movement began relative to the beginning of the target. The final gaze position was determined at the time when the velocity returned to ≤2 SDs of the baseline mean velocity. For trials with no saccade, caused either by perception at the front center or by failure of localization or detection, the start and the end of the “saccade” were chosen to be 200 and 200 ms (see Discussion for the limitation of the duplex theory and the possible role of envelope ITD cue). The overall sound level of the BB noise masker was fixed at 50 dB SPL. The overall sound level of the BB or BL target varied between 20 and 80 dB SPL. All the noise tokens (the masker or the signal, BB or BL) were independent samples but were kept frozen for each condition. To verify that the major results were not associated with special tokens of frozen noise, several conditions were repeated using different frozen noise samples.

With the use of the empirical data set, the SD of the residuals of the fitted function was computed to measure the distribution of behavioral responses about the mean gain. This value, referred to as $\delta$, provides a numerical estimate of the precision (or consistency) of the localization responses (see also Moore et al. 2008). Bootstrapping was also used to obtain an estimate of the 95% confidence intervals of $\delta$. To statistically determine whether or not forward masking had an effect on the localization accuracy or precision, a slightly different algorithm of bootstrapping (Moore et al. 2002) was used to test the significance of the change in the gain ($g_{\text{FM}} = \text{control gain} - \text{FM gain}$) and $\delta$ ($\delta_{\text{FM}} = \text{abs(FM gain} - \text{control}\delta)$) when the forward masker was added. As shown in Results, the gain can only decrease under forward conditions.

---

1 When there was no gaze movement, we did not know how the subject perceived the sound. One possibility was that the subject recognized the presence of a sound but could not localize it. Another possibility is that the subject heard and localized the sound but did not want to make a gaze shift due to other behavioral factors.

---

![Fig. 2. Scatter plots of the final gaze positions (top) and gaze shifts vs. motor error (bottom) for a broadband target (25 ms, 35 dB SPL) obtained with cat 33. A: no forward masker was presented (control). B: a 50-dB forward masker was presented during fixation, and there was a 15-ms gap between the end of the masker and the beginning of the target (FM15). C: a 50-dB forward masker was presented, and the target was presented immediately after the termination of the masker (FM0). The large open symbols at top represent target locations, and the small filled symbols represent gaze positions with the same color and shape as the corresponding targets. The gain and precision ($\delta$) derived from the regression analysis (straight lines, bottom) are indicated by the 2 values in brackets computed separately for horizontal (1st value; Hori) and vertical (2nd value; Vert) gaze movements. Only 11 of the 17 target locations are plotted for clarity.](https://jneurophyys.org/article/10.1152/jn.00255.2013){/lightbox}
masking, whereas \( \delta \) may increase or decrease. Specifically, a null hypothesis was constructed by pooling all the control trials together with the forward-masking trials. For each bootstrapping iteration, two new sets of trials were randomly selected (with replacement) from the pool, and the difference between the two new gains or \( \delta \) values was computed. A distribution of the gain/\( \delta \) difference was formed after 1,000 iterations. If the actual gain/\( \delta \) change (\( g_2 \) or \( \delta_2 \)) fell on the tail of this distribution (\( P < 0.05 \); a rare event), the effect of forward masking was considered significant. If \( g_2/\delta_2 \) fell on the main body of this distribution (\( P > 0.05 \)), the effect of forward masking was not significant (i.e., the null hypothesis could not be rejected).

RESULTS

Forward masking affected localization of a broadband target at all sound levels. Figure 2A, top, shows the final gaze positions in response to a BB signal (35 dB SPL) at various spatial locations (different colors) without a masker. Although some degree of scatter was observed, the final gaze positions (filled symbols) were generally close to the true target positions (open symbols). Without a masker, the localization gains derived from comparing the gaze shift and the motor error were close to but less than 1.0 for both horizontal and vertical dimensions (Fig. 2A, bottom). When a BB forward masker (50 dB SPL, >600 ms) was presented at the front center (0°, 0°), the target positions were overall underestimated; that is, the final gaze positions were more toward the fixation center and further away from the actual target locations in the presence of a forward masker (Fig. 2, B and C, top). Consequently, the linear regressions had shallower slopes (Fig. 2, B and C, bottom), and the gain/accuracy decreased. In addition, \( \delta \) increased, indicating lower precision, i.e., more scatter in the gaze response as a result of larger trial-by-trial variability (Fig. 2, B and C, top). The masking effect was larger when the target was presented immediately after the termination of the masker (Fig. 2C, FM0) compared with when there was a 15-ms interstimulus gap (Fig. 2B, FM15).

Figure 3 summarizes the masking effect on localization accuracy for all target sound levels (20–80 dB SPL). Although the masking effect was largest at the lowest levels (20 or 30 dB SPL), the effect persisted at higher sound levels when the target was close to but less than 1.0 for both horizontal and vertical dimensions (Fig. 2). When a BB forward masker (50 dB SPL, >600 ms) was presented immediately after the termination of the masker, the presence of a BB forward masker generally led to more scatter in the gaze response for all but the lowest sound level (Fig. 4A). In other words, for relatively loud sound levels, the gaze shift was in the direction of the target but not precise, and it varied trial by trial. As was true with gain, for most of the sound levels, \( \delta \) obtained in the FM15 condition lay between that of the FM0 and the control (Fig. 4A). However, at the lowest sound level (20 dB SPL), \( \delta \) for FM0 was significantly smaller than \( \delta \) for FM15. The scatter plots of the final gaze positions (in the format of Fig. 2C, top, but not shown) indicated that the cats did not move their gaze away from (0°, 0°) for most of the trials. It is possible that the signal was completely masked by the forward masker with no interstimulus gap at this sound level or that the sound level (20 dB SPL) was close to the masked detection threshold. There was also the possibility that the signal was detected but the location was undetermined. However, since we did not do a detection study to measure the threshold, nor could we communicate with the subjects, we were unable to make any judgment of the signal perception at low levels.

There was a negative correlation between the gain and the \( \delta \) when all the conditions (masked or unmasked) and levels for either horizontal or vertical gaze movement were combined (Fig. 4B). This correlation was significant (\( P < 0.05 \)) except for cat 36’s vertical movement (−0.38). In general, larger scatter was associated with lower gains for localization of a BB target, and the localization performances in azimuth (blue) and elevation (red) were comparable.

The presence of a forward masker also increased the response/gaze latency, especially at relatively low target sound levels (Fig. 5). For targets in the horizontal plane, we only computed the latency for the horizontal component of gaze movement (blue), and for targets in the sagittal plane, only the vertical gaze movement (red). Latencies were computed along both dimensions for the four targets in diagonal (±20°, ±20°). The latency change potentially reflected difficulties in the subject’s decision making for either detection or localization of the signal.
Localization of low-frequency targets was highly affected by forward masking. According to the duplex theory, horizontal localization of low-frequency sound depends on the ITD cue. The example in Fig. 6 shows that, following a BB masker, the localization of an 80-dB SPL low-frequency BL target (0.2–2 kHz) became highly inaccurate (compressed gains) (Fig. 6B) compared with that without a masker (Fig. 6A). This was the case for all the sound levels tested (50, 65, and 80 dB SPL) and for both cats (Fig. 7A). We did not test sound levels lower than 50 dB SPL.

However, forward masking did not affect the precision for either animal (Fig. 8A) except for one condition with cat 36 (Fig. 8A, bottom, 80 dB SPL). Forward masking tended to increase the latency of the response for cat 33, but not for cat 36 (Fig. 9A). Nevertheless, because the perceived locations were highly compressed toward the front center as measured by the gain (Fig. 7A), it is fair to conclude that the binaural mechanism based on ITDs is highly susceptible to forward masking.

It is believed that the mid-frequency notch introduced by head and pinna filtering properties is critical for localization in elevation (7–15 kHz for targets within ±30°; Musicant et al. 1990; Tollin and Koka 2009a, 2009b). Here an interesting finding is that, even for the frequency range (0.2–2 kHz) far below the notch frequencies, the cats were sometimes able to perceive the vertical location of the target (Figs. 6 and 7A, red). A similar finding has been observed with some of our cats in previous experiments with long-duration (100 ms) BL sound (see Fig. 3 in Tollin et al. 2013). However, this possible low-frequency spectral cue became highly unreliable under forward masking. For example, cat 36 typically responded to straight left or right with no movement in elevation for the four targets in diagonal (±20°, ±20°) in the FM0 and FM15 conditions.
conditions (not shown). For cat 33, both horizontal and vertical localizations were disrupted (Fig. 7A, top).

Forward masking had a small effect on localization of high-frequency targets. In contrast, when high-frequency (7–15 kHz) BL targets were likely to be well detected (78 dB SPL), the localization accuracy was barely affected by forward masking (Fig. 7C). Although for one cat (cat 33) the decrease in gain for FM0 at 65 dB SPL was significant (Fig. 7C, top, blue), the change was certainly not as large as what was observed for low-frequency targets. There was also little effect on the precision (Fig. 8C) and latency (Fig. 9C) at levels ≥35 dB SPL for both horizontal and vertical dimensions. At 25 dB SPL, both cats showed significant decreases in gain for FM0 (Fig. 7C). This decrease of gain was also accompanied by a large increase of δ (Fig. 8C, bottom), implying that the target was still detectable. Nevertheless, the detection mechanisms, rather than the localization mechanisms, are likely to play the major role at this sound level.

Overall, these data indicate that the ILD cue, as well as the spectral cue including the important spectral notches (7–15 kHz), is more robust than the ITD cue under forward masking.

Localization of mid-frequency targets can be affected by forward masking. For cat 33, the localization accuracy of mid-frequency (2–7 kHz) targets was significantly affected by forward masking at all sound levels (Fig. 7B, top). For the other cat, the change of gain was less consistent (Fig. 7B, bottom). δ decreased at the lowest level tested (25 dB SPL) for both cats, and also at 65 dB SPL for cat 33 (Fig. 8B). The latency increased with the forward masker for low and medium sound levels (Fig. 10). Overall, the binaural localization mechanisms, possibly a mixture of ITD and ILD cues, were affected by forward masking. Again, some degree of elevational information was perceived by the cats when there was no masker (Fig. 7B). This subtle elevational cue was not robust under forward masking.

Correlations between localization accuracy and precision. Figure 10 shows the correlations between gain (accuracy) and δ (response scatter) for different BB and BL conditions. As described earlier (Fig. 4B), there were negative correlations between gain and δ (i.e., positive correlations between accuracy and precision) for BB targets, since a lower gain was usually accompanied by larger scatter except at the lowest sound levels. In contrast, there was no consistent correlation for the BL conditions, although generally speaking, for the horizontal dimension the correlation was more likely to be negative (solid bars), and the vertical dimension, positive (shaded bars). In other words, a small δ may occur for very accurate (such as the controls in azimuth) or very poor (such as the masked performances in elevation) localization.

Localization Accuracy for Band-Limited Targets

![Fig. 7. Localization accuracy measured with the gain for BL targets (25 ms) as a function of target sound level, in the same format as Fig. 3. Target levels plotted are the overall sound levels. An 80-dB SPL overall sound level corresponds to a spectrum level of 47, 43, and 41 dB SPL, for the low-, mid-, and high-frequency conditions, respectively; n > 68 trials for each condition (masked or unmasked).](image-url)
Note that the above results were obtained with one token of frozen noise for each of the BB or BL condition fixed at all sound levels. To verify that the different observations were not due to special noise tokens, we repeated the experiment at one target level (50 dB SPL) with a new sample of low-frequency BL noise and a new high-frequency noise. For cat 36, the same conclusion can be derived with the new frozen noise that forward masking affected the ITD but not the ILD cue (Fig. 11, 1st and 2nd conditions). For cat 33, the localization of the new low-frequency noise was also highly affected (Fig. 11, 3rd condition). Although we did observe a significant decrease in the horizontal performance for the high-frequency noise (Fig. 11, 4th condition), it was less dramatic compared with its low-frequency performance. Note that this cat always showed larger influence of masking than the other cat did, for both BB (Fig. 3) and high-frequency BL targets (Fig. 7C).

**DISCUSSION**

*Forward masking affects mammalian localization mechanisms.* The present study found that forward masking had a significant effect on the localization of broadband and low- or mid-frequency band-limited sound over a large range of sound levels. We believe that the masking effect observed here was not an artifact, since we observed temporal release of masking when a 15-ms interstimulus gap was added between the masker
and the signal. In addition, we often observed systematically increased gaze latencies when the masker was added and the gap was removed, reflecting more difficult decision making.

Our results obtained with the BL targets indicated that ITD information carried by low-frequency sound becomes highly unreliable under forward masking, whereas the ILD and spectral cues are relatively unaffected. One possibility is that the ability of the subject to integrate neural information varied with different cues. For example, comparing the localization performance for FM0 and FM15, we found that the masking effect can be notably reduced even after a 15-ms gap. For the ILD and high-frequency spectral cue, it is possible that the subject can make the judgment based on the latter part of ongoing information for the 25-ms target, thereby creating large temporal decay of masking. In contrast, if the binaural mechanism for ITD weighs heavily on the onset, which suffers the largest masking effect, it will not be surprising to observe a pronounced degradation of localization performance at low frequencies. This so-called “onset dominance” has been demonstrated by many sound lateralization studies using pulse trains, trains of frozen noise, etc. (Balakrishnan and Freyman 2002; Freyman et al. 1997, 2010; Saberi and Perrott 1995). However, dominance of ongoing ITDs has also been shown for noise stimuli (Best et al. 2004; Freyman et al. 2010).

To test whether the large masking effect observed at low frequencies was due to onset ITD dominance, we doubled the target duration to 50 ms (Fig. 12, right, top bar) for cat 36. If the localization of the low-frequency target following a forward masker was only determined by the onset ITD of the target, we expect to see similar performance with the FM0 (50-ms duration) and the FM0 (25-ms duration; Fig. 12, right, 2nd bar). However, the localization accuracy obtained with the long-duration target was considerably higher than that of the short-duration sound. The localization accuracy of the FM0 (50 ms) was also better than the accuracy of the FM25 (25 ms) (Fig. 12, left, blue), indicating that the cat was able to integrate ITD information along the whole stimulus duration, rather than focusing on the end where the masking effect was the least. Therefore, we believe that the large masking effect observed at low frequencies was due to disruptions of both onset and ongoing ITD cues, rather than the subject’s incapability of using ongoing ITD cues.

However the mechanism is affected, our results are in agreement with the general findings of binaural sluggishness (e.g., Grantham and Wightman 1978, 1979; Kollmeier and Gilkey 1990), which is typically examined by using headphone stimulations with ITDs. For example, in the presence of a continuous masker, an estimated “binaural minimum integration time” of more than 40 ms is required to detect the ITD of a signal. In the present study, the signal was 25 ms in duration following a long (>600 ms) masker. If similar integration time is needed for free-field sound localization based on ITDs, then we would expect that a large masking effect would be observed with the 25-ms signal but not with the 50-ms signal.

In addition, previous localization studies found that simultaneous masking does not significantly affect localization for large SNRs (Good and Gilkey 1996; Lorenzi et al. 1999). Because these studies use long-duration signals, the binaural sluggishness is presumably not an issue. It is also likely that the auditory system can better counteract the masking effect by analyzing multiple “looks” of the ongoing localization information (Gai et al. 2013; Hofman and Van Opstal 1998). To test whether the discrepancy between our finding and the observation with simultaneous masking is due to the intrinsic nature of different masking types or different sound durations, future localization studies should repeat the simultaneous masking experiments with brief signals and maskers.

Despite the long held belief that ILD is the major cue for localization of high-frequency sound, many studies have pointed...
out that localization at high frequencies may also utilize the envelope ITD cue (e.g., Bernstein and Trahiotis 1994; Henning 1974; Klumpp and Eady 1956; Nuetzel and Hafer 1981). However, listeners’ ability to follow the envelope fluctuation decreases at high modulation rate (Bernstein and Trahiotis 1994; Nuetzel and Hafer 1981). The bandwidths of the BL targets in the present study were much broader than the critical bands. Although we cannot eliminate the possibility that some envelope ITD cue was present and used by the cats to localize high-frequency sound, future studies should test truly narrow-band stimuli so that the envelope modulation is slow enough for the subjects to follow for the purpose of encoding ITDs carried by the envelope.

*Localization accuracy and precision can vary independently under forward masking.* Localization accuracy (quantified by the gain) measures how close the mean responses are to the actual target locations. Localization precision (quantified by $\delta$) measures the consistency of trial-by-trial responses. Although, generally speaking, poor localizations are usually accompanied by inaccurate response that varies trial by trial, Tollin et al. (2005), along with Heffner and Heffner (2005; author reply), argue that theoretically these two measurements can be completely independent of each other. For most of the BB and low/mid-frequency BL conditions, we found that the localization accuracy changed systematically when a masker was presented with a 15-ms gap and further when the temporal gap was removed. For BB targets, this decreased accuracy (i.e., undershooting the true target locations) was usually accompanied by larger trial-by-trial variations, except at low sound levels. It seems that when all three cues could be used for localization, there was consistency in the localization behavior across sound levels and subjects. For BL targets, the systematic change in the gain created by forward masking was not accompanied by a systematic change in the precision. In general, we suspect there are three types of situations. First, the target location can be easily identified. For most of the control conditions with the BB targets, or some of the control conditions with the BL targets located in the horizontal plane, from trial to trial the subjects consistently (a small $\delta$) and accurately (a high gain) responded to the true location or its vicinity. The second situation occurred when a masker was presented but the target was likely to be well detected with abundant localization cues, such as the case of BB targets at high levels. In each trial, the subject may have had a definite perception of the target location, but it varied from trial to trial because of the influence of the masker. This would result in a low gain and a large $\delta$ (a negative correlation). However, there was also a possibility that the target location, or some dimension of the target location, was ambiguous to the subject. For example, for low- or mid-frequency BL targets located in the sagittal plane, the most important elevational cues were deprived and the perception of target locations may have been strange. The subjects were found to generate little vertical movements in these cases, resulting in a low gain and a small $\delta$ (a positive correlation). This situation may also have occurred when the target was close to or below the masked detection threshold. This could be the case for the BB targets at the lowest two levels (Fig. 4A), but since we did not measure the detection threshold, we are uncertain whether or not this was true.

In addition to localization accuracy and precision, acuity is a measure of relative localization, which is normally assessed by examining the smallest angle between two sources that can be discriminated. Intuitively, acuity should be related to precision. However, Moore et al. (2008) found that in some cases when the accuracy is poor, the acuity is poorly correlated to precision. Future studies should examine the effect of forward masking on sound-source discrimination to see whether a systematic change can be observed (like the accuracy) or not (like the precision). In addition, we believe that neither the accuracy nor the precision is always related to the correlation coefficient between the responses and target locations. For example, if the response always undershoots (low accuracy/gain) but is proportional to the target location, there can still be a high correlation coefficient between the responses and target locations.

*Localization aftereffect.* The “localization aftereffect” describes the phenomenon of perceiving a second sound away from its true location after being exposed to a previous sound from the same location (Getzmann 2003; Kawashima and Sato 2012). Since the masker was fixed at the front center in the present study, we could only examine those trials when the targets were also presented at the front center. For cat 36, in the absence of the masker, the unsigned error (i.e., the distance between the gaze position and the central target) remained relatively small ($\sim 6^\circ$) and invariant with sound level (not shown). Under forward masking (FM0 or FM15), the error increased with decreasing target level up to $15-20^\circ$. At low sound levels, when the cat looked away it turned to look at a default position, which was the lower left of the front center (see Tollin et al. 2013 about the default positions). There were times when the cat looked to the left or the right, but never to the upper hemifield. For cat 33, no trend was observed with the error for either sound levels or masking conditions.

*Spatial release from masking.* Studies on the binaural release from forward masking (Bekhterev et al. 2002; Deatherage and Evans 1969; Kollmeier and Gilkey 1990; Zwicker and Zwicker 1984) suggest that, at near-threshold levels, localization mechanisms can be utilized to facilitate signal detection. In those headphone studies on binaural release from forward masking, the detection threshold decreases when the signal changes from diotic (in phase across 2 ears; NoSo) to dichotic (anti-phase across 2 ears; NoSs). When we examined the effect of the BB masker on the localization of BB or BL targets at relatively high sound levels, it was obvious that under forward masking, the localization of peripheral targets (close to $50^\circ$ in either direction) was no more accurate than the localization of targets that were closer to the central masker (see example shown in Fig. 2, B and C). At low sound levels, the cats could not localize the majority of targets, whether central or peripheral. Therefore, spatial release from forward masking was not observed in terms of localization or detection.

There are several possible explanations for our finding on the lack of spatial release. First, only targets in the frontal hemifield within a rather limited range, $\pm 50^\circ$ in azimuth and $\pm 30^\circ$ in elevation, were tested. The observation may be different when larger angles (such as up to $\pm 90^\circ$ in the front or even in the back hemifield) are included. The observation may also be different if the masker is located off center so that the masker and the target can be on the same side or different relative to the subject. Of particular interest is the masking effect on localization of low-frequency BL noise or tones to match those headphone studies.
Second, when sound is delivered through headphones, the subject may detect the signal ($S^*$) by comparing the sound waveforms across ears. In other words, the detection can be done by examining any interaural difference, i.e., an alteration of the masker because of the addition of the signal. It does not require that the signal is separable from the masker. This strategy cannot be used in free-field localization, for which the source of the signal has to be separable.

Last, in a free-field condition where sounds originate from different locations, the subject’s attention may be drawn to the masker and its vicinity, which makes the detection or localization of a faraway signal less efficient. Studies have shown that listeners can modulate spatial attention in a detection task by attending to expected target locations and withdrawing attention from expected masker locations (Allen et al. 2011), and thus spatial release from masking may only occur for attended location (Allen et al. 2009). In the present study, because the masker was always presented at the front center and preceding the signal, the cat’s attention was very likely to be drawn to the front center, thereby showing no release for peripheral targets.

ACKNOWLEDGMENTS
We thank J. Sekulski and Y. Yan for helping with computer programming and A. Jones for animal care and training. D. Tollin, D. Oertel, V. Kotak, and L. Carney provided useful input.

GRANTS
This work was supported by National Institute of Deafness and Communicative Disorders Grant DC-07177.

DISCLOSURES
No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS
Y.G., J.L.R., and T.C.T.Y. conception and design of research; Y.G. and J.L.R. performed experiments; Y.G. and J.L.R. analyzed data; Y.G., J.L.R., and T.C.T.Y. interpreted results of experiments; Y.G. prepared figures; Y.G. and J.L.R. analyzed data; Y.G., J.L.R., and T.C.T.Y. approved the final version of manuscript.

REFERENCES


