

Research Note

The Effect of Remote Masking on the Reception of Speech by Young School-Age Children

Carla L. Youngdahl,^a Eric W. Healy,^a Sarah E. Yoho,^a
Frédéric Apoux,^a and Rachael Frush Holt^a

Purpose: Psychoacoustic data indicate that infants and children are less likely than adults to focus on a spectral region containing an anticipated signal and are more susceptible to remote masking of a signal. These detection tasks suggest that infants and children, unlike adults, do not listen selectively. However, less is known about children's ability to listen selectively during speech recognition. Accordingly, the current study examines remote masking during speech recognition in children and adults.

Method: Adults and 7- and 5-year-old children performed sentence recognition in the presence of various spectrally remote maskers. Intelligibility was determined for each remote-masker condition, and performance was compared across age groups.

Results: It was found that speech recognition for 5-year-olds was reduced in the presence of spectrally remote noise, whereas the maskers had no effect on the 7-year-olds or adults. Maskers of different bandwidth and remoteness had similar effects.

Conclusions: In accord with psychoacoustic data, young children do not appear to focus on a spectral region of interest and ignore other regions during speech recognition. This tendency may help account for their typically poorer speech perception in noise. This study also appears to capture an important developmental stage, during which a substantial refinement in spectral listening occurs.

Children often experience poorer speech perception in noise, relative to adults (e.g., Buss, Hall, Grose, & Dev, 1999; Hall, Grose, Buss, & Dev, 2002; Leibold & Neff, 2007; Leibold & Werner, 2006; Leibold, Yarnell Bonino, & Buss, 2016; Oh, Wightman, & Lutfi, 2001; Papso & Blood, 1989). Many studies have shown that children require higher signal-to-noise ratios (SNRs) to recognize speech in noise at the same performance level as adults (e.g., Elliott et al., 1979; Litovsky, 2005; Wightman & Kistler, 2005). However, there is limited understanding

of why this difference exists. These difficulties in understanding speech in noise are of particular importance for children, because children must attend to important auditory information in noisy classroom environments. A comparison between children and adults on the effect of spectrally remote maskers on speech recognition might serve to clarify some of the abilities and limitations of child listeners.

Evidence from psychophysical studies indicates that adults are able to focus on a frequency at which a signal is expected to occur, as a strategy to improve detection (e.g., Dai, Scharf, & Buus, 1991; Scharf, Quigley, Aoki, Peachey, & Reeves, 1987; Schlauch & Hafter, 1991). In contrast, infants appear to employ a broader "listening strategy" and not attend specifically to a spectral region of interest (see Bargones & Werner, 1994). A common measure of frequency-selective listening is a modification of the probe-signal method (Greenberg & Larkin, 1968). In this method, a target signal tone is presented in a masker over many trials. In a small proportion of trials, an off-frequency probe—one having a frequency different from that of the target—replaces the target signal. Detection accuracy of these probes as a function of their relative frequency provides a measure of frequency-selective listening. Using this method, adults have demonstrated a highly selective listening strategy that closely follows the shape of the psychophysical

^aDepartment of Speech and Hearing Science, The Ohio State University, Columbus

Carla L. Youngdahl is now at the Department of Communicative Sciences and Disorders, Saint Mary's College, University of Notre Dame, IN.

Sarah E. Yoho is now at the Department of Communicative Disorders & Deaf Education, Utah State University, Logan.

Frédéric Apoux is now at the Department of Otolaryngology-Head & Neck Surgery, The Ohio State University, Columbus.

Correspondence to Carla L. Youngdahl: cyoungdahl@saintmarys.edu, and Eric W. Healy: healy.66@osu.edu

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Jennifer Lentz

Received April 3, 2017

Revision received September 1, 2017

Accepted September 7, 2017

https://doi.org/10.1044/2017_JSLHR-H-17-0118

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

auditory filters (Dai et al., 1991; Greenberg & Larkin, 1968; Ison, Virag, Allen, & Hammond, 2002; Scharf et al., 1987). This similarity in tuning suggests that adults are able to attend to the specific spectral region where a target is anticipated by focusing on the output of single auditory filters. However, this same selective listening strategy decreases detection for probe frequencies that are distant from the anticipated region.

Existing data suggest that the psychophysical filters that originate in the auditory periphery become adultlike early in life (Hall & Grose, 1991; Irwin, Stillman, & Schade, 1986; Schneider, Morrongiello, & Trehub, 1990; Soderquist, 1993), with estimates of peripheral maturity ranging as young as 5 months old (Olsho, 1985). In contrast, frequency-selective listening may develop more slowly with age, along with the developing central auditory system. There is limited information regarding when this frequency-selective listening strategy develops. Bargones and Werner (1994) showed that infants listen less selectively than adults. Using an off-frequency probe method similar to that described above, it was found that 7- to 9-month-olds displayed similar performance across targets and all off-frequency probes. These results suggest that, in sharp contrast to adults, infants are not primarily monitoring an anticipated frequency and instead appear to be monitoring a broad spectral region, despite having adultlike peripheral psychophysical auditory filters. In an earlier study also using a probe-signal method, Greenberg, Bray, and Beasley (1970) measured frequency-selective listening in five school-age children (6–8 years). Detection of a target signal (1000 Hz) and four probe signals (850, 925, 1075, and 1150 Hz) was measured in noise. Results showed that one child demonstrated adultlike performance, whereas the other four demonstrated a broader listening bandwidth than the adult controls—a pattern more similar to that of infants.

Related psychophysical work has shown that infants and children, but not adults, are susceptible to masking by noise that is remote in frequency to a signal. Werner and Bargones (1991) showed that 4- to 10-kHz noise could elevate thresholds for a 1-kHz tonal signal by as much as 10 dB in infants, whereas adult thresholds were unaffected. Leibold and Neff (2011) extended these results to children. They found that 4- to 6-year-olds were susceptible to remote masking, with thresholds elevated relative to quiet. However, by the age of 7–9 years, the children performed more similarly to adults, with no threshold elevation in the presence of the remote masker. These authors suggested that their results might reflect a reduced ability to selectively attend to the signal frequency on the part of the youngest children. More recently, Leibold and Buss (2016) examined factors that may influence the psychophysical remote masking observed in children but not adults. Large increases in detection threshold were found for a 2-kHz tonal signal in the presence of a remote noise masker for children younger than 7 years, in accord with prior work and supporting the idea that young children have difficulty selectively attending to frequency regions containing a target signal. The most masking was found to be produced by a narrow-band masker that was

proximate in frequency to the signal. In addition, maskers that were gated with the signal produced more remote masking, in accord with a view that simultaneous sounds are more difficult to segregate than asynchronous-onset sounds.

This psychophysical work involving the detection of tonal signals at anticipated and unanticipated frequencies, or in the presence of remote noise maskers, can be related to the results of experiments involving the detection or recognition of speech stimuli in the presence of remote maskers. Polka, Rvachew, and Molnar (2008) examined the ability of infants to attend to speech sounds in the presence of irrelevant sounds. Infants discriminated phonemes either in isolation or mixed with a distractor signal (bird or cricket song). The distractor was high-pass filtered above 5 kHz, whereas the phonemes were low-pass filtered below 4 kHz. The distractor negatively influenced the infants' ability to discriminate the phonemes, despite the fact that the distractor did not spectrally overlap with speech. Newman, Morini, and Chatterjee (2013) tested the ability of infants to attend to their name in either spectrally overlapping or non-overlapping noise. It was found that infants listened longer to their own name than a different name in the presence of nonoverlapping noise, but not in overlapping noise. This result suggests the complementary finding that infants do display some ability to take advantage of the remoteness of a masker in a preference task.

Overall, the results of these psychophysical studies employing infants and children, and some of the speech data involving infants, together suggest that the developing auditory system is less able to selectively focus on a frequency region containing a signal and more susceptible to noise when the signal and masker are spectrally distinct, than is the fully developed system. However, few data exist regarding these frequency-selective listening effects for speech perception by young school-age children, and so, the developmental time course for this effect is unknown. Therefore, the aim of this study was to examine the effect of remote masking noise on speech recognition by young school-age children and adults. This extends the speech results involving infants (Newman et al., 2013; Polka et al., 2008) in an attempt to better map out the development of frequency-selective listening for speech and to potentially help explain the more general findings involving children's increased susceptibility to noise during speech recognition. In the current study, a wide-band remote masker was employed along with four narrow-band remote maskers having various frequency separations from the target speech, thus allowing an examination of remote-masker bandwidth and spectral proximity effects on speech intelligibility.

Method

Participants

Ten adult participants (all female, aged 19–25 years, mean age = 21 years) and 18 child participants were employed. All were native speakers of American English having no prior exposure to the sentences used. Child participants

were divided into subgroups of 5- and 7-year-olds, with nine in each group. The 5-year-old group (seven girls, two boys) had an age range from 5;0 to 5;10 years;months, with a mean age of 5;6 years;months. The 7-year-old group (two girls, seven boys) had an age range from 7;1 to 7;11 years;months, with a mean age of 7;6 years;months. All listeners had pure-tone audiometric thresholds of 20 dB HL or better at octave frequencies from 500 through 8000 Hz and 25 dB HL or better at 250 Hz (American National Standards Institute, 2004, 2010). The exception was one participant in the 5-year-old group who had thresholds on the day of test of 25 dB HL at 2000 and 4000 Hz and 30 dB HL at 8000 Hz, all in the right ear only. A subsequent examination indicated that this child's data fell well within those of his counterparts. Child participants demonstrated typical speech and language development as indicated by the Token Test for Children—Second Edition (McGhee, Ehrler, & DiSimoni, 2007) and informal observation by a licensed speech-language pathologist (author C. L. Y.). Participants were recruited from The Ohio State University and the surrounding Columbus, Ohio, metropolitan area and were compensated with course credit, money, and/or children's books for participating.

Stimuli

Target stimuli consisted of 120 Bamford–Kowal–Bench sentences (Bench, Kowal, & Bamford, 1979). Each sentence contains three to five key words, which were scored to represent intelligibility. Sentence stimuli (44.1 kHz, 16-bit resolution) were first bandpass filtered from 100 to 1500 Hz using a 2000-order finite-duration impulse response digital filter, implemented in MATLAB. The total root-mean-square level of each filtered sentence was then equated. Speech-shaped noise (SSN) was created by shaping a noise to match the long-term average amplitude spectrum (65,536-point Hanning-windowed fast Fourier transform with 97% temporal overlap) of the filtered, concatenated, and equated sentences. The noise was shaped using a 500-order arbitrary-response FIR digital filter in MATLAB. Filtered speech and matched SSN were mixed at 0-dB SNR for adults and at 5-dB SNR for children. This was done to approximately equate recognition performance across groups. For both groups, the filtered speech-plus-SSN mixture was set to 55 dBA.

A control condition included this filtered speech-plus-SSN mixture. Five additional conditions were created by combining the control stimulus with a spectrally remote noise. Five remote maskers were employed: a one-octave band (2500–5000 Hz) and four 1/3-octave bands (3000–3780, 4000–5040, 5000–6300, and 6000–7560 Hz). These remote maskers had steep filter slopes (2000-order FIR filters) and lower cutoff frequencies selected to ensure no overlap of peripheral excitation between the speech band and the remote masker, as indicated by the excitation pattern calculations of Moore, Glasberg, and Baer (1997). The lower-frequency cutoff of the lowest 1/3-octave band was set to 3000 Hz, rather than the 2500-Hz lower bound used for the octave band, because the narrower bandwidth but equal presentation

level of the 1/3-octave band resulted in an increased spectrum level and increased spread of excitation. Despite that all remote maskers had a bandwidth of at least 780 Hz, low-noise noise was employed for the maskers (see Hartmann & Pumplin, 1988; Pumplin, 1985) to minimize the differing amplitude modulations that can result from steep-filtered noise having differing bandwidths. The low-noise noise was created by iteratively dividing a Gaussian noise by its Hilbert envelope 100 times (as described in Healy & Bacon, 2006; Kohlrausch et al., 1997) in MATLAB. Remote maskers were each set to 75 dBA and were gated and mixed with the filtered speech-plus-SSN mixture.

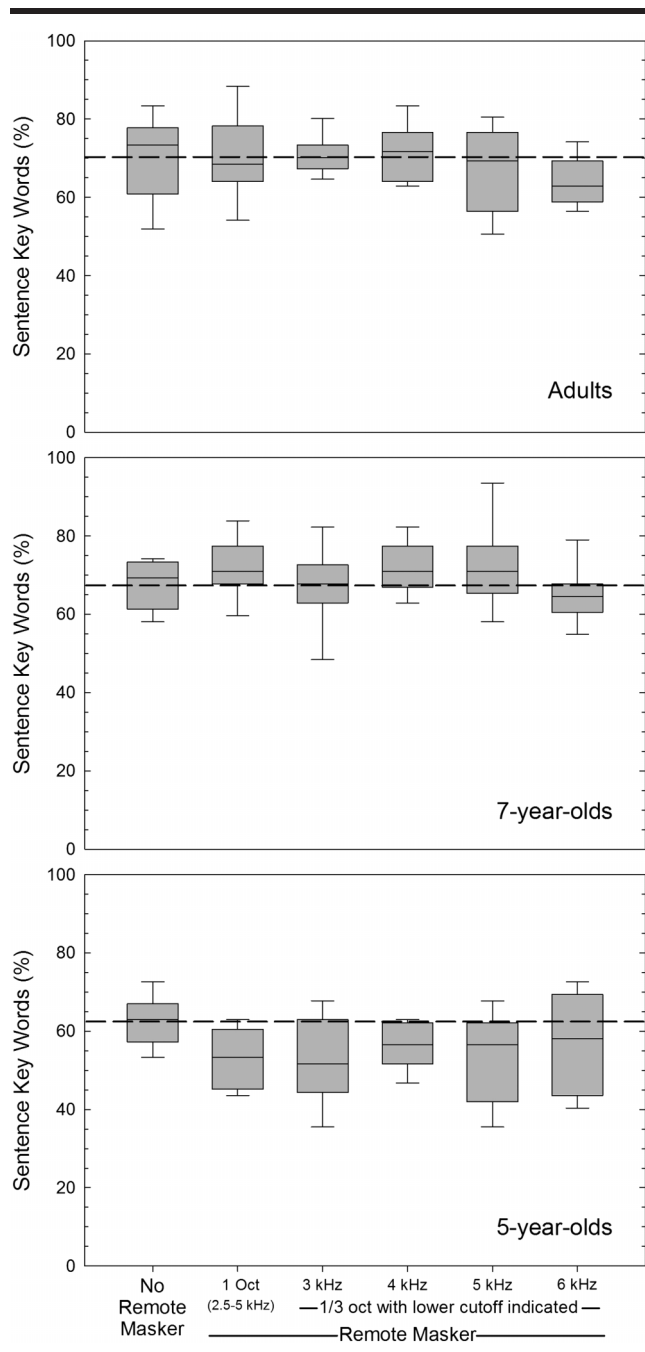
Procedure

Each block in this experiment was composed of the six conditions (control plus the five remote-masker conditions), with 10 sentences per condition. Participants heard two such blocks, with a new randomization of condition order for each block. Sentence list-to-condition correspondence was also balanced across listeners to control for possible differences in the difficulty of the 10-sentence filtered lists. Participants were instructed to listen to each sentence and repeat what was heard. Each sentence was preceded by the word “ready” to indicate that the listener should begin attending. Testing took place in a double-walled sound booth. Responses were collected by an examiner seated in the booth with the participant. Stimuli were played back from a PC, converted to analog form using Echo Digital Audio Gina 3G digital-to-analog converters, and presented diotically over Sennheiser HD280 headphones. Headphone levels were calibrated using a Larson Davis sound-level meter and flat-plate headphone coupler (Models 824 and AEC 101).

Results and Discussion

Figure 1 displays sentence intelligibility in each condition for each of the three participant groups. Adult performance (top panel) in the control condition averaged 70.2% ($SE = 3.4\%$). Performance in the remote masker conditions was within 2.4 percentage points of this value, with the exception of the most remote masker (6000–7560 Hz), which was 6.0 percentage points below performance in the control condition. The average for the pooled remote-masker conditions was within 1.2 percentage points of the control-condition mean. Mean 7-year-old performance (center panel) in the control condition was 67.6% ($SE = 2.0\%$). Performance in the remote-masker conditions all fell within 4.7 percentage points of this value, and the average for the pooled remote-masker conditions was within 2.0 percentage points of the control-condition mean. Mean 5-year-old performance (bottom panel) in the control condition was 62.5% ($SE = 2.0\%$). In contrast to what was seen in the adults and 7-year-olds, performance was reduced in the presence of all remote maskers. Mean speech intelligibility in remote maskers was reduced by 5.6–10.2 percentage points, relative to the control-condition mean. The average for the

Figure 1. Sentence intelligibility based on percent correct keyword recognition. Scores are shown for adults (top), 7-year-olds (center), and 5-year-olds (bottom) in a control condition containing no remote masker and in each of the five remote masker conditions. The boxes represent the 25th and 75th percentiles and contain the median, and the whiskers represent the 10th and 90th percentiles. The dashed line in each panel represents mean performance with the remote masker absent.



pooled remote-masker conditions was 8.1 percentage points below the control-condition mean.

An initial analysis was performed to compare baseline scores across age groups. Subject means were normalized

for variance using the rationalized arcsine unit transform (Studebaker, 1985), and a one-way analysis of variance was performed on the control-condition scores. No significant effect of age group was observed, $F(2, 25) = 2.30, p = .12$, indicating that baseline scores did not differ significantly across the three groups. To examine if differences existed in performance across the remote-masker conditions, one-way repeated-measures analyses of variance were performed on rationalized-arcsine-unit-transformed scores for each age group in the remote-masker conditions. No significant effect of remote-masker condition was observed for adults, $F(4, 36) = 1.30, p = .29$, 7-year-olds, $F(4, 32) = 1.70, p = .17$, or 5-year-olds, $F(4, 32) = 0.60, p = .67$. This result suggests that scores were equivalent irrespective of the bandwidth or spectral location of the remote masker. A planned paired t test was conducted for each age group to compare performance in the control condition (filtered speech plus SSN), with that averaged across the five remote-masker conditions. There was no significant difference in performance for the adults, $t(9) = 0.49, p = .64$, or the 7-year-olds, $t(8) = 1.03, p = .33$, suggesting that these groups were not affected by remote masking. The effect size (d) describing the standardized difference between means for the pooled remote maskers and control was -0.11 for the adults and 0.33 for the 7-year-olds. Thus, both effects were “small” (Cohen, 1988). However, the 5-year-old group did display a significant difference in performance between the control condition and that averaged across the five remote maskers, $t(8) = 2.58, p = .03$. The effect size for this group ($d = -1.31$) was “large” (Cohen, 1988) or “very large” (Sawilowsky, 2009). The negative effect-size value indicates poorer performance in the presence of the remote maskers.

These results suggest that speech recognition is susceptible to masking by off-frequency noise in 5-year-old children but not in 7-year-old children or adults. This result is in direct accord with the psychoacoustic data of Leibold and Neff (2011) and Leibold and Buss (2016), who found that 4- to 6-year-olds displayed elevated pure-tone detection thresholds in the presence of a remote noise masker but that older children and adults generally did not. When combined with these and other psychophysical results, it appears that the current remote-masking susceptibility of speech displayed by the youngest child participants may be due to their tendency to monitor across a broader range of auditory filters, which makes them less able to focus on spectral regions containing the signal and ignore spectral regions containing noise.

General Discussion

The aim of the current study was to assess remote masking of speech in young school-age children. The general hypothesis tested was that masking components producing no overlap of peripheral excitation with the speech signal will not affect adults but will be detrimental to young children, due to their inability to focus on a specific spectral region and exclude other regions. This hypothesis was based on prior psychoacoustic results demonstrating (a) the

considerable spectral tuning that adult listeners demonstrate when detecting tonal signals and the conversely broad monitoring displayed by infants and (b) the increased tone-detection thresholds in the presence of remote maskers displayed by infants and young children relative to adults. However, the link between these psychoacoustic data and the more complex task of speech recognition has not been made. More specifically, the prior results involving remote masking of speech have been constrained to infants, to our knowledge.

In the current study, it was found that masking by a spectrally remote noise band hindered the speech recognition performance of children aged 5 years, in apparent accord with some data on infants (Polka et al., 2008). However, remote masking did not hinder the performance of 7-year-old children, who displayed adultlike performance. It appears that an emerging narrowing of the listening bandwidth proved sufficient to overcome the influence of the remote maskers in older children. A similar interpretation is that young children require more of the speech signal (i.e., additional bandwidth) to understand it, and so for this reason, they necessarily employ a broader spectral listening range, making them unable to reject frequency regions containing noise, even if those regions are remote.

The current data suggest a substantial refinement in listening skills in 7-year-olds relative to 5-year-olds. These results involving a transition in speech recognition behavior from an infantlike listening pattern to an adultlike pattern around the age of 6–7 years converge well with the psychoacoustic data of Leibold and Neff (2011) and Leibold and Buss (2016), who found that remote masking increased tone-detection thresholds for children below this age but not generally above it. Together, the current speech-recognition results and the earlier psychoacoustic data suggest an important developmental period between 6 and 7 years old in which young children's listening patterns undergo marked changes that result in a more adultlike pattern.

Figure 1 shows that all remote maskers produced similar amounts of masking (or lack thereof) regardless of bandwidth or spectral separation from the target speech. This stands in contrast to the recent psychophysical results of Leibold and Buss (2016), who found that a narrow-band remote masker produced more masking of a tonal signal than a wide-band masker having the same lower frequency bound. The current results also contrast with the Leibold and Buss finding that narrow-band remote maskers that were more proximate in frequency to the signal produced more masking than maskers that were more distant. Methodological differences between the current speech study and this prior psychophysical study are considerable and could potentially account for the differing patterns of results. One fundamental difference involves the current use of suprathreshold speech recognition, rather than the detection of tonal signals.

Figure 1 also shows that the modification to SNR for the children versus adults had the desired effect of approximately equating intelligibility in the control (no-remote-masker) condition. Of course, it is difficult to equate sentence

intelligibility across different subject-group types with precision, so the differences in baseline intelligibility observed currently between the younger children and the older children (mean scores within 5.1 percentage points) and between the younger children and the adults (mean scores within 7.7 percentage points) can be considered slight. Statistically, they did not differ. However, it is still possible that the numerically lower baseline scores of the younger children made them more susceptible to corrupting influences (i.e., their speech perception might be more fragile generally due to lower scores) and that the effects observed currently are not specific to remote masking. However, this would require that the psychometric function relating intelligibility to information content be substantially steeper at 63% correct than it is at 68% or 70% correct, and existing functions generally do not show this characteristic (e.g., American National Standards Institute, 1969).

An eventual goal of this work is to better understand the overall speech recognition in noise abilities and limitations of children. The current results may potentially be related to the general finding that children experience poorer speech perception in noise relative to adults. More specifically, there may be a relationship between the inability to focus on a spectral region of interest and to ignore energy outside that region and overall poorer speech understanding in background noise. The connection comes from the glimpsing model of speech perception in noise (Cooke, 2006), in which the auditory system extracts and integrates speech information from relatively clean time–frequency portions of a speech-plus-noise mixture to understand speech in noise. Apoux and Healy (2009) showed that adults can process auditory filter outputs with considerable independence when recognizing speech in noise. In that study, speech and noise were filtered into 30 contiguous 1-ERB_N (Glasberg & Moore, 1990) frequency bands. When the speech and noise were present in spectrally interleaved and nonoverlapping bands, the noise had little effect on phoneme recognition. This suggests that the fully developed normal auditory system is able to process portions of the spectrum that are dominated by target speech with considerable effectiveness and essentially ignore those portions that are dominated by noise. In contrast, an inability to focus on portions of the spectrum that are dominated by speech and ignore other regions would likely impair this glimpsing process, thus impairing everyday speech recognition in noise.

However, the current remote-masking data cannot entirely capture the complexity of everyday noisy speech perception. The developmental trajectory for speech perception in noise is strikingly protracted with improvements observed until at least 10 years old and possibly through the teenage years (Eisenberg, Shannon, Martinez, Wygonski, & Boothroyd, 2000; Elliott & Katz, 1980; Elliott, Longinotti, Clifton, & Meyer, 1981; Johnson, 2000; Leibold et al., 2016). This trajectory may also depend on noise type, with evidence that adultlike performance occurs earlier for SSN than for two-talker maskers (Corbin, Bonino, Buss, & Leibold, 2016). Thus, the time course for everyday noisy speech perception differs from that observed currently for remote

masking of speech (and in psychoacoustic work), which appears to become adultlike by the age of 7 years. Furthermore, the SNR in the current study was more favorable for both groups of child participants than for adults. So, despite the adultlike remote masking pattern displayed by the 7-year-olds, other hindrances to everyday speech-in-noise perception remain.

One possible explanation for the different time courses observed for remote masking versus everyday speech-in-noise perception is that the spectral listening effects observed currently represent an early stage of inhibition development, which contribute to but do not entirely account for children's increased susceptibility to noise. There is evidence from electrophysiological studies that the efferent or neural inhibitory auditory pathways continue to develop throughout childhood (e.g., Ubiali, Sanfins, Borges, & Colella-Santos, 2016; van Dinteren, Arns, Jongsma, & Kessels, 2014). Immaturity in these efferent pathways may potentially reduce the ability to disregard or inhibit processing of extraneous masking noise. These effects could even be related to more general neurocognitive inhibition. This is part of the umbrella concept of executive function, which is centered in the prefrontal system and contains brain circuits associated with selective attention, working memory, and emotion processing. Compared with other areas of the brain, the prefrontal system has a protracted postnatal course of development, maturing well into adolescence and early adulthood (e.g., Ciccia, Meulenbroek, & Turkstra, 2009). The role of neurocognitive function in the perception of degraded speech is supported by correlations between the two, as observed for adults (e.g., Clayton et al., 2016) and children (e.g., Roman, Pisoni, Kronnberger, & Faulkner, 2017).

In accord with a possible relationship to general neurocognitive processes, the current remote-masking effects may be related to a modality-independent development of attention. Plebanek and Sloutsky (2017) had 4- to 5-year-old children and adults perform a visual task involving shape-change detection. The adults outperformed the children on shapes that had been cued before each trial, but the children outperformed the adults on uncued shapes. In a visual search task, the memory for search-relevant and search-irrelevant information was assessed. The children outperformed the adults on remembering search-irrelevant features of the visual diagrams. It was suggested that adults employ selective visual attention, whereas children appear to display distributed attention. Furthermore, it was pointed out that selective attention has clear benefits, but it also has critical costs.

There are important implications for understanding the listening strategies employed by young children. Children are often asked to learn and perform in noisy environments. On average, there are 22 children in a typical kindergarten classroom (National Center for Education Statistics, 1993). The age range from 5 to 7 years is critical for formal reading instruction in schools, which often targets phonemic awareness and phonics skills as the building blocks to reading. These are auditory activities that may be negatively impacted by noisy environments, among other things (e.g.,

hearing loss, impoverished phonological encoding; see, e.g., Nittrouer, Sansom, Low, Rice, & Caldwell-Tarr, 2014). An understanding that preschool and young school-age children are unable to use adultlike focused and efficient listening strategies motivates the improvement of the acoustic learning environment. These apparent, normally occurring developmental limitations support the implementation of improved room acoustics, classroom amplification systems, and decreased class sizes as techniques to improve learning and reduce listening fatigue. It should also be noted that the current results are from children with typical hearing and language skills. Any impairment in these areas could interact with normal decreased listening efficiency to produce an even greater negative impact of noise on speech recognition.

Conclusions

The results of the current study may contribute to our understanding of speech and noise processing and how it develops in the early school-age years. Young children, perhaps like infants, seem unable to disregard noise during speech recognition, even when that noise does not overlap in the frequency domain with the target speech. Instead, they may monitor more broadly over many auditory filters, including when doing so is not advantageous.

The current comparison between adults and young children with regard to speech recognition in remote masking noise cannot entirely account for but may help explain why children typically display poorer speech recognition performance in noise. As with the broader peripheral auditory filters associated with sensorineural hearing impairment, children's broader spectral attending may lead to a less effective rejection of background noise. Consequently, this allows the background to interfere to a greater extent with detection and recognition of sounds, including speech. Finally, the current results seem to capture what may be an important developmental period, as the ability to ignore spectrally remote noise appears to transition toward adultlike performance between the relatively narrow age range of 5–7 years.

Acknowledgments

This work formed a portion of a dissertation submitted by the first author, under the direction of the second author, in partial fulfillment of degree requirements for the PhD in Speech and Hearing Science. This research was supported in part through grants from the National Institutes of Health (NIDCD R01DC008594 and R01 DC015521 to author Eric W. Healy and R01 DC014956 to author Rachael Frush Holt). Additional funding came from The Ohio State University Alumni Grants for Graduate Research and Scholarship Program to author Carla L. Youngdahl. Portions were presented at the 2015 Annual American Speech-Language-Hearing Association Convention and the 169th Meeting of the Acoustical Society of America. The authors thank Brittney Carter for assistance in collecting the data and Jordan Vasko for conducting the excitation pattern calculations and assisting with manuscript preparation.

References

- American National Standards Institute.** (1969). *Methods for the calculation of the articulation index* (ANSI S3.5-1969). New York, NY: Author.
- American National Standards Institute.** (2004). *Methods for manual pure-tone threshold audiometry* (ANSI S3.21-2004, R2009). New York, NY: Author.
- American National Standards Institute.** (2010). *Specification for audiometers* (ANSI S3.6-2010). New York, NY: Author.
- Apoux, F., & Healy, E. W.** (2009). On the number of auditory filter outputs needed to understand speech: Further evidence for auditory channel independence. *Hearing Research*, *255*, 99–108.
- Bargones, J. Y., & Werner, L. A.** (1994). Adults listen selectively; infants do not. *Psychological Science*, *5*, 170–174.
- 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.
- Buss, E., Hall, J. W., III, Grose, J. H., & Dev, M. B.** (1999). Development of adult-like performance in backward, simultaneous, and forward masking. *Journal of Speech, Language, and Hearing Research*, *42*, 844–849.
- Ciccia, A. H., Meulenbroek, P., & Turkstra, L. S.** (2009). Adolescent brain and cognitive developments: Implications for clinical assessment in traumatic brain injury. *Topics in Language Disorders*, *29*, 249–265.
- Clayton, K. K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., & Kidd, G., Jr.** (2016). Executive function, visual attention and the cocktail party problem in musicians and non-musicians. *PLoS One*, *11*, 1–17.
- Cohen, L.** (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Routledge.
- Cooke, M.** (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*, *119*, 1562–1573.
- Corbin, N. E., Bonino, A. Y., Buss, E., & Leibold, L. J.** (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear and Hearing*, *37*, 55–63.
- Dai, H. P., Scharf, B., & Buus, S.** (1991). Effective attenuation of signals in noise under focused attention. *The Journal of the Acoustical Society of America*, *89*, 2837–2842.
- Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., & Boothroyd, A.** (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America*, *107*, 2704–2710.
- Elliott, L. L., Connors, S., Kille, E., Levin, S., Ball, K., & Katz, D.** (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *The Journal of the Acoustical Society of America*, *66*, 12–21.
- Elliott, L. L., & Katz, D. R.** (1980). Children's pure-tone detection. *The Journal of the Acoustical Society of America*, *67*, 343–344.
- Elliott, L. L., Longinotti, C., Clifton, L., & Meyer, D.** (1981). Detection and identification thresholds for consonant-vowel syllables. *Perception & Psychophysics*, *30*, 411–416.
- Glasberg, B. R., & Moore, B. C. J.** (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*, 103–138.
- Greenberg, G. Z., Bray, N. W., & Beasley, D. S.** (1970). Children's frequency-selective detection of signals in noise. *Perception & Psychophysics*, *8*, 173–175.
- Greenberg, G. Z., & Larkin, W. D.** (1968). Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: The probe-signal method. *The Journal of the Acoustical Society of America*, *44*, 1513–1523.
- Hall, J. W., III, & Grose, J. H.** (1991). Notched-noise measures of frequency selectivity in adults and children using fixed-masker-level and fixed-signal-level presentation. *Journal of Speech and Hearing Research*, *34*, 651–660.
- Hall, J. W., III, 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.
- Hartmann, W. M., & Pumplin, J.** (1988). Noise power fluctuations and the masking of sine signals. *The Journal of the Acoustical Society of America*, *83*, 2277–2289.
- Healy, E. W., & Bacon, S. P.** (2006). Measuring the critical band for speech. *The Journal of the Acoustical Society of America*, *119*, 1083–1091.
- Irwin, R. J., Stillman, J. A., & Schade, A.** (1986). The width of the auditory filter in children. *Journal of Experimental Child Psychology*, *41*, 429–442.
- Ison, J. R., Virag, T. M., Allen, P. D., & Hammond, G. R.** (2002). The attention filter for tones in noise has the same shape and effective bandwidth in the elderly as it has in young listeners. *The Journal of the Acoustical Society of America*, *112*, 238–246.
- Johnson, C. E.** (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research*, *43*, 144–157.
- Kohlrausch, A., Fassel, R., van der Heijden, M., Kortekaas, R., van de Par, S., Oxenham, A. J., & Püschel, D.** (1997). Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations. *Acta Acustica United With Acustica*, *83*, 659–669.
- Leibold, L. J., & Buss, E.** (2016). Factors responsible for remote-frequency masking in children and adults. *The Journal of the Acoustical Society of America*, *140*, 4367–4377.
- Leibold, L. J., & Neff, D. L.** (2007). Effects of masker-spectral variability and masker fringes in children and adults. *The Journal of the Acoustical Society of America*, *121*, 3666–3676.
- Leibold, L. J., & Neff, D. L.** (2011). Masking by a remote-frequency noise band in children and adults. *Ear and Hearing*, *32*, 663–666.
- 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*, 3960–3970.
- Leibold, L. J., Yarnell Bonino, A., & Buss, E.** (2016). Masked speech perception thresholds in infants, children, and adults. *Ear and Hearing*, *37*, 345–353.
- Litovsky, R. Y.** (2005). Speech intelligibility and spatial release from masking in young children. *The Journal of the Acoustical Society of America*, *117*, 3091–3099.
- McGhee, R. L., Ehrler, D. J., & DiSimoni, F.** (2007). *Token Test for Children—Second Edition*. Austin, TX: Pro-Ed.
- Moore, B. C. J., Glasberg, B. R., & Baer, T.** (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, *45*, 224–240.
- National Center for Education Statistics.** (1993). *Characteristics of public school kindergarten students and classes, in public school kindergarten teachers' views on children's readiness for school*. Retrieved from <http://nces.ed.gov/surveys/frss/publications/93410/index.asp?sectionid=5>
- 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*, EL377–EL383.

- Nittrouer, S., Sansom, E., Low, K., Rice, C., & Caldwell-Tarr, A.** (2014). Language structures used by kindergarteners with cochlear implants: Relationship to phonological awareness, lexical knowledge and hearing loss. *Ear and Hearing, 35*, 506–518.
- Oh, E. L., Wightman, F., & Lutfi, R. A.** (2001). Children's detection of pure-tone signals with random multitone maskers. *The Journal of the Acoustical Society of America, 109*, 2888–2895.
- Olsho, L. W.** (1985). Infant auditory perception: Tonal masking. *Infant Behavior and Development, 8*, 371–384.
- Papso, C. F., & Blood, I. M.** (1989). Word recognition skills of children and adults in background noise. *Ear and Hearing, 10*, 235–236.
- Plebanek, D. J., & Sloutsky, V. M.** (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science, 28*, 723–732.
- Polka, L., Rvachew, S., & Molnar, M.** (2008). Speech perception by 6- to 8-month-olds in the presence of distracting sounds. *Infancy, 13*, 421–439.
- Pumplin, J.** (1985). Low-noise noise. *The Journal of the Acoustical Society of America, 78*, 100–104.
- Roman, A. S., Pisoni, D. B., Kronnberger, W. G., & Faulkner, K. F.** (2017). Some neurocognitive correlates of noise-vocoded speech perception in children with normal hearing: A replication and extension of Eisenberg et al. (2002). *Ear and Hearing, 38*, 344–356.
- Sawilowsky, S. S.** (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods, 8*, 597–599.
- Scharf, B., Quigley, S., Aoki, C., Peachey, N., & Reeves, A.** (1987). Focused auditory attention and frequency selectivity. *Perception and Psychophysics, 42*, 215–223.
- Schlauch, R. S., & Hafter, E. R.** (1991). Listening bandwidths and frequency uncertainty in pure-tone signal detection. *The Journal of the Acoustical Society of America, 90*, 1332–1339.
- Schneider, B. A., Morrongiello, B. A., & Trehub, S. E.** (1990). Size of critical band in infants, children, and adults. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 642–652.
- Soderquist, D. R.** (1993). Auditory filter widths in children and adults. *Journal of Experimental Child Psychology, 56*, 371–384.
- Studebaker, G. A.** (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research, 28*, 455–462.
- Ubiali, T., Sanfins, M. D., Borges, L. R., & Colella-Santos, M. F.** (2016). Contralateral noise stimulation delays P300 latency in school-aged children. *PLoS One, 11*, e0148360.
- van Dinteren, R., Arns, M., Jongsma, M. L., & Kessels, R. P.** (2014). P300 development across the lifespan: A systematic review and meta-analysis. *PLoS One, 9*, e87347.
- Werner, L. A., & Bargones, J. Y.** (1991). Sources of auditory masking in infants: Distraction effects. *Perception & Psychophysics, 50*, 405–412.
- 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*, 3164–3176.