Research Article

The Influence of Target and Masker Characteristics on Infants' and Adults' Detection of Speech

Monika-Maria Oster^a and Lynne A. Werner^a

Purpose: Several investigators have compared infants' detection of speech in speech and nonspeech maskers to evaluate developmental differences in masking. Such comparisons have produced contradictory results, possibly because each investigation used different stimuli. The current study examined target and masker effects on infants' and adults' detection of speech.

Method: An observer-based procedure was used to compare infants' and adults' detection of the vowel /a/ and the word "baby" in a 2-talker speech masker and matched speech-spectrum noise. The measure of performance was *d'*. A total of 43 7-month-old infants and 41 young adults were randomly assigned to 1 target-masker combination condition, and mean performance was compared across conditions at each age.

Results: Adults' detection was influenced by an interaction between the target and the masker: Adults detected the vowel better in the 2-talker masker than in speech-spectrum noise but detected the word equally well in the 2 maskers. In contrast, infants detected both targets better in speech-spectrum noise than in the 2-talker masker.

Conclusions: The relative effects of the masker on target detection by infants and adults depend on the target to be detected. Thus, conclusions drawn about differences between infants and adults in the mechanisms responsible for masking will depend on the stimuli. Standardization of speech stimuli in developmental research would help clarify the nature of infants' segregation difficulties.

Supplemental Material: <https://doi.org/10.23641/asha.5613139>

It is well known that infants have greater difficulties than adults separating speech from competing sounds (e.g., Buss, Hall, & Grose, 2012; Leibold, Bonino, & Buss, 2016; Newman, 2005; Nozza, Rossman, Bond, & Miller, 1990; Werner, 2013). Given that infants spend much of their time in noisy environments and are faced with the task of learning language in them, it is not surprising that research in recent years has focused on the question of why infants require a greater target-to-masker ratio (TMR) than adults to successfully separate sounds (e.g., Barges & Werner, 1994; Leibold et al., 2016; Leibold & Werner, 2006; Polka, Rvachew, & Molnar, 2008; Werner, 2013). The TMR at which infants successfully detect and discriminate speech in competing sounds varies widely between studies, ranging from -3 dB to $+15$ dB (e.g.,

Leibold et al., 2016; Newman, Morini, & Chatterjee, 2013; Nozza et al., 1990; Polka et al., 2008). It is not entirely clear why infants require a greater TMR than adults or why infant results vary so dramatically between studies. Separating speech from competing sounds requires several processes. These include auditory coding, the representation of the sound complex; sound source segregation, the process by which sound components are assigned to a source; and auditory attention, the selection of a particular source for additional processing. Given that infants' ability to encode sounds appears to be mature (Buss et al., 2012), sound source segregation and auditory attention are potentially the most likely contributors to infants' greater difficulties.

Across-study differences in the maskers used may explain some of the differences observed in infant masked thresholds. In the adult literature, differences in maskers have in fact been identified as one source of variability in performance (e.g., Freyman, Balakrishnan, & Helfer, 2004). This may not be surprising, given that masking is a complicated process. Competing sounds may interfere with the ability to resolve the frequency components of the incoming sound in the periphery, a mechanism known as *energetic*

^aDepartment of Speech and Hearing Sciences, University of Washington, Seattle

Correspondence to Monika-Maria Oster: mmoster@uw.edu

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Daniel Fogerty

Received December 23, 2016

Revision received April 27, 2017

Accepted July 28, 2017

https://doi.org/10.1044/2017_JSLHR-H-16-0464

Disclosure: The authors have declared that no competing interests existed at the time of publication.

masking. The intrinsic modulations in competing sounds can also result in a type of interference known as *modulation masking* (Stone & Canavan, 2016; Stone, Füllgrabe, & Moore, 2012). Furthermore, competing sounds may interfere with sound source segregation and selective auditory attention; that sort of interference is commonly referred to as *informational masking*. When the competing sound is speech, linguistic interference is also possible (Hawley, Litovsky, & Culling, 2004). Different maskers are thought to engage different masking mechanisms. For example, in adults, noise appears to limit performance due to energetic and modulation masking (Stone & Canavan, 2016; Stone et al., 2012), whereas competing speech is thought to limit performance primarily via informational masking (Brungart, Chang, Simpson, & Wang, 2006; Brungart, Chang, Simpson, & Wang, 2009; Leibold & Buss, 2013; Wightman & Kistler, 2005). While spectrotemporal variability is associated with high levels of informational masking, spectrotemporal variability in some competing sounds, notably speech, also provides opportunities for listeners to “glimpse” the target speech in low-energy segments of the masker, resulting in a release from masking. Glimpsing is observed at negative TMRs and is thought to require listeners to segregate the target from the masker (Bernstein & Grant, 2009; Brungart, 2001).

To the extent that infants differ from adults in susceptibility to informational masking or in the ability to glimpse a target in a spectrotemporally variable masker, it would be predicted that infants and adults would exhibit different patterns of masker effects. However, the few studies that have compared infant performance in different maskers have produced contradictory results. For example, Newman (2009) reported that 4- to 6-month-old infants did not recognize their own name in the presence of a single-talker masker, but were able to do so in the presence of a multitalker (i.e., noiselike) masker at the same TMR. The author concluded that infants experienced informational masking in the single-talker masker, but not in the multitalker masker. In contrast, Leibold et al. (2016) found that 8- to 10-month-old infants’ detection thresholds for bisyllabic words were the same in a two-talker speech masker and speech-spectrum noise. However, infants’ thresholds in both maskers were higher (worse) than adults’ in the same study, consistent with prior research showing that infants require a more favorable TMR. The authors conclude on the basis of these results and the fact that infants have mature frequency resolution that infants must experience greater informational masking than adults do in speech-spectrum noise. Results on infants’ ability to use glimpsing in the segregation of target from competing speech are similarly contradictory. While Newman concluded that infants were unable to glimpse their names in “dips” in the single-talker masker, Werner (2013) reported that infants could detect and discriminate between vowels better in a modulated noise masker than in an unmodulated noise masker at the same TMR.

In contrast to masker differences, the influence of the chosen target speech material on masking has received

little attention (but see Buss et al., 2016). When different targets, such as vowels, syllables, words, or sentences, are presented in the same masker, the spectrotemporal structure of the targets may influence the ability to separate the target from the competing sounds. This has indeed been demonstrated in adults. For example, Liu and Eddins (2008) found that adult listeners’ detection thresholds for vowels in speech-spectrum noise were dependent on the specific vowel. Given that infant studies have used a variety of target speech ranging from syllables to words and phrases (Barker & Newman, 2004; Hollich, Newman, & Jusczyk, 2005; Newman, 2011; Newman & Hussain, 2006), differences in target speech are another possible explanation for the observed inconsistencies in infant results. However, the influence of the specific target on infant performance has not been investigated systematically (but see Leibold et al., 2016). On the basis of this sparse evidence, it can be hypothesized that the target might influence performance differently for infants than for adults.

Because comparisons between age groups are regularly used to describe the development of auditory skills, it is important to determine whether the specific target and masker used influences performance differently for infants and adults. To this end, the current study evaluated the detection of a vowel and a bisyllabic word in a two-talker speech masker and matched speech-spectrum noise in infant and adult listeners.

Methods

Subjects

A total of 43 infants (28 male, mean age = 31 weeks, $SD = 3.5$ weeks) and 41 adults (15 male, mean age = 23.2 years, $SD = 3.8$ years) participated in this study. Inclusion criteria included (a) being generally healthy and developing normally; (b) normal hearing as indicated by parent- or self-report, passing newborn hearing screening, negative family history of hearing loss, and no risk factors for hearing loss; (c) normal middle ear function as indicated by no treatment for otitis media within 2 weeks prior to participation, no more than two prior episodes of otitis media, and passing a tympanometry screen on each test day; (d) English as the dominant language spoken at home; and (e) no prior participation in psychoacoustic experiments. Subjects were recruited through flyers and the Communication Studies Participant Pool of the University of Washington. Subjects were paid for their participation. An additional 23 infants who otherwise met inclusion criteria were excluded: nine infants were too fussy or tired to be tested, five infants did not reach training criteria, and nine infants reached training criteria but were unable to complete testing.

Stimuli

The stimuli were designed to be similar to those found in the existing infant literature to allow for meaningful comparisons between studies. The choice of specific maskers and speech targets was based on a pilot experiment that

assessed adult masked sensitivity to several different targets in two-talker maskers and speech-spectrum noise (see Supplemental Material for a detailed description). The stimuli were recorded by three female monolingual English speakers; two recorded the stimuli for the two-talker masker, and the third recorded the speech targets. The stimuli were recorded in a sound-attenuated booth, using a high-quality condenser microphone (AKG C1000S) placed 6–8 in from the talker's mouth. All stimuli were digitized at a resolution of 32 bits and a sampling rate of 44.1 kHz.

The two-talker masker consisted of two different passages from the children's book "Cloudy with a chance of meatballs" (Barrett & Barrett, 1978), read as if reading to an adult. Before combining the two speech streams, the passages were manually edited to remove silences longer than 300 ms, cut to the same length, and balanced for root-mean-square amplitude. The resulting two-talker masker was 1.21 min long. The spectral envelope of the two-talker masker was extracted using linear predictive coding. Gaussian noise was then filtered to create matched speech-spectrum noise. The spectrogram of the two-talker speech masker and the spectrum of the speech-spectrum noise are shown in Figure 1.

The speech targets were the vowel / Λ / and the spondee "baby," recorded in an adult-directed speaking style. The two targets were similar in duration (/ Λ /: 462 ms, "baby": 482 ms) and matched for root-mean-square amplitude. The spectrograms of the targets are shown in Figure 2. The results of pilot testing suggested that adults would detect / Λ / better in the two-talker masker than speech-spectrum noise, whereas they detected "baby" equally well in the two maskers.

Stimuli were presented via an ER2 (Etymotic Research) insert earphone to the subject's right ear, calibrated using a Zwislocki coupler (Knowles Electronics) and flat weighting. The two-talker maskers and targets were calibrated to the level of matched speech-spectrum noise. Subjects were tested

in a sound-attenuated booth. Each subject was randomly assigned to be tested in one target-masker combination.

Procedure

The observer-based psychoacoustic procedure (Werner, 1995) was used to assess infants' masked sensitivity. In this procedure, an observer decides whether a target was presented on the basis of the infant's behavior. During testing, the infant was seated on the caregiver's lap in a sound-attenuated booth. An assistant, seated to the left of the caregiver and infant, manipulated quiet toys in front of the infant to keep her engaged and facing forward. Both caregiver and assistant wore headphones with masking sounds to ensure that they were unable to hear the stimuli. Mechanical toys and a small TV screen were located to the right of the infant, used for visual reinforcement. The observer was seated outside the booth where she was able to see the infant, caregiver, and assistant through the observation window or through two video cameras.

The masker was presented continuously throughout the testing session at 55 dB SPL. When the infant appeared calm and was facing forward, the observer initiated a trial. The target was presented randomly on some trials; on the remaining no-target trials, the masker simply continued. The observer was blind as to whether a target had been presented. The observer judged on each trial whether or not a target had been presented, based only on the infant's behavior. The observer received feedback after every trial. The infant was reinforced by the activation of a mechanical toy or a video whenever the observer correctly identified a target trial.

Each testing session progressed through three phases: a demonstration phase, a training phase, and a testing phase. The demonstration phase served to familiarize subjects

Figure 1. The spectrogram of a 2-s segment of the two-talker speech masker (left) and the spectrum of the corresponding time segment of speech-spectrum noise (right).

Figure 2. The spectrogram of the vowel /ʌ/ (left) and the word “baby” (right).

with the association between the target and the reinforcers. The reinforcer was activated whenever the target was presented, regardless of the observer's decision. During the demonstration phase, target trials were 4 times more likely to occur than no-target trials. This phase was completed once the observer responded correctly on one no-target trial and four of the last five target trials. During the following training phase, the target and no-target trials were equally likely and presented in random order. The reinforcer was only activated when the observer correctly identified a target trial. The training phase was completed when the observer correctly responded to four of the last five target trials and four of the last five no-target trials. If an infant did not reach this training criterion on one visit, testing was terminated and attempted again on a second visit. During the testing phase, the infant was presented with two blocks of eight target and eight no-target trials. The target and no-target trials within each block were presented in random order. Testing was terminated if the infant became sleepy or fussy, or when all 32 testing trials were completed. If a session had to be terminated before the testing phase was completed, a few reminder trials were presented at the beginning of the next visit before the testing phase was presented again. All infants completed the testing phase in a single visit. The target level was 75 dB SPL during all phases, resulting in a TMR of +20 dB. This level was chosen on the basis of pilot testing to produce a d' of approximately 1 in speech-spectrum noise during the testing phase. Infants completed the experiment in one to three visits.

Adult subjects were tested using a similar procedure. Adults were seated in the sound booth alone and were instructed to raise their hand whenever they heard the “sound that makes the toys come on.” The observer was seated outside the booth and recorded the subject's responses.

The experimental phases and criteria for completion were the same for adults and infants. The masker was presented continuously at 55 dB SPL. The target level in the demonstration and training phases was chosen to allow adults to reach the criterion; it ranged from 60 to 40 dB SPL, depending on the condition. The target level in the testing phase was 35 dB SPL in all conditions, resulting in a TMR of -20 dB. This level was chosen on the basis of pilot testing to produce a d' of approximately 1 in speech-spectrum noise. Adults completed testing in one visit.

Results

The measure of performance for both age groups was d' in the test phase. Higher values of d' represent better performance. Perfect hit or false-alarm rates of 1 and 0 were adjusted by $1/n$, where n is the number of trials.

Figure 3 shows a box plot of d' in each experimental condition for infants and adults. For adults, the detection of /ʌ/ was much better in the two-talker masker than in speech-spectrum noise, whereas the detection of “baby” was similar in the two maskers. In the infant group, it appears that the detection of both targets is better in speech-spectrum noise than in the two-talker masker.

To statistically evaluate these observations, an Age (2) $\times$ Target (2) $\times$ Masker (2) analysis of variance of d' was conducted. The model included the three main effects, all possible two-way interactions and the three-way interaction. The main effects of age and target were significant, Age: $F(1, 76) = 16.64, p = .0001, \eta^2 = .18$; Target: $F(1, 76) = 12.93, p = .0006, \eta^2 = .15$, whereas the main effect of masker was not, $F(1, 76) = 2.73, p = .1024$. All of the two-way interactions were significant, Target $\times$ Masker: $F(1, 76) = 7.73, p = .0068, \eta^2 = .09$; Target $\times$ Age: $F(1, 76) = 18.66, p < .001, \eta^2 = .2$; Age $\times$ Masker: $F(1, 76) = 14.75, p = .003$,

Figure 3. d' in speech-spectrum noise (SSN) and in a two-talker masker in infants and adults. Higher values of d' represent better performance. The upper and lower lines on the box correspond to the 25th and the 75th percentile, respectively, whereas the line inside the box represents the median. Whiskers represent maximum and minimum d' obtained, whereas single dots represent outliers. Infants detected both targets better in SSN than in the two-talker masker. In contrast, adults detected / Λ / better in the two-talker masker than in SSN, whereas the detection of “baby” was similar in the two maskers.

$\eta^2 = .16$. The three-way interaction was also significant, $F(1, 76) = 9.34$, $p = .0031$, $\eta^2 = .11$.

Follow-up Target $\times$ Masker analysis of variance of d' was conducted separately for infants and adults. For infants, the effect of masker was significant, $F(1, 39) = 3.98$, $p = .05$, $\eta^2 = .09$. The effect of target and the interaction between target and masker failed to reach significance, Target: $F(1, 39) = 0.44$, $p = .513$; Target $\times$ Masker: $F(1, 39) = 0.06$, $p = .8024$. Thus, infants' performance was better in speech-spectrum noise than in the two-talker masker for both targets. For adults, the effect of target, $F(1, 37) = 37.78$, $p < .001$, $\eta^2 = .37$, the effect of masker, $F(1, 37) = 10.49$, $p = .0025$, $\eta^2 = .22$, and their interaction, $F(1, 37) = 11.84$, $p = .0015$, $\eta^2 = .24$, reached significance. Post hoc pairwise comparisons using Bonferroni adjustment confirmed that for adults, the detection of / Λ / was significantly better in the two-talker masker than in speech-spectrum noise ($p < .001$, $\eta^2 = .61$), whereas the detection of “baby” did not differ between the two maskers ($p = .901$).

Discussion

The results indicate that the effects of the masker and the target on speech detection are different for infants and adults. Adults' detection depended on both the target and the masker: They detected / Λ / better in the two-talker masker than in speech-spectrum noise, and they detected “baby” equally well in the two maskers. These results indicate that adults experienced release from masking through glimpsing, but only for the vowel / Λ /. One possible explanation for this difference are the differences in the spectrotemporal properties of the targets. The spectrum of / Λ /

is relatively constant over the duration of the vowel, so that a short glimpse at any point in time would be enough to identify the target. For the target “baby,” such short glimpses would not provide as much information. Thus, for adults, glimpsing the vowel may have been more informative than glimpsing “baby.” Furthermore, the lack of spectrotemporal variation in the vowel would also make it perceptually distinct from the masker, allowing for improved segregation. The target “baby” would not have that advantage.

In contrast to adults, infants detected both speech targets better in speech-spectrum noise than in the two-talker masker. One explanation for the infants' results is that for them, informational masking is stronger in the two-talker masker than in the speech-spectrum noise condition. Infants did not appear to use glimpsing to gain release from informational masking. A likely explanation for this result is the large positive TMR, which always provided a more favorable target than masker signal (Bernstein & Grant, 2009). Furthermore, these results indicate that infants were unable to use the difference in spectrotemporal patterns between / Λ / and the two-talker masker to gain release from informational masking.

Comparison of infants' and adults' detection of / Λ / suggests that the relative importance of the various mechanisms involved in masking is quite different for the two age groups, because infants perform much more poorly than adults in the two-talker speech masker, even when their performance is equated in noise. The effect of age is in the same direction for the target “baby,” that is, infants perform worse than the adults in the two-talker speech masker, but the age-related effect is much smaller. Although their performance is about the same for the two targets, infants seem less mature in one condition than the other, because adult performance is differentially affected by target. These results support our original hypothesis that age-dependent target and masker effects are likely responsible for at least some of the contradictory results in the infant masking literature.

The pattern of results in this experiment is similar to that of Newman (2009), who found that 4- to 6-month-old infants apparently require a higher TMR to recognize their own name presented in a single-talker masker than they do in a multitalker masker. The current results obtained with “baby” may appear to be inconsistent with those of Leibold et al. (2016), who observed that infants' thresholds for detecting bisyllabic words, including “baby,” were the same in speech-spectrum noise and in a two-talker masker. However, when the effects of bias on thresholds are considered, the results are more similar. The dependent measure in the Leibold et al. study was adaptive threshold; in a single-interval paradigm, adaptive thresholds depend on both sensitivity and response bias. Leibold et al. reported that threshold did not differ between masking conditions for either infants or adults. However, they also reported that for infants, d' at threshold was greater for the noise masker than for the two-talker masker, implying that infants were less sensitive in the two-talker speech masker

than in noise, as observed here. For adults, Leibold et al. reported that d' at threshold was greater for the two-talker masker than for the noise masker. Their analysis of d' showed that the Age $\times$ Masker interaction, which would indicate differential effects of masker for infants and adults, just missed significance ($p = .06$). The direction of this trend is consistent with the results of the current study for the target “baby.” Taken together, these results indicate that informational masking is a greater problem for infants in a speech masker than in noise.

In the current study, TMRs were chosen to equate infant and adult sensitivity to targets in speech-spectrum noise. However, the fact that the TMRs required to do so were 40 dB higher for infants than for adults is notable. Leibold et al. (2016) reported a difference between infant and adult adaptive thresholds for masked word detection of about 25 dB, which, as the authors noted, underestimates the threshold difference because of the age-related differences in response bias. Newman and her colleagues (e.g., Newman, 2005, 2009; Newman & Jusczyk, 1996; Newman et al., 2013) also reported that infants require TMRs in the range of +5 to +15 dB to detect words, but the method used to assess infant performance in those studies precludes comparison to adult performance. The use of such different TMRs for infants and adults may provide different cues to sound source segregation to the two age groups. At the same time, infant–adult differences in the threshold TMR for consonant or vowel discrimination in other studies (e.g., Nozza et al., 1990; Werner, 2013) are much smaller than those reported here and in the studies cited above. Thus, differences in the cues available at different TMRs cannot fully explain differences in infant performance across studies. There is, however, consensus that infants are more susceptible to informational masking than adults, which has implications for infants’ auditory learning in natural environments.

In addition to contributing to the growing body of evidence regarding informational masking in infants, the results of this experiment highlight the importance of stimulus selection in developmental research. The conclusions reached about age-related differences in the ability to separate a target from competing sounds can depend on the specific target and competing speech chosen for study. Standardized and well-characterized speech tests, consisting of sentences or word lists, are widely used in hearing research in adults and in young children (for some examples, see Bent & Atagi, 2015; Fogerty, 2011; Meyer & Pisoni, 1999; Wightman & Kistler, 2005). Researchers of hearing in infants should consider working together to develop a set of uniform stimuli that have been acoustically and psychoacoustically well characterized that would be suitable for use in infant studies. For many years, developmental psychoacoustics has focused on designing rigorous procedures for assessing performance in infants and children in a way that allows for comparison to adult performance (e.g., Aslin, Pisoni, Hennessy, & Perey, 1981; Browning, Buss, & Leibold, 2014; Schneider & Trehub, 1992; Trehub, Schneider, Thorpe, & Judge, 1991; Werker & Lalonde, 1988; Werner,

1995). As the interest in understanding infants’ auditory processing in complex sound environments increases, it is imperative that we apply the same effort to the characterization of the sounds we use.

In summary, for the stimuli used here, target detection in adults was influenced by an interaction of the target and the masker. Infants’ performance, on the other hand, was influenced only by the masker. Estimates of infants’ susceptibility to informational masking compared with adults would, as a result, depend on the target chosen in a particular study. The differential influence of stimulus characteristics on performance by infants and adults makes the generalization of results across studies difficult and suggests a need for standardized speech stimuli in developmental research.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grants R01 DC000396 awarded to Lynne A. Werner, T32 DC000033 awarded to Monika-Maria Oster, F31 DC016209 awarded to Monika-Maria Oster, and P30 DC004661 awarded to Edwin W. Rubel. We thank Lori Leibold for her advice, the members of the Infant Hearing Lab for their assistance in subject recruitment and data collection, and Brandon Warren for software development and assistance.

References

- Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development*, 52(4), 1135–1145. <https://doi.org/10.2307/1129499>
- Bargones, J. Y., & Werner, L. A. (1994). Adults listen selectively; infants do not. *Psychological Science (Wiley-Blackwell)*, 5(3), 170–174.
- Barker, B. A., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, 94(2), B45–B53. <https://doi.org/10.1016/j.cognition.2004.06.001>
- Barrett, R., & Barrett, J. (1978). *Cloudy with a chance of meatballs* (1st ed.). New York, NY: Atheneum Books for Young Readers.
- Bent, T., & Atagi, E. (2015). Children’s perception of nonnative-accented sentences in noise and quiet. *The Journal of the Acoustical Society of America*, 138(6), 3985–3993. <https://doi.org/10.1121/1.4938228>
- Bernstein, J. G. W., & Grant, K. W. (2009). Auditory and auditory–visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 125(5), 3358. <https://doi.org/10.1121/1.3110132>
- Browning, J., Buss, E., & Leibold, L. J. (2014). Preliminary evaluation of a two-interval, two-alternative infant behavioral testing procedure. *The Journal of the Acoustical Society of America*, 136(3), EL236–EL241. <https://doi.org/10.1121/1.4893333>
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109(3), 1101–1109. <https://doi.org/10.1121/1.1345696>
- Brungart, D. S., Chang, P. S., Simpson, B. D., & Wang, D. (2006). Isolating the energetic component of speech-on-speech masking with ideal time–frequency segregation. *The Journal of the*

- Acoustical Society of America*, 120(6), 4007–4018. <https://doi.org/10.1121/1.2363929>
- Brungart, D. S., Chang, P. S., Simpson, B. D., & Wang, D. L.** (2009). Multitalker speech perception with ideal time–frequency segregation: Effects of voice characteristics and number of talkers. *The Journal of the Acoustical Society of America*, 125(6), 4006–4022. <https://doi.org/10.1121/1.3117686>
- Buss, E., Hall, J. W., & Grose, J. H.** (2012). Development of auditory coding as reflected in psychophysical performance. In L. A. Werner, R. R. Fay, & A. N. Popper (Eds.), *Human auditory development* (pp. 107–136). New York, NY: Springer.
- Buss, E., Leibold, L. J., & Hall, J. W.** (2016). Effect of response context and masker type on word recognition in school-age children and adults. *The Journal of the Acoustical Society of America*, 140(2), 968. <https://doi.org/10.1121/1.4960587>
- Fogerty, D.** (2011). Perceptual weighting of individual and concurrent cues for sentence intelligibility: Frequency, envelope, and fine structure. *The Journal of the Acoustical Society of America*, 129(2), 977–988. <https://doi.org/10.1121/1.3531954>
- Freyman, R. L., Balakrishnan, U., & Helfer, K. S.** (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America*, 115(5), 2246–2256. <https://doi.org/10.1121/1.1689343>
- Hawley, M. L., Litovsky, R. Y., & Culling, J. F.** (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *The Journal of the Acoustical Society of America*, 115(2), 833–843. <https://doi.org/10.1121/1.1639908>
- Hollich, G., Newman, R. S., & Jusczyk, P. W.** (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development*, 76(3), 598–613. <https://doi.org/10.1111/j.1467-8624.2005.00866.x>
- Leibold, L. J., Bonino, Y. A., & Buss, E.** (2016). Masked speech perception thresholds in infants, children, and adults. *Ear and Hearing*, 37(3), 345–353. <https://doi.org/10.1097/AUD.0000000000000270>
- Leibold, L. J., & Buss, E.** (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. *Journal of Speech, Language, and Hearing Research*, 56(4), 1144–1155.
- Leibold, L. J., & Werner, L. A.** (2006). Effect of masker-frequency variability on the detection performance of infants and adults. *The Journal of the Acoustical Society of America*, 119(6), 3960–3970. <https://doi.org/10.1121/1.2200150>
- Liu, C., & Eddins, D. A.** (2008). Categorical dependence of vowel detection in long-term speech-shaped noise. *The Journal of the Acoustical Society of America*, 123(6), 4539–4546.
- Meyer, T. A., & Pisoni, D.** (1999). Some computational analyses of the PBK test: Effects of frequency and lexical density on spoken word recognition. *Ear and Hearing*, 20(4), 363–371.
- Newman, R. S.** (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology*, 41(2), 352–362. <https://doi.org/10.1037/0012-1649.41.2.352>
- Newman, R. S.** (2009). Infants' listening in multitalker environments: Effect of the number of background talkers. *Attention Perception & Psychophysics*, 71(4), 822–836. <https://doi.org/10.3758/app.71.4.822>
- Newman, R. S.** (2011). 2-year-olds' speech understanding in multitalker environments. *Infancy*, 16(5), 447–470. <https://doi.org/10.1111/j.1532-7078.2010.00062.x>
- Newman, R. S., & Hussain, I.** (2006). Changes in preference for infant-directed speech in low and moderate noise by 4.5- to 13-month-olds. *Infancy*, 10(1), 61–76. https://doi.org/10.1207/s15327078in1001_4
- Newman, R. S., & Jusczyk, P. W.** (1996). The cocktail party effect in infants. *Perception & Psychophysics*, 58(8), 1145–1156. <https://doi.org/10.3758/bf03207548>
- Newman, R. S., Morini, G., & Chatterjee, M.** (2013). Infants' name recognition in on- and off-channel noise. *The Journal of the Acoustical Society of America*, 133(5), EL377–EL383. <https://doi.org/10.1121/1.4798269>
- Nozza, R. J., Rossman, R. N. F., Bond, L. C., & Miller, S. L.** (1990). Infant speech-sound discrimination in noise. *The Journal of the Acoustical Society of America*, 87(1), 339–350.
- Polka, L., Rvachew, S., & Molnar, M.** (2008). Speech perception by 6- to 8-month-olds in the presence of distracting sounds. *Infancy*, 13(5), 421–439.
- Schneider, B. A., & Trehub, S. E.** (1992). *Sources of developmental change in auditory sensitivity*. Washington, DC: American Psychological Association.
- Stone, M. A., & Canavan, S.** (2016). The near non-existence of “pure” energetic masking release for speech: Extension to spectrotemporal modulation and glimpsing. *The Journal of the Acoustical Society of America*, 140(2), 832–842. <https://doi.org/10.1121/1.4960483>
- Stone, M. A., Füllgrabe, C., & Moore, B. C. J.** (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *The Journal of the Acoustical Society of America*, 132(1), 317. <https://doi.org/10.1121/1.4725766>
- Trehub, S. E., Schneider, B. A., Thorpe, L. A., & Judge, P.** (1991). Observational measures of auditory sensitivity in early infancy. *Developmental Psychology*, 27(1), 40–49. <https://doi.org/10.1037/0012-1649.27.1.40>
- Werker, J. F., & Lalonde, C. E.** (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, 24(5), 672–683. <https://doi.org/10.1037/0012-1649.24.5.672>
- Werner, L. A.** (1995). Observer-based approaches to human infant psychoacoustics. In G. M. Klump, R. J. Dooling, R. R. Fay, & W. C. Stebbins (Eds.), *Biomethods: Methods in comparative psychoacoustics* (Vol. 6, pp. 135–146).
- Werner, L. A.** (2013). Infants' detection and discrimination of sounds in modulated maskers. *The Journal of the Acoustical Society of America*, 133(6), 4156–4167. <https://doi.org/10.1121/1.4803903>
- Wightman, F. L., & Kistler, D. J.** (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *The Journal of the Acoustical Society of America*, 118(5), 3164–3176. <https://doi.org/10.1121/1.2082567>