Discrimination of static and dynamic spectral patterns by children and young adults in relationship to speech perception in noise

Hanin Rayes, Stanley Sheft, Valeriy Shafiro

Department of Communication Disorders and Sciences, Rush University Medical Center, Chicago, IL, USA

Abstract

Past work has shown relationship between the ability to discriminate spectral patterns and measures of speech intelligibility. The purpose of this study was to investigate the ability of both children and young adults to discriminate static and dynamic spectral patterns, comparing performance between the two groups and evaluating within-group results in terms of relationship to speech-in-noise perception. Data were collected from normal-hearing children (age range: 5.4-12.8 years) and young adults (mean age: 22.8 years) on two spectral discrimination tasks and speech-in-noise perception. The first discrimination task, involving static spectral profiles, measured the ability to detect a change in the phase of a low-density sinusoidal spectral ripple of wideband noise. Using dynamic spectral patterns, the second task determined the signal-to-noise ratio needed to discriminate the temporal pattern of frequency fluctuation imposed by stochastic low-rate frequency modulation (FM). Children performed significantly poorer than young adults on both discrimination tasks. For children, a significant correlation between speech-in-noise perception and spectral-pattern discrimination was obtained only with the dynamic patterns of the FM condition, with partial correlation suggesting that factors related to the children's age mediated the relationship.

Correspondence: Hanin Rayes, Center of Hearing and Balance, Jeddah, Saudi Arabia.
Cell: +966.50.555.4720. E-mail: haninrayes@gmail.com

Key words: spectral patterns, frequency modulation, children, speech-in-noise perception.

Acknowledgments: this work was supported by Grant R15 DC011916 from the National Institute on Deafness and Other Communication Disorders (NIDCD) to Stanley Sheft. We thank George Spanos and Monika Tiido for data collection from the young adults. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIDCD or the National Institutes of Health.

Contributions: the authors contributed equally.

Conflicts of interests: the authors declare no potential conflicts of interest.

Funding: the work was supported by NIH grant no. R15 DC011916.

Conference presentation: part of this paper was presented at the annual meeting of American Academy of Audiology, 2013, Apr 4-6, Anaheim, CA, USA.

Received for publication: 13 February 2014.
Revision received: 27 June 2014.
Accepted for publication: 29 July 2014.

Audiology Research 2014;4:101
doi:10.4081/audiores.2014.101

Introduction

Speech perception is dependent on spectral and temporal processing, with this processing often assessed through psychoacoustic measures of detection and discrimination performance for auditory stimuli other than speech. Understanding the relationship between speech and psychoacoustic abilities in children is important not only for demonstrating developmental aspects of speech processing, but for also suggesting approaches for clinical evaluation and rehabilitation strategies. While many studies have investigated developmental aspects of psychoacoustic abilities (for a recent review, see Buss *et al.*[1]), far fewer have evaluated the relationship of these abilities to speech perception in children.

Due to manner of production, speech can be characterized by spectral patterns that vary over time. Processing of theses spectral patterns involves the ability to discriminate static spectral patterns as exemplified by vowel formant structure, and also sensitivity to dynamic variations in the spectral patterns with formant transitions and modulation of voicing fundamental as examples of these changes in speech. Allen and Wightman[2] evaluated static spectral discrimination in children 3.5 to 9.5 years old, comparing their performance to young adults. Conditions measured the depth of spectral modulation needed to detect a sinusoidal ripple of the amplitude spectrum of multi-component tonal stimuli. Low ripple densities were used corresponding to spectral amplitude peaks every three or four critical bandwidths. In both quiet and in a noise masker, performance improved with age across children with mean thresholds similar to young adults by age 9. A similar pattern of results was obtained for discrimination of spectral patterns associated with isolated vowel and consonant sounds. When controlling for age, modest but significant correlations were obtained between performance levels for discrimination of speech sounds and ripple detection thresholds in quiet.

Dawes and Bishop[3] evaluated spectral-ripple detection in children 6-10 years old and adults, using iterated rippled noise (IRN). A harmonic spectral ripple can be generated through a process in which a wideband stimulus is delayed then added back to itself. For IRN, the delay-add process is repeated to progressively steepen the spectral slopes of the stimulus.[4] Using eight iterations, the ripple delay in stimulus generation in the Dawes and Bishop study was 10 ms, corresponding to a

100-Hz spacing of spectral amplitude peaks, much narrower than used by Allen and Wightman.[2] The linear peak spacing of the IRN used by Dawes and Bishop leads to percept of a complex pitch, often attributed to temporal rather than spectral processing. Their results showed adult-like performance from children across the age range tested. Dawes and Bishop also assessed processing of dynamic stimuli, measuring thresholds for detecting frequency modulation (FM) at rates of 2, 40, and 240 Hz. In these conditions, there was an effect of age with most children not achieving adult levels of performance until age 9 for 2-Hz FM, while performing at adult levels by age 7 for 40- and 240-Hz FM. With age-standardized scoring of psychoacoustic performance, relatively low but significant correlations were found between psychoacoustic thresholds and the intelligibility of filtered and masked speech as measured by SCAN-C.[5] Results from other studies on developmental aspects in FM detection are in general agreement with the findings of Dawes and Bishop. Across studies, within-group variability of younger children is often high with many to most children exhibiting adult performance levels by the age of roughly 8 to 10 years old.[6-8]

Analysis of the FM spectrum of speech shows a lowpass characteristic, highlighting the importance of low-rate speech FM.[9] Past work with adults has shown involvement of low-rate FM in speech processing through enhancement of speech coherence,[10] aiding segmentation at the word and syllable boundaries of the speech stream,[11] and effect on speech intelligibility in the presence of competing interference.[12] However for children, significant correlation between low-rate FM detection thresholds and speech-in-noise processing has not been found.[13,14]

These past studies of processing by children of spectral patterns and variation measured detection thresholds. The depth of many of the spectral modulations of speech are far above detection threshold,[15] suggesting variation in discrimination rather than detection ability as a determinant of relationship to speech processing. Sheft *et al.*[16] evaluated the ability to discriminate both static and dynamic spectral patterns in relationship to speech perception. Intended for clinical use, the procedures measured thresholds for discriminating a change in the phase of a low-rate spectral ripple of wideband noise, and the signal-to-noise ratio (SNR) for discriminating low-rate stochastic patterns of FM of a tonal carrier. For adult listeners, results showed an effect of aging and relationship to the perception of distorted speech and speech in noise.

We are unaware of any past studies of either static or dynamic spectral discrimination in children, especially involving stimuli chosen for relevance to speech. Using the discrimination conditions developed by Sheft *et al.*,[16] there were two main goals in the present study. The first was to evaluate how well children process, compared to young adults, complex auditory signals in tasks that have been shown to bear some relationship to speech-in-noise perception. The second goal was to determine whether children show the same pattern of relationship between psychoacoustic and speech results as do young adults. The motivation of this study was to better understand the overall developmental aspects of speech processing, and to suggest approaches to improve speech testing for children in clinical settings through the utilization of non-verbal stimuli that past work suggests may predict children's speech perception in noise.

Along with developmental factors, bilingualism affects the speech-in-noise abilities of children with bilingual children exhibiting greater deficits due to masking.[17,18] On the other hand, based on the evoked brainstem response, Krizman *et al.*[19] reported enhanced encoding of fundamental frequency in bilingual adolescents with this enhancement related to superior frequency discrimination and speech-in-noise performance by adults.[20] We are unaware of any studies that have evaluated the psychoacoustic abilities of bilingual children. Therefore, a secondary goal of the present study was to offer initial investigation of psychoacoustic performance of bilingual children. By evaluating the relationship between static and dynamic spectral discrimination and speech-in-noise ability, the intent was to gain a better understanding of the speech difficulties associated with bilingual children.

Overall, this study presents a novel approach, using a recently developed clinically feasible psychophysical protocol to predict children's speech perception in noise without the use of verbal stimuli. Such a tool could be helpful to assess children with special needs, children with limited verbal language, and children who are non-native speakers or have an auditory processing disorder.

Materials and Methods

Participants

Two groups of listeners participated in the study. The first consisted of 20 monolingual young adult women (age range: 19-27 years; mean: 22.8 years) who had normal audiometric thresholds (≤15 dB HL re: ANSI[21]) for the octave frequencies between 0.25-8.0 kHz. The second group consisted of 20 children (9 girls and 11 boys, age range: 5.4-12.8 years, mean 9.2 years).

All children passed a hearing screening at 20 dB HL for the octave frequencies between 0.5-4.0 kHz. Eight of the 20 children were monolingual. The remaining 12 children were bilingual with a second language of either Arabic (10 participants) or Russian (2 participants). English was the language of all monolingual children and the primary language of all bilingual children. The primary language was identified as the language that is spoken at school and home, and the language that the child uses to communicate his or her needs on a daily basis. Nine of the bilingual children were simultaneous bilinguals, meaning they were exposed to both languages from birth. The remaining three bilingual children began learning English at the age of 2-3 years old. According to the questionnaire, which was filled by the child's parents, all bilingual children were reported to be more fluent in English than their other language, expect for two children who were equally fluent in both languages. Children were recruited from a private elementary school after approval of the school management. Informed consent was signed by a parent of each child, with monetary compensation given for study participation. Young adults were all speech and hearing students at Rush University Medical Center, recruited through class notice. The young adults received extra credit in their class for participation in the study. Experimental protocol was approved by the Institutional Review Board of Rush University Medical Center and was conducted in accordance with the Helsinki declaration.

Psychoacoustic stimuli and procedure

The ability to discriminate spectral patterns was evaluated in two conditions. In the first, the patterns were sinusoidal ripples of the long-term stimulus spectrum. Consequently, patterns were constant or static over the stimulus duration. The second condition utilized FM to generate dynamic spectral patterns whose spectral content varied over time.

Spectral-ripple condition

Discrimination of static spectral patterns was assessed using wideband stimuli (0.2-8.0 kHz) whose amplitude spectra were sinusoidally rippled in terms of the logarithms of both frequency and amplitude. Ripple density was 3.0 cycles per octave (cpo) with a peak-to-trough difference of 30 dB. In the cued two-interval forced-choice (2IFC) task, the phase of the sinusoidal spectral ripple of the standard stimulus was randomized each trial with the task to detect in which of the two observation intervals the starting phase of the ripple was changed (Figure 1). Thresholds were the just-detectable change in phase. Using random component phases selected from a uniform distribution, stimuli were generated with ¼-Hz resolution of an inverse fast Fourier transform.

With a 600-ms inter-stimulus interval (ISI), the 500-ms rippled stimuli were shaped with a 50-ms rise/fall time and passed through a speech-shape filter. Based on the combined male and female speech spectrum reported in Table 2 of Byrne *et al.*,[22] a digital finite impulse-response filter emphasizing the mid frequencies (roughly 200-500 Hz) was used for speech-shape filtering. The filter had a steep roll-off in the low frequencies of over 20-dB per octave and a gradual roll-off of roughly 3- to 6-dB per octave in the high frequencies. Since filter input was spectrally rippled stimuli with a large constant peak-to-trough difference of 30 dB, filtered stimuli deviated from the average long-term speech spectrum. Stimuli were presented to listeners at 70 dB SPL.

Frequency modulation condition

Dynamic spectral-pattern discrimination was evaluated in terms of the ability to discriminate 1-kHz pure tones frequency modulated by different samples of 5-Hz lowpass noise. A consequence of the modulation is that the instantaneous frequency of the stimulus follows the amplitude pattern of the noise modulator. The bandwidth of the noise modulator determines the average rate of FM. For 5-Hz lowpass noise modulators, the average rate is roughly 4 Hz. Modulator peak amplitude determines ΔF, the maximum frequency excursion of the FM stimulus. In the present work, ΔF was fixed at 400 Hz for all stimuli to approximate formant characteristics of speech. With ΔF fixed and a common sampling distribution of noise modulators, discrimination can rely on only the temporal pattern of frequency deviation (Figure 2), rather than change in average or peak stimulus statistics. The 500-ms modulated stimuli were temporally centered in 1000-ms maskers with thresholds measured in terms of the SNR needed to just discriminate the pattern of frequency fluctuation. To have modulation characteristics similar to speech, maskers were speech-shaped wideband noise which was processed to include slow random variations in local fine-structure periodicities and loudness. Speech-shape filtering was as described above for the spectral-ripple condition. The fine-structure periodicities were introduced through an iterative delay-add process in which delay time was dynamically varied between 0.75-3.0 ms by the time structure of 15-Hz lowpass noise. The loudness variations were achieved by comodulating the maskers with 2.5-Hz lowpass noise. Signals and maskers were separately shaped with a 50-ms rise/fall time with a 750-ms ISI separating the three stimulus presentations of each cued 2IFC trial. In the task, masker level was fixed at 80 dB SPL with the level of the FM tones varied to estimate the threshold SNR. Masker level was selected to match level used in our past work to allow for comparison across studies.

Procedure

In the cued 2IFC procedure, the cue was the second stimulus presentation with listeners indicating their selection of which observation interval differed from the cue. The test procedure used a modified descending method of limits. Thresholds were derived from performance on a 42-trial block, cycling seven times from high to low through six levels of the independent variable (i.v.), either delta ripple phase or FM SNR. For ripple phase, the starting delta was 2.36 radians with each subsequent delta smaller by a factor of 0.56. In the FM condition, the six values of SNR ranged from -18 to 12 dB with 6 dB between adjacent levels. The 2IFC psychometric function ranges between 50 and 100% correct. Assuming a stable underlying function with function slope symmetric about threshold at 75% correct, threshold can be arithmetically derived if levels of the i.v. are evenly spaced and at least minimally bracket the threshold point. Specifically, threshold is:

$$high + step/2 - step*(2*p - num) \quad (1)$$

where *high* is highest level of the i.v., *step* is the decrement between successive levels of the i.v., *num* is the number of levels used, and *p* is the sum of the correct-response probabilities across all levels. A final assumption used in threshold derivation is that no response probability can be below chance performance. In the ripple condition, logarithmic values of the variables *high* and *step* were used in threshold estimation, while values in the FM condition were from SNRs in dB.

A single threshold estimate was derived for each listener in each condition. For both the spectral-ripple and FM SNR conditions, a single 42-trial stimulus set was used so that all participants from both subject groups (*i.e.*, children and young adults) were tested with the same stimuli. The stimulus sets, developed for clinical use, were generated through random sampling of the appropriate noise distributions as described above. Thus, though values of the independent variables would be the same in additional stimulus sets, the actual stimuli would differ. Test procedure differed between children and young adults. For children, a laptop computer was used for experimental control. The tests were presented in the form of a computer game via a child-friendly interface with animated graphics marking observation and response intervals, and providing feedback regarding correct response. In accord with the game-like graphics of the procedure, children were instructed to *help the mouse find the cheese* by listening carefully and deciding whether the first or last sound was different than the second sound. Children indicated their response by clicking on the appropriate screen graphic with the task unpaced in that response time was unrestricted. Young adults were tested using a protocol developed for clinical use. For both conditions, trial blocks were recorded on a CD for subsequent playback with the carrier word *ready* spoken by an adult male preceding each trial. Listeners were instructed to verbally indicate, during a 3.5 s response interval that followed each trial, whether the first or last sound was different than the second sound. Correct-response feedback was not provided. For both subject groups, a five-trial version of each task was used for familiarization before data collection began. If the

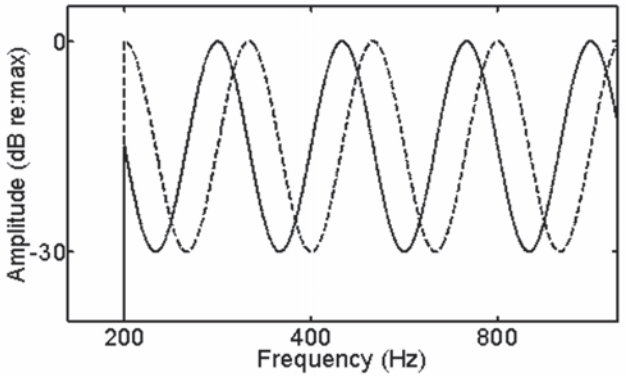

Figure 1. Schematic illustration of the spectral-ripple condition showing the contrasting amplitude spectra of a discrimination trial with difference due to change in starting phase of the spectral ripple.

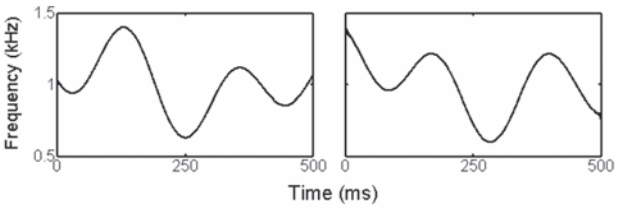

Figure 2. Schematic illustration of stochastic frequency modulation showing the contrasting instantaneous frequency functions of two stimuli of a discrimination trial.

experimenter deemed that a listener was unclear on procedure, the familiarization was repeated. Extensive training in a given listening task is often provided in psychoacoustic studies. Intended as clinical measures, training apart from the brief familiarization was not incorporated in the present protocols. Consequently, results from neither subject group represent performance levels that might be attainable with training on the listening tasks.

Young adults were tested in a double-walled sound-proof booth. Testing of children was conducted in a quiet room (ambient noise level less than 40 dBA) with close proximity of the experimenter to redirect a non-attentive child to the task. All stimuli were presented diotically, with Sennheiser HD 280 Pro headphones used with children and Etymiotic ER-3A insert earphones with young adults. The equipment used in testing the young adults was professionally calibrated as part of the routine maintenance of the audiology clinics at Rush University. Calibration of the laptop computer used in testing children was done by matching calibration tone levels measured through a Knowles electronic manikin for acoustic research (KEMAR) to levels obtained with the clinical setup also measured through KEMAR.

Speech measures

Since speech results were intended to be used only for within-group evaluation of relationship to psychoacoustic findings, different age-appropriate speech tests were used for the children and young adults. The Bamford-Kowal-Bench speech-in-noise test (BKB-SIN[23]) was used to measure children's speech perception in the presence of Auditec[24] four-talker speech babble. The sentences, approximately at a first-grade reading level, were originally constructed from language samples taken from young hearing-impaired children.[25] Performance was assessed based on responses to four sentence lists with a presentation level of 70 dB HL as specified in the standard BKB-SIN test protocol. Spoken by a male talker, each list consists of ten sentences. Between sentences of each list, the SNR decreased in 3-dB steps from 21 dB for the first sentence to -6 dB for the last. Based on the number of key words correctly repeated, results were converted to the metric SNR Loss, the estimated SNR needed for 50% correct relative the performance of normal-hearing listeners.[23] To ensure that the children understood the task, a separate practice list was administered before scored testing. For the young adults, speech perception was evaluated in terms of the intelligibility of sentences from the quick speech-in-noise test (QuickSIN[26]) in the presence of Auditec[24] four-talker speech babble. Presented at 70 dB HL, two scored lists of six sentences were used along with a single practice list. Across each list, the SNR decreased in 5-dB steps from 25 to 0 dB with results again converted to SNR Loss.[26] Both BKB-SIN and QuickSIN testing was conducted with diotic presentation. Manner of signal presentation and calibration was as described above for the psychoacoustic testing.

Results

Results from the ripple-phase and FM SNR conditions are shown as box plots in Figures 3 and 4, respectively, with data from monolingual

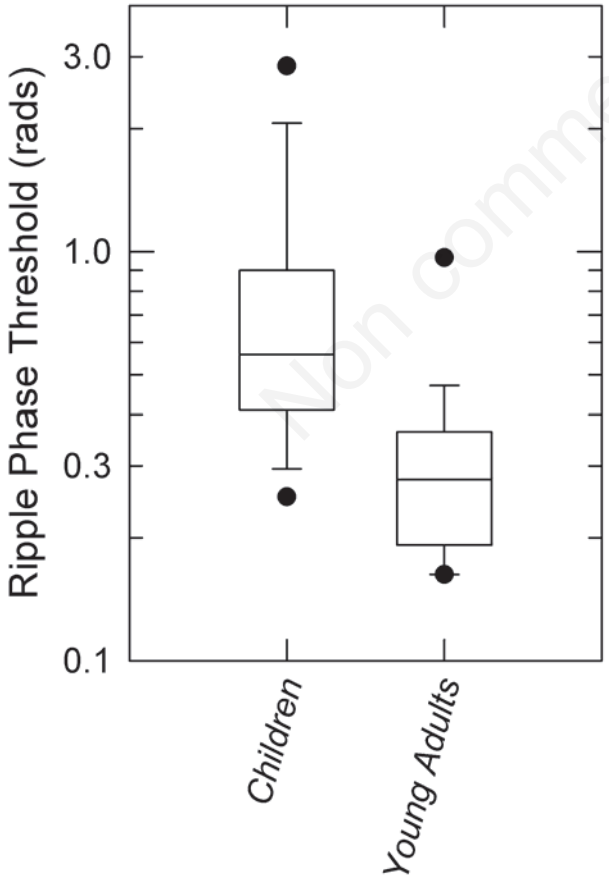

Figure 3. Results from the ripple-phase condition for each subject group shown as a box plot.

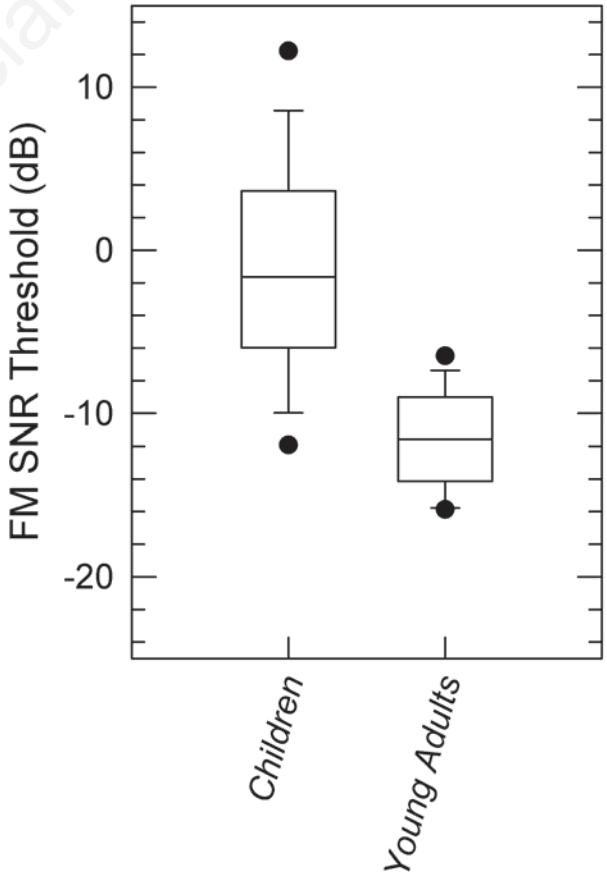

Figure 4. Results from the frequency modulation signal-to-noise ratio (FM SNR) condition for each subject group shown as a box plot.

and bilingual children grouped together. In both cases, the distribution of thresholds from the children was elevated compared to those from the young adults and also covered a wider range. From lowest to highest, the children's ripple-phase thresholds varied by a factor of 11.6 with the factor for young adults 6.2. The children's FM SNR thresholds varied by over 24 dB while the thresholds from the young adults varied by slightly less than 10 dB. Significantly better performance by young adults than children was confirmed by independent-samples t tests on logarithmically transformed ripple-phase thresholds [$t(38)$=-8.95, P<0.001, d=2.83] and FM SNR thresholds in dB [$t(26.13)$=-6.31, P<0.001, d=1.99]. In the FM SNR analysis, degrees of freedom were adjusted due to significance of Levene's test for equality of variances. When the children were partitioned either by gender or as monolingual versus bilingual, there were no significant differences between subgroups in performance on either test. For speech testing of young adults, the mean QuickSIN SNR Loss was 0.83 dB [standard deviation (SD)=1.20]. A one-sample t test showed that this result was significantly different than 0 dB, the normalized QuickSIN SNR Loss [$t(19)$=3.09, P=0.006, d=0.69]. With BKB-SIN testing of children, the mean SNR loss of 0.71 dB (SD=1.75) was also slightly above 0 dB, but in this case without significant difference. Partitioning of children by gender showed no significant difference. As anticipated, the SNR loss of monolingual children (M=-0.56, SD=1.56) was significantly lower than for bilingual children [M=1.55, SD=1.33, $t(18)$=3.24, P=0.005, d=1.46]. Among study participants, bilingual children were on average younger (M=8.60 yrs, SD=2.21) than monolingual children (M=10.16 years, SD=1.32). The better speech-in-noise performance of monolingual children, however, remained significant when using an analysis of covariance to control for the effect of age among the children [F (1, 17)=5.88, P=0.027, η_p^2=0.26].

Scatter plots of the relationships between speech-in-noise performance and each psychoacoustic measure for young adults and children are shown in Figures 5 and 6, respectively. For each subject group, Pearson product-moment correlation analysis was used to evaluate the relationships between the various measures. With no known result or model indicating an adverse relationship between the current measures, one-tailed significance testing was used. Among the young adults, there was a significant correlation between performance levels on the two psychoacoustic tasks (r=0.52, P=0.01). The scatter plots of Figure 5, however, illustrate that for young adults, results from neither the ripple-phase (r=0.10, P=0.35) nor FM SNR (r=0.12, P=0.31) task showed a significant relationship to speech perception as measured by QuickSIN.

A different pattern of relationship was obtained in the results from children. Similar to young adults, the ripple-phase results of children (Figure 6, left panel) were distributed across the range of BKB-SIN thresholds, with the relationship between the two measures not significant (r=0.08, P=0.37). In contrast, a significant result with children was obtained for the relationship between FM SNR and BKB-SIN performance (r=0.55, P=0.006; Figure 6, right panel). Full evaluation of relationships for children included age as a variable with results shown in Table 1.

Unlike young adults, the relationship between psychoacoustic measures was not significant (r=0.37, P=0.06). The scatter plots of Figure 7 illustrate the relationships between age and either ripple-phase (r=-0.36, P=0.06) or FM SNR (r=-0.56, P=0.005) performance. The significant negative correlation between age and FM SNR indicates lower thresholds (*i.e.*, better performance) with increasing age. A significant negative correlation to age was also obtained with the BKB-SIN SNR Loss (r=-0.64, P=0.001). Thus, for the significant relationship between FM SNR and BKB-SIN in children, both variables showed a significant negative correlation to age. When using partial correlation to control for the effect of age, the correlation between BKB-SIN and FM SNR performance dropped to 0.31 and was no longer significant (P=0.10), sug-

Table 1. For children only, pair-wise Pearson correlations between experimental measures with the one-tailed P value in parentheses.

	BKB-SIN	Ripple phase	FM SNR
Age	-0.64 (0.001)	-0.36 (0.06)	-0.56 (0.005)
BKB-SIN	-	0.08 (0.37)	0.55 (0.006)
Ripple phase	-	-	0.37 (0.06)

BKB-SIN, Bamford-Kowal-Bench speech-in-noise test; FM SNR, frequency modulation signal-to-noise ratio.

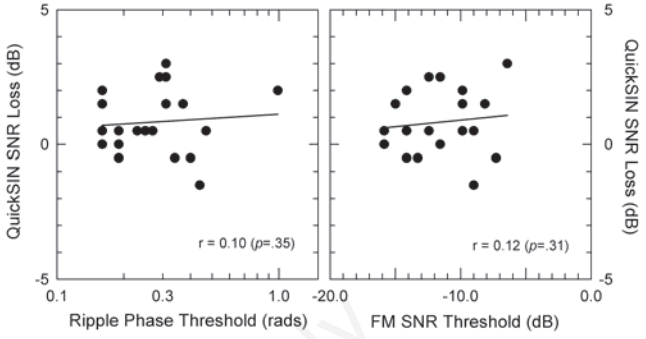

Figure 5. For young adults, scatter plots showing the relationship between QuickSIN signal-to-noise ratio (SNR) Loss and either ripple-phase (left panel) or frequency modulation (FM) SNR threshold (right panel). The solid lines are linear regressions of the data with correlation listed in the lower right of each panel.

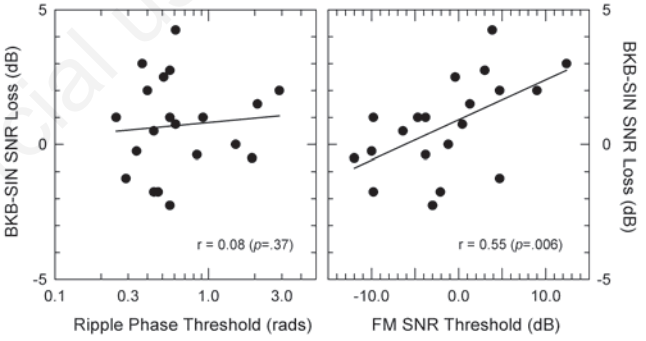

Figure 6. For children, scatter plots showing the relationship between BKB-SIN, Bamford-Kowal-Bench speech-in-noise test (BKB-SIN) signal-to-noise ratio (SNR) Loss and either ripple-phase (left panel) or frequency modulation (FM) SNR threshold (right panel). The solid lines are linear regressions of the data with correlation listed in the lower right of each panel.

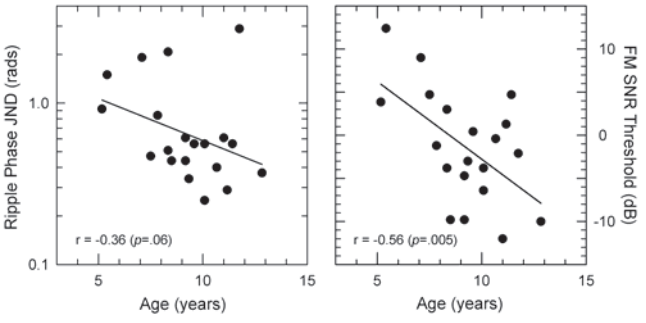

Figure 7. For children, scatter plots showing the relationship between age and either ripple-phase (left panel) or frequency modulation signal-to-noise ratio (FM SNR) threshold (right panel). The solid lines are linear regressions of the data with correlation listed in the lower left of each panel.

gesting that the relationship between the two measures was mediated by factors related to the children's age.

Age, bilingualism, and thresholds from the two psychoacoustic measures were used in a stepwise multiple regression analysis to predict BKB-SIN performance. The final model was significant [F (2, 17)=10.71, P=0.001], containing only age and bilingualism as the first and second predictors, respectively, with the psychoacoustic variables removed in stepwise analysis. The two predictors explained 51% of the data variance (adjusted R^2=0.51) with age significantly predicting BKB-SIN SNR Loss (β=-0.471, P=0.015) as did bilingualism (β=-0.425, P=0.027). For each predictor, the percent of data variance uniquely explained as estimated by squared semipartial correlation was 0.40 for age and 0.15 for bilingualism.

Discussion

The ability to discriminate both static and dynamic spectral patterns was investigated for both children and young adults with results evaluated in terms of relationship to speech-in-noise perception. Discrimination of static profiles was measured in terms of the just-detectable phase shift of a sinusoidal spectral ripple of wideband noise, while processing of dynamic patterns was evaluated as the SNR needed to discriminate the temporal pattern of frequency fluctuation imposed by stochastic low-rate FM. Compared to young adults, children performed more poorly on both discrimination tasks. The two subject groups differed in terms of gender with the young adults exclusively female. We are not aware of any published evidence indicating that psychoacoustic performance is affected by the gender of young normal-hearing adults. As described in the *Materials and Methods* section, identical procedures were not used for assessing the discrimination abilities of children and young adults. With self-pacing of trials and response feedback (procedural aspects that could benefit performance) incorporated only in the protocol used with children, it is unlikely that procedural differences could account for the poorer performance of the children. The young-adult participants were all monolingual while 12 of the 20 children were bilingual. We are aware of no evidence that bilingualism affects performance of normal-hearing adults on forced-choice psychoacoustic tasks. Coupled with the absence of a significant difference between the monolingual and bilingual children on either psychoacoustic measure in the current work, it is unlikely that bilingualism contributed to the group difference between the children and young adults.

In terms of relationships of discrimination ability to speech-in-noise perception, the only significant correlation involved the thresholds from children on the FM discrimination task. However, this relationship appeared mediated by factors related to the age of the children. As found in past studies,[17-18] the speech-in-noise thresholds of bilingual children were elevated compared to monolingual children. Stepwise multiple regression confirmed the importance of the factors age and bilingualism in predicting children's speech-in-noise performance with the psychoacoustic variables removed from the final model. Unlike most past studies of bilingual children, English was the primary language of the bilingual children in the current study, extending the effect of bilingualism beyond second-language learners. However, there was no significant difference between the monolingual and bilingual children on either static or dynamic spectral-pattern discrimination, offering no evidence of a psychoacoustic basis for their differing speech abilities.

Two developmental effects appear in the results. The first is the poorer performance of children on both psychoacoustic tasks when compared to young adults, and the second is the significant correlations of only the FM SNR thresholds, and not ripple-phase performance, to both age and speech-in-noise performance for children. The low correlation between children's performance on the two psychoacoustic tasks along with the finding that only one showed a significant correlation to age suggest that neither the elevated thresholds of children nor the relationship between their FM SNR and BKB-SIN performance was due solely to a global factor related to psychoacoustic procedure.

Dawes and Bishop[3] obtained adult-like performance from children for detecting the pitch associated with IRN, leading to speculation that complex pitch perception as indexed by IRN is an early developing skill. In the present study, the ripple-phase condition used a ripple density of 3.0 cpo which coincides with an augmented chord in music (*e.g.*, C E G#). These stimuli evoke a strong musical chord percept as if played on an organ so that thresholds represent the just-detectable change in the fundamental of a musical chord. Unlike the pitch result of Dawes and Bishop, adult-like performance was not obtained in the ripple-phase condition of the current work. There are important differences between the two paradigms, notably the distinction between the detection task used by Dawes and Bishop and the discrimination metric of the current work. Bishop *et al.*[6] reported that discrimination of complex pitch based only on upper harmonics was quite variable among children (age 8-10 yrs) with performance of some failing to converge on a threshold.

Across studies of frequency discrimination in children, large effects have been observed for a variety of stimulus and procedural parameters, affecting the size of the performance difference between children and adults.[27-29] Among these parameters, a significant deleterious effect has been associated with frequency uncertainty when introduced as a frequency rove. The ripple-phase condition of the present study incorporates stimulus uncertainty, randomizing ripple starting phase on each trial. Stimulus uncertainty is also a central aspect of the FM SNR condition in which listeners discriminated between stochastic patterns of FM on each trial. Past studies thus suggest that the elevated discrimination thresholds of children compared to young adults in the current work may, at least in part, reflect greater impact of stimulus uncertainty on performance.

Uncertainty was incorporated into the current stimulus sets by Sheft *et al.*[16] with the intent to enhance potential relationship between psychoacoustic performance and speech perception. In that work, significant relationships to speech-in-noise ability were found for both the ripple-phase and FM SNR conditions mediated by aging among adult listeners. For normal-hearing young adults, a relationship was obtained involving the ripple-phase metric when the intelligibility of speech-in-noise (SPIN-R) sentences[30] was measured at a fixed SNR. In a separate study with young adults, significant correlation between FM SNR performance and the intelligibility of both SPIN-R sentences and monosyllabic words was obtained when the speech material was vocoded.[31] In these past studies, the correlations between psychoacoustic and speech performance varied from roughly 0.5 to 0.7. Based on these results, *a priori* power analysis for the current work indicated 11 to 21 as the requisite number of subjects. With 20 young-adult listeners in the present study, the absence of significant correlation between psychoacoustic and speech performance does not appear to be due to sample size, but rather suggests a role in the relationship for stimulus parameters (*i.e.*, masking or vocoding) and intelligibility metric (*i.e.*, SNR Loss or percent correct).

The FM SNR thresholds of children in the current work did show significant correlation to BKB-SIN SNR Loss with a relationship to speech absent for their ripple-phase performance. To account for the differing effects of children's age on IRN and FM detection, Dawes and Bishop[3] suggested possible involvement of temporal focus of auditory attention. Their argument was that attention lapses would have greater impact for low-rate FM stimuli which vary during the observation interval than for static IRN. Assuming that the frequency and extent of attention lapses are on average inversely related to a child's age, a similar argument could account for the current finding of significant correlation of

age to only discrimination of low-rate stochastic FM. Alternatively, based on results for detection of amplitude and frequency modulation, Banai *et al.*[7] suggested the possibility of different developmental trajectories for task-specific short-term auditory memory. Keller and Cowan[27] demonstrated developmental differences in pitch-memory persistence. The memory requirements of the current ripple-phase and FM SNR conditions differ, with the former presenting a single fundamental of a musical chord during an observation interval and the later needing retention of the distinctive features of a dynamic pitch pattern. Age-dependent differences in memory capabilities could then possibly contribute to task-dependent effects of age on psychoacoustic performance.

The relationship between FM SNR and BKB-SIN thresholds was no longer significant when controlling for the effect of age. This age-mediated relationship may be due either to a single common factor, such as ability to temporally focus attention, or concurrent maturation of multiple factors. The developmental trajectory of children's speech-in-noise ability is affected by both masker type and speech material, indicating involvement of multiple factors.[32-34] It is thus unlikely that maturation of a single factor in children mediated the relationship between the ability to discriminate random patterns of low-rate FM and speech-in-noise perception.

Conclusions

The ability of monolingual and bilingual children and young adults to discriminate static and dynamic spectral patterns was evaluated in terms of relationship to speech-in-noise perception. Compared to young adults, children performed more poorly on both discrimination tasks. As found in past studies, the speech-in-noise thresholds were higher for bilingual than monolingual children. With no significant difference between the monolingual and bilingual children on either discrimination task, no evidence was found of a psychoacoustic basis for their differing speech abilities. In terms of relationships of discrimination ability to speech-in-noise perception, the only significant correlation involved discrimination of dynamic spectral patterns. However, partial correlation suggested that this relationship was mediated by factors related to the children's age. The effect of children's age on the discrimination of complex stimuli and its relationship to speech thus indicated involvement of factors selective to temporally dynamic rather than static stimuli.

References

1. Buss E, Hall JW, Grose JH. Development of auditory coding as reflected in psychophysical performance. In: Werner LA, Fay RR, Popper ANF, eds. Human auditory development. New York: Springer Science and Business Media; 2012. pp 107-36.
2. Allen PA, Wightman F. Spectral pattern discrimination by children. J Speech Hear Res. 1992;35:222-33.
3. Dawes P, Bishop DVM. Maturation of visual and auditory temporal processing in school-aged children. J Speech Lang Hear Res 2008;51:1002-15.
4. Yost WA, Patterson RD, Sheft S. A time domain description for the pitch strength of iterated rippled noise. J Acoust Soc Am 1996;99: 1066-78.
5. Keith RW. SCAN-C test for auditory processing disorders in children-revised. San Antonio, TX: Psychological Corporation; 2000.
6. Bishop DVM, Carlyon RP, Deeks JM, Bishop SJ. Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. J Speech Lang Hear Res 1999;42:1295-310.
7. Banai K, Sabin AT, Wright BA. Separable developmental trajectories for the ability to detect amplitude and frequency modulation. Hear Res 2011;280:219-27.
8. Moore, DR, Cowan, JA, Riley A, Edmondson-Jones AM, Ferguson MA. Development of auditory processing in 6- to 11-yr-old shildren. Ear Hear 2011;32:269-85.
9. Sheft S, Shafiro V, Lorenzi C, McMullen R, Farrell C. Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception. Ear Hear 2012;33:709-20.
10. Dorman MF, Cutting JE, Raphael LJ. Perception of temporal order in vowel sequences with and without formant transitions. J Exp Psych Hum Percept Perf 1975;104:121-9.
11. Spitzer SM, Liss JM, Mattys SL. Acoustic cues to lexical segmentation: a study of resynthesized speech. J Acoust Soc Am 2007;122:3678-87.
12. Binns C, Culling JF. The role of fundamental frequency contours in the perception of speech against interfering speech. J Acoust Soc Am 2007;122:1765-76.
13. Poelmans H, Luts H, Vandermosten M, Boets B, Ghesquière P, Wouters J. Reduced sensitivity to slow-rate dynamic auditory information in children with dyslexia. Res Development Dis 2011;32:2810-19.
14. Vandewalle E, Boets B, Ghesquière P, Zink I. Auditory processing and speech perception in children with specific language impairment: relations with oral language and literacy skills. Res Development Dis 2012;33:635-44.
15. Drullman R, Festen JM, Houtgast T. Effect of temporal modulation reduction on spectral contrasts in speech. J Acoust Soc Am 1996;99:2358-64.
16. Sheft S, Risley R, Shafiro V. Clinical measures of static and dynamic spectral-pattern discrimination in relationship to speech perception. In: Dau T, Jepsen ML, Poulsen T, Dalsgaard JC, eds. Speech perception and auditory disorders. Ballerup: The Danavox Jubilee Foundation; 2012. pp 481-8.
17. Crandell CC, Smaldino, JJ. Speech perception in noise by children for whom English is a second language. Am J Audiol 1996;5:47-51.
18. Nelson P, Kohnert K, Sabur S, Shaw D. Classroom noise and children learning through a second language: double Jeopardy? Lang Speech Hearing Serv Schools 2005;36:219-29.
19. Krizman J, Marian V, Shook A, Skoe E, Kraus, N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc Natl Acad Sci 2012;109:7877-81.
20. Parbery-Clark A, Skoe E, Lam C, Kraus N. Musician enhancement for speech-in-noise. Ear Hear 2009;30:653-61.
21. ANSI. Specifications for audiometers. [ANSI S3.6-2004]. New York: American National Standards Institute; 2004.
22. Byrne D, Dillon H, Tran K, Arlinger S, Wilbraham K, Cox R, et al. An international comparison of long-term average speech spectra. J Acoust Soc Am 1994;96:2108-20.
23. Etymōtic Research. Bamford-Kowal-Bench speech-in-noise test [Audio CD, version 1.03]. Elk Grove Village, IL: Etymōtic Research; 2005.
24. Auditec. Four-talker babble. St. Louis, MI: Auditec; 1971.
25. Bench J, Kowal A, Bamford J. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. Br J Audiol 1979;13:108-12.
26. Etymōtic Research. Quick speech-in-noise test [Audio CD, version 1.3]. Elk Grove Village, IL: Etymōtic Research; 2001.
27. Keller TA, Cowan N. Developmental increase in the duration of memory for tone pitch. Develop Psych 1994;30:855-63.

28. Sutcliffe P, Bishop D. Psychophysical design influences frequency discrimination performance in young children. J Exp Child Psych 2005;91:249-70.
29. Halliday LF, Taylor JL, Edmondson-Jones AM, Moore DR. Frequency discrimination learning in children. J Acoust Soc Am 2008;123:4393-402.
30. Elliott LL. Verbal auditory closure and the speech perception in noise (SPIN) test. J Speech Hear Res 1995;38:1363-76.
31. Shafiro V, Sheft S, Gygi B, Ho K. The influence of environmental sound training on the perception of spectrally degraded speech and environmental sounds. Trends Amplif 2012;16:83-101.
32. Johnson CE. Children's phoneme identification in reverberation and noise. J Speech Lang Hear Res 2001;43:144-57.
33. Hall JW, Grose JH, Buss E, Dev MB. Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. Ear Hear 2002;23:159-65.
34. Wightman FL, Kistler DJ. Informational masking of speech in children: effects of ipsilateral and contralateral distractors. J Acoust Soc Am 2005;118:3164-76.