The Development of Auditory “Spectral Attention Bands” in Children

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Carla L. Youngdahl, M.A.

Graduate Program in Speech and Hearing Science

The Ohio State University

2015

Dissertation Committee:

Eric W. Healy, Advisor

Rachael Frush Holt

Allison Bean Ellawadi

Abstract

This study seeks to further our understanding of auditory development by investigating “spectral attention bands” (spectral region of attention for an expected target) and the ability to integrate or segregate information across frequency bands in children. The ability to attend to a target signal and discriminate speech from noise is of special importance in children. On a daily basis children must listen and attend to important auditory information in noisy classroom environments. A comparison of spectral attention bandwidth in children and adults might clarify where aspects of processing/listening efficiency breaks down. The current three experiments investigate the shape of spectral attention bands in children age 5 to 8 as compared to adults and indicate that the spectral attending listening strategy may effect understanding speech in noise. This study indicates that children do in fact listen differently than adults, using less efficient listening strategies that may lead them to be more susceptible to noise. This study also shows that between the ages of 5 and 8, enough substantial refinements in listening strategies occur to see a change to more adult-like performance in the older child age range.

Dedication

Dedicated to my parents, my daughter Greta, and Michael.

Through all your support this was possible.

Acknowledgements

I would like to especially thank my advisor and mentor, Eric W. Healy. You have given me indispensable mentorship, critical guidance, candid reassurance, and provided me with every opportunity to succeed throughout my PhD studies.

I would like to acknowledge the assistance from my dissertation committee Rachael Frush Holt and Allison Bean Ellawadi for keeping me on track with working with children. I would also like to acknowledge the support from the Speech and Hearing Science Department and The Ohio State University.

I would like to give special thanks to Fred Apoux for constructive criticism and assistance with MATLAB coding, Brittney Carter for helping with running experiments and proofreading, and to Jordan Vasko for being my unofficial statistics consultant.

I would also like to give a sincere thank you to all the research participants, especially the parents for bringing in their children to participate.

Lastly, a very special and sincere thank you to Sarah Yoho Leopold, who has served every role imaginable for me and this project: co-worker, PhD cohort, research design and analysis, editor, and friend.

This research was supported in part through a grant from the National Institutes of Health (NIH / NIDCD R01 DC008594) and a grant from The Ohio State University Alumni Grants for Graduate Research and Scholarship Program.

Vita

January 25, 1981……………………………………Born

2003…………………………………………………B.S. Communication Disorders,
Music minor
Truman State University
Kirksville, MO

2005…………………………………………………M.A. Speech Language Pathology,
Ohio University
Athens, OH

2005 - 2006…………………………………………Speech-Language Pathologist
Enduracare Therapy Mgmnt., Inc.
OH

2006 - 2009…………………………………………Speech-Language Pathologist
South-Western City School District
OH

2007 - 2014…………………………………………Speech-Language Pathologist
Extendicare-Progressive Step Corp.
OH

2012…..…..………………………………………….Speech-Language Pathologist
2014-2015
Fairfield County Educational Service
Center, OH

2009 - 2015…………………………………………Graduate Research Associate
Speech Psychoacoustics Laboratory
The Ohio State University
Columbus, OH

Publications

Healy, E. W., Youngdahl, C. L., & Apoux, F. (2014). Evidence for independent time-unit processing of speech using noise promoting or suppressing masking release. *Journal of the Acoustical Society of America*, 135, 581-584.

Healy, E.W., Yoho, S.E., Youngdahl, C.L., & Apoux, F. (2013). Talker effects in speech band importance functions. *Proceedings of Meetings on Acoustics*, 19, 050066, pp. 1-6.

Apoux, F., Yoho, S.E., Youngdahl, C.L., &. Healy, E.W. (2013). Can envelope recovery account for speech recognition based on temporal fine structure? *Proceedings of Meetings on Acoustics*, 19, 050072, pp. 1-6.

Apoux, F., Yoho, S.E., Youngdahl, C.L., & Healy, E.W. (2013). Role and relative contribution of temporal envelope and fine structure cues in sentence recognition by normal-hearing listeners. *Journal of the Acoustical Society of America*, 134, 2205-2212.

Cochran, P.S., Ellman, K., Imes, S., Kelly, K., Fuemmeler, R., Youngdahl, C. (2003). Simplifying Civics with Technology. The Catalyst - The Western Center for *Microcomputer Applications in Special Education*, Summer 2003, 19(4).

Field of Study

Major Field: Speech and Hearing Science

Table of Contents

List of Tables

List of Figures

Introduction

The aim of this research is to further our understanding of development of the specific spectral auditory attention listening strategy. This study investigates the development of spectral auditory attention bands in a group of 5- to 8-year-old children and their ability to integrate or segregate information across frequencies. Attending to a specific spectral region allows adult listeners to focus on where they expect a target to occur as a strategy to maximize detection. Overall, there is limited information about how this listening strategy develops. The ability to attend to a target signal and discriminate speech from noise is of special importance for children. On a daily basis children must listen and attend to important auditory information in noisy classroom environments. A comparison of spectral attention bandwidth in children to that of adults, and to peripheral auditory filters, might clarify where some processing/listening efficiency breaks down.

An important differentiation is that the peripheral auditory filters are associated with particular physical locations along the basilar membrane. In contrast, the concept of attention bands is based on the ability to use listening strategies to attend to certain spectral regions. The peripheral auditory system can be considered as a bank of overlapping band-pass filters (Fletcher, 1940). These peripheral filters partition incoming sounds into a series of frequency bands known as critical bands. This partitioning allows the auditory system to "ignore" bands that may contain large amounts of noise that would impede the detection or understanding of the target signal. In accord with a "glimpsing" view (Cooke, 2005), speech recognition in noise most likely relies on the combination of a limited number of auditory filter outputs, which may be sparsely distributed across the spectrum.

Due to the nature of peripheral filters and narrow attention bandwidth, adults can detect a signal in noise as long as the signal and noise are separated in frequency by more than a critical bandwidth (Fletcher, 1940). Studies have also demonstrated that processing of speech in bandwidths that are equivalent to peripheral filter bandwidths is not substantially affected by the presence of complementary, largely non-overlapping bands of noise in adults (Allen *et al.*, 1989; Hall & Grose, 1991). Further, adults have demonstrated similar performance for consonant identification both when non-adjacent speech bands are presented in quiet and when presented with interleaved noise (Apoux & Healy, 2009; 2010). This suggests that speech recognition in noise relies somewhat on the independence of auditory channels and that speech recognition is based on the auditory filters centered on the speech bands (Apoux & Healy, 2009).

Children may not process information from critical bands in the same way as adults, as many studies have documented differences in performance on psychoacoustic tasks between adults and children. Elevated thresholds in quiet, noise, and tone complexes have been documented from infancy to up to 10-years old (Elliott & Katz, 1980; Leibold & Neff, 2007; Olsho *et al.* 1988; Schneider *et al.*, 1986; Trehub *et al.*, 1980; 1988). Explanations for this difference in performance range from continued growth of the external auditory canal to immature neural coding. Additionally, some evidence suggests decreased frequency resolution in early childhood, with improvements to adult-like capabilities in the 6-10 year age range (Allen *et al.*, 1989; Hall & Grose, 1991; Irwin *et al.*, 1986; Olsho, 1985; Schneider *et al.*, 1990). While other studies suggest frequency resolution maturing before 6 years of age (Soderquist, 1993). The degree of frequency resolution that can be achieved by the peripheral auditory system is critical: the narrower the bandwidth of the auditory filters, the more noise the system can reject.

Extensive research on developmental changes in speech perception in noise has shown that improvement in these abilities continues until at least 10 years of age and possibly through the teenage years (Elliott & Katz, 1980; Elliott *et al.*, 1981, Johnson, 2000). Many studies have already shown that children require a higher signal-to-noise ratio (SNR) to recognize speech in noise at the same performance level as adults (e.g., Elliott *et al.*, 1979; Litovsky, 2005; Wightman & Kistler, 2005). Also, there is evidence that children integrate information over a larger number of auditory filters (Oh *et al.*, 2001). The combination of requiring a more favorable SNR and integrating information across more filters emphasizes the importance of listening strategies for understanding speech in noise. Depending on how listening strategies develop, children's performance on some tasks may or may not be adult-like.

While these psychophysical filters in the peripheral auditory system are adult-like by the age of six months (Hall & Grose, 1991; Olsho, 1985; Schneider *et al.*, 1990), the attention band may differ while the auditory cortex is still developing in childhood. Adult listeners are able to focus on individual auditory filters to optimize listening to a target in noise (e.g., Ison *et al.*, 2002; Scharf *et al.*, 1987; Schlauch & Hafter, 1991). Adults demonstrate narrow attention bands that closely follow the shape of the psychophysical auditory filters (Dai *et al.*, 1991; Ison *et al.*, 2002; Scharf *et al.*, 1987). One study with a limited number of children has indicated increased variability in attention bandwidths in children, with some appearing adult-like (Greenberg *et al.*, 1970). In contrast, young infants show an elevated, flat performance across a wide range of frequencies, indicating either extremely broad attention bands, or a lack of this listening skill in general (Bargones & Werner, 1994; Olsho, 1985).

Basic Pyschoacoustics

Whereas there are anatomical changes occurring postnatally that may affect hearing abilities, these cannot fully account for some developmental differences seen in performance on basic psychoacoustic tasks between infants, children, and adults. Infants from birth to one year old experience the most rapid growth in peripheral auditory structures. After this time the rate of growth for the ear canal and middle ear structures declines, but continues more slowly until puberty. This continuing development in late school-age years may still have some small effect on resonant frequencies, particularly in mid-ranges (Keefe & Abdala, 2007). Probably a more impactful developmental effect on psychoacoustic tasks is the continued development of the auditory cortex. Evidence suggests that the auditory cortex continues to develop up to 12 years old (Moore, 2002; Moore & Guan, 2001). Additionally, pathways to the auditory cortex appear to mature along different timelines, with some paths maturing as early as 2 years old and others all the way into the late teen years.

This development of the auditory system can be broadly defined in three stages (Werner, 2007). The first stage occurs in early infancy (within the first 6 months) and involves the maturation of neural mechanisms for coding sound. During this stage, infants show immature frequency discrimination and resolution (Spetner & Olsho, 1990; Werner & Gillenwater, 1990). The second stage lasts from infancy through early school-age and involves the development of increased specificity of spectral acoustic cues used in making perceptual decisions. This specificity refers to the use of newer, more nuanced, acoustic cues to distinguish sounds, such as improved sound-source determination and tone discrimination in low frequencies. The third stage may begin to develop in late school-age (around 8 or 9 years old) and may last well into adolescence. In this last stage children become more adult-like and are capable of choosing the most beneficial acoustic cues to identify sounds. This in a flexible way to maximize information received (Werner, 2007). With these developmental stages in mind, it is the development of the central auditory processing system that may more largely account for immature performance in psychoacoustic tasks.

A) Absolute and Masked Thresholds

Studies have indicated elevated absolute and masked thresholds in infants and children, with detection thresholds increasing for tones in quiet with age from early infancy to early school-age (Elliott & Katz, 1980; Leibold & Neff, 2007; Olsho *et al.* 1988; Schneider *et al.*, 1986; Trehub *et al.*, 1980; 1988). The study by Olsho *et al.* (1988) investigated auditory sensitivity for infants, 3 to 12 months old. Due to the young age of participants an observer-based psychoacoustic approach was used. A train of tone bursts at octave intervals from 250 to 8000 Hz was presented monaurally over an earphone to the infant while an observer reported behavioral changes that indicated whether the infant heard the tones. The behavioral changes involved responding to visual reinforcement that accompanied the tones. The youngest infants, 3 months old, had thresholds elevated 15 to 30 dB above those of adults, with the larger discrepancy found at higher frequencies. Between three and six months of age, detection thresholds improved for all frequencies (except 250 Hz) by approximately 10 to 20 dB. There was no significant change in detection threshold for a 250 Hz tone between the ages of three to six months and additionally, no further overall increase in sensitivity for all frequencies tested at 12 months old.

Trehub, Schneider, and colleagues completed a series of studies that investigated auditory sensitivity from infancy through late school-age using band-restricted noise presented over speakers. The initial study in this series (Trehub *et al.*, 1980) measured detection thresholds in quiet for infants 6 to 18 months old. In this study, infants listened to an octave band noise presented over one of two speakers placed in corners of a room.

This two-alternative forced choice was visually reinforced when the infant turned towards the correct speaker playing the tone. The octave-band noises were centered at 250, 500, 1000, 2000, and 4000 Hz. Results showed detection thresholds in quiet at lower frequencies to be elevated 20-30 dB above that of adults and to be nearing adult-like thresholds for higher frequencies. For low frequencies, 6-month-old infants showed detection thresholds elevated 5-8 dB higher than 12- and 18-month-old infants.

Using similar methodology, these results were later extended by Schneider *et al.* (1986) to children 3 to 5 years old. Octave and 1/3-octave-wide bands of noise were presented over one of two speakers placed in corners of the room. Behavioral measures for younger children, aged three and four, were head turns toward the speaker, while 5-year-olds pushed buttons corresponding to the speaker playing the noise. In this study, detection thresholds for all frequency bands continued to improve with increasing age. While 5-year-old's detection thresholds were lower than their younger counterparts, their thresholds were still elevated 5 to 15 dB above adult performance (Schneider *et al.*, 1986). In Figure 1, you can see results of detection thresholds measured in quiet for infants and preschool-age children from the data of Schneider *et al.* (1986) and Trehub *et al.* (1980).

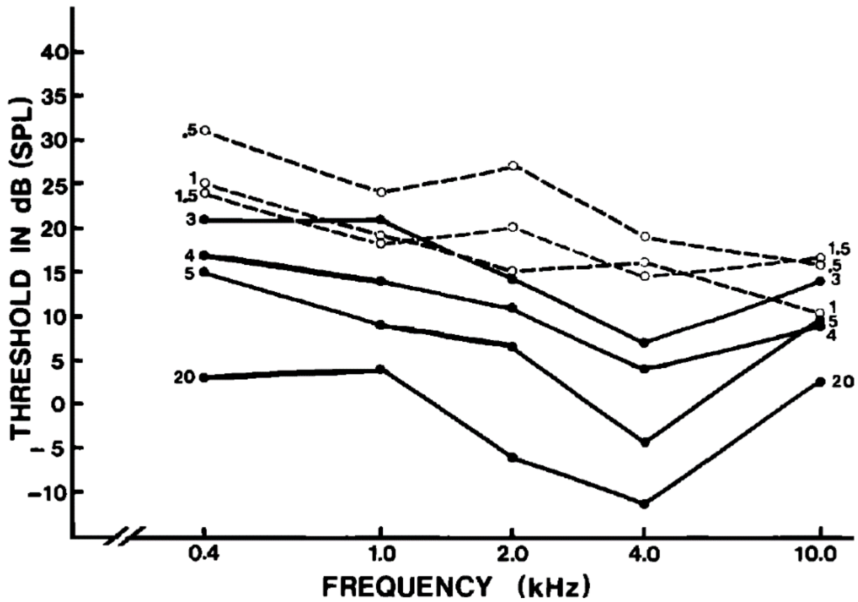

Figure 1. (From Schneider *et al.* 1986) Thresholds for children 3-5 years of age and adults (age 20 yrs) as a function of frequency of octave-band noises (solid lines). Similar data for infants 0.5, 1, and 1.5 years of age from Trehub *et al.* (1980; dashed lines). Number on either side of the lines indicate subject age in years.

Trehub *et al.* (1988) followed up these results with still older children, ages 6 to 16 years old. This study used the same two-alternative forced choice paradigm, and response buttons corresponding to the speaker playing the noise. Listeners were presented with an octave-wide noise centered on 400, 1000, 2000, 4000, 10000 Hz and 1/3-octave band noise centered on 10 and 20 kHz. Thresholds were shown to continue to improve up to 10 years old, with lower frequencies reaching adult-like thresholds later than higher frequencies.

Additionally, studies using different methodologies have found similar improvements in sensitivity beyond the preschool age range. The study by Elliott and Katz (1980) used a three-alternative forced choice adaptive procedure with 6- and 10-year-old children. In this study, listeners were asked to identify which interval contained a tone or narrow-band noise. Results are in concordance with findings from Trehub *et al.* (1988), with 6-year-olds showing elevated detection thresholds compared to 10-year-olds and adults. Additionally, while 10-year-old detection thresholds were lower compared to 6-year-olds, they were still not quite at adult-like performance.

The study by Leibold and Neff (2007) also found decreasing thresholds with age from 5 to 10 years old. In this study a two-interval forced choice paradigm was used in an adaptive procedure. Listeners were presented with a 1000-Hz tone prior to trials, they then indicated in which of two following intervals the signal occurred. Results, as seen in Figure 2, show that by the age of 10 years, absolute thresholds in quiet are only slightly elevated above those of adults (2 to 4 dB). What is also apparent in these results is that within age groups there is significant variability in detection thresholds, which is commonly seen in psychoacoustic measures in children.

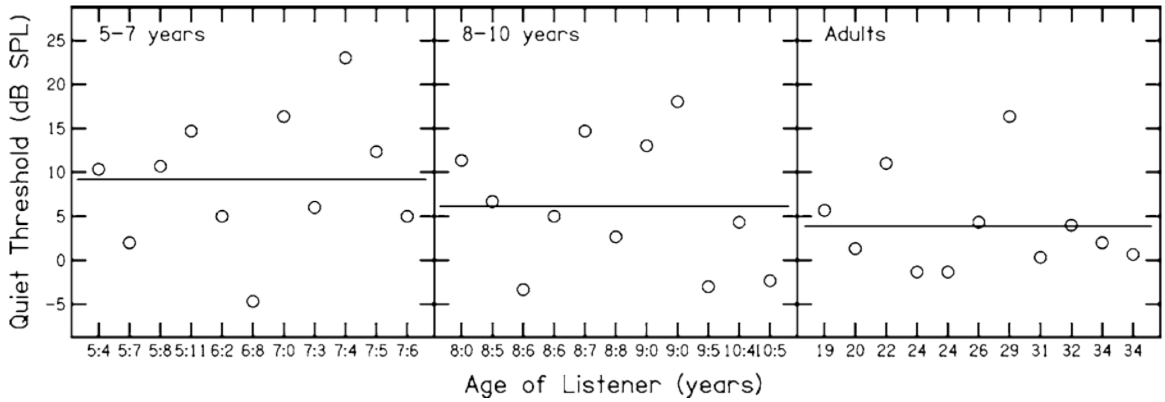

Figure 2. (From Leibold & Neff 2007) Quiet thresholds for the 1000-Hz signal are plotted for individual listeners, with listeners ordered by their age in years: months. Data across panels are for younger children (5-7 years), older children (8-10 years), and adults. Solid horizontal lines indicate the mean group threshold.

Similar to results for thresholds in quiet, infants and children demonstrate elevated thresholds in noise compared to adults. For both tone and noise stimuli, masked thresholds are typically elevated by about 5 to15 dB in 6-month-old infants, gradually decreasing to adult levels by approximately 10 years of age (Allen & Wightman, 1994; Leibold & Neff, 2007; Oh *et al.*, 2001; Schneider *et al.*, 1989).

A study by Allen and Wightman (1994) examined detection thresholds in noise for children 3 to 5 years old along with psychometric function slopes to better track differences in performance. Listeners heard two intervals containing a Gaussian noise masker over an earphone. One of the intervals also contained a tone at 501, 1000, or 2818 Hz. Listeners indicated which interval contained the tone in an adaptive procedure. Similar to other studies, mean thresholds for children were consistently higher than adult thresholds (10 to 20 dB). Some children displayed psychometric function slopes shallower than those of adults, which could indicate wider filters, which allows more noise to be masking the signal leading to elevated thresholds (Allen & Wightman, 1994). However, another possible explanation is that the auditory periphery is mature, but higher cognitive processes essential to detection are not.

Schneider, Trehub, and colleagues investigated detection thresholds in noise for infants to late childhood (Schneider *et al.*, 1989). Using similar procedures to their previous work determining detection thresholds in quiet (Schneider *et al.*, 1986; Trehub *et al.*, 1980), they determined detection thresholds for 1-octave and 1/3-octave noise bands centered at 400, 1000, 2000, 4000, and 10,000 Hz presented in broadband noise. Results showed a systematic decline in masked thresholds with increasing age, with detection thresholds being close to adult-like across all frequencies by the age of 10 years. In contrast to these results, the previously discussed study by Elliot and Katz (1980) found little to no difference between child and adult thresholds in noise. In that study, thresholds were compared when the masker was a narrow-band noise to when it was a wide-band noise. Overall, there was a 2- to 4-dB difference between adults and 6-year-olds for detection thresholds at 2000 Hz in both the wide-band and narrow-band noise.

A study by Oh *et al.* (2001) investigated tone detection in multi-tonal and Gaussian noise maskers. This study employed an adaptive two-interval forced choice procedure similar to Allen and Wightman (1994). Maskers comprised of 10 to 40 tone components produced 30 to 60 dB of masking in some listeners, but less in other listeners. Those same maskers produced larger amounts of masking, up to 70 to 83 dB in many of the preschool children, again with individual variability. As seen in Figure 3, and similar to findings by Allen and Wightman (1994), psychometric function slopes varied among children, with some appearing adult-like, but others being shallower. However, nearly all psychometric functions reach 100% accuracy, indicating that motivation and attention to task were likely not factors.

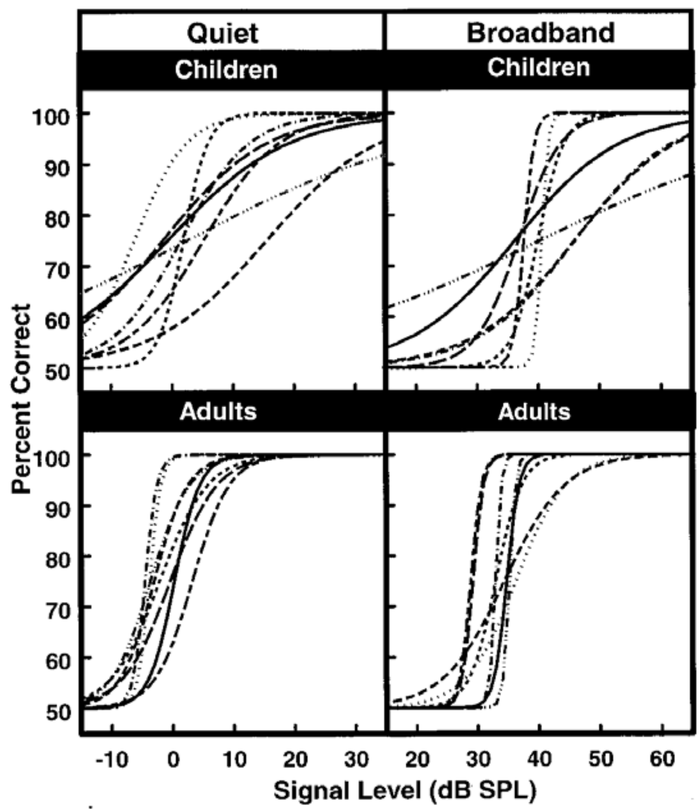

Figure 3. (From Oh et al, 2001) Fitted psychometric functions for individual children (top panels) and adults (bottom panels). The label ‘‘Quiet’’ indicates the condition where the signal was presented in quiet, and the label ‘‘Broadband’’ indicates the condition where the signal was simultaneously presented with a broadband noise.

Similar elevated masked thresholds have been reported with multi-tonal maskers. Leibold and Neff (2007) reported data from school-age children, which showed increased thresholds for a 1000 Hz tone presented in the context of spectrally remote, fixed-frequency maskers. For this study, the fixed-frequency tonal masker was contrasted with another tonal masker with randomly varied components in each trial (random). As seen in Figure 4, in all masker types (broadband noise, random-tone complex, and fixed-tone complex) younger children demonstrated elevated thresholds compared to older children and adults. Older children, while showing decreased masked thresholds, still did not reach adult-like performance in any masker type.

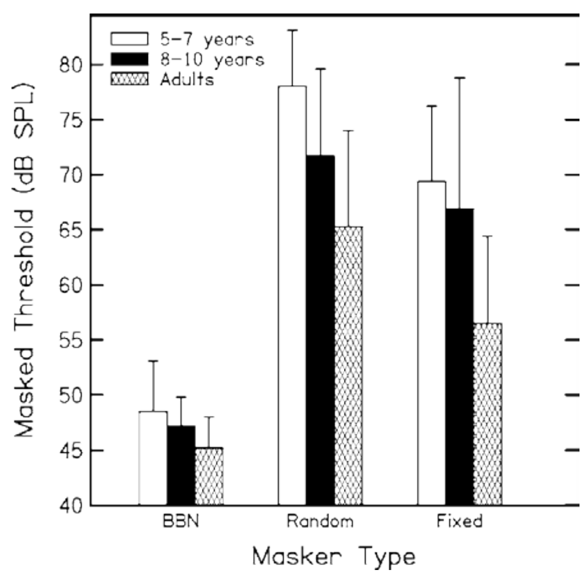

Figure 4. (From Leibold & Neff, 2007) Average masked thresholds (with SDs) across listeners for each of the three age groups (open bars for younger children, solid bars for older children, and hatched bars for adults) are presented for the three types of maskers: broadband noise (BBN), ten-tone random frequency (Random), and ten-tone fixed frequency (Fixed). Masker level was 60 dB SPL.

These combined and similar findings, across different methodologies, indicate that auditory sensitivity in quiet and in noise continues to improve well into school-age. These age-related improvements exist across the frequency range, though there is some discrepancy whether sensitivity improves earlier for high frequency versus low frequency.

Many suggestions have been made for the explanation of elevated thresholds for infants and young children. As mentioned earlier, motivational factors have been mostly ruled out due to a lack of consistent difference in psychometric function slopes (Allen & Wightman, 1994; Oh *et al.*, 2001). One possible explanation for these age-related changes could be the growth and development of the external auditory canal and middle ear structures. Small differences in conductive efficiency (e.g., size of ear canal) and changes in the resonant frequency of the external auditory canal and auricle could explain some improvements with age in the mid- and low-frequency regions. And, while the structures of the middle ear may be fully developed at a young age, their mechanical properties may continue to mature into the school-age years, affecting auditory sensitivity. However, immature neural coding may also be a factor. Electrophysiological and otoacoustic emissions studies have shown evidence suggesting that the developmental differences in detection thresholds that persist into the early school years are not likely to arise from cochlear or neural factors associated with the sensitivity of the auditory periphery (Abdala & Folsom, 1995; Bargones & Burns 1988).

B) Frequency Selectivity and Resolution

Frequency selectivity and frequency resolution are abilities that go hand in hand to enable the auditory system to extract specific cues and crucial information from the auditory environment. Frequency selectivity reflects the auditory system's ability to separate out different frequencies in a complex mixture. Frequency resolution is the degree of spectral detail that can be achieved from the signal. Immaturity in both frequency selectivity and resolution may lead to decreased spectral details in complex sounds, such as speech, an unclear representation of the signal in the auditory system leading to poorer perception of speech in noise. Several studies have investigated frequency selectivity and frequency resolution in the developing auditory system.

Olsho (1985) found adult-like frequency selectivity in infant listeners 4 to 8 months old by measuring psychophysical tuning curves. Four probe frequency stimuli (500, 100, 2000, and 4000 Hz) presented as tone bursts in quiet and five different masker frequencies (50, 200, and 400 Hz above the probe). Infant responses were measured with visually reinforced head turns. As seen in Figure 5, infant tuning curves, widths, and slopes were similar to those of adults, although there were differences in the positions of the curves. Across all frequencies infants showed an increased susceptibility to masking when compared to adults. This result is not surprising considering the elevated masked thresholds discussed earlier for infants and children.

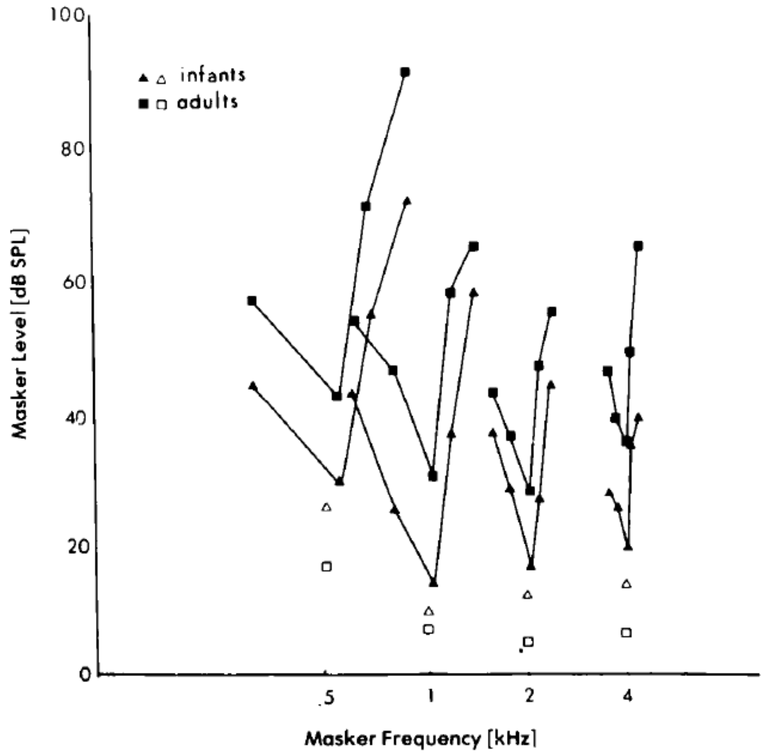

Figure 5. (From Olsho, 1985) Composite tuning curves for infants and adults in four masking conditions. Absolute (unmasked) thresholds are plotted with open symbols.

Additionally, a study by Schneider *et al.* (1990), measured thresholds in 6-month-old, 2- and 5-year-old children for an 800 Hz tone in different narrow-band masker bandwidths. Masked thresholds increased with bandwidth for adults, children, and infants, but did not change after a critical width. As shown in Figure 6, differences between child and adult thresholds measured 8 to 15 dB, while the width of the critical band did not show much change from infancy to adult.

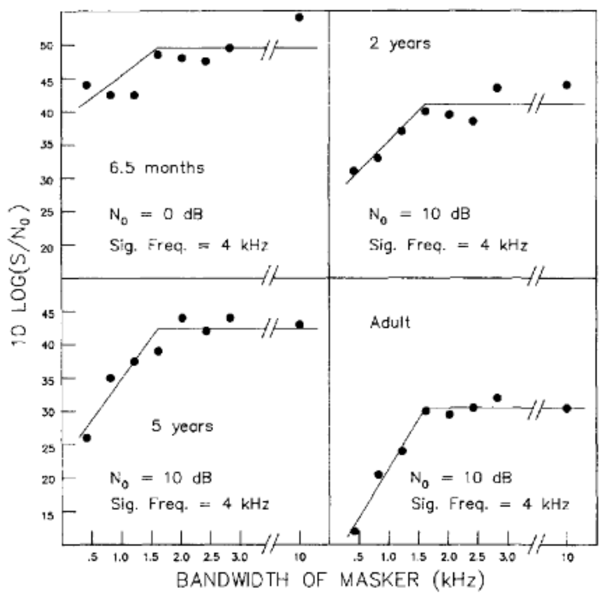

Figure 6. (From Schneider *et al.*, 1990) Thresholds at four different ages for 4-kHz signal as a function of effective rectangular bandwidth.

Some studies have concluded that frequency resolution is adult-like by 5 years of age (Allen *et al.*, 1989; Hall & Grose, 1991). In the study by Allen *et al.* (1989), child listeners were asked to detect tones at 500, 2000, or 4000 Hz in Gaussian notched-noise of varying bandwidths. The method to determine thresholds was a three-alternative forced choice paradigm. Results indicated that frequency-resolving abilities continue to improve from 2 to 6 years old for higher frequencies (2000 and 4000 Hz). At 500 Hz, 6-year-olds did not improve performance more than 4-year-olds and overall adult-like performance did not occur until 10 years old. Similar to detection threshold measures, there was large variability in performance across child subjects, which is understandable as the ability may be developing across such a wide age range of child subjects.

Hall and Grose (1991) found relatively poor frequency selectivity in 4-year-old listeners using a notched-noise measure. In this method, the detection threshold for a pure-tone signal presented in a masking noise was determined as a function of the width of a band-stop region (notch) centered on the signal frequency. Frequency selectivity was measured for 500 and 2000 Hz with three notched-noise maskers having three varying notch widths. Results indicated reduced frequency selectivity in 4-year-olds increasing to adult-like selectivity by the age of 6 years old (Hall & Grose, 1991).

Irwin *et al.* (1986) found reduced frequency resolution still present in 6-year-olds. This study cited the rationale that, because young children have decreased temporal resolution, they should have increased frequency resolution. To test this, they measured auditory-filter widths in 6- and 10-year-old children. Pure tones at 500, 1000, and 3000 Hz were presented in six notched noises of differing widths with a two-alternative, forced-choice, adaptive procedure. Results showed that auditory filters of 6-year-olds were wider than those of 10-year-olds and adults, indicating that younger children have poorer spectral resolution.

As demonstrated by the study by Soderquist (1993), some of the discrepancies between these studies measuring critical bandwidth and frequency selectivity/resolution could be explained by the types of psychoacoustic measures used. To alleviate this confusion, a study by Soderquist (1993) used both measures, a notched-noise and psychophysical tuning curves with the same subjects - children aged 6 to 10 years old.

For the psychophysical procedure they used a three-alternative forced-choice task in conjunction with a 2-down, 1-up adaptive procedure.

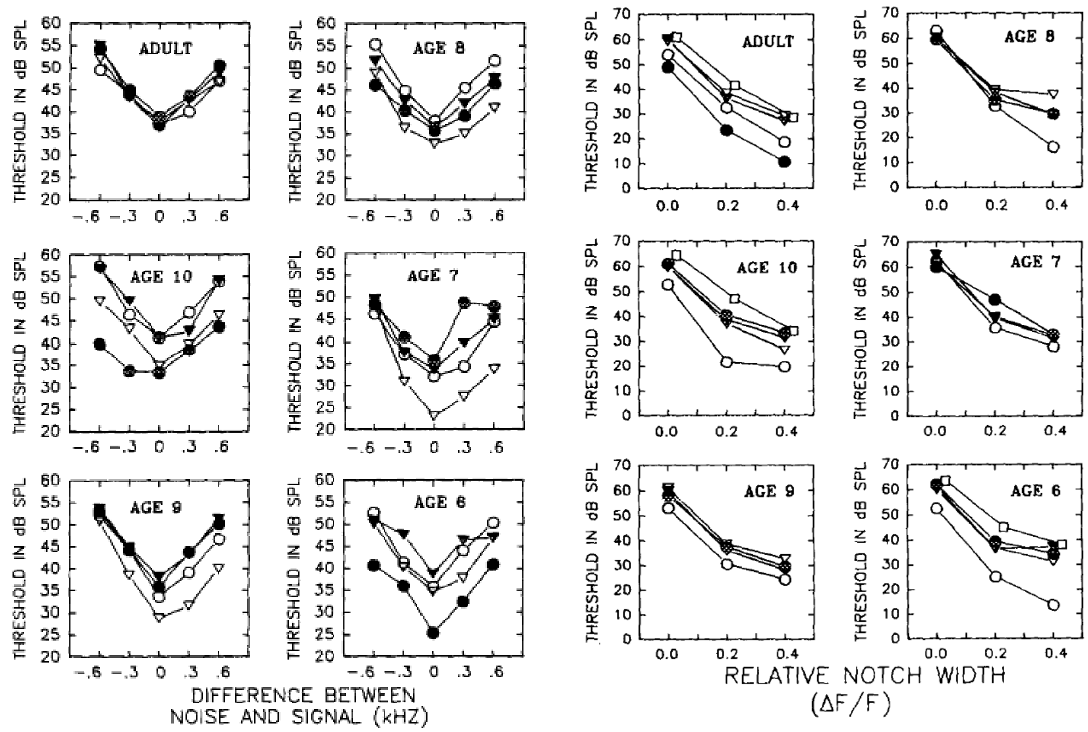

Figure 7. (From Soderquist, 1993) Individual psychophysical tuning curves (left) and notch-noise tuning curves (right) for six groups of subjects. The data show individual performance as a function of the difference, in kHz, between the signal and the nearest edge of the masking noise for tuning curves and as a function of relative notch-width.

As seen in Figure 7, the data from this study showed that all children had auditory filter widths comparable to those of adults as measured with psychophysical tuning curves and notched-noise functions. This supports the idea that peripheral filters and frequency selectivity mature at an age earlier than 6 years old and that infant and children's inability to extract auditory cues and identify target words might be related to other factors.

Overall, previous studies have reported conflicting data on the maturity of frequency selectivity with some studies reporting adult-like frequency selectivity as early as infancy (Olsho, 1985; Schnedier *et al.*, 1990) while others reporting frequency selectivity as still immature in the preschool to early school-age range of 4- to 10-years-old (Allen et al, 1989; Hall & Grose, 1991; Irwin *et al.*, 1986). However, Soderquist (1993) used different types of psychoacoustic measures with the same groups of children to show that frequency selectivity was adult-like by the youngest age tested, 6-years-old. This suggests that difference between child and adult performance do not lie in the peripheral filter bandwidths but might be in the efficiency of their respective listening strategies.

C) Susceptibility to Masking

A large body of evidence shows that children have decreased performance in noise compared to adults. Both detection of tones and recognition of speech in noise result in elevated thresholds in children compared to adults (e.g., Buss *et al.*, 1999; Hall *et al.*, 2002; Oh *et al.*, 2001; Papso & Blood, 1989). This may be in large part due to an increased susceptibility to masking or limited ability to obtain masking release. Children have generally been shown to have fully developed adult-like auditory filters by 6 years old (Irwin *et al.*, 1986; Soderquist, 1993); therefore energetic masking should have a similar effect on children and adults. However, this is not what is typically seen with masking measures. Further explanation for children's increased susceptibility to masking or lack of masking release is therefore required.

Energetic or peripheral masking occurs because some, or all, of a masker's energy falls within the same auditory filter as the signal energy, thus reducing audibility or resulting in a degraded representation of the signal in the auditory system (Arbogast *et al.*, 2002; Brungart, 2001). This degraded signal may be due to 'excitation,' in which neural activity is swamped by masker stimulation, with neural excitation spreading to the location of the target signal thus overwhelming the neural activity (Fletcher, 1940). This assumption is used in many psychophysical models of masking (Glasberg & Moore, 1990; Rosen & Baker, 1994; Zwicker, 1970). Another theory to account for masking, known as suppression, is more controversial. Suppression is a reduction in response to a target signal because of the introduction of another signal, the masker. Neural activity resulting from the target may become indistinguishable from spontaneous activity

(Oxenham & Plack, 1998). These two theories of masking mechanisms may not be mutually exclusive. A further assumption with energetic masking is that only masking components falling within the critical bandwidth of the target signal filter will have an effect towards masking the signal.

The upward spread of masking is a phenomenon in which a low frequency masker will mask a higher frequency target signal more than a lower frequency target (Bilger & Hirsh, 1956). For adults, thresholds increase roughly linearly for target signals near the masker frequency, but increase more rapidly for target signals that are well above the masker frequency region. Studies have shown evidence that suppression accounts for masking target signals that occur above the masker frequency, particularly for target signals much higher in frequency (Delgutte, 1990; Oxenham & Plack, 1998).

There is also evidence that noise outside of an auditory filter can contribute to masking. For example, remote masking occurs when a high frequency band of noise at high intensity masks a lower frequency (at least an octave apart) target signal (Bilger & Hirsh, 1956). In a study with infants, Werner and Bargones (1991) presented data indicating that infants were susceptible to remote masking of a 1000 Hz signal by a noise-band at 4000 to 10,000 Hz. The presence of the remote masker elevated thresholds by as much as 10 dB in infants, while adult thresholds were unaffected.

A further study by Leibold and Neff (2011) found that some younger children are susceptible to remote masking in similar conditions to Werner and Bargones' (1991) study. Thresholds were measured for a 1000 Hz tone with a bandpass remote noise masker at 4000 to 10,000 Hz using a two-interval forced-choice procedure. As seen in

Figure 8, 4- to 6-year-old children were susceptible to remote masking, with thresholds elevated over 3 dB from quiet. By the age of 7 to 9 years, data suggests children perform more similarly to adults because they are no longer susceptible to the remote masker. The authors suggest that these results are not likely due to immature representation of the signal but possibly a reduced ability to attend appropriately to the signal frequency.

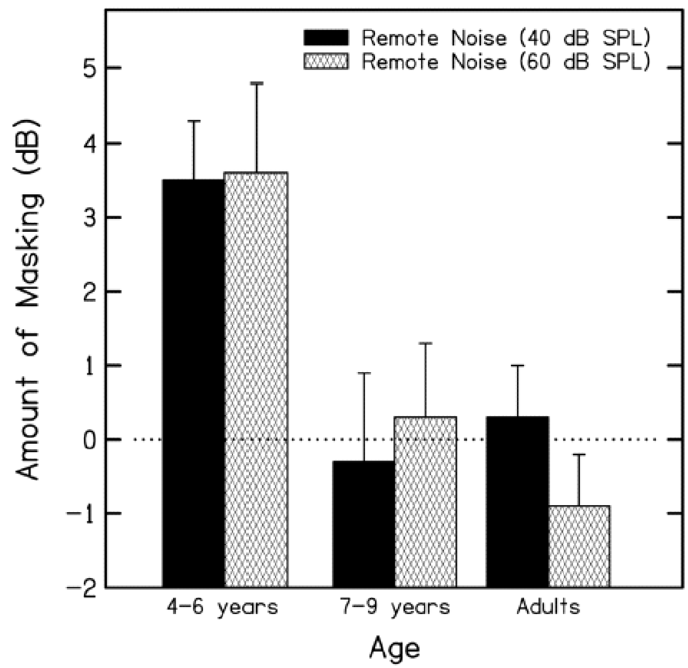

Figure 8. (From Leibold & Neff, 2011) Average amount of masking (difference in threshold between masked and quiet conditions) for 4- to 6-year-olds, 7- to 9-year-olds, and adults in the presence of remote-frequency noise at 40 dB SPL (solid bars or 60 dB SPL (hatched bars). Error bars represent +1 SE. Data falling at or below the dotted horizontal line indicate no masking.

Informational masking reflects the listener's remaining inability to segregate a target and masker after energetic masking is already taken into account. Most commonly, informational masking has been defined by many researchers as masking that occurs beyond that which can be attributed to energetic masking. While energetic masking effects can be measured in the auditory periphery, it is believed that informational masking operates in the central processes of the auditory system (Arbogast *et al.*, 2002; Moore, 2007). In informational masking, the target and masker may both be clearly audible, however, the listener is unable to segregate or distinguish the two. There are two factors believed to play a large role in informational masking: stimulus uncertainty (Lutfi, 1989; Neff & Callaghan, 1988; Watson & Nichols, 1976) and spectral and temporal similarities between masker and target signals (Brungart, 2001; Kidd *et al.*, 2002; Oh & Lutfi, 2000). A lack of difference in spectral and intensity components of the masker and target can result in the auditory systems' inability to discriminate between the two separate signals. Many studies have manipulated target and masker similarity to determine the extent of its influence in information masking.

Stimulus manipulations have shown that promoting perceptual segregation between target and masker can reduce the amount of informational masking (Neff, 1995). Using single-burst thresholds, this study manipulated grouping and segregation principles such as amplitude modulation of the target signal, target and masker onset, and perceptual quality between target signal and masker. In these cases, changes in signal and masker parameters decreased their similarity. Reduction of target and masker similarity resulted in a considerable reduction in informational masking. Other studies have also examined target-masker similarity with similar findings. Decreasing target and masker similarity through step-glides, constant- and alternating-frequency signal bursts, and frequency-modulated glides also decreases informational masking (Durlach *et al.*, 2003; Kidd *et al.*, 2002).

When dealing with speech as the target signal with a speech masker, voice characteristics of the respective speakers have an effect on the amount of informational masking, which in turn affects intelligibility. Studies have shown that different sex talkers used for the masker and the target create the most spectral difference and allow for better performance on speech intelligibility tasks (Brungart, 2001). Further, some studies have indicated that school-age children are able to use some acoustic cues (e.g. temporal, spatial, binaural) similar to adults to achieve a release from informational masking. However, there are limits to the extent and efficiency with which children can use these acoustic cues, resulting in increased susceptibility to informational masking and decreased performance when compared to adults.

Hall *et al.* (2005) and Wightman *et al.* (2003) found that a masker that produced significant informational masking when it was presented in the same ear as the target produced little or no masking when it was presented to the contralateral ear for both adults and older children. However, younger children continued to exhibit informational masking with the contralateral masker. In a study by Wightman and Kistler (2005), an ipsilateral single talker created more than 15 dB greater masking in young children than it did in adults. Even children as old as 16 years old showed increased informational masking when compared to adults. Overall, these results suggest that children have a relatively poor ability to benefit from informational masking release cues.

Another factor contributing to informational masking is stimulus uncertainty. Stimulus uncertainty refers to the configuration of the target signal and the masker such that there is an unexpected or changing element. When frequency patterns of a stimulus were fixed, resulting in low-uncertainty, a smaller amount of informational masking was observed. In contrast, when a large amount of variability was introduced to stimuli, larger amounts of masking were observed (e.g., Watson & Nichols, 1976). Other studies have shown similar results. Some examples of informational masking resulting from stimulus uncertainty include a multi-component masker with frequencies either at fixed frequencies or selected randomly from a range of frequencies that change from trial to trial. Subjects demonstrated increased signal-detection thresholds in the random masker condition compared to the fixed condition (Kidd *et al.*, 2002; Neff & Callaghan, 1988).

While studies have shown target-masker similarity and stimulus uncertainty as separate components accounting for informational masking, they are not necessarily mutually exclusive. There is an interaction between target-masker similarity and stimulus uncertainty. When target-masker similarity is high, stimulus uncertainty via masker randomizations has limited effect on listener performance (Durlach *et al.*, 2003). Furthermore, energetic and informational masking cannot always be taken into account separately. Many informational masking studies take great care to minimize energetic masking by creating a protected spectral region, where no masker frequency can occur. These regions are a certain bandwidth surrounding the frequency of the target signal

(Kidd *et al.*, 2002; Neff & Callaghan, 1988). However, energetic masking may still play a role in informational masking. Several studies have compared the contributions of energetic and informational masking in adults. Measurement of the relative contribution of energetic and informational masking suggests that the amount of informational masking is related to the number of auditory filters over which the information is spread (Oh & Lutfi, 1998).

Overall, studies have also shown that infants and children are more susceptible to informational masking than adults in the presence of multi-tonal maskers, even when the masker's spectrum does not seem to cause uncertainty (Leibold & Neff, 2007; Leibold & Werner, 2006). In Leibold and Werner (2006), adults and 7- to 9-month-old infants detected a brief 1000 Hz tone in different masking conditions, a broadband noise, a remote two-tone complex with fixed frequencies, and a remote two-tone complex with random frequencies. Both adults and infants demonstrated a masking release effect with the fixed two-tone complex and elevated thresholds with the random two-tone complex, but overall masking effects for infants were more detrimental. Compared to broadband noise, adults obtained masking release of 22 dB SPL in the fixed masking condition, while infants only benefitted by 10 dB SPL. Further, adults had thresholds elevated by 18 dB SPL in the random two-tone masking conditions, while infants had thresholds elevated by around 23 dB SPL. It has been suggested, especially in quiet conditions, that children experience an increase in informational masking due to internal noise. This internal noise may be the results of increased physiological noise (e.g. blood flow, body movement) or variability of neural representation (Buss *et al.*, 2009; Schneider *et al.*, 1986). Another explanation for the increased susceptibility to informational masking observed for children relative to adults is that children may have immature sound source segregation and/or selective attention abilities.

In summary, infants and young children often demonstrate poorer performance when compared to adults across many basic psychoacoustic performance tasks, with some auditory abilities still developing into late childhood. Tasks measuring absolute threshold, frequency resolution, temporal resolution, sound segregation, and speech perception have shown that infants and children frequently do not perform at adult-like levels despite near adult-like cochlear function. In general, data support decreases in thresholds (both in quiet and noise) and increases in detection and discrimination tasks with increasing age.

Speech Psychoacoustics

Speech perception tasks are inherently more complex than basic psychoacoustic detection tasks. Speech is a complex auditory signal with many spectral and temporal components along with linguistic content. Studies have shown that children continue to develop speech recognition abilities through the first 10 to 12 years of life (Boothroyd, 1968; Elliot *et al.*, 1979; Hnath-Chisolm *et al.*, 1998; Papso & Blood, 1989). When specifically addressing deficits in speech perception tasks by children, there are several possible contributing factors (i.e., vocabulary, phonemic categorization, decision-making, and/or articulation skills). One of the forefront considerations for speech measures is the vocabulary and linguistic content of the target stimuli used for assessing speech perception in children. However, even when these factors are taken into account and controlled for in speech stimuli, developmental differences in speech recognition continue to remain (Elliott *et al.*, 1979).

An important consideration in speech recognition is the cognitive load associated with processing the complex speech signal. Common manipulations, such as degrading the signal through filtering or addition of noise, serve to increase the perceptual load. However, it is also important to consider the cognitive load associated with the particular speech signal employed. When higher-level information, such as lexical-semantic context, is not available, the auditory system will likely rely more heavily on low-level acoustic cues (Cooke, 2006). This indicates that the balance between top-down and bottom-up processing, and the associated cognitive loads, can vary depending upon the stimulus employed.

When lexical-semantic cognitive load of the acoustic stimulus is controlled for, other aspects of processing in the developing auditory system may still account for some of this discrepancy in performance between adults and children. As discussed earlier, while the anatomical structures of the auditory system is mostly developed at birth, the human auditory cortex continues to develop well until adolescence (*e.g.*, Moore, 2002). This continual development may partially explain age-related improvements in speech recognition. Immaturity of central auditory processes and the adoption of inefficient listening strategies during development may lead to decreases in performance when compared to adults (Allen & Wightman, 1994; Lutfi *et al.*, 2003).

Several studies have shown reduced ability of children to understand speech in noise (e.g., Elliott *et al.*, 1979; Litovsky, 2005; Wightman & Kistler, 2005). Evidence suggests that adults can understand sentences in noise at SNR levels of approximately 0 dB, reflecting equal levels of speech and noise. On the other hand, young children require speech-in-noise conditions that are more favorable (e.g., SNR of +4 dB; Nilsson *et al.*, 1994). A hallmark study by Elliott (1979) measured speech perception scores for children 9 to 17 years old on Speech Perception in Noise (SPIN) sentences. Results from this study indicate that even to the age of 13, children are not performing at adult-like levels. As seen in Figure 9, there was continuous improvement with age, with 9-year-olds performing more poorly in speech in noise tasks than 11-year-olds.

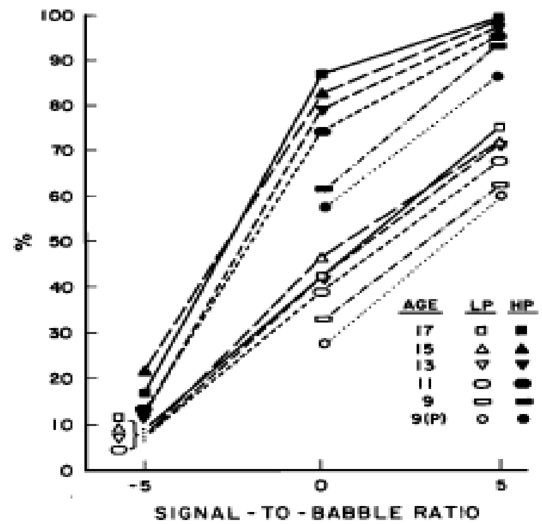

Figure 9. (From Elliott, 1979) Children's performance (in percent correct) on the SPIN test as a function of signal-to-babble ratio and children's age for high predictability (filled symbols) and low predictability (unfilled symbols) sentences.

Similarly, in a study by Elliott *et al.* (1979), younger children showed slightly elevated thresholds for recognizing monosyllabic nouns in quiet when compared to older children and adults. Six-year-olds demonstrated thresholds 5 to 7 dB above adults. As seen in Figure 10, overall performance became more adult-like with age in quiet conditions. However, there was no significant difference between child and adult speech recognition thresholds in a noise or babble background.

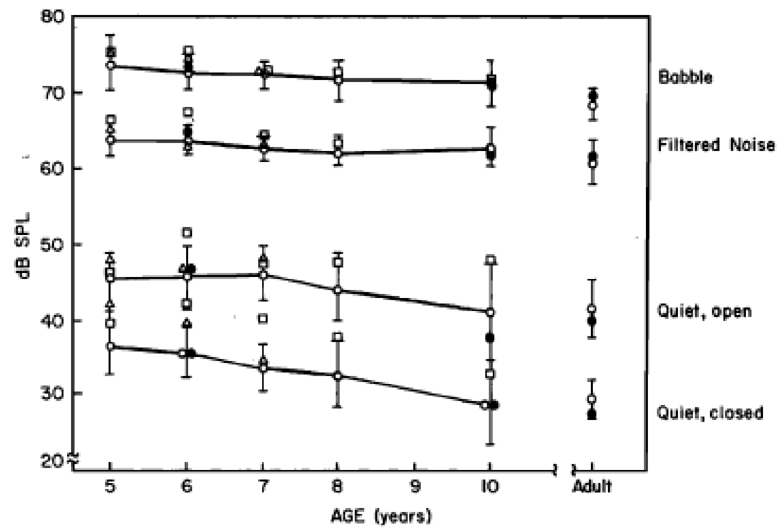

Figure 10. (From Elliott *et al.*, 1979) Mean SPL for 71% performance for groups of different ages. Lines connect points representing mean performance for "normally progressing" children. Squares and triangles represent mean performance of children with "learning problems" and developmental "articulation problems," respectively.

In a study by Papso and Blood (1989), 4- to 6-year-old children achieved word recognition scores in quiet comparable to adults, with both groups at ceiling performance of 100% for adults and 94% for children. For adults, the addition of noise or babble results in virtually unchanged word recognition scores. However, for children, the addition of noise dropped recognition scores down to approximately 77%, while the addition of babble dropped scores even further to 67%. This illustrates an increase of masking for children that is not seen in adults.

It is possible that many acoustic cues used by adults for these psychoacoustic tasks are available for children, but they have not learned to encode or use them in an efficient manner. Auditory processing at subcortical levels may improve as a result of top-down processing, as a child learns the usefulness and importance of certain acoustic cues and information. With development comes an improved ability to utilize the most optimal acoustic cues.

In dichotic listening tasks, the listener must pay attention to one of two different concurrent signals, with each delivered to only one ear over stereo earphones. This task is often used as a measure of selective attention. Studies have shown that up to 8 years old, children generally perform more poorly than adults in dichotic listening tasks. In a study by Wightman and Kistler (2005), children ranging from 4 to 16 years old were asked to listen to a speech signal presented in one ear. In the other ear, the listener received either quiet, competing speech, or noise. Conclusions from this study were that all children, even at 16 years old, had more difficulty than adults selectively attending to the target speech. For the youngest children, 4 to 5 years old, masking was elevated more than 15 dB above adult performance. However, the ability to selectively attend was evident in children. When masking speech was produced by a speaker opposite in sex to the target speech, all children had an 8 dB release from masking, which cannot be fully attributed to energetic masking.

Many changes take place in performance on speech psychoacoustic tasks from infancy to adulthood. Reduced processing efficiency in children seems to often be a contributing factor to performance, as different measures can yield different performance across similar age groups. Age seems to have little effect on the peripheral encoding of sound; however, central factors related to the processing of peripheral information are likely to be less effective in children than in adults. The concept of listening efficiency could potentially provide a reasonable explanation for the reduced processing ability in children, but the nature of such inefficiency is unclear.

Attentional Factors

Another factor that may play a role in differences between infant/young child and adult performance is attention. As suggested by Schneider *et al.* (1990), deficits in child performance on listening tasks could be linked to motivation and attention; however, the definition of attention can be vague.

Children undergo important developmental changes to attention between toddlerhood and the early school years. Attention can be defined by selectivity and those processes involved with that selectivity. With age, children develop an increased ability to focus attention (Weissberg, *et al.*, 1990). They also become more aware of noise as a distractor that can affect attention (Miller & Zalenski, 1982). The prefrontal cortex is responsible for inhibition, planning, and goal-directed attention, which are all important for overall attention. This area of the prefrontal cortex is at the height of development at the end of one year old and beyond (Diamond, 1991). Generally, there appears to be a dramatic increase during the preschool age range, before 5 years old, in children's ability for self-monitoring, self-regulation, inhibitory control in problem solving, memory tasks, error detection and error correction (Gerardi-Caulton, 2000; Jones *et al.*, 2003; Miller, 1990; Reed *et al.*, 1984). We might anticipate that the development of auditory attention, and spectral attention in particular, might parallel or rely on this general development of attention during the preschool years.

General attentiveness has been shown to account for some decreased performance in children when compared to adults, but does not offer enough evidence to fully explain the differences (Buss *et al.*, 2009). Overall, researchers have concluded that general developmental factors nonspecific to auditory processing, such as attention; fail to explain developmental effects observed in children. For example, Schneider *et al.* (1989) argued that motivation and inattention play little, if any, role in the poorer performance of children. They cited that the stability of thresholds over time and small effects of changing task/reward structure showed that inattention and motivation were not contributing factors to performance. As mentioned previously, a deficit in differential attention would also be reflected in psychometric function slopes, which did not appear (Schneider *et al.*, 1989).

More likely, it is the combination of sensory cues that play a role in the reduced performance of young children under some listening conditions that can be modified by selective auditory attention. Frequency-selective listening strategies have been shown to improve adult performance in detecting an expected frequency. However, this same selective listening decreases detection for frequencies that are not in the region of attention. Adults have demonstrated narrow attention bands that closely follow the shape of psychophysical auditory filters (Dai *et al.*, 1991; Greenberg & Larkin, 1968; Ison *et al.*, 2002; Scharf *et al.*, 1987). This similarity in shape and tuning indicates that adults are able to focus on single auditory filters to optimize listening and maximize detection by attending to the specific spectral regions where a target is expected. A common measure of this selective auditory attention is a modification of the probe-signal method, first introduced by Greenberg and Larkin (1968). In this two-alternative forced-choice measure, a target signal (tone) is presented in a noise masker over many trials. In a small proportion of trials an off-frequency probe, a tone with varying distance in frequency from the target, replaces the target signal. In the modified procedure, detection of the probes as a function of their relative frequency location provides a measurement of the attention band.

Using this off-frequency probe method, attention bands in adults have been measured to be similar in width to the auditory filter at 1000, 2000, and 4000 Hz but half as wide at 250 and 500 Hz (Dai, *et al.*, 1991). As seen in Figure 11, adults were able to detect the target with 90% accuracy (Scharf *et al.*, 1987). However, accuracy declined to 85% when probes were 25 Hz from the target and further fell to 65% when the probe was 75 Hz from the target. Probes at a distance of two critical bands from the target were not detected, with scores being similar to chance.

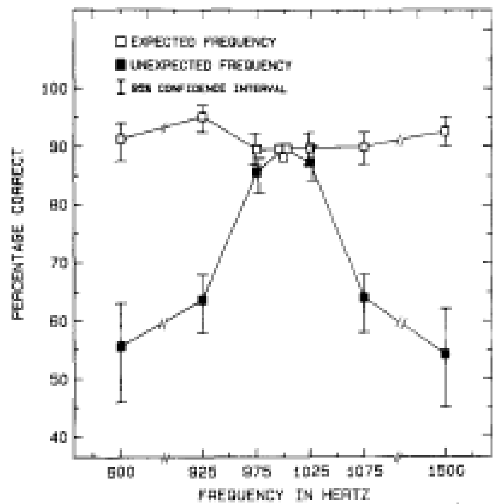

Figure 11. (From Scharf *et al.*, 1987) Percentage of trials that adult subject reported correctly the interval that contained the signal. Percentage correct is plotted as a function of the frequency of the pure-tone signal.

As seen in Figure 12, similar results were found in a study by Dai, Scharf, & Buus (1991), but with an additional observation that at low frequencies (500 Hz) the attention band function is steeper when probes occurred at frequencies higher than the target. Also, at higher frequencies (4000 Hz) the attention band was broader than the auditory filter. Similarly, Ison *et al.* (2002) found that for targets at 800 Hz and 1200 Hz target detection fell from 70% accuracy to 58% for probes 25 Hz from the target and to 40-45% for probes 50 Hz from the target.

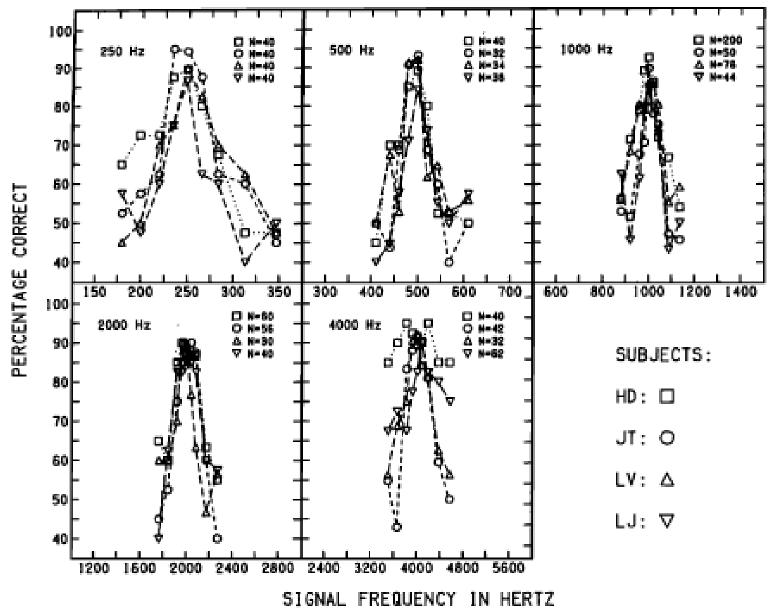

Figure 12. (From Dai *et al.*, 1991) Percentage of correct responses as a function of probe frequency for four listeners at five target frequencies. Each panel shows individual data at one target frequency.

In contrast, studies have shown that infants and children may listen less selectively than adults. Young infants have shown elevated, flat performance across a wide range of frequencies, indicating either extremely broad attention bands, or a lack of this listening skill (Bargones & Werner, 1994). This study used a similar off-frequency probe method used for adults to measure attention bands in infants. Seven- to nine-month-old infants had similar performance at 75% to 85% accuracy across targets and all off- frequency probes. As seen in Figure 13, even for remote probes 360 Hz from the target 1000-Hz tone, infants displayed accurate detection. These results reflect an inability to selectively attend to an auditory filter for an expected signal frequency, which suggests that infants and young children use a broad-band listening strategy. A consequence of this broad listening strategy is what can be classified as informational masking, or interference from sounds that typically would have little to no effect on adults.

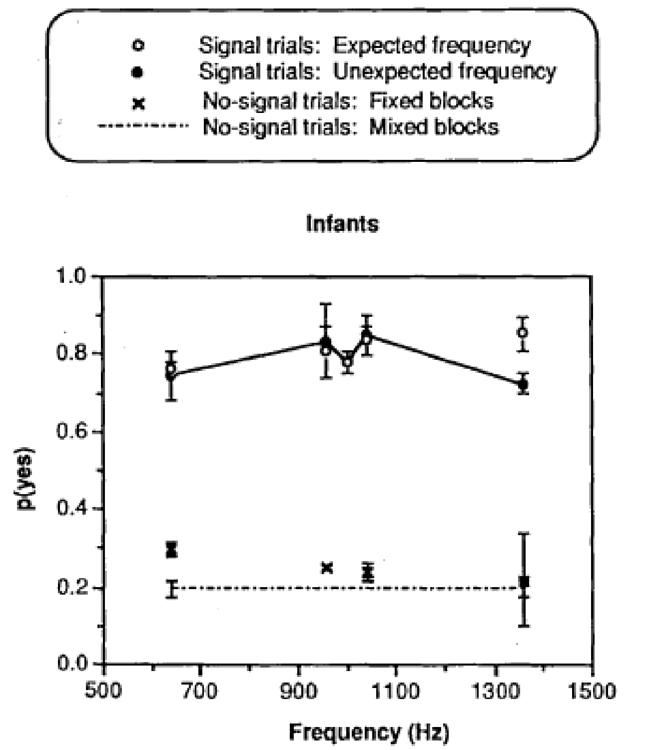

Figure 13. (From Bargones & Werner 1994) group listening bands for infants.

Important to the current study, there is limited information about how or when a selective attention listening strategy develops. A study by Greenberg, Bray, and Beasley (1970) measured attention bandwidths in five children ages 6 to 8 years old. This study measured detection for a target signal (1000 Hz) and four probe signals (850, 925, 1075, and 1150 Hz) in noise. The intensity level of probe frequencies were adjusted to produce equal detectability across all frequencies. As seen in Figure 14, results showed that one child demonstrated adult-like performance while the other four demonstrated a broader attention bandwidth than the adult controls. This limited data set suggests that between infancy and childhood, the listening strategy of spectral attention bandwidths may be developing.

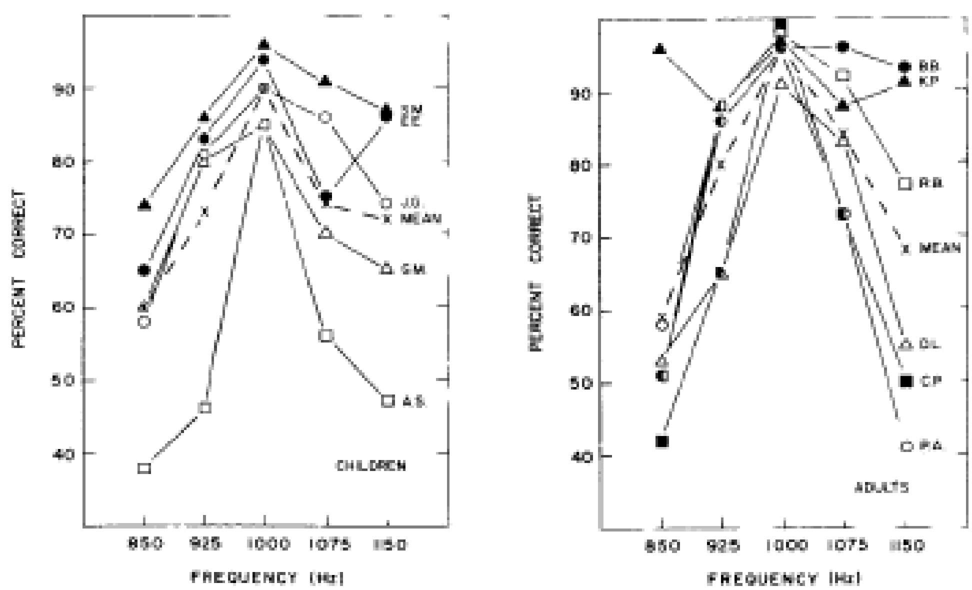

Figure 14. (From Greenberg *et al.*, 1970) Mean percent correct detections of individual subjects for tones in noise. Group means represented by a dashed line. Right panel children left panel adults.

Overall, this difference in performance between children and adults for attending to specific filters may indicate a still developing, or less efficient, listening strategy. Immaturity in listening strategies, such as this, may partially account for the overall susceptibility to noise that children demonstrate across basic speech psychoacoustic listening tasks and speech listening tasks.

General Methods

To summarize, children often do not perform at the same levels as adults on psychoacoustic and speech recognition tasks. Infants and young children have elevated thresholds for tone detection tasks compared to adults, despite a mostly mature peripheral auditory system. This has been shown in quiet, in noise, and in complex tone maskers (Elliott & Katz, 1980; Leibold & Neff, 2007; Olsho *et al.,* 1988; Schneider *et al.*, 1986; Trehub *et al.*, 1980; 1988). These detriments decrease child performance on psychoacoustic tasks compared to adults and persist despite evidence that the peripheral auditory filters and frequency selectivity are adult-like before 6 years old. For more complex tasks, children have decreased performance on speech recognition in noise compared to adults (Elliott & Katz, 1980; Elliott *et al.*, 1981, Johnson, 2000). Children continue to develop speech recognition abilities up to late childhood (Boothroyd, 1968; Elliot *et al.*, 1979; Hnath-Chisolm *et al.*, 1998; Papso & Blood, 1989), which is a developmental timecourse similar to that of the auditory cortex (Moore, 2002; Moore & Guan, 2001). In early childhood, the preschool age range, a dramatic increase in abilities occurs for many attentional tasks that may aid in performing auditory tasks.

The concept of spectral auditory attention bands impacts various areas, including tone detection, susceptibility to masking, speech recognition, and attention. Using the spectral attention-band listening strategy, adults can attend to a specific spectral region to maximize detection and suppress irrelevant noise outside of where the target is expected. Infants show an absence of this listening strategy by detecting tones at the target location and at other spectral locations equally well (Bargones & Werner, 1994). What remains unknown is the extent to which children's spectral attention band listening strategy is similar to adults. With the peripheral auditory system maturing before the age of 5, many acoustic cues used by adults are available to children, but they may not have learned to encode or use them in an efficient manner. A selective-attention listening strategy may develop along with higher auditory cognitive function. What is also unknown is how spectral attention bands may play a role in more common everyday listening such as understanding speech in noise.

Overall, it is unclear how spectral attention bandwidth develops and this study seeks to contribute to that understanding. This study incorporates basic psychoacoustic tasks to measure attention bands, and then seeks to investigate how these attention bands may play a role in speech-recognition tasks. Experiments are designed to investigate what role attention bands might play as the spectral location of maskers changes (i.e., spectrally remote and interleaved). Experiment 1 uses a basic psychoacoustic tone-detection task to measure spectral attention bandwidth in children and adults. Experiment 2 expands this knowledge into recognition of sentences in a remote noise. Experiment 3 investigates recognition of words in overlapping and non-overlapping noise. The results of this study are analyzed and discussed with regard to benefits and effectiveness of listening strategies across development.

The following methods are consistent across experiments unless otherwise stated.

Subjects

Children (5 to 8 years) and adults (18 to 40 years) participated in the study. This narrow age range for children was selected based on two main factors. First, as discussed in Attentional Factors, important attention processes appear to mature rapidly before the age of 5. This ability to attend to tasks in a more general manner allows for this age range of child subjects to complete tasks in a similar manner to adult subjects. Secondly, many previous studies have shown that this younger school-age range seems to show a more dynamic range of performance on psychoacoustics tasks when compared to adults and older children.

The results of various psychoacoustic studies indicate that the auditory abilities of children aged 5 to 8 years differ from those of older children or adults. As discussed previously, thresholds in quiet continue to improve to the age of ten, with more elevated thresholds in 4- to 6- year-olds (Leibold & Neff 2007). Eight- to ten-year-old children demonstrated thresholds lower than 5- to 7-year-old children for tone detection in a multi-tonal masker (Leibold & Bonino, 2009; Leibold & Neff, 2007). Leibold and Neff (2011) found that thresholds in remote noise maskers were elevated for a 4- to 6-year-old group relative to a 7- to 9-year-old or adult groups. Other work has shown that younger children are more susceptible to both energetic and remote masking (Allen & Wightman, 1995; Elliot & Katz, 1980; Leibold & Neff, 2011; Oh *et al.*, 2001) and may show greater release from masking when provided with additional cues. Finally, Bonino, Leibold, and Buss (2013) demonstrated that visual cues along with the auditory signal can improve thresholds substantially for children in the 5 to 7 year age range.

All listeners employed currently had normal hearing as indicated by a screening at 20 dB HL at octave frequencies from 500 to 8000 Hz (ANSI, 2004; 2010) and 25 dB HL at 250 Hz. Child participants demonstrated typical speech and language development as indicated by a language screening with the Token Test for Children, 2nd edition, and informal observation by a licensed speech-language pathologist. Subjects were recruited from The Ohio State University and surrounding Columbus metropolitan area and were compensated with course credit, money, and/or books for participating.

Procedures

Subjects were seated in a double-walled sound booth with the examiner. Stimuli were converted from digital to analog form using a personal computer (PC) and Echo Gina 3G digital-to-analog converters and were presented diotically over Sennheiser HD 280 circumaural headphones. Stimulus levels were calibrated using a Larson Davis sound-level meter (824) and flat plate headphone coupler (AEC 101). Child listeners had breaks and re-instruction to tasks as needed and completed all aspects of the experiment (language screening, hearing screening, experiment, and breaks) in one or two sessions, lasting no more than 2 hours each. The amount of time spent listening to experimental stimuli and responding was no more than 45 minutes per session for all participants.

Results

Raw data for individual subjects is presented in appendices for each experiment. When appropriate, individual data are presented in figures. Data for adults are always plotted separately from children due to the fact some parameter values differed between the groups in order to equate baseline performance across adults and children. Child groups were divided into sub-groups based on age and averaged.

Experiment 1. Attention Bandwidth in Children and Adults

The specific goal of Experiment 1 was to provide a measure of attention bandwidth in young school-age children. The inability to attend to specific spectral regions has been proposed as an explanation for children's poorer performance in noise compared to adults. However, there is little information regarding spectral attention bandwidths in children. Experiment 1 sheds light on how this effective listening strategy used by adults may develop in children.

Experiment 1 Methods

Subjects

Subjects were as described in the General Methods. This experiment included ten adult subjects (all female, aged 20 to 34, mean 22 years) and fourteen child subjects. Child subjects were divided into subgroups of 6- to 8-year-olds and 5-year-olds, with seven subjects in each group. The 6- to 8-year-old group (1 female) had an age range from 6;8 to 8;0 with a mean age of 7;6. The 5-year-old group (5 female) had an age range from 5;0 (years;months) to 5;11 with a mean age of 5;6. Eight of the child subjects completed the experimental tasks across two sessions.

Preliminary Measure

Prior to the experimental trials, each subject's masked threshold at 1000 Hz was measured. This preliminary measure served as a reference for presentation levels during the experiment and familiarized the subjects with the general task. The method of constant stimuli and a two-alternative forced choice design was used due to the limited number of trials required and similarity to the experimental task. Each trial consisted of two intervals, each containing a 500-ms broadband masker at 70 dBA, separated by a 300-ms silent inter-stimulus interval (ISI). In either the first or second interval, a 1000-Hz pure-tone signal 300-ms in duration (10 ms raised cosine rise/fall) was presented 100 ms after masker onset. Tone occurrence was equally and randomly distributed between the two intervals. The listeners indicated which interval contained the signal orally or through visual gesture.

Subjects listened to multiple rounds of trials. A round consisted of 10 trials at each of three to five signal intensities presented in random order. In the initial round, adult subjects were presented with four to five signal intensities in the 40- to 54-dB range. Children were presented with three to five signal intensities in the 44- to 60-dB range. Based on performance in this initial round, signal levels were adjusted on subsequent rounds to obtain three intensity levels that were below, at, and above threshold. Testing was discontinued after at least three rounds (30 trials) of data collected at the threshold intensity level. Cubic polynomial regression (Eq. 1) was used to fit the psychometric function relating percent correct detection to signal level for each individual subject. Threshold determined by the signal intensity at which the curve crossed the 70% accuracy point.

$$f = \frac{a}{1+e^{-\left(\frac{x-x_0}{b}\right)}}$$

Eq. 1

Experimental Stimuli & Procedure

This experiment employed a two-alternative forced-choice design, with a preceding priming interval to make three intervals total (3I, 2AFC, See Figure 15). Subjects were instructed that they would hear a priming tone in noise then hear another tone in one of two following intervals. Each interval contained a 500-ms broadband (100 to 8000 Hz) white noise masker at 70 dBA. A 300-ms silent ISI separated the maskers. The priming interval also contained a 1000-Hz priming tone, which was 300 ms in duration including 10-ms raised cosine rise/fall and was temporally centered in the masker. The priming tone was presented above the individual subject's masked threshold.

Either in the second or third interval (Interval A or B), one of nine target or probe tones was presented at or above threshold. The target tone was at 1000 Hz and probe tones were at distances of 1/4, 1/3, 1/2, 1, and 2 equivalent rectangular bands (ERBs) above and below 1000 Hz (967/1033, 957/1044, 935/1067, 874/1140, or 762/1294 Hz). These tones were 300 ms in duration with a 10-ms raised cosine rise/fall and were temporally centered in the masker.

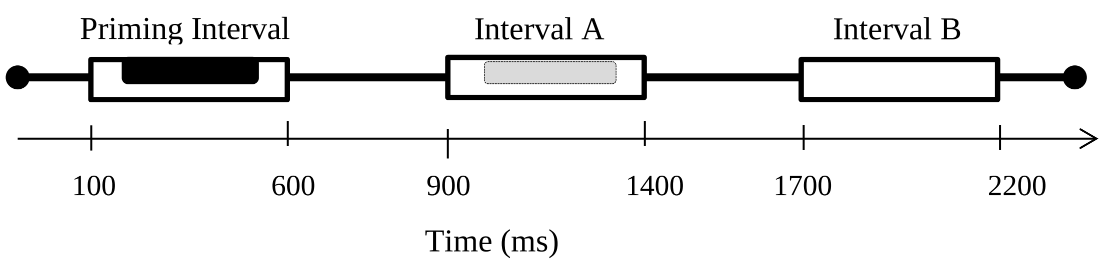

Figure 15. Schematic of experiment intervals for Experiment 1.

Subjects indicated if they heard the tone in Interval A or B by pointing to visual representations of the intervals on a computer monitor or by verbally stating which interval, depending on the most reliable measure for each individual subject. Child subjects were presented with images for intervals, see Figure 16.

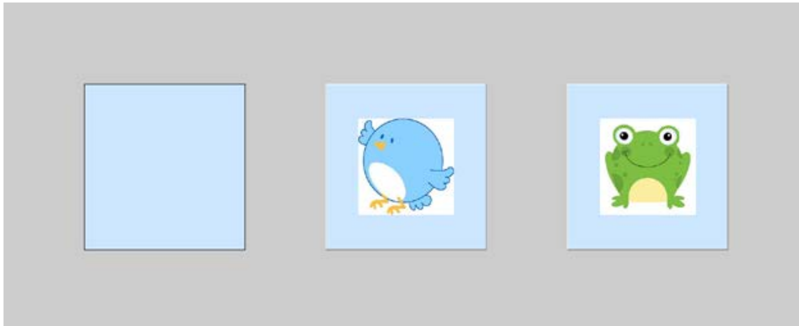

Figure 16. Example of the user interface for children in Experiment 1. Blue boxes changed color to green during the interval presented. The first blank blue box was the priming interval, while the following blue boxes with images represented Intervals A and B. Adult subjects had a similar user interface with pictures changed to letters A and B.

Instructions for the task were purposefully minimal so not to over or under emphasize the importance of the priming tone. Subjects were instructed that they would always hear a tone (a “beep”) in the first interval, and their job was to indicate which of the two other intervals (A or B) contained a tone. At times, child subjects were unsure which interval contained a tone. In these instances they were encouraged to make their best guess. Some other subjects, both child and adult, became aware of the off-frequency probes. In these situations, the examiner’s response was to confirm that this was correct and that the subject should simply select the interval in which a tone was heard. Two adult subjects (A1 and A2) were experienced listeners with extensive knowledge of the study design.

Each presentation round consisted of 24 target trials and 8 probe trials. As illustrated in Table 1, each round contained the target 1000 Hz tone during 75% of the trials. The remaining 25% of trials contained one of the eight probe tones: 762, 874, 935, 967, 1067, 1140, 1294 Hz. The order of target and probe conditions were randomized for each listener in each round and equated to occur randomly but equally in Interval A or B. All adult subjects listened to 12 rounds for a total of 384 trials. Child subjects listened to at least 10 rounds, with the exception of subject C1 who listened to 7 rounds due to decreased tolerance for the listening activity (see Table A.2).

Table 1. Frequency of occurrence for target and probe tones in Experiment 1.

	ERB Distance				
Probe Frequency (Hz)	0	1/4	1/2	1	2
Low	1000	967	935	874	762
High		1033	1067	1140	1294
% Occurrence	75	25			

Table 2 displays masked thresholds and presentation levels for the child subjects. Priming tones were presented in the 5-9 dB range above individual subject’s masked threshold for adults and in the 3-13 dB range above threshold for children. This range was selected to be above masked threshold to ensure audibility and encourage the listener to consider 1000 Hz to be the expected frequency for the following intervals. Probe and target tones occurring in either Interval A or B were presented at a level 1-5 dB above threshold for adults and 0-9 dB above threshold for children, with the exception of C13, whose obtained masked threshold appeared to be elevated compared to his actual abilities. For this subject, target/probe tones were presented +/- 1 dB in reference to the obtained masked threshold.

Table 2. Individual child-subject masked thresholds and intensity-level ranges for priming tone and probe/target tone presentation.

Listener, Age (yr; mo)	Threshold (dB SPL)	Priming Tone Intensity Range (dB re: masked threshold)	Probe Tone Intensity Range (dB re: masked threshold)
C 1 5;0	55	7 - 13	3 - 9
C 2 5;3	51	7 - 9	3 - 5
C 3 5;4	49	7 - 9	3 - 5
C 4 5;5	56	6 - 8	2 - 4
C 5 5;9	51	5 - 7	1 - 3
C 6 5;10	55	5 - 7	1 - 3
C 7 5;11	51	5 - 11	1 - 7
C 8 6;8	50	4 - 6	0 - 2
C 9 7;6	51	5 - 7	1 - 3
C 10 7;6	50	6 - 8	2 - 4
C 11 7;7	52	4 - 6	0 - 2
C 12 7;8	53	5 - 7	1 - 3
C 13 7;8	53	3 - 5	-1 - 1
C 14 8;0	49	5	1

Experiment 1 Results and Discussion

Figure 17 presents the masked thresholds for individual adults, which ranged from 43-47 dB SPL. Figures 18 and 19 show masked thresholds for individual 6-to 8-year olds and 5-year-olds, respectively. Child subjects are numbered in ascending order of chronological age. Masked thresholds for children ranged from 49-56 dB SPL. This difference in performance between adults and children is consistent with previous masked thresholds reported by Leibold and Neff (2007), who showed threshold in a 60 dBA broadband masker to be approximately 45 dB SPL for adults and approximately 48 dB for 5-to 7-year-olds.

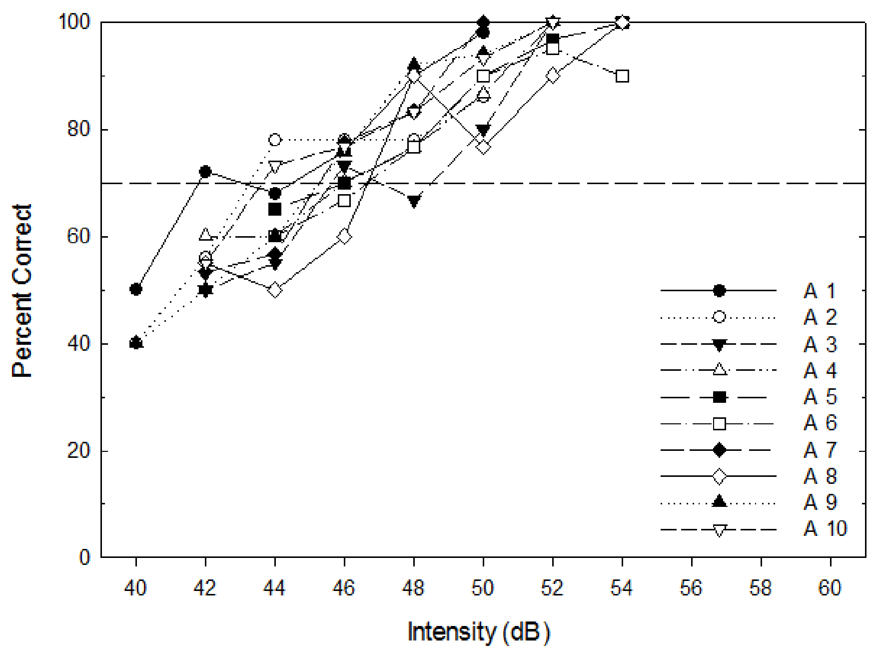

Figure 17. Adult's percent correct detection of a 1000 Hz tone in a 70 dBA white noise, as a function of tone level (dB SPL). The dashed line marks the 70% criterion, indicating threshold level.

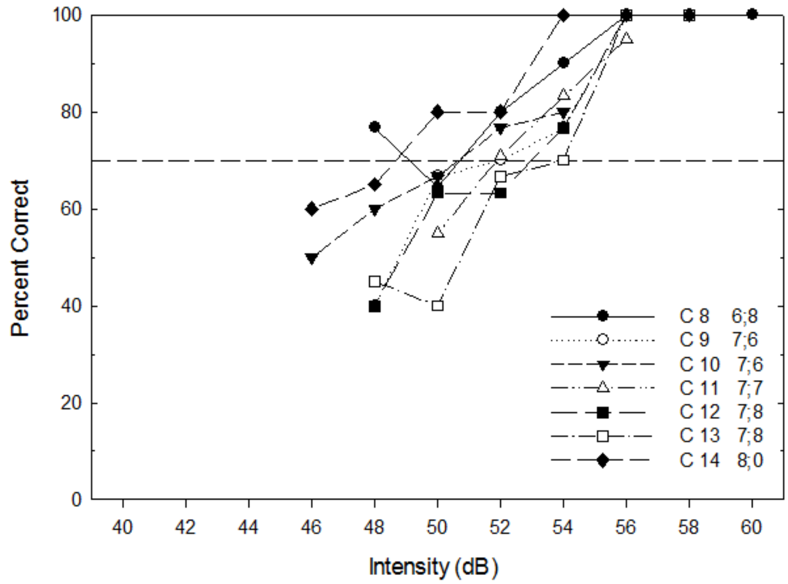

Figure 18. Six- to eight-year-old group percent correct detection of a 1000 Hz tone in a 70 dBA white noise, as a function of tone level (dB SPL). Age in years and months is provided next to subject number. The dashed line marks the 70% criterion, indicating threshold level.

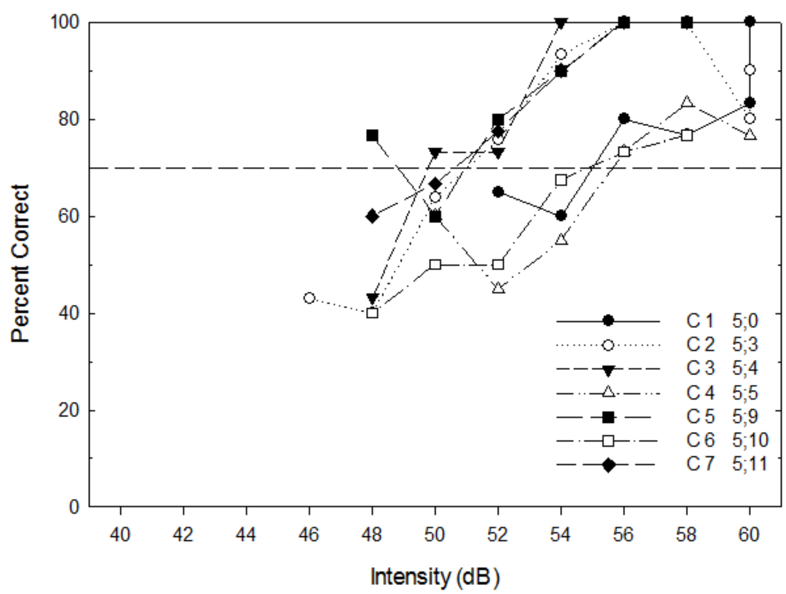

Figure 19. Five-year-old group percent correct detection of a 1000 Hz tone in a 70 dBA white noise, as a function of tone level (dB SPL). Age in years and months is provided next to subject number. The dashed line marks the 70% criterion, indicating threshold level.

Figure 20 shows percent correct tone detection as a function of target-/probe-tone frequency for the adult subjects. A3, A6, and A8, demonstrated a highly asymmetrical detection of remote probes. In these cases, probe tones one to two ERBs below the priming tone were detected with 25-50 %- points more accuracy than similar probe distances above the 1000-Hz priming tone. This may indicate that these listeners were attending to the lower frequency region in addition to the target.

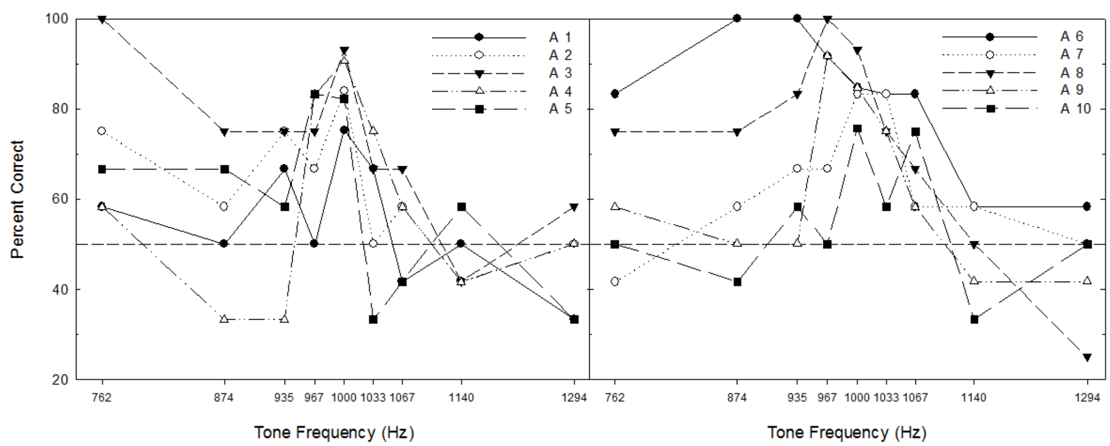

Figure 20. Mean percent-correct detection by individual adult subjects for tones in 70 dBA broadband noise. The dashed line indicates chance level at 50% accuracy.

As expected, there was increased variability in performance of children (Figures 21 and 22) as compared to adults. In general, it appears that many 6- to 8-year-olds and most 5-year-old subjects and demonstrated relatively accurate detection of probe tones, with performance on par with detection of the target 1000-Hz tone. These results suggest that subjects, especially those in the younger age range, are monitoring with equal net effectiveness spectral regions outside of the location of an expected target.

As seen in Figure 21, Subject C11's performance dropped drastically below chance levels, even to 0% accuracy for the 935-Hz probe tone. This suggests that a reversal occurred for this subject, in which he/she was selecting the interval *not* containing the tone. Due to this apparent reversal, Subject C11's data was not included in the means shown in Figure 23 or in any of the statistical analyses.

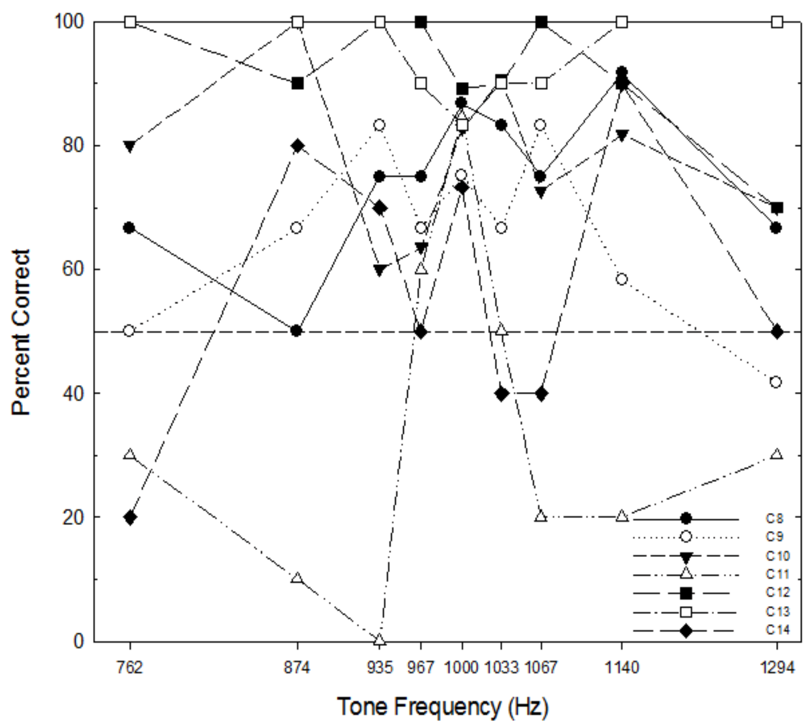

Figure 21. Mean percent correct detection for individual child subjects in the 6- to 8-year-old group for tones in 70 dBA broadband noise. The dashed line indicates chance level at 50% accuracy.

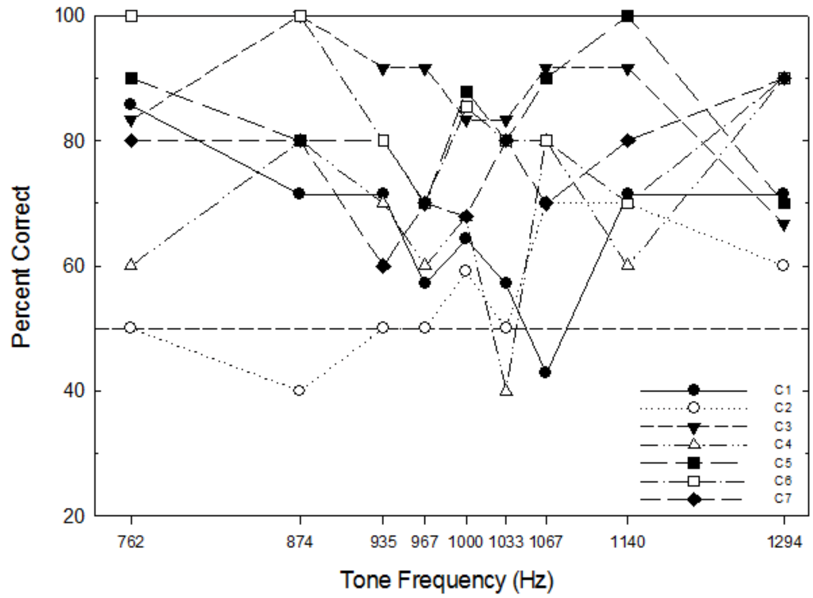

Figure 22. Mean percent correct detection for individual child subjects in the 5-year-old group for tones in 70 dBA broadband noise. The dashed line indicates chance level at 50% accuracy.

Figure 23 shows group mean target-/probe-tone detection by the three listener groups. Apparent from the figure is that the adults displayed the narrowest attention bands, followed by the older children, and the youngest children displayed the broadest attention tuning. Auditory attention-band tuning was sufficiently broad in the youngest children to be essentially flat.

In accord with our adjustment to target/probe presentation level for the adult and child subjects, all three subject groups displayed generally similar mean accuracy in detection of the 1000-Hz target. The oldest children produced scores within 3%-points of those of the adults (81.8% and 84.7% respectively), and the youngest children produced scores that were 11%-points lower (73.7%). The target-tone accuracy displayed by the adults and the general decline in performance as probe tones were farther away in frequency from the target is consistent with previous findings (Dai *et al.*, 1991; Ison *et al.*, 2002; Scharf *et al.*, 1987). However, one notable difference is the asymmetrical detection of probe tones above and below the target with steadily decreasing accuracy for the higher-frequency probe tones as they became more remote. A shallower decrease in performance is observed for the lower-frequency probe tones. Previous findings have suggested an increase in detectability for higher-frequency probe tones, which is the opposite of what is observed here.

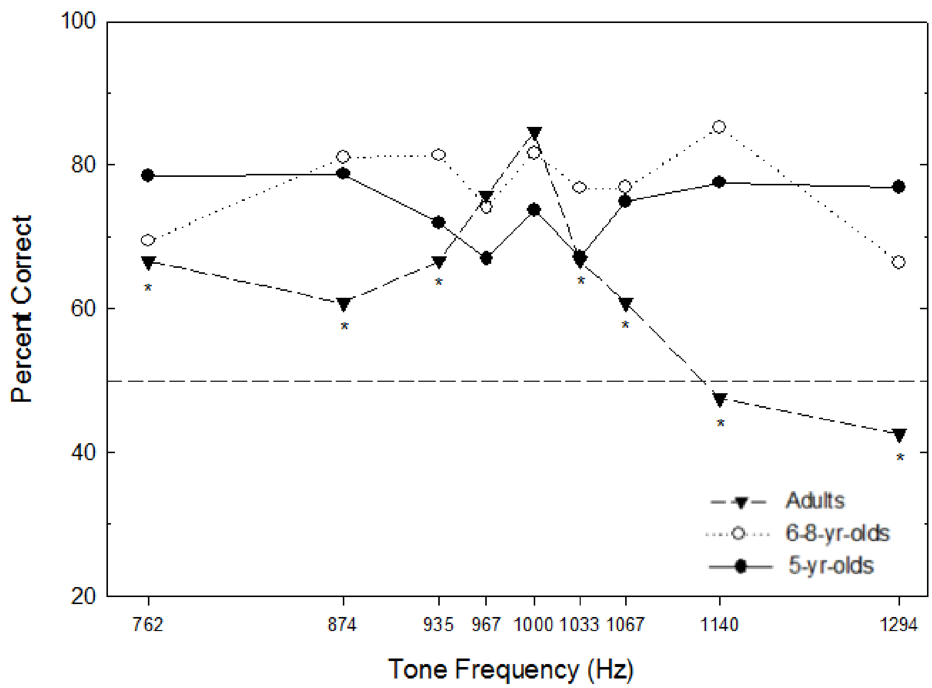

Figure 23. Mean percent correct detections by child and adult groups for tones in 70 dBA broadband noise. The dashed line indicates chance level at 50% accuracy. Significant differences between probe tones and the target 1000 Hz tone are indicated by an asterisk.

Mean detection scores were normalized for variance using the rationalized arcsine transformation (RAU; Studebaker, 1985), and all statistical analyses employed arcsine-transformed units.

A two-way repeated-measures ANOVA was performed having factors age and frequency. Of interest was the interaction between these factors, and this interaction between age and frequency was significant [$F(16,160) = 3.63$, $p < .001$]. To investigate performance as a function of signal frequency (the shape of the curve) for each age group, simple-effects testing involving three separate one-way repeated-measures ANOVAs was performed, one for each subject age group.

A one-way repeated-measures ANOVA was performed for each subject group separately. A significant main effect of target-/probe-tone frequency was observed only for adults [$F(17, 72) = 9.57$, $p < .001$)]. Post hoc *t* tests, using a Bonferroni-correction for multiple pair-wise comparisons, revealed that significant differences existed between detection of the 1000-Hz target versus all probe tones ($p \leq .036$) except for the most adjacent tone on the low-frequency side (967 Hz). For the 6- to 8-year-old group and the 5-year-old group, the one-way repeated-measures ANOVAs yielded no significant main effect of tone frequency [$F(13,40) = 1.01$, $p = .446$; $F(14,48) = 1.49$, $p = .185$, respectively].

Most adults and some older children appear able to attend to specific spectral regions where they expect a target to appear as a strategy to maximize detection. However, this strategy appears to be far less prominent in the younger children tested here. In fact, the current data suggest that auditory attention is not at all specific in frequency for children in this younger age group. These data support the idea that adults employ narrow spectral attention bands as a listening strategy to increase detection of expected frequencies, but that this listening strategy is not fully developed in 5- to 8-year-old children.

Bargones and Werner (1994) provided some possible explanations for infants' equal performance across the target and near and far probes. These included an inability to attend to a selected filter range, possible monitoring of a broad range of filters, misplacing attention filters (monitor remote filter, close but not right), or not monitoring any particular filters. The current results suggest that 5-to 8-year-olds may be using these same inefficient strategies. Although determined largely by target- and probe-tone presentation level, detection accuracy at 1000 Hz was lowest for the youngest children tested currently. This decrease in detecting the 1000-Hz target by the 5-year-olds might reflect these inefficiencies.

All age groups tested here, even the youngest, demonstrated a drop in performance at the closest most proximal (1/4-ERB distance) probe tones. This indicates that there might be some selective attending to the 1000-Hz target, even by the youngest subjects. However, for both child groups, there was not a continued decrease in performance at more remote probe tones. Five-year-olds performed at near target level for both 1/2-ERB distance probes, while 6- to 8-year-olds performed at near target level for one 1/2-ERB distance probe. The 6- to 8-year-old group displayed a drop in performance at both 2-ERB distances, which may indicate an emergent narrowing of spectral attention. The ability to attend to a focused spectral region where a signal is expected to occur appears to develop over time, with ages 5 to 8 years representing an early transitional stage in the manifestation of this listening strategy. In Experiments 2 and 3, we explore how this broader spectral attending may be detrimental to an essential skill -- understanding speech in noise.

Experiment 2: Effects of Remote Maskers on Speech Recognition

The specific goals of Experiment 2 were to transition from a non-speech psychoacoustic task to one involving speech stimuli and to examine directly the effect of remote masking noise on speech recognition for adults and young school-age children. For tone-detection tasks, there is evidence that young children (4 to 6 years old) are susceptible to a remote masker, whereas older children (7 to 9 years old) perform more similarly to adults, who are not susceptible to remote masking. A suggested explanation for young children's greater susceptibility to noise is a reduced ability to attend to a signal frequency (Leibold & Neff, 2011), which is also supported by the results of the current Experiment 1. Further, extensive research has found that speech recognition is more susceptible to noise for children up to the age of ten and through the teenage years (Elliott & Katz, 1980; Elliott *et al.*, 1981; Johnson, 2000). Speech is a far more complex signal than pure tones with many redundant acoustic and auditory cues to support intelligibility. Experiment 2 extends the results of psychoacoustic studies in an attempt to link the limited ability to attend to a signal frequency to the more general findings of increased susceptibility to noise during speech recognition.

Experiment 2 Methods

Subjects

Subjects were the same as described in the General Methods, with the exceptions noted. This experiment included 10 adult subjects (all female, aged 19 to 25, mean 21 years) and 18 child subjects. Child subjects were divided into subgroups of 7-year-olds and 5-year-olds, with 9 subjects in each group. The 7-year-old group (2 females) had an age range from 7;1 to 7;11 with a mean age of 7;6. The 5-year-old group (7 females) had an age range from 5;0 to 5;10 with a mean age of 5;6. Child Subject 13 was enrolled in speech therapy for articulation errors at time of testing and was included due to scores within the average range on the language-screening measure. Child Subject 9 had hearing thresholds of 25 dB HL at 2000 and 4000 Hz and 30 dB HL at 8000 Hz in the right ear only.

Stimuli

Target stimuli consisted of Bamford-Kowal-Bench sentences (Bench *et al.*, 1979). Each sentence contained three to five keywords, which were scored to measure intelligibility. Audio files of the sentences had a sample rate of 44,100 Hz at 16-bit resolution. Sentence stimuli were first band-pass filtered from 100-1500 Hz using a 2000-order FIR filter, implemented in MATLAB (fir1). The total RMS level of each filtered sentence was equated within 1 dB. Speech-shaped noise (SSN) was created by concatenating the equated sentences and measuring with a 65,000 point fast Fourier transform having a Hanning window and 97% overlap. The resultant frequency-amplitude values were used to shape a white noise using an arbitrary-response digital filter (500 order fir2 in MATLAB). Speech and SSN were mixed at a 0 dB signal to noise ratio (SNR) for adults and at a +5 dB SNR for children. This was done to approximately equate recognition performance across groups. For both groups, the speech and noise mixture was set to 55 dBA, which was then presented as a control condition.

For the experimental conditions, the control stimulus was combined with a spectrally remote noise masker. Five remote maskers were created: a one-octave band (2500-5000 Hz) and four 1/3-octave bands (3000-3780 Hz; 4000-5040 Hz; 5000-6300 Hz; and 6000-7560 Hz). Remote maskers were created using low-noise noise (LNN, see Pumplin, 1985; Hartmann & Pumplin, 1988). The use of LNN ensured that modulations of the masker envelopes would be minimal and more similar for all narrow-band maskers. The LNN was created by dividing a Gaussian noise by its Hilbert envelope 100 times using a MATLAB script (as described in Healy & Bacon, 2006; Kohlrausch *et al.*, 1997). Lower cut-offs of these remote maskers were 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 *et al.* (1997). The lower cut-off of the lowest 1/3-octave band was set to 3000 Hz rather than 2500 Hz as for the octave band, because the narrower bandwidth resulted in an increased spectrum level. Remote maskers were each set to 75 dBA and mixed with the speech and SSN mixture (see Figure 24). To reduce the excess masking associated with masker onset, often termed “overshoot” with tone in noise maskers (Bacon & Liu, 2000; Zwicker, 1965), a minimum of 500 ms masker fringe preceded and followed each sentence.

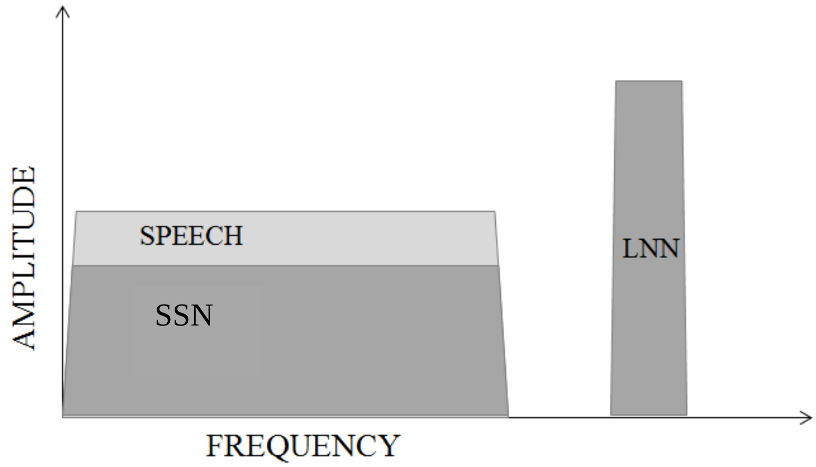

Figure 24. Schematic of stimuli for Experiment 2. Speech in speech-shaped noise (SSN) was presented along with a low-noise noise (LNN) remote masker.

Procedure

Each participant heard the filtered speech and SSN only (CONTROL) along with the five remote-masker conditions (CONTROL + remote masker), resulting in a total of six conditions. Each round in this experiment was composed of the six conditions, with 10 sentences per condition. Subjects heard two such rounds, with a new randomization of condition order for each round. Sentence list-to-condition correspondence was balanced across listeners to control for possible differences in the difficulty of the lists. At the end of all trials, subjects were presented 10 sentences of only the band-pass speech in quiet. Subjects 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. Subjects did not have a fixed time restriction for response.

Experiment 2 Results and Discussion

As shown in Figure 25, mean adult performance in CONTROL (speech +SSN) was 70.2% (SE = 3.4). Mean speech-recognition performance in remote masker conditions were within 2.5%-points of this value, with the exception of the most remote masker (6000-7560 Hz), which was 6.6%-points below performance in CONTROL. The average for the pooled remote maskers was within 1.2%-points of the CONTROL mean.

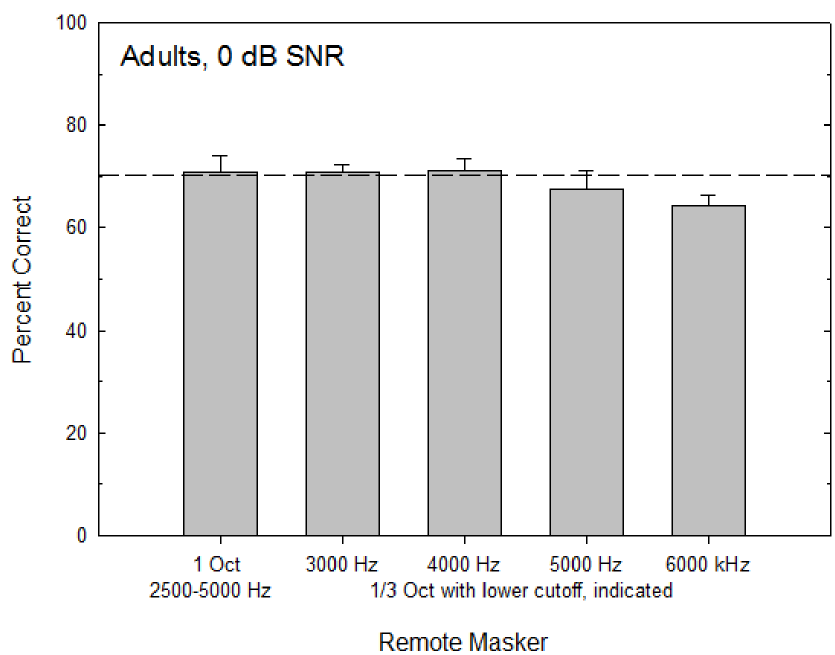

Figure 25. Mean sentence recognition scores (and standard errors) for ten adult subjects in the presence of remote- noise maskers. The dashed line indicates average performance in the control condition, in which remote noise was absent.

As shown in Figure 26, mean 7-year-old performance in the control condition (speech +SSN) was 67.6% (SE = 2.0). Mean speech-recognition performance for each of the remote maskers fell within 5%-points of this value. The average for the pooled remote masker conditions was within 2%-points of the CONTROL mean.

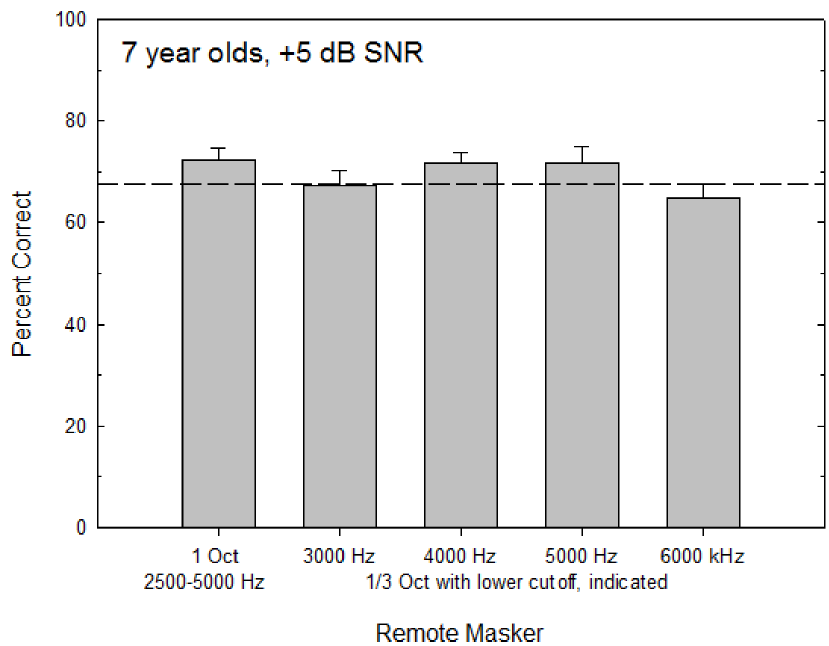

Figure 26. Mean sentence recognition scores (and standard errors) for nine 7-year-old subjects in the presence of remote-noise maskers. The dashed line indicates average performance in the control condition, in which remote noise was absent.

As shown in Figure 27, mean 5-year-old performance in the control condition (speech +SSN) was 62.5% (SE = 2.0). In contrast to what was seen in adults and 7-year-olds, performance was reduced in the presence of all remote maskers. Mean speech-recognition performance in remote maskers fell below CONTROL by 5.5 to 10%-points. The average for the pooled remote masker conditions was 8%-points below CONTROL.

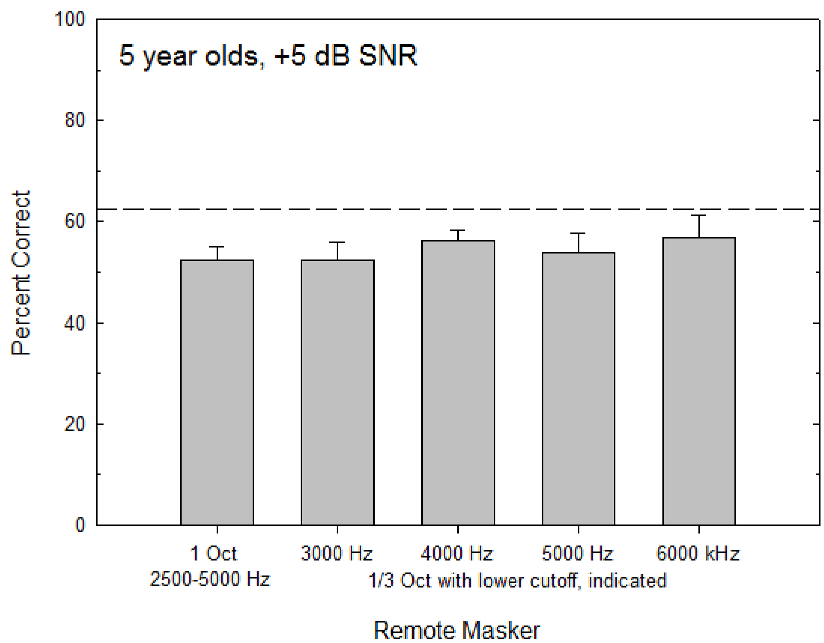

Figure 27. Mean sentence recognition scores (and standard errors) for nine 5-year-old subjects in the presence of remote-noise maskers. The dashed line indicates average performance in the control condition, in which remote noise was absent.

Overall, the younger group of children demonstrated reduced sentence recognition in the presence of remote maskers. However, the spectral distance from the target did not seem to play a role, with a similar decrease in speech recognition across all remote maskers. To further examine the similarity in performance for all remote maskers, one-way repeated-measures ANOVAs were calculated for each age group on performance in the remote masker conditions. Mean sentence recognition scores were normalized for variance using the rationalized arcsine transformation (RAU; Studebaker, 1985), and all statistical analyses employed arcsine-transformed units. No significant main effect of remote-masker condition was observed for adults [F (13,36) = 1.30, p = .289], 7-year-olds [F (12,32)] = 1.70, p = .174, or 5-year-olds [F (12,32) = 0.60, p = .668]. This lack of significant differences between maskers indicates that the observed effects of remote masking were irrespective of bandwidth or spectral location of the masker. Further, it enables the pooling of performance across all remote maskers. A planned paired-samples t-test was conducted for each listener group to compare performance in CONTROL with that of the pooled remote maskers. There was not a significant difference in performance for adults [t(9) = 0.49, p = .637] or for the 7-year-olds [t(8) = -1.03, p = .334], indicating that those groups were not affected by remote masking. However, the 5-year-old group did display a significant difference in performance between CONTROL and the pooled remote maskers [t(8) = 2.58, p = .033], indicating a reduction in sentence recognition in the presence of the remote maskers.

Figure 28 shows the amount of remote masking (difference between performance in CONTROL and that in the pooled remote maskers) as a function of age. Adult mean remote masking was 1.2 %-points, indicated by the dashed line. A Pearson correlation was conducted to determine the relationship between amount of remote masking and child-listener age. There was a large, negative correlation (r_s = -0.61, p = .007) with the older children showing decreased remote masking with age.

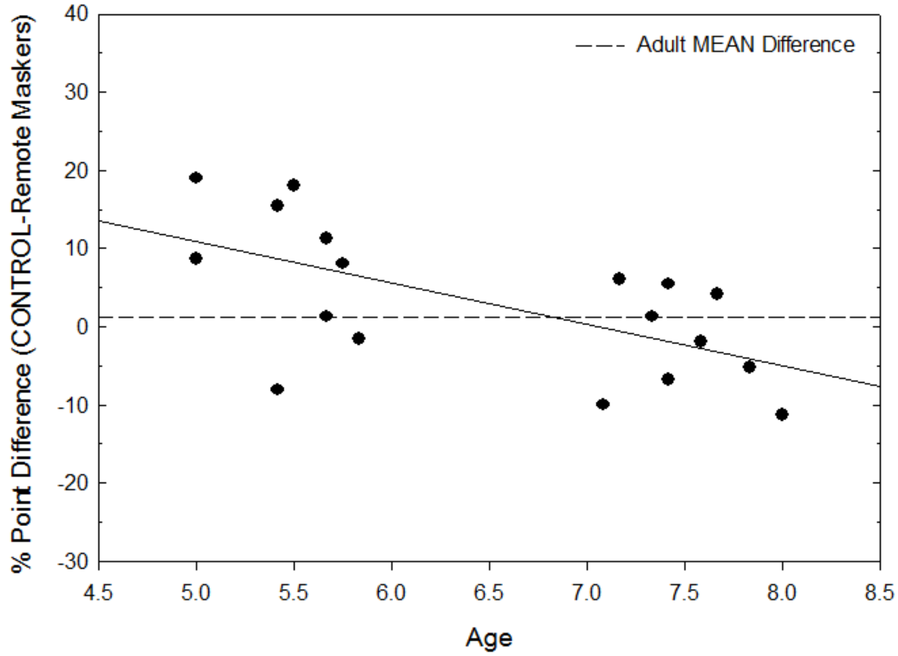

Figure 28. Amount of remote masking as a function of child-subject age. Shown are percentage point differences between performance in the control condition and performance in the pooled remote-masker conditions for each subject. The dashed line represents mean remote masking for the adults. (r_s = -0.61, p = .007)

Experiment 2 provides evidence that 5-year-old children are more susceptible than 7- to 8-year-old children to masking by spectrally remote noises. When combined with the results of Experiment 1, it is concluded that this remote masking susceptibility may be due to the younger subjects' tendency to monitor across a broader range of auditory filters. The current results also seem to capture what may be a critical developmental period, as listening strategies seem to increase toward adult-like performance between the relatively narrow age range of 5 to 8 years.

Experiment 3: Spectrally Interleaved Speech and Noise

The specific goal of Experiment 3 was to examine the effect of adjacent and overlapping masking noise on speech recognition for adults and young school-age children. Experiment 1 suggests that young children, aged 5 to 8 years, have broader auditory attention filters than adults. Experiment 2 suggests that young children, roughly aged 5 years, are more susceptible to remote masking than adults or older children. However, the conditions employed in those experiments are considerably removed from what is typically encountered during everyday speech recognition. Due to this limitation, these conditions fall short of providing an explanation for children's poorer speech recognition in noise and the need for increased SNRs, relative to adults. Experiment 3 takes a step closer to a more plausible scenario of everyday speech recognition in noise, especially in light of the "glimpsing" model of how this process is performed (Apoux & Healy, 2009; 2010; Cooke, 2005).

Apoux and Healy (2010) found that normal-hearing adult listeners were little affected when bands of noise were introduced between bands of speech. Consonants and noise were filtered in complementary fashion and the two were added at various ratios. Consonant recognition remained essentially intact so long as the SNR was 0 or above. It was concluded that adults possess the ability to strictly isolate frequency regions containing clean speech even when surrounding frequencies contain only noise. This in turn supports the notion that the output of auditory filters can be processed with substantial independence (Apoux & Healy, 2009; 2010; also see Kidd *et al.*, 2005).

If, in contrast, the broader attention filters of younger listeners prevent them from being able to perceptually isolate specific frequency regions, then those listeners may be hindered more than adults by spectrally adjacent noise. In the current experiment, speech recognition in the presence of spectrally complementary noise is examined in adults and children. In doing so, this experiment mimics in some ways the process of "glimpsing" spectro-temporal regions of clean speech for noise backgrounds. The current experiment also provides an indication of the independence of auditory channels in children relative to adults.

Experiment 3 Methods

Subjects

Subjects were the same as described in the General Methods. This experiment included 10 adult subjects (all female, aged 19 to 22, mean 20 years) and 18 child subjects. Child subjects were divided into subgroups of 6- to7-year-olds and 5-year-olds, with 9 subjects in each group. The 7-year-old group (3 females) had an age range from 6;7 to 7;11 with a mean age of 7;6. The 5-year-old group (5 females) had an age range from 5;0 to 5;11 with a mean age of 5;6. Child Subject 4 had hearing thresholds of 25 dB HL at 500 Hz and 35 dB HL at 8000 Hz in the left ear only.

Stimuli

Target stimuli were the Phonetically Balanced Word Lists - Kindergarten (PBK; Haskins, 1949) spoken by a male talker. Audio files of the words with carrier phrases had a sample rate of 44,100 Hz at 16-bit resolution. Stimuli were band-pass filtered from 80-7563 Hz, then divided into 30 contiguous 1-ERB-width bands using two cascaded 12th-order digital Butterworth filters (Apoux & Healy, 2009; 2010). Stimuli were filtered in both the forward and reverse directions to produce zero phase distortion. The total RMS value for each word in the combined 30 bands were equated to be within 1 dB. A simplified SSN was generated using a first-order Butterworth filter, resulting in a constant spectrum level below 800 Hz and 6 dB/oct roll-off above 800 Hz. The SSN was also filtered into 30 1-ERB-width bands in an identical manner to the speech stimuli (See Figure 29).

Fifteen ERB-wide bands of speech and fifteen ERB-wide bands of noise were mixed, as alternating bands, to create an interleaved speech and noise condition, with limited spectral overlap between the speech and noise (OFF condition). The same fifteen ERB-wide bands of speech were mixed with fifteen on-frequency ERB-wide bands of noise to create complete spectral overlap (ON condition). A third condition was the fifteen ERB-wide bands of speech combined with the SSN masker (all 30 1-ERB-width SSN bands; broadband condition BB). These three conditions were presented at +5 dB SNR for child subjects and at 0 dB SNR for adult subjects, to roughly equate performance between children and adults in the BB condition. A baseline condition consisting of the fifteen ERB-wide bands of speech in quiet was also presented, yielding a total of four conditions. Each individual speech stimulus was calibrated such that the 30 bands of speech with the 30 bands of SSN at the appropriate SNR was 70 dBA. All other conditions were based on this calibrated stimulus so that spectrum levels remained constant across conditions.

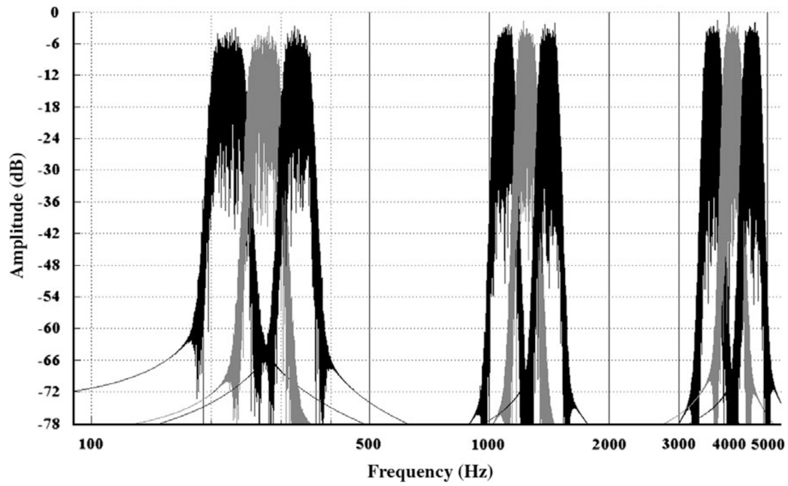

Figure 29. (From Apoux & Healy, 2009) Example of long-term average spectra and amount of energy overlap for the output of select filters in low, mid, and high frequency ranges of the 30 1-ERB-band filtering. Shown are filters 4, 5, 6; 14, 15, 16; and 24, 25, 26 when the input signal was 60 seconds of white noise. Grey and black are used to differentiate adjacent filters.

Procedures

Subjects were instructed to listen to the talker and repeat what was heard. For PBK stimuli, each word was preceded by the carrier phrase "Say the word…" to indicate that the listener should begin attending. The examiner marked correct responses and transcribed incorrect responses. The carrier phrase was not scored. Each listener was presented with two 20-word blocks each for conditions OFF, ON, and BB and one 20-word block for the quiet condition. Each subject was presented conditions in two rounds (only one round of QUIET) with a new randomization of condition order for each round. Word-to-condition correspondence was balanced across listeners to control for possible differences in the difficulty of the word blocks.

Experiment 3 Results and Discussion

As shown in Figure 30, group-mean intelligibility as percent correct was plotted in each condition (BB, ON, OFF, and QUIET) for each subject group. Mean word recognition scores were normalized for variance using the rationalized arcsine transformation (RAU; Studebaker, 1985), and all statistical analyses employed arcsine-transformed units. A one-way repeated-measures ANOVA calculated on the adult data indicated a significant main effect [F (12, 27) = 209.93, $p < .001$]. Post hoc t tests, using a Bonferroni-correction for multiple pair-wise comparisons, revealed significant differences between all conditions ($p < .001$), except ON and BB ($p = .097$).

A one-way repeated-measures ANOVA on the performance of 6-to 7-year-olds indicated a significant main effect [F (11, 24) = 89.32, $p < .001$]. Post hoc t tests, using a Bonferroni-correction for multiple pair-wise comparisons, revealed significant differences between all masker configurations ($p \leq .002$) except ON and BB ($p = .290$).

A one-way repeated-measures ANOVA on the performance of 5-year-olds also indicated a significant main effect [F (11, 24) = 85.18, $p < .001$]. Post hoc t tests, using a Bonferroni correction for multiple pair-wise comparisons, revealed a significant difference between QUIET and all other masker configurations, as well as between OFF and BB ($p < 0.001$). Importantly, for 5-year-olds, there was no significant difference between OFF and ON ($p = 0.154$). This indicates that even noise that was mostly non-overlapping interfered with speech recognition as much as noise that was on frequency.

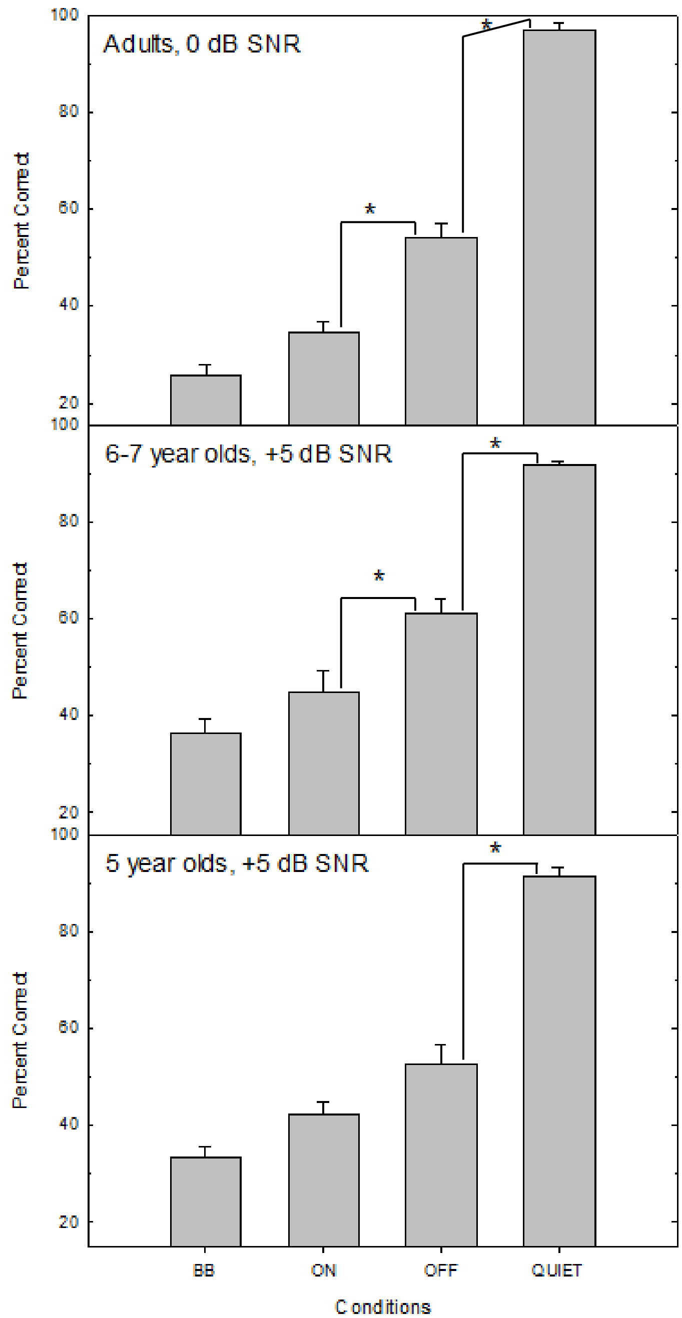

Figure 30. Mean performance of adult, 6-to 7-year-old, and 5-year-old groups (top, middle, and bottom panel, respectively) in different noise conditions. Standard errors are given and significant differences between adjacent conditions are marked with asterisks.

Apoux and Healy (2010) found similar performance between broadband and on-frequency maskers, under conditions involving adults and consonant stimuli, but otherwise similar to those employed here. In accord with these previous results, performance in the ON condition did not differ significantly from the BB condition for any of the groups. The amount of masking was similar across these two conditions, indicating that the on-frequency noise components, those with overlapping excitation patterns with the speech, are responsible for much of the masking observed in the BB condition. With regard to the scores of children in BB versus ON, it is not clear why the additional noise in off-frequency channels in BB did not significantly reduce performance relative to performance in ON.

The critical comparison in the current experiment was between OFF and ON conditions. In the former, noise was present only in the frequency bands that did not contain speech, and so speech and noise were spectrally interleaved. In the latter, noise was present only in the frequency bands that also contained speech, and so speech and noise were always spectrally overlapping. Adults can isolate frequency regions containing clean signals and therefore display large differences between these conditions. This should be anticipated and was found by Apoux and Healy (2010) for consonant stimuli. In the current experiment, the children displayed smaller differences between OFF and ON, and the differences for youngest children were small enough to be statistically equivalent. These results support the existing psychoacoustic evidence that children integrate information over a larger number of auditory filters (Oh *et al.*, 2001). They also support the results and conclusions of Experiments 1 and 2, that children possess broader auditory attention filters and that these broader filters cause them to be more susceptible to maskers that are separated from the signal frequency.

General Discussion

The first aim of this study was to provide a measure of the spectral attention bandwidth in groups of children and adults. In Experiment 1, the specific hypothesis tested was that children differ in their ability to attend to a specific spectral region where a target is expected to occur. The second aim of this study was to measure how these spectral attention bandwidths play a role in listening for speech in noise. The specific hypothesis tested across Experiments 2 and 3 was that, because children differ in their ability to attend to specific spectral regions, a masking component in a frequency region different from that containing speech would still be detrimental. This series of experiments provides evidence that the attention band develops while the auditory system is still developing in childhood. Further, this inefficient listening strategy may account for some of the difficulty children have understanding speech in noise.

Adult listeners were presumably able to focus on a single auditory filter to optimize detection of a target in noise, but at the detriment to probe frequencies presented in a range of 1/4-to 2-ERBs away. Children, however, appeared to monitor broadly across many filters. These data reveal that the attentional filter and the frequency selectivity that it provides is not fully developed in 5- to 8-year-old children. The youngest child subjects showed similar performance across maskers and probes, as found in infants (Bargones & Werner, 1994). The older child subjects also demonstrated elevated performance for probe frequencies, with a decrease in performance at the most distant probes, indicating that there may be some narrowing of spectral attention bands with increasing maturity. The results of Experiment 1 also suggest that some of the informational masking observed in speech-recognition tasks might be in fact due to broader attentional filters or misplaced filter monitoring.

Performance on speech recognition in noise tasks relies on independence of auditory channels and the ability to attend to them (Apoux & Healy, 2009; 2010; Kidd *et al.*, 2005). The reduced selectivity (broader attentional bandwidths) in children may provide an explanation for their increased susceptibility to noise. In Experiments 2 and 3, it was found that spectrally remote maskers and interleaved noise conditions with limited spectral overlap (OFF) had a larger detrimental effect on speech recognition performance for these groups of 5-year-olds, relative to their older counterparts.

More specifically, in Experiment 2, groups of adults, 7-year-olds, and 5-year-olds were presented with sentences in SSN and high-intensity, spectrally remote maskers. The addition of remote maskers made no difference in performance for adult and 7-year-old groups, however it decreased performance for the 5-year-old group. This suggests that the 5-year-old group did not have the same ability to parse the speech signal from the competing background noise, even when it was spectrally remote.

One question that arises is why the 7-year-old group was not affected by the remote maskers, even though they also showed an inability to attend to a specific frequency region in Experiment 1, similar to the 5-year-old group. One possible explanation is that by the age of 7 years, the ability to monitor filters in a more narrow manner is emerging. Evidence for this can be seen by looking at the shape of the spectral attention bands seen in Figure 23. At the furthest probe distance (2-ERBs, 762 and 1294 Hz), a trend toward decreased detection can be seen. This may indicate a narrowing of the filters being monitored to within a 2-ERB distance from the expected target. In Experiment 2, the expected target was low-pass filtered speech with an upper cutoff of 1500 Hz whereas the closest remote masker had a lower cutoff of 2500 Hz, which was well over 4-ERBs away. This suggests that for the 7-year-old group, even a slight narrowing in filter monitoring may have proved beneficial in the remote-noise masking situations.

Experiment 3 investigated the effects of complementary, largely non-overlapping, bands of noise with these same age groups. This experiment differed in that the noise was presented in adjacent ERBs. As adults demonstrated a narrow spectral listening strategy, with detection significantly decreasing at a 1/2-ERB distance, the expected results were observed involving reduced performance in ON relative to OFF masking conditions. Adults were apparently able to disregard the irrelevant noise and focus on the available speech signal.

In Figure 23, there are observed dips in detection performance for both the 5-year-old group and 6- to 7-year-old group at the 1/4-ERB probe distances, though non-significant for both. Interestingly, the subjects in the 6- to 7-year-old group were able to disregard enough of the off-frequency noise to achieve a significant increase in performance in OFF relative to ON conditions, just as the adults were. A possible explanation is an increased ability to attend to the ERB bands where speech was present and to discard non-relevant information. Conversely, the 5-year-old group may be unable to disregard the auditory filter outputs containing noise. These results indicate a refinement in listening skills by the 7-year-old group. These combined results across experiments help to emphasize the importance of developmental listening strategies for understanding speech in noise.

As mentioned in the Speech Psychoacoustics chapter, speech-recognition tasks such as those in Experiment 2 and 3, require a higher cognitive load than more basic tone-detection tasks, such as that in Experiment 1. These experiments therefore form a continuum of cognitive processing demands. Experiment 1 used tones as signals, a cognitive load-free task which relies primarily on basic bottom-up acoustic cues for processing. Experiment 3 used words in a carrier phrase, which limits linguistic context cues for cognitive processing. However, there are phonotactic rules, coarticulation, and probability cues that the listener can employ. Lastly, Experiment 2 used sentences, which allow for maximizing top-down processing with semantic context.

In order to identify a word, it must be correctly matched to the subject's existing mental lexicon. There is some variability in the amount of cognitive load this entails based on the neighborhood density of the perceived word, or the amount of acoustically similar words to the target stimuli speech. Well-accepted models of spoken word recognition, such as the Neighborhood Activation Model (NAM; Luce & Pisoni, 1998), follow a specific process. This first involves the perception of a spoken word activating that item's lexical entry as well as its acoustically similar neighbors. Next, the perceived word is selected from the lexical neighborhood, while its neighbors are inhibited. In this model, acoustically similar words create neighborhoods that are similar to the target word through difference of one phoneme, whether that be by deletion, addition, or substitution of that single phoneme (Greenberg & Jenkins, 1964). These neighborhoods can be measured with regard to their density, how many words are included in the neighborhood.

For adults, words occurring in high-density neighborhoods are classified more accurately than words in low-density neighborhoods (Luce & Pisoni, 1998), but there have been few measures on children's speech corpora. From the existing studies, it has been shown that children can achieve both benefits and detriments based on lexical inventory. Storkel and Hoover (2010) computed neighborhood density for nouns and compared them for children and adults. In this study, neighborhood density was lower, due to the overall decreased lexicon size of children compared to adults (Charles-Luce & Luce, 1990, 1995; Coady & Aslin, 2003).

As one example of further analyses that could be performed to probe the influence of higher-level cognitive functions on the mechanisms observed here, an analysis of neighborhood density was performed. Experiment 3 measured speech recognition with words, which contained fewer semantic cues than the sentences employed in Experiment 2. This likely leads to a decreased cognitive load and more reliance on bottom-up cognitive factors compared with sentence stimuli. The stimuli used in Experiment 3 consisted of the PBK words. An analysis of these 150 words based on the Storkel and Hoover's (2010) child corpus yielded neighborhood densities ranging from 0 to 34 (mean = 11.99). These values indicate that the overall neighborhood density for the corpus was quite low. Stimulus words in each condition (BB, ON, OFF, QUIET) were sub-divided into 2 groups, those with a lower neighborhood density (12 and below) and those with a higher density (above 12). Performance for these higher- and lower-density words were then compared within each condition. For the younger-child group a paired t-test yielded no significant difference between lower-density and higher-density stimuli on three of the conditions [BB $t(8) = 1.13$, $p = .292$; OFF $t(8) = .78$, $p = .461$; QUIET $t(8) = -.24$, $p = .817$], but a significant difference for ON [$t(8) = 3.05$, $p = .016$]. In this ON condition, performance was better for stimuli occurring in lower density neighborhoods. Older-child performance for higher- and lower-density words was compared in a similar manner. A paired t-test yielded no significant difference between performance with lower-density and higher-density stimuli on three conditions [BB $t(8) = .32$, $p = .756$; ON $t(8) = 1.24$, $p = .249$; QUIET $t(8) = 1.51$, $p = .170$] but a significant difference for OFF [$t(8) = 2.81$, $p = .023$]. Similar to the younger children, performance was better for stimuli occurring in lower-density neighborhoods.

In contrast to what was found here, it might be expected that high neighborhood density words would yield higher recognition accuracy. However, the overall low neighborhood density and low word frequency of the current stimuli may account for these seemingly contradictory results. Word frequency is another factor that interacts with processing and neighborhood activation, with high frequency words associated with processing advantages (Luce & Pisoni, 1998). Similar to neighborhood density, word frequency for these word stimuli were low, ranging from 1.00 to 5.28 (mean = 3.04). The non-significant differences found for three out of the four conditions for each age group are then not surprising considering that overall the neighborhood density and word frequency for all words was low.

The implications for understanding listening strategies employed by young children are important. Children are often asked to learn and perform in noisy environments. On average, there are 22 children in a typical kindergarten classroom (NCES, 1993). The age range of 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 could potentially be negatively impacted by noisy environments. An understanding that preschool and young school-age children are unable to use adult-like focused and efficient listening strategies supports the need to better manage the classroom. Improved room acoustics and decreased class sizes could facilitate a better learning environment that takes into consideration these apparent developmental limitations.

It should be noted that the current implications are for children with typical hearing and language skills. Any impairment in these areas could interact with decreased listening efficiency to produce an even greater negative impact of noise on speech-recognition. The current results therefore help to emphasize how crucial it is to take noise into consideration for the learning environment of typical and special populations of children.

The results of these studies contribute to our understanding of auditory spectral attention as a listening strategy and how it may be developing in the early school-age years. As with broader auditory filters seen in many sensorineural hearing-impaired listeners, children's broader spectral attending leads to a less effective rejection of background noise. This then allows the background to interfere with detection and discrimination of sounds, including speech. The current data confirm that adults are able to attend to specific filters to maximize detection and disregard noise. However, young children are not obtaining similar performance, suggesting that they are monitoring broadly over many filters, and further suggesting that they are employing a less effective listening strategy. This comparison between adults and young children in spectral attention bandwidth along with speech-recognition tasks in remote and interleaved noise, provides evidence to better explain why children typically require more favorable SNRs than adults to achieve similar levels of speech performance in noise.

References

Abdala, C. & Folsom, R. C. (1995). The development of frequency resolution in humans as revealed by the auditory brain-stem response recorded with notched-noise masking. *Journal of the Acoustical Society of America*, 98, 921-930.

Allen, P. & Wightman, F. (1994). Psychometric functions for children’s detection of tones in noise. *Journal of Speech and Hearing Research*, 37, 205-215.

Allen, P., Wightman, F., Kistler, D., & Dolan, T. (1989). Frequency resolution in children. *Journal of Speech and Hearing Research*, 32, 317-322.

American National Standard Inst. (2004). ANSI S3.21 (R2009). American National Standard Methods for Manual Pure-Tone Threshold Audiometry (American National Standards Inst., New York).

American National Standard Inst. (2010). ANSI S3.6-2010. American National Standard Specification for Audiometers (American National Standards Inst., New York).

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.

Apoux, F. & Healy, E. W. (2010). Relative contribution of off-and on-frequency spectral components of background noise to the masking of unprocessed and vocoded speech. *Journal of the Acoustical Society of America*, 128, 2075-2084.

Arbogast, T., Mason, C., & Kidd, G. (2002). The effect of spatial separation on informational and energetic masking of speech. *Journal of the Acoustical Society of America*, 112, 2086-2098.

Bacon, S. P., & Liu, L. (2000). Effects of ipsilateral and contralateral precursors on overshoot. *Journal of the Acoustical Society of America*, 108, 1811-1818.

Bargones, J. Y. & Burns, E. M. (1988). Suppression tuning curves for spontaneous otoacoustic emissions in infants and adults. *Journal of the Acoustical Society of America*, 83, 1809-1816.

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.

Bilger, R.C. & Hirsch, I.J. (1956). Masking of tones by bands of noise. *Journal of the Acoustical Society of America*, 28, 623-630.

Bonino, A. Y., Leibold, L. J., & Buss, E. (2013). Effect of signal-temporal uncertainty in children and adults: Tone detection in noise or a random-frequency masker. *Journal of the Acoustical Society of America*, 134, 4446-4457.

Boothroyd, A. (1968). Developments in speech audiometry. *British Journal of Audiology*, 2, 3-10.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America*, 109, 1101-1109.

Buss, E., Hall, J. W., & Grose, J. H. (2009). Psychometric functions for pure tone intensity discrimination: Slope differences in school-aged children and adults. *Journal of the Acoustical Society of America*, 125, 1050-1058.

Buss, E., Hall, J. W., 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.

Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, 17, 205-215.

Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language*, 22, 727-735.

Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language*, 30, 441-469.

Cooke, M. (2005). “Making sense of everyday speech: A glimpsing account,” in Speech Separation by Humans and Machines, edited by P. Divenyi (Kluwer Academic, Dordrecht), pp. 305-314.

Cooke, M. (2006). A glimpsing model of speech perception in noise. *Journal of the Acoustical Society of America*, 119, 1562-1573.

Cooke, M., Lecumberri, M. G., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *Journal of the Acoustical Society of America*, 123, 414-427.

Dai, H., Scharf, B., & Buus, S. (1991). Effective attenuation of signals in noise under focused attention. *Journal of the Acoustical Society of America*, 89, 2837-2842.

Delgutte, B. (1990). Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers. *Journal of the Acoustical Society of America*, 87, 791-809.

Diamond, A. (1991). “Neuropsychological insights into the meaning of object concept development,” in The Epigenesis of Mind: Essays on Biology and Knowledge, edited by S. Carey & R. Gelman, pp. 67-110. Hillsdale, NJ: Erlbaum.

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd Jr, G. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *Journal of the Acoustical Society of America*, 114, 368-379.

Elliott, L. L. (1979). Performance of children aged 9 to 17 on a test of speech intelligibility in noise using sentence material with controlled word predictability. *Journal of the Acoustical Society of America*, 66, 651-653.

Elliott, L. L., Conners, S., Kille, E., Levin, S., Ball, K., & Katz, D. (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *Journal of the Acoustical Society of America*, 66, 12-21.

Elliott, L. L. & Katz, D. R. (1980). Children's pure tone detection. *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.

Fletcher, H. (1940). Auditory patterns. *Review of Modern Physics*, 12, 47-65.

Gerardi-Caulton, G. (2000). Sensitivity to spatial conflict and the development of self-regulation in children 24–36 months of age. *Developmental Science*, 3, 397-404.

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, J. H., & Jenkins, J. J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, 20, 157-177.

Greenberg, G. Z. & Larkin, W. D. (1968). Frequency-response characteristics of auditory observers detecting signal of a single frequency in noise: The probe-signal method. *Journal of the Acoustical Society of America*, 44, 1513-1523.

Hall, J. W., Buss, E., & Grose, J. H. (2005). Informational masking release in children and adults. *Journal of the Acoustical Society of America*, 118, 1605-1613.

Hall, J. W. & 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., 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. *Journal of the Acoustical Society of America*, 83, 2277-2289.

Haskins, H. (1949**)**. A phonetically balanced test of speech discrimination for children, Unpublished master's thesis, Northwestern University, Evanston, IL.

Healy, E. W., & Bacon, S. P. (2006). Measuring the critical band for speech. *Journal of the Acoustical Society of America*, 119, 1083-1091.

Hnath-Chisolm, T. E., Laipply, E., & Boothroyd, A. (1998). Age-related changes on a children's test of sensory-level speech perception. *Journal of Speech Language and Hearing Research*, 41, 94-106.

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. *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.

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science*, 6, 498-504.

Keefe, D. H. & Abdala, C. (2007). Theory of forward and reverse middle-ear transmission applied to otoacoustic emissions in infant and adult ears. *Journal of the Acoustical Society of America*, 121, 978-993.

Kidd, G., Arbogast, T. L., Mason, C. R., & Walsh, M. (2002). Informational masking in listeners with sensorineural hearing loss. *Journal of the Association for Research in Otolaryngology*, 3, 107-119.

Kidd Jr, G., Mason, C. R., & Gallun, F. J. (2005). Combining energetic and informational masking for speech identification. *Journal of the Acoustical Society of America*, 118, 982-992.

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. & Bonino, A. Y. (2009). Release from informational masking in children: Effect of multiple signal bursts. *Journal of the Acoustical Society of America*, 125, 2200-2208.

Leibold, L. J. & Neff, D. L. (2007). Effects of masker-spectral variability and masker fringes in children and adults. *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. *Journal of the Acoustical Society of America*, 119, 3960-3970.

Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America*, 117, 3091-3099.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing*, 19, 1-36.

Lutfi, R. A. (1989). Informational processing of complex sound. I: Intensity discrimination. *Journal of the Acoustical Society of America*, 86, 934-944.

Lutfi, R. A., Kistler, D. J., Callahan, M. R., & Wightman, F. L. (2003). Psychometric functions for informational masking. *Journal of the Acoustical Society of America*, 114, 3273-3282.

Miller, P. H. (1990). “The development of strategies of selective attention,” in Children’s Strategies: Contemporary Views of Cognitive Development, edited by D. F. Bjorklund, pp 157-184. Hillsdale, NJ: Erlbaum.

Miller, P. H., & Zalenski, R. (1982). Preschoolers' knowledge about attention. *Developmental Psychology*, 18, 871-875.

Moore, B. C. J. (2007). Cochlear Hearing Loss: Physiological, Psychological and Technical Issues. 2nd Edition. England: John Wiley & Sons Ltd.

Moore, B. C., 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.

Moore, J. K. (2002). Maturation of human auditory cortex: Implications for speech perception. *Annals of Otology Rhinology and Laryngology Supplement*, 189, 7-10.

Moore, J. K. & Guan, Y. L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology*, 2, 297-311.

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. http://nces.ed.gov/surveys/frss/publications/93410/index.asp?sectionid=5 Accessed on April, 4, 2015.

Neff, D. L. (1995). Signal properties that reduce masking by simultaneous, random-frequency maskers. *Journal of the Acoustical Society of America*, 98, 1909-1920.

Neff, D. L. & Callaghan, B. P. (1988). Effective properties of multicomponent simultaneous maskers under conditions of uncertainty. *Journal of the Acoustical Society of America*, 83, 1833-1838.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*, 95, 1085-1099.

Oh, E. L. & Lutfi, R. A. (1998). Nonmonotonicity of informational masking. *Journal of the Acoustical Society of America*, 104, 3489-3499.

Oh, E. L. & Lutfi, R. A. (2000). Effect of masker harmonicity on informational masking. *Journal of the Acoustical Society of America*, 108, 706-709.

Oh, E. L., Wightman, F., & Lutfi, R. A. (2001). Children's detection of pure-tone signals with random multitone maskers. *Journal of the Acoustical Society of America*, 109, 2888-2895.

Olsho, L. W. (1985). Infant auditory perception: Tonal masking. *Infant Behavior & Development*, 8, 371-384.

Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *Journal of the Acoustical Society of America*, 84,1316-1324.

Oxenham, A.J. & Plack, C.J. (1998). Suppression and the upward spread of masking. *Journal of the Acoustical Society of America*, 104, 3500-3510.

Papso, C. F. & Blood, I. M. (1989). Word recognition skills of children and adults in background noise. *Ear and Hearing*, 10, 235-236.

Pumplin, J. (1985). Low-noise noise. *Journal of the Acoustical Society of America*, 78, 100-104.

Reed, M. A., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly*, 30, 131-147.

Rosen, S. & Baker, R. J. (1994). Characterising auditory filter nonlinearity. *Hearing Research*, 73, 231-243.

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. *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.

Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1986). Auditory sensitivity in preschool children. *Journal of the Acoustical Society of America*, 79, 447-452.

Schneider, B. A., Trehub, S. E., Morrongiello, B. A., & Thorpe, L. A. (1989). Developmental changes in masked thresholds. *Journal of the Acoustical Society of America*, 86, 1733-1742.

Soderquist, D. R. (1993). Auditory filter widths in children and adults. *Journal of Experimental Child Psychology*, 56, 371-384.

Spetner, N. B., & Olsho, L. W. (1990). Auditory frequency resolution in human infancy. *Child Development*, 61, 632-652.

Storkel, H. L., & Hoover, J. R. (2010). An on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English. *Behavior Research Methods*, 42, 497–506.

Studebaker, G. A. (1985). A rationalized arcsine transform. *Journal of Speech, Language, and Hearing Research*, 28, 455-462.

Trehub, S. E., Schneider, B. A., & Endman, M. (1980). Developmental changes in infants' sensitivity to octave-band noises. *Journal of Experimental Child Psychology*, 29, 282-293.

Trehub, S. E., Schneider, B. A., Morrongiello, B. A., & Thorpe, L. A. (1988). Auditory sensitivity in school-age children. *Journal of Experimental Child Psychology*, 46, 273-285.

Watson, C. S. & Nichols, T. L. (1976). Detectability of auditory signals presented without defined observation intervals. *Journal of the Acoustical Society of America*, *59*, 655-668.

Weissberg, R., Ruff, H. A., & Lawson, K. R. (1990). The usefulness of reaction time tasks in studying attention and organization of behavior in young children. *Journal of Developmental & Behavioral Pediatrics*, 11, 59-64.

Werner, L. A. (2007). Issues in human auditory development. *Journal of Communication Disorders*, 40, 275-283.

Werner, L. A. & Bargones, J. Y. (1991). Sources of auditory masking in infants: Distraction effects. *Perception and Psychophysics*, 50, 405-412.

Werner, L. A. & Gillenwater, J.M. (1990). Pure-tone sensitivity of 2-to 5-week-old infants. *Infant Behavior and Development*, 13, 355-375.

Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child Development*, 611-624.

Wightman, F. L., Callahan, M. R., Lutfi, R. A., Kistler, D. J., & Oh, E. (2003). Children's detection of pure-tone signals: Informational masking with contralateral maskers. *Journal of the Acoustical Society of America*, 113, 3297-3305.

Wightman, F. L. & Kistler, D. J. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *Journal of the Acoustical Society of America*, 118, 3164-3176.

Zwicker, E. (1965). Temporal effects in simultaneous masking by white-noise bursts. *Journal of the Acoustical Society of America*, 37, 653-663.

Zwicker, E. (1970). Masking and psychological excitation as consequences of the ear's frequency analysis. In: Plomp & Smoorenburg (Eds). Frequency Analysis and Periodicity Detection in Hearing. Leyden: Sijthoff.

Appendix A

Raw Individual Data for Experiment 1

Table 3. Experiment 1 Individual Adult Data: Percent-Correct Detection.

	Target/Probe Frequency (Hz)								
Subject	762	874	935	967	1000	1033	1067	1140	1294
A 1	58.30	50.00	66.70	50.00	75.30	66.70	41.70	50.00	33.30
A 2	75.00	58.30	75.00	66.70	84.00	50.00	58.30	41.70	50.00
A 3	100	75.00	75.00	75.00	93.10	66.70	66.70	41.70	58.30
A 4	58.30	33.30	33.30	83.30	90.60	75.00	58.30	41.60	50.00
A 5	66.70	66.70	58.30	83.30	82.30	33.30	41.70	58.30	33.30
A 6	83.30	100	100	91.70	84.70	83.30	83.30	58.30	58.30
A 7	41.60	58.30	66.70	66.70	83.30	83.30	58.30	58.30	50.00
A 8	75.00	75.00	83.30	100	93.10	75.00	66.70	50.00	0.25
A 9	58.30	50.00	50.00	91.70	84.70	75.00	58.30	41.70	41.70
A 10	50.00	41.70	58.30	50.00	75.70	58.30	75.00	33.30	50.00

Table 4. Experiment 1 Individual Child Data: Percent-Correct Detection.

			Probe Location (Hz)								
Subject	Age (yrs; mo)	# of Rounds	762	874	935	967	1000	1033	1067	1140	1294
C 1	5;0	7	85.70	71.40	71.40	57.10	64.30	42.90	42.90	71.40	71.40
C 2	5;3	10	50.00	40.00	50.00	50.00	59.10	50.00	70.00	70.00	60.00
C 3	5;4	12	83.30	100	91.70	91.70	83.30	83.30	91.70	91.70	66.70
C 4	5;5	10	60.00	80.00	70.00	60.00	67.90	40.00	80.00	60.00	90.00
C 5	5;9	10	90.00	80.00	80.00	70.00	87.90	80.00	90.00	100	70.00
C 6	5;10	10	100	100	80.00	70.00	85.40	80.00	80.00	70.00	90.00
C 7	5;11	10	80.00	80.00	60.00	70.00	67.90	80.00	70.00	80.00	90.00
C 8	6;8	12	66.70	50.00	75.00	75.00	86.80	83.30	75.00	91.70	66.70
C 9	7;6	12	50.00	66.70	83.30	66.70	75.20	66.70	83.30	58.30	41.70
C 10	7;6	10+ partial	80.00	100	60.00	63.60	82.70	90.90	72.70	81.80	70.00
C 11	7;7	10	30.00	10.00	0.00	60.00	84.60	50.00	20.00	20.00	30.00
C 12	7;8	10	100	90.00	100	100	89.20	90.00	100	90.00	70.00
C 13	7;8	10	100	100	100	90.00	83.30	90.00	90.00	100	100
C 14	8;0	10	20.00	80.00	70.00	50.00	73.30	40.00	40.00	90.00	50.00

Appendix B

Raw Individual Data for Experiment 2

Table 5. Experiment 2 Raw Individual Data for Adult Subjects: Percent-Correct Sentence Recognition in each Condition.

				1/3-octave-width (lower cut-off)			
Subject	Quiet	CONTROL (LP Speech + SSN)	2500-5000 Hz	3000 Hz	4000 Hz	5000 Hz	6000 Hz
Subj 1	100.00	54.84	64.52	80.65	62.90	50.00	66.13
Subj 2	96.77	75.81	67.74	64.52	72.58	66.13	67.74
Subj 3	90.32	69.35	75.81	75.81	72.58	56.45	61.29
Subj 4	100.00	75.81	69.35	69.35	64.52	66.13	56.45
Subj 5	100.00	83.87	75.81	70.97	62.90	80.65	74.19
Subj 6	90.32	77.42	62.90	67.74	79.03	75.81	64.52
Subj 7	100.00	79.03	85.48	70.97	75.81	79.03	74.19
Subj 8	96.77	51.61	53.23	72.58	66.13	72.58	61.29
Subj 9	90.32	70.97	88.71	69.35	70.97	56.45	59.68
Subj 10	96.77	62.90	64.52	66.13	83.87	74.19	56.45

Table 6. Experiment 2 Raw Individual Data for Child Subjects: Percent-Correct Sentence Recognition in each Condition.

				1/3-octave width (lower cut-off)				
Subject	Age (yr;mo)	Quiet	CONTROL (LP Speech + SSN)	2500-5000 Hz	3000 Hz	4000 Hz	5000 Hz	6000 Hz
Subj 1	7;4	87.10	74.19	83.87	70.97	77.42	66.13	66.13
Subj 2	5;0	70.97	72.58	46.77	40.32	61.29	61.29	58.06
Subj 3	7;7	93.55	72.58	69.35	61.29	82.26	93.55	66.13
Subj 4	8;0	100.00	61.29	77.42	66.13	67.74	72.58	79.03
Subj 5	5;0	77.42	58.06	58.06	51.61	54.84	35.48	46.77
Subj 6	7;8	93.55	66.13	59.68	64.52	66.13	64.52	54.84
Subj 7	5;5	90.32	53.23	62.90	59.68	54.84	56.45	72.58
Subj 8	7;5	90.32	74.19	70.97	67.74	62.90	79.03	62.90
Subj 9	5;5	74.19	56.45	43.55	35.48	48.39	37.10	40.32
Subj 10	5;6	83.87	69.35	43.55	48.39	56.45	67.74	40.32
Subj 11	7;2	83.87	70.97	67.74	48.39	70.97	75.81	61.29
Subj 12	5;8	83.87	64.52	62.90	67.74	62.90	56.45	66.13
Subj 13	5;8	87.10	62.90	53.23	50.00	58.06	46.77	50.00
Subj 14	5;9	93.55	64.52	46.77	53.23	46.77	62.90	72.58
Subj 15	7;1	93.55	61.29	75.81	74.19	70.97	70.97	64.52
Subj 16	7;10	93.55	69.35	77.42	82.26	77.42	66.13	69.35
Subj 17	7;5	90.32	58.06	67.74	69.35	69.35	58.06	59.68
Subj 18	5;10	83.87	61.29	58.06	66.13	62.90	61.29	66.13

Appendix C

Raw Individual Data for Experiment 3

Table 7. Experiment 3 Raw Individual Data for Adults: Percent-Correct Word Recognition in each Condition

	Masker Condition			
Subject	BB	ON	OFF	QUIET
Subj 1	30	42.5	45	100
Subj 2	32.5	45	55	100
Subj 3	27.5	37.5	57.5	100
Subj 4	22.5	37.5	52.5	95
Subj 5	12.5	32.5	57.5	85
Subj 6	30	32.5	42.5	100
Subj 7	25	32.5	70	95
Subj 8	32.5	25	47.5	95
Subj 9	30	37.5	67.5	100
Subj 10	15	22.5	45	100

Table 8. Experiment 3 Raw Individual Data for Children: Percent-Correct Word Recognition in each Condition

		Masker Condition			
Subject	Age (yr;mo)	BB	ON	OFF	QUIET
Subj 1	5;0	30.00	47.50	35.00	82.50
Subj 2	6;7	45.36	62.50	62.78	90.48
Subj 3	7;11	30.00	35.00	62.50	90.00
Subj 4	7;11	47.50	45.00	57.50	90.00
Subj 5	5;5	32.50	47.50	45.00	85.00
Subj 6	5;8	35.00	35.00	57.50	95.00
Subj 7	7;5	37.50	40.00	67.50	95.00
Subj 8	5;7	37.50	32.50	37.50	90.00
Subj 9	7;7	37.50	37.50	77.50	95.00
Subj 10	7;5	25.00	20.00	55.00	90.00
Subj 11	5;3	37.50	37.50	62.50	90.00
Subj 12	7;11	40.00	62.50	65.00	90.00
Subj 13	5;3	37.50	37.50	52.50	95.00
Subj 14	5;11	20.00	37.50	70.00	90.00
Subj 15	7;9	42.50	57.50	52.50	95.00
Subj 16	5;9	25.00	55.00	47.50	95.00
Subj 17	7;4	22.50	42.50	50.00	90.00
Subj 18	5;6	45.00	50.00	65.00	100.00