

Research Article

Linguistic Masking Release in School-Age Children and Adults

Lauren Calandruccio,^a Lori J. Leibold,^b and Emily Buss^c

Purpose: This study assessed if 6- to 8-year-old children benefit from a language mismatch between target and masker speech for sentence recognition in a 2-talker masker.

Method: English sentence recognition was evaluated for English monolingual children (ages 6–8 years, $n = 15$) and adults ($n = 15$) in an English 2-talker and a Spanish 2-talker masker. A regression analysis with subject as a random variable was used to test the fixed effect of listener group and masker language and the interaction of these two effects.

Results: Thresholds were approximately 5 dB higher for children than for adults in both maskers. However, children

and adults benefited to the same degree from a mismatch between the target and masker language with approximately 3 dB lower thresholds in the Spanish than the English masker.

Conclusions: Results suggest that children are able to take advantage of linguistic differences between English and Spanish speech maskers to the same degree as adults. Yet, overall worse performance for children may indicate general cognitive immaturity compared with adults, perhaps causing children to be less efficient when combining glimpses of degraded speech information into a meaningful sentence.

Children often have more difficulty understanding speech in noisy acoustic environments than adults (e.g., Elliott, Connors, Kills, & Levin, 1979). This child–adult difference tends to be larger when the background is a small number of competing talkers compared with nominally steady noise (e.g., Hall, Grose, Buss, & Dev, 2002; Leibold & Buss, 2013). For example, Hall et al. (2002) measured word recognition thresholds in 5- to 10-year-olds and adults; when the masker was continuous, children’s thresholds were 3 dB higher than adults’ in steady noise and almost 7 dB higher than adults’ in two-talker speech. This is an important finding as children’s learning often takes place in noisy environments (Bhardwaj et al., 2013; Bradley & Sato, 2008; Knecht, Nelson, Whitelaw, & Feth, 2002; Shield & Dockrell, 2004) in which multiple people may be speaking at the same time. These data are also indicative that children’s ability to recognize masked speech matures at different rates for noise and speech maskers. In both children and adults, the ability to understand speech in a speech masker is affected by the relative spatial location (Johnstone & Litovsky, 2006), the presence or absence of

a carrier phrase (Bonino, Leibold, & Buss, 2013; Lynn & Brotman, 1981), the inclusion of visual cues (at least for children older than 9 years of age; Wightman, Kistler, & Brungart, 2006), and the speaking style of the target talker (Baker, Buss, Jacks, Taylor, & Leibold, 2014; Pittman & Wiley, 2001). Another factor shown to affect adults’ performance is if the target and masker are spoken in the same or different languages; adult listeners have been shown to take advantage of mismatched target and masker language combinations to improve their speech recognition performance (e.g., Freyman, Balakrishnan, & Helfer, 2001; Van Engen & Bradlow, 2007). The purpose of this experiment was to investigate whether school-age children (6- to 8-year-olds) are sensitive to target/masker language mismatches when trying to understand target speech embedded in competing speech.

It is not uncommon for people, adults and children alike, to have difficulty understanding speech in complex listening environments. This difficulty is often thought to be due to a combination of both energetic and informational masking: Performance is limited by overlapping excitation patterns in the auditory periphery associated with the target and competing auditory inputs (Miller, 1947) and by difficulties “hearing out” the target signal due to the confusion associated with the presence of multiple streams of auditory input (Bregman, 1990; Carhart, Tillman, & Greetis, 1968; Durlach, Mason, Kidd, et al., 2003; Watson, 2005). For speech-on-speech listening conditions, adult listeners with normal hearing show reduced masking for

^aCase Western Reserve University, Cleveland, OH

^bCenter for Hearing Research, Boys Town National Research Hospital, Omaha, NE

^cThe University of North Carolina at Chapel Hill

Correspondence to Lauren Calandruccio: lauren.calandruccio@case.edu

Editor and Associate Editor: Larry Humes

Received September 10, 2015

Revision received November 16, 2015

Accepted December 8, 2015

DOI: 10.1044/2015_AJA-15-0053

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

target and masker speech combinations that are mismatched in language (e.g., Van Engen & Bradlow, 2007).

Initial studies aimed at investigating the effect of target/masker language mismatches used a foreign (unfamiliar to the listener) language in the mismatched masker condition (Dirks & Bower, 1969; Freyman et al., 2001; Garcia Lecumberri & Cooke, 2006; Rhebergen, Versfeld, & Dreschler, 2005; Tun, O’Kane, & Wingfield, 2002; Van Engen & Bradlow, 2007). For example, Freyman et al. (2001) evaluated the masker effectiveness of two-talker Dutch and two-talker English maskers for listeners who spoke English but not Dutch. That experiment used perceived spatial separation (Freyman, Helfer, McCall, & Clifton, 1999) to determine differences in masker effectiveness due to energetic and informational masking contributions. For an English sentence recognition task, the Dutch masker was less effective than the English masker provided the target and masker were colocated. However, in the perceived spatial separation condition, the two maskers were equally effective. The authors concluded that the difference in effectiveness between the two maskers in the colocated condition was due to informational rather than energetic masking contributions.

Early studies showing a reduced effectiveness associated with foreign language maskers were consistent with the hypothesis that the benefit of a target/masker linguistic mismatch was due to the fact that only the target is understandable (Garcia Lecumberri & Cooke, 2006; Rhebergen et al., 2005). However, later studies showed that the benefit of a target/masker language mismatch could be observed even when listeners spoke both languages fluently (Brouwer, Van Engen, Calandruccio, & Bradlow, 2012; Calandruccio & Zhou, 2014; Van Engen, 2010). Calandruccio and Zhou (2014) reported data for monolingual and bilingual listeners in a target/masker mismatch experiment that included English target speech and competing speech spoken in either English or Greek. The monolingual listener group had no experience with the Greek language, whereas the bilingual group was fluent in both English and Greek. Listeners in the bilingual group were considered simultaneous bilinguals—that is, they acquired both of their languages prior to age 2 (Bialystok, 2001). The monolingual and simultaneous bilingual listeners benefitted from target/masker language mismatches to a similar degree. Results are different for sequential bilinguals, who learn their second language after the acquisition of their first language (Bialystok, 2001). For sequential bilinguals, there is reduced benefit (Brouwer et al., 2012; Van Engen, 2010) or no benefit (Garcia Lecumberri & Cooke, 2006) of a target/masker language mismatch when the target is presented in their second language. This effect could be related to poorer overall performance for masked speech perception for the second language learned by a sequential bilingual (e.g., Rogers, Lister, Febo, Besing, & Abrams, 2006).

In combination, the data described above indicate that informational masking is greater when the target and masker languages are matched. Further, it appears that the listeners’ experience with the target language can modulate

the degree to which they can benefit from a target/masker language mismatch. However, it also appears that the understandability of the masker speech is not the sole factor responsible for the target/masker mismatch language benefit. Both psychophysical and speech perception data indicate that the more perceptually dissimilar two competing streams are, the easier it will be to separate them (Festen & Plomp, 1990; Moore & Gockel, 2012). It is possible that the benefit of a target/masker language mismatch in adults is due to enhanced stream segregation due to the acoustic and/or phonetic differences between the two languages. Following the work of Bregman (1990), these differences could be primitive, relying on low-level acoustic cues, or schema-based, relying on higher-level linguistic knowledge. For the case of speech-on-speech recognition, it is likely that both mechanisms could play a role in the benefit of target/masker language mismatches.

The finding that nonnative speakers of the target language obtain less benefit of a target/masker language mismatch than native speakers is consistent with the idea that “schema-based” segregation plays a substantial role in this effect. It is therefore possible that young children could benefit to a lesser degree than adults due to their relative linguistic inexperience and the fact that they are still learning about speech and language. To evaluate the role of auditory development with respect to the ability to utilize a target/masker language mismatch cue, we tested young, school-aged children (ages 6–8 years) on an open set, sentence-recognition task. This age range was chosen so that we could use similar methodology as used in previous studies (Brouwer et al., 2012; Calandruccio & Zhou, 2014; Van Engen & Bradlow, 2007) while ensuring reliable scoring of the open set speech productions from our younger listeners.

Methods

Participants

Thirty listeners participated in this experiment, 15 adults and 15 children. The 15 adults who participated (11 women, four men) ranged in age from 20 to 35 years old. The 15 children who participated (six girls, nine boys) ranged in age from 6.10 to 8.02 years old. All participants were monolingual, native speakers of American English. All children were typically developing and had normal speech and language development by parent report. None of the participants reported familiarity with Spanish. All participants had audiometric thresholds at octave frequencies between 250 and 8000 Hz within normal limits (equal to or less than 20 dB HL) bilaterally (American National Standards Institute, 2010). All participants provided informed consent in accordance with the institutional review board at the University of North Carolina at Chapel Hill and were paid for their participation.

Stimuli

The target stimuli were recordings of the Revised Bamford-Kowal-Bench (BKB) Standard Sentence Test (Bench, Kowal, & Bamford, 1979) spoken by a female talker. The

BKB sentences were originally developed using a lexicon derived from the speech of 240 children, ages 8 to 15 (Bench et al., 1979), making these materials appropriate for pediatric testing. The BKB corpus includes 21 lists of 16 sentences, each with three to four key words, for a total of 50 key words/list. For each list, two of the 16 sentences have an even number (four) of key words. The remaining 14 sentences have three key words each. An example of a three-key word BKB sentence is “The CLOWN had a FUNNY FACE” (key words in capital letters). The talker, a native speaker of American English, was instructed to speak in a natural style of speech as if she were having a conversation with a friend. She produced the BKB sentences one at a time as they appeared on a computer screen, speaking into a Shure SM81 Condenser microphone attached to a MOTU Ultralight A/D convertor. Sentences were recorded in a double-walled, sound-treated booth using a sampling rate of 22 kHz with 24-bit resolution at Northwestern University, and each sentence was saved to a .wav file (as used in Calandruccio, Van Engen, Dhar, & Bradlow, 2010; Van Engen, 2010). These recordings were root-mean-square normalized using Praat (Boersma & Weenick, 2012).

Masker stimuli included recordings of two female talkers each speaking in either English or Spanish. Both talkers were simultaneous English and Spanish speakers who grew up in Spanish-English bilingual households (for a more complete description, see Calandruccio, Gomez, Buss, & Leibold, 2014). The rationale for using the same two talkers to create both the English and Spanish masker stimuli was to minimize the spectral differences between maskers in the two languages. Both talkers consistently used both English and Spanish in their daily lives and reported being equally proficient in their reading, speaking, and listening abilities in English and Spanish. The English and Spanish masker stimuli were composed of passages from the story *Jack and the Beanstalk* and the Spanish translation of this story, *Juan y los Frijoles Mágicos* (Walker, 1999a, 1999b), respectively. The two talkers were recorded separately, each reading a different passage from each book; the selection of different passages for each talker prevented repetition of text when the two single-talker streams were subsequently summed. Silent periods greater than 300 ms were digitally edited using SoundStudio audio software and reduced to 100 ms. The recordings were then root-mean-square normalized using Praat. The recordings in each language were summed, resulting in an English two-talker masker and a Spanish two-talker masker.

Although the same two talkers were used to create the English and Spanish masker stimuli, there was still a visually observable (albeit slight) difference in the long-term-average speech spectra (LTASS) between the two maskers between approximately 3500 and 4500 Hz (see figure 1 in Calandruccio, Gomez, et al., 2014). Even though this difference was not perceptually salient, it has been shown that spectral differences between linguistic maskers can affect performance in two-talker masker conditions (Calandruccio, Dhar, & Bradlow, 2010). Therefore, the LTASS of the two maskers were normalized using MATLAB. This was

completed by determining the LTASS of the English and Spanish two-talker maskers, using a fast Fourier analysis on 2,048-point Hamming-windowed samples, and then computing the average magnitude spectrum of each masker. These data were then used to compute a grand average LTASS for each masker. The grand average was then used to normalize both the English and Spanish two-talker masker magnitude spectra (see Brouwer et al., 2012).

Temporal differences between the English and Spanish two-talker maskers were also evaluated because it has also been shown that differences in amplitude modulation patterns can cause differences in masker effectiveness regardless of the linguistic content of the masker speech (Calandruccio, Dhar, & Bradlow, 2010). The cumulative distribution of the filtered envelope values of the English and Spanish two-talker maskers was shown to be nearly identical. The cumulative distribution values were based on the Hilbert envelopes of the two maskers, which were low-pass filtered using a second-order Butterworth filter with a 40-Hz cutoff. This provided a quantitative evaluation of the masker envelope minima that were available to the listener because differences in the proportion of relatively low envelope values would indicate variance in the opportunity for “dip-listening” across the two maskers (Festen & Plomp, 1990).

Procedure

All participants were seated comfortably in a sound-isolated room. They were instructed to repeat the sentence they heard spoken by the target talker while trying to ignore the speech of the competing talkers. All listeners were familiarized with the task and completed several practice tracks to ensure familiarity with both the task and the voice of the target talker. Participants listened to each of the two maskers prior to testing (in a random order) during the familiarization phase, which included two threshold estimates in each of the maskers. For both practice and test tracks, the starting level of the target speech was +15 dB SNR and +5 dB SNR for children and adults, respectively. The selection and presentation of the test stimuli were controlled by a custom MATLAB program. Stimuli were mixed digitally (TDT RZ6) and presented diotically via Sennheiser HD25 II supra-aural headphones. The two-talker masker was 65 dB SPL, and the level of the target sentence was adjusted on the basis of the listener’s responses. The two-talker masker began 500 ms prior to the start of each sentence and ended at least 500 ms after the target sentence.¹ A simple up-down adaptive track estimated key word identification thresholds corresponding to 50% correct identification. An examiner, blind to the experimental hypothesis, scored listener responses. Each word was scored correct if it was repeated exactly as written in the stimulus materials.

¹Due to a programming error, the masker was gated on and off abruptly with no temporal smoothing. The associated splatter was not perceptually salient. Given the relatively long time interval between onset of the masker and presentation of the target (500 ms), it seems unlikely that omission of masker gating was of consequence for the results.

Any deviation from the stimulus transcript (e.g., addition or omission of a plural morpheme, tense change, etc.) resulted in the word being scored as incorrect. A sentence was considered correct when more than half of the key words were repeated correctly and incorrect when fewer than half of the key words were repeated correctly. If the listener responded correctly to exactly half of the key words, the program randomly categorized the sentence as either correct or incorrect with equal probability; this occurred infrequently. Level adjustments of the target sentences were made in 4-dB steps for the first two track reversals; 2-dB steps were used thereafter. Eight reversals were obtained for each track. Thresholds were estimated as the average level of the target sentence for the last six reversals. No sentences were repeated during testing.

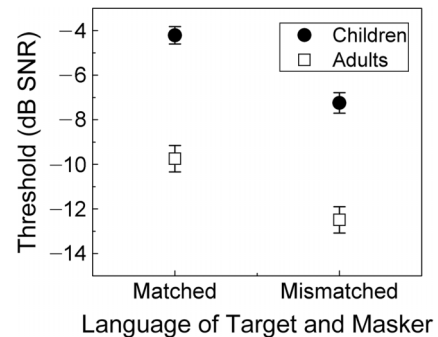
Both adults and children completed two conditions: (a) English target sentences in a two-talker English masker, and (b) English target sentences in a two-talker Spanish masker. The order of the masker language was randomized across listeners. For both the English and Spanish two-talker maskers, thresholds were estimated on the basis of the average of two tracks. If the thresholds of the two tracks were more than 2-dB different, a third track was completed, and the threshold estimate was based on the two tracks that had the most similar thresholds. For the English two-talker masker, a third track was collected for 20% of threshold estimates for both the adult and children listeners. For the Spanish two-talker masker, a third track was collected for 46% and 40% of threshold estimates for adult and children listeners, respectively.

Results

Results are reported on the basis of the estimated signal-to-noise ratio (SNR) of the target relative to the masker speech needed to obtain 50% correct. A regression analysis with subject as a random variable was conducted, testing the fixed effect of listener group and masker language and the interaction of these two effects. Results indicated a significant effect of group, $F(1, 28) = 67.40, p < .0001$, and of masker language, $F(1, 28) = 79.76, p < .0001$, but no significant interaction, $F(1, 28) = 0.20, p = .6589$. The data shown in Figure 1 indicate that children needed a more advantageous SNR than adult listeners in both conditions to achieve 50% correct recognition (mean of -11.11 dB SNR [$SD = 2.66$] and -5.72 dB SNR [$SD = 2.24$] for adults and children, respectively). Further, for both groups of listeners, the Spanish two-talker masker was less effective than the English two-talker masker (mean of -6.98 dB SNR [$SD = 3.41$] and -9.87 dB SNR [$SD = 3.34$] for the English and Spanish masker, respectively), resulting in lower thresholds in the Spanish masker condition. Mean data for both listener groups are shown in Figure 1.

For 13 of the 15 adults, thresholds were lower in the Spanish than the English masker. Thresholds in the two maskers differed by 2.75 dB on average (range = -0.67 to 7.17 dB). This is consistent with previously observed effects of target/masker language mismatch in adults (Rhebergen

Figure 1. Mean signal-to-noise ratio associated with 50% correct English sentence recognition thresholds for children (filled circles) and adults (unfilled squares). Data are shown for matched (English target and English masker) and mismatched (English target and Spanish masker) target/masker language conditions. Error bars represent one standard error of the mean.



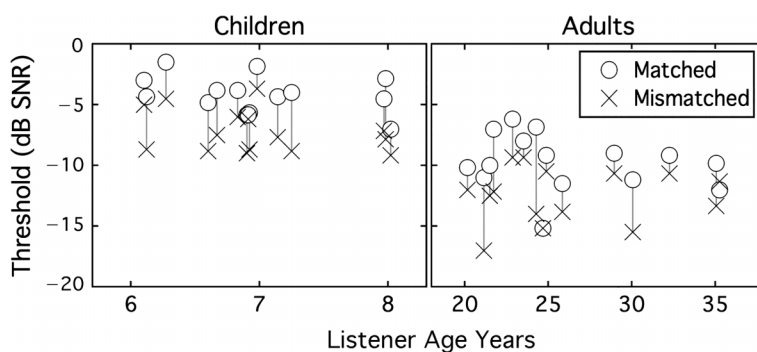
et al., 2005). For all 15 children, thresholds were lower in the Spanish than the English masker with a mean difference between the masker conditions of 3.03 dB (range = 0.34 to 5.00 dB; see Figure 2 for individual data). On average, children's thresholds were 5.54 dB higher than adults' for the English masker and 5.25 dB higher than adults' for the Spanish masker. This child-adult threshold difference is consistent with previously reported data for masked sentence recognition (Hall, Buss, Grose, & Roush, 2012; note, however, that Hall et al. used a modulated noise instead of a speech-based masker). The data of both child and adult listeners highlight the individual variability often observed for listening tasks characterized by informational masking (Kidd, Mason, Deliwala, Woods, & Colburn, 1994).

Discussion

The present study was used to evaluate child-adult differences in the benefit associated with introducing a target/masker language mismatch for a speech-on-speech recognition task. The targets were English sentences, and the two-talker masker was either English or Spanish. For 6- to 8-year-old children, thresholds were approximately 3 dB lower for the Spanish (mismatched language) than the English (matched language) masker. This benefit, associated with the target/masker language mismatch, was comparable to that obtained for adults. As expected, children's overall performance was worse than adults', an effect of approximately 5 dB for both maskers. Common to many reports on informational masking, large individual differences were observed (e.g., Durlach, Mason, Shinn-Cunningham, et al., 2003; Kidd et al., 1994) with some listeners improving by more than 5 dB when a target/masker language mismatch was introduced and others showing similar thresholds between masker conditions.

The present results are consistent with the idea that children are as proficient as adults at taking advantage of linguistic differences between English and Spanish speech

Figure 2. Individual signal-to-noise ratio thresholds associated with 50% correct English sentence recognition for children (left panel) and adult (right panel) listeners. Data for the matched (English) masker are shown using os, and data for the mismatched (Spanish) masker are shown using xs. The lines between the two data points indicate individual benefit of having the target and masker language mismatched.



streams. However, overall worse performance for children may reflect general cognitive immaturity, making children less efficient at combining degraded speech information into a coherent message (see also Hall et al., 2012). Further research is needed to better understand which aspects of cognitive maturation are important for this type of listening task. For example, children may be less able to inhibit attention to the competing speech, causing overall worse performance; support for this possibility comes from studies of the developmental trajectory of executive function (Reetzke, Maddox, & Chandrasekaran, 2016). It is also possible that children's reduced linguistic experience with the target language reduces their ability to use syntactical cues to improve their overall sentence recognition score. The latter possibility seems somewhat unlikely as large child–adult differences in two-talker maskers have also been observed for closed set, word-identification tasks that require less linguistic comprehension (Hall et al., 2002) and because the BKB sentences used in this study are linguistically age appropriate for our listener group (Bench et al., 1979).

It remains unclear why children are able to take advantage of the language mismatch as efficiently as adult native listeners, yet adult nonnative speakers of English generally show less benefit with this type of target/masker mismatch (Calandruccio, Bradlow, & Dhar, 2014; Garcia Lecumberri & Cooke, 2006; Van Engen, 2010; Van Engen & Bradlow, 2007). It has been suggested that speech-on-speech recognition can be affected not only by high-level cues (e.g., lexicon, syntax; Mayo, Florentine, & Buus, 1997; van Wijngaarden, Steeneken, & Houtgast, 2002), but also by low-level cues (e.g., phonetic or vocalic context; Cutler, Garcia Lecumberri, & Cooke, 2008; Garcia Lecumberri & Cooke, 2006). Perhaps by the age of 6, monolingual, normal-hearing children have a well-defined representation of low-level acoustic speech cues and enough high-level cues to efficiently separate English target speech from a Spanish-language masker whereas, for nonnative adult listeners, their native language sound system may limit the extent to which they can distinguish the target from the mismatched language masker speech (Best, McRoberts, & Goodwell,

2001). Further research is needed to explore how children develop the auditory skills needed to efficiently separate linguistically mismatched target/masker speech and whether the time course of development for this ability differs for monolingual and multilingual children.

Gross spectral and temporal differences between the two different maskers were controlled in this experiment (see Calandruccio, Gomez, et al., 2014), reducing the possibility that differences in energetic masking and “dip-listening” opportunities (Festen & Plomp, 1990) were responsible for the observed masker effects. In fact, Calandruccio, Gomez, et al. reported preliminary data using these specific English and Spanish maskers and showed that the two maskers were equally effective maskers for speech recognition tasks with linguistically matched targets (i.e., English targets with English masker and Spanish targets with Spanish masker). Nevertheless, other temporal differences between the two maskers may have contributed to the reduced masking effectiveness of the Spanish masker in combination with the English target speech (e.g., difference in syllable structure and/or rhythmic patterns between the two languages; see Reel & Hicks, 2012). If listeners were indeed benefitting from differences between the rhythms of different languages when the target and masker languages were mismatched, we would predict that languages that are similar in rhythm would be associated with less target/masker language mismatch benefit relative to languages that are more distinct in rhythm (see Calandruccio, Brouwer, Van Engen, Dhar, & Bradlow, 2013, for a preliminary exploration of this question).

Primitive segregation, defined as segregation on the basis of cues that do not need to be learned (Bregman, 1990), could benefit listeners in a speech-on-speech recognition task in which the language of the target and masker speech are perceptually different. For example, newborns have been shown to distinguish between differences in rhythmic properties between languages (Mehler, Jusczyk, Lamsertz, & French, 1988; Nazzi, Bertoncini, & Mehler, 1998). If listeners rely on a primitive rhythmic difference between the target and masker language to enhance stream

segregation, we would predict that children and infants would benefit from the target and masker mismatch to a similar extent as adults. However, primitive grouping cues that facilitate the separation of target and masker speech in the mismatched conditions are not likely the only contributing factors as data from nonnative speakers of the target language indicate less benefit in these mismatched conditions (Brouwer et al., 2012; Van Engen, 2010). These results suggest that learned properties of the language, such as vocabulary, syntactic structures, and prosodic intonation, could also facilitate separating the two streams. Further research is needed to explore the time span over which the ability to use schema-based cues in speech-on-speech segregation develops and how listeners are able to combine both primitive and schema-based auditory cues to improve speech recognition and reduce informational masking.

Acknowledgment

Support provided by the National Institutes of Health Grant R01 DC011038 (awarded to Lori J. Leibold).

References

- American National Standards Institute.** (2010). *Specifications for audiometers* (ANSI S3.6-2010). New York, NY: Author.
- Baker, M., Buss, E., Jacks, A., Taylor, C., & Leibold, L. J.** (2014). Children's perception of speech produced in a two-talker background. *Journal of Speech, Language, and Hearing Research, 57*, 327–337.
- Bench, J., Kowal, A., & Bamford, J.** (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology, 13*, 108–112.
- Best, C. T., McRoberts, G. W., & Goodwell, E.** (2001). Discrimination of nonnative consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America, 109*, 775–794.
- Bhardwaj, M., Baum, U., Markevych, I., Mohamed, A., Weinmann, T., Nowak, D., & Radon, K.** (2013). Are primary school students exposed to higher noise levels than secondary school students in Germany? *The International Journal of Occupational and Environmental Medicine, 4*, 2–11.
- Bialystok, E.** (2001). *Bilingualism in development*. Cambridge, United Kingdom: Cambridge University Press.
- Boersma, P., & Weenick, D.** (2012). Praat: Doing phonetics by computer [computer program]. Retrieved from <http://www.praat.org>
- Bonino, A. Y., Leibold, L. J., & Buss, E.** (2013). Release from perceptual masking for children and adults: Benefit of a carrier phrase. *Ear and Hearing, 34*, 3–14.
- Bradley, J. S., & Sato, H.** (2008). The intelligibility of speech in elementary school classrooms. *The Journal of the Acoustical Society of America, 123*, 2078–2086.
- Bregman, A. S.** (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: The MIT Press.
- Brouwer, S., Van Engen, K. J., Calandruccio, L., & Bradlow, A. R.** (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. *The Journal of the Acoustical Society of America, 131*, 1449–1464.
- Calandruccio, L., Bradlow, A. R., & Dhar, S.** (2014). Speech-on-speech masking with variable access to the linguistic content of the masker speech for native and nonnative English speakers. *Journal of the American Academy of Audiology, 25*, 355–366.
- Calandruccio, L., Brouwer, S., Van Engen, K. J., Dhar, S., & Bradlow, A. R.** (2013). Masking release due to linguistic and phonetic dissimilarity between the target and masker speech. *American Journal of Audiology, 22*, 157–164.
- Calandruccio, L., Dhar, S., & Bradlow, A. R.** (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. *The Journal of the Acoustical Society of America, 128*, 860–869.
- Calandruccio, L., Gomez, B., Buss, E., & Leibold, L.** (2014). Development and preliminary evaluation of a pediatric Spanish-English speech perception task. *American Journal of Audiology, 23*, 1–15.
- Calandruccio, L., Van Engen, K., Dhar, S., & Bradlow, A. R.** (2010). The effectiveness of clear speech as a masker. *Journal of Speech, Language, and Hearing Research, 53*, 1458–1471.
- Calandruccio, L., & Zhou, H.** (2014). Increase in speech recognition due to linguistic mismatch between target and masker speech: Monolingual and simultaneous bilingual performance. *Journal of Speech, Language, and Hearing Research, 57*, 1089–1097.
- Carhart, R., Tillman, T. W., & Greetis, E. S.** (1968). Perceptual masking in multiple sound backgrounds. *The Journal of the Acoustical Society of America, 45*, 694–703.
- Cutler, A., Garcia Lecumberri, M. L., & Cooke, M.** (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *The Journal of the Acoustical Society of America, 124*, 1264–1268.
- Dirks, D. D., & Bower, D. R.** (1969). Masking effects of speech competing messages. *Journal of Speech and Hearing Research, 12*, 229–245.
- Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G.** (2003). Note on informational masking. *The Journal of the Acoustical Society of America, 113*, 2984–2987.
- Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd, G.** (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *The Journal of the Acoustical Society of America, 114*, 368–379.
- Elliott, L. L., Connors, S., Kills, E., & Levin, S.** (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *The Journal of the Acoustical Society of America, 66*, 12–21.
- Festen, J. M., & Plomp, R.** (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America, 88*, 1725–1736.
- Freyman, R. L., Balakrishnan, U., & Helfer, K. S.** (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America, 109*(5, Pt. 1), 2112–2122.
- Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K.** (1999). The role of perceived spatial separation in the unmasking of speech. *The Journal of the Acoustical Society of America, 106*, 3578–3588.
- Garcia Lecumberri, M. L., & Cooke, M.** (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America, 119*, 2445–2454.
- Hall, J. W., Buss, E., Grose, J. H., & Roush, P. A.** (2012). Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear and Hearing, 33*, 340–348.
- Hall, J. W., Grose, J. H., Buss, E., & Dev, M. B.** (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing, 23*, 159–165.

- Johnstone, P., & Litovsky, R. Y.** (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *The Journal of the Acoustical Society of America*, *120*, 2177–2189.
- Kidd, G., Mason, C. R., Deliwalla, P. S., Woods, W. S., & Colburn, H. S.** (1994). Reducing informational masking by sound segregation. *The Journal of the Acoustical Society of America*, *95*, 3475–3480.
- Knecht, H. A., Nelson, P. B., Whitelaw, G. M., & Feth, L. L.** (2002). Background noise levels and reverberation times in unoccupied classrooms: Predictions and measurements. *American Journal of Audiology*, *11*, 65–71.
- 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*, 1144–1155.
- Lynn, J. M., & Brotman, S. R.** (1981). Perceptual significance of the CID W-22 carrier phrase. *Ear and Hearing*, *2*, 95–99.
- Mayo, L. H., Florentine, M., & Buus, S.** (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research*, *40*, 686–693.
- Mehler, J., Jusczyk, P., Lamsertz, G., & French, F.** (1988). A precursor to language acquisition in young infants. *Cognition*, *29*, 143–178.
- Miller, G. A.** (1947). The masking of speech. *Psychological Bulletin*, *44*, 105–129.
- Moore, B. C. J., & Gockel, H. E.** (2012). Properties of auditory stream formation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 919–931.
- Nazzi, T., Bertoncini, J., & Mehler, J.** (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 756–766.
- Pittman, A. L., & Wiley, T. L.** (2001). Recognition of speech produced in noise. *Journal of Speech, Language, and Hearing Research*, *44*, 487–496.
- Reel, L. A., & Hicks, C. B.** (2012). Selective auditory attention in adults: Effects of rhythmic structure of the competing language. *Journal of Speech, Language, and Hearing Research*, *55*, 89–104.
- Reetzke, R., Maddox, W. T., & Chandrasekaran, B.** (2016). The role of age and executive function in auditory category learning. *Journal of Experimental Child Psychology*, *142*, 48–65.
- Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A.** (2005). Release from informational masking by time reversal of native and non-native interfering speech. *The Journal of the Acoustical Society of America*, *118*, 1274–1277.
- Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B.** (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, *27*, 465–485.
- Shield, B., & Dockrell, J. E.** (2004). External and internal noise surveys of London primary schools. *The Journal of the Acoustical Society of America*, *115*, 730–738.
- Tun, P. A., O'Kane, G., & Wingfield, A.** (2002). Distraction by competing speech in young and older adult listeners. *Psychology and Aging*, *17*, 453–467.
- Van Engen, K. J.** (2010). Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble. *Speech Communication*, *52*, 943–953.
- Van Engen, K. J., & Bradlow, A. R.** (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America*, *121*, 519–526.
- van Wijngaarden, S. J., Steeneken, H. J. M., & Houtgast, T.** (2002). Quantifying the intelligibility of speech in noise for non-native talkers. *The Journal of the Acoustical Society of America*, *112*, 3004–3013.
- Walker, R.** (1999a). *Jack and the Beanstalk*. Cambridge, MA: Barefoot Books.
- Walker, R.** (1999b). *Juan y los Frijoles Mágicos*. Cambridge, MA: Barefoot Books.
- Watson, C. S.** (2005). Some comments on informational masking. *Acta Acustica United with Acustica*, *91*, 502–512.
- Wightman, F., Kistler, D., & Brungart, D.** (2006). Informational masking of speech in children: Auditory-visual integration. *The Journal of the Acoustical Society of America*, *119*, 3940–3949.