

Sensitivity of school-aged children to pitch-related cues

Mickael L. D. Deroche,^{a)} Danielle J. Zion, Jaclyn R. Schurman, and Monita Chatterjee
Cochlear Implants and Psychophysics Lab, Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742

(Received 2 June 2011; revised 9 February 2012; accepted 9 February 2012)

Two experiments investigated the ability of 17 school-aged children to process purely temporal and spectro-temporal cues that signal changes in pitch. Percentage correct was measured for the discrimination of sinusoidal amplitude modulation rate (AMR) of broadband noise in experiment 1 and for the discrimination of fundamental frequency (F0) of broadband sine-phase harmonic complexes in experiment 2. The reference AMR was 100 Hz as was the reference F0. A child-friendly interface helped listeners to remain attentive to the task. Data were fitted using a maximum-likelihood technique that extracted threshold, slope, and lapse rate. All thresholds were subsequently standardized to a common d' value equal to 0.77. There were relatively large individual differences across listeners: eight had relatively adult-like thresholds in both tasks and nine had higher thresholds. However, these individual differences did not vary systematically with age, over the span of 6–16 yr. Thresholds were correlated across the two tasks and were about nine times finer for F0 discrimination than for AMR discrimination as has been previously observed in adults. © 2012 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.3692230>]

PACS number(s): 43.66.Hg, 43.66.Ba, 43.75.Cd [EB]

Pages: 2938–2947

I. INTRODUCTION

Natural variations in the pitch of spoken utterances convey important linguistic and emotional information to the listener. It has been shown that infants, from a few days to 9 months of age, prefer attending to infant-directed speech than adult-directed speech, infant-directed speech being characterized by exaggerated intonation contours and by exhibited expression of emotion (Fernald, 1985; Fernald and Kuhl, 1987; Cooper and Aslin, 1994; Cooper *et al.*, 1997; Trainor *et al.*, 2000). Given this preference, some researchers have proposed that prosodic cues might somehow facilitate the early stages of language acquisition (Crystal, 1979; Mehler *et al.*, 1988). One of these early stages could involve, for instance, the development of the brain structures responsible for processing voices, which in adults correspond to the superior temporal left and right cortices. Grossmann *et al.* (2010) used near-infrared spectroscopy to show an increased activity in these cortices in response to the human voice as opposed to non-vocal sounds for 7-month-old infants but not for 4-month-old infants. In a second experiment, 7-month-old infants were presented with words spoken with happy, angry, or neutral prosody. The voice-processing region was activated more in response to emotional than neutral prosody, as well as the right inferior frontal cortex, which is associated with emotion perception. Thus variations in voice pitch, sometimes exaggerated when expressing an emotion, might help infants to categorize important aspects of speech. For young children speaking a tonal language, rapid changes in voice pitch within syllables can carry important linguistic information. For these children, therefore, the ability to hear

subtle differences in voice pitch is even more critical during the early years as they acquire their native language. Despite this potential functionality of pitch processing in spoken language, sensitivity of school-aged children to subtle cues signaling voice pitch has been little investigated. The present research aims to fill this gap.

A. Pitch perception in children

Pitch is a two-dimensional percept, consisting of tone height and tone chroma. Tone height refers to that aspect of pitch that continues to get higher as frequency increases. Tone chroma refers to the cyclical property of pitch, i.e., the fact that two tones, one octave apart, sound similar even though they are quite distinct with respect to tone height (Bachem, 1950). Interestingly, infants seem to perceive not only tone height but also tone chroma. For instance, when 3-month-old infants are presented with two successive melodic sequences of pure tones, the second sequence being a distorted version of the first one, infants display novelty reactions when the distortion consists in shifting the frequency of some tones through a seventh and a ninth but display little reaction when the distortion consists in octave shifts (Demany and Armand, 1984). Because chroma is relatively irrelevant for speech and perhaps because it must be reinforced by musical training, the sensitivity to chroma is lost for some adults; as a result tone height often dominates the sensitivity to pitch. Note that the official definition of pitch, i.e., “that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high” (ANSI, 1994), is itself very much embodied in the notion of tone height. The present study also refers to “pitch” in the sense of height, not chroma.

Pitch perception has been, and is still currently, largely investigated in adults. A robust finding is that pitch primarily arises from the region where partials are spectrally resolved

^{a)}Author to whom correspondence should be addressed. Present address: Department of Otolaryngology, Johns Hopkins University School of Medicine, 818 Ross Research Building, 720 Rutland Avenue, Baltimore, MD 21205. Electronic mail: mderoch2@jhmi.edu

in the auditory periphery (Carlyon and Shackleton, 1994; Shackleton and Carlyon, 1994; Gockel *et al.*, 2005; Ives and Patterson, 2008). Hence, it is of interest to know whether children can resolve partials in the same way as adults do, i.e., whether their hypothetical auditory filters have similar bandwidth as those of adults. Measures of auditory filters' bandwidth differed somewhat depending on the experimental procedure. Using the notched noise paradigm, Irwin *et al.* (1986) showed evidence for significant improvement in frequency resolution between 6 and 10 yr of age. Using variations of the same paradigm, Allen *et al.* (1989) and Hall and Grose (1991) found that frequency resolution of children was adult-like by 6 yr of age. Using psychological tuning curves, Olsho (1985) did not find evidence that frequency resolution of infants was different from that of adults. Overall, despite some small discrepancies in the data, school-aged children (6 yr of age and above) are likely to have adult-like auditory filters.

Children's ability to discriminate frequency differences has also been extensively investigated (Maxon and Hochberg, 1982; Jensen and Neff, 1993; Cooper, 1994; Thompson *et al.*, 1999). Overall the results suggest that young children have larger difference limens (DLs) than 8-yr-old (y.o.) children and adults. Data on the children's perception of pure tones, however, offer a very limited view of complex pitch processing. At the neurophysiological level, evidence from animal work indicates that complex pitch processing must involve cortical networks that are not required for pure-tone-based pitch perception. For instance, cats can be trained to discriminate sounds on the basis of the pitch of the missing fundamental (Heffner and Whitfield, 1976). Lesions of auditory cortex impair the cat's ability to derive the pitch of the missing fundamental (Whitfield, 1980) but spare the ability to make simple frequency discrimination (Elliott and Trahiotis, 1972). Thus sensitivity to complex pitch may well differ from sensitivity to frequency differences. Given that the auditory pathways in the cortex develop well into the late teenage years in humans (Moore and Linthicum, 2007), one might expect some developmental effects in complex pitch perception of school-aged children.

Complex pitch processing starts developing very early in life. Clarkson and Clifton (1985) showed that 7-month-old infants could categorize harmonic complexes with different partial numbers according to their respective F0 (160 or 200 Hz) and could perform the same categorization when F0 was missing. Thus the sensation of a single pitch caused by the grouping of simultaneous harmonic frequencies already exists by the first year of age. More interestingly, Clarkson and Rogers (1995) showed that 7-month-old infants succeeded in categorizing the respective pitch of complexes made of resolved partials (in which F0 was missing) but not the pitch of complexes made of unresolved partials. This result suggests that by the first year of age, the developing auditory system of infants is also equipped with a "dominant" region for pitch. Only a few studies, however, have investigated children's sensitivity to finer changes in F0. Recently, Stalinski *et al.* (2008) measured children's psychometric functions for the direction of F0 changes,

ranging from 4 to 0.1 semitones. Performance improved from 5 to 8 yr of age at which sensitivity was adult-like. However, this study focused more on musical pitch than voice pitch because their stimuli were complex tones synthesized with a piano timbre, shifted in F0 from a reference of 880-Hz F0. Thus it is difficult to know how relevant their result is to the processing of F0s in the range of the human voice. To our knowledge, no study to date has investigated F0 discrimination of simple harmonic complexes in school-aged children.

B. Processing of temporal envelope in children

Studies related to the processing of temporal envelope cues by normal hearing (NH) children have focused on detection of amplitude modulation (AM). Hall and Grose (1994) and Lorenzi *et al.* (2000) quantified temporal modulation transfer functions (TMTFs): Although thresholds were overall higher for children than for adults, changes with modulation frequency were similar to those in adults. In other words, children are relatively inefficient in processing AM but have TMTFs that are adult-like in shape. These studies investigated children's ability to *detect* temporal envelope cues; in contrast, few studies have focused on the ability of children to *discriminate* between such cues, e.g., in the simple form of an AM rate discrimination task. Rocheron *et al.* (2002) examined AM rate discrimination (as well as AM detection) to examine the differences between dyslexic and normal children in their ability to process temporal envelopes relevant for speech. Their result is further discussed relative to the data collected in experiment 1.

Today many young children, profoundly hearing impaired or deaf, receive cochlear implants (CIs) early in life. Several studies (e.g., Gilley *et al.*, 2008; Niparko *et al.*, 2010) suggest that earlier implantation provides children with an advantage over those who are implanted later. It seems clear that auditory stimulation and brain plasticity during the critical period (about 7 yr of age) are important for shaping the responsiveness of the auditory system to various aspects of speech. CIs transmit complex pitch cues primarily via the temporal envelope within specific frequency bands, resulting in pitch discrimination that is barely sufficient for gender and intonation recognition and not suitable for music perception. The greater neural plasticity in the critical period may help early-implanted CI children to overcome these device limitations and maximize their reception of F0-related information from the periodicity cues present in the temporal envelope. To evaluate, in future studies, whether early implanted children have an enhanced sensitivity to the temporal envelope cues delivered by CIs, a normative baseline is required for comparison with NH children of the same age.

C. Protocol targeted to a young population

The aim of the present experiments is to fill an apparent gap in the literature on children's auditory abilities by measuring their psychophysical sensitivity to cues signaling changes in amplitude modulation rate (AMR) and changes in F0.

Several studies have acknowledged that the poor performance of young children in pitch-related tasks might partly be due to the type of responses they are asked to provide. For instance, [Johnsrude et al. \(2000\)](#) showed evidence that labeling the direction of a pitch change is cognitively more demanding than discriminating a pitch difference. They had access to patients who had undergone surgical resection from the left or right temporal lobe for the relief of epilepsy. While these patients did not differ from normal listeners in their ability to discriminate pure tones, the patients who had excisions of the right (but not the left) temporal lobe had more difficulties than normal listeners to label the direction of pitch changes. It is also possible that at an early stage of development, the brain is not equipped to label a pitch change even though it can detect a pitch difference. In line with this view, [Cooper \(1994\)](#) and [Sergeant and Boyle \(1980\)](#) reported that children can discriminate pure tones without necessarily knowing whether this change was higher or lower in pitch. In addition to cognitive limitations, children might also face semantic and linguistic limitations. When asked to verbalize changes in pitch, young children often use, erroneously, terms associated with loudness when describing changes in pitch ([Andrews and Diehl, 1970](#); [Hair, 1981](#); [Van Zee, 1976](#)). In English, the words “high” and “low” are not restricted to pitch description, they also have spatial, emotional, and loudness connotations. In other languages, like Spanish and French, the descriptives of pitch have a single meaning, which results in a better ability to label pitch changes ([Costa-Giomi and Descombes, 1996](#)). Thus in a task that requires labeling the direction of a pitch change, part of the developmental effect observed with children is sometimes due to a better understanding of the pitch concept, not necessarily to a better perceptual ability. A task that does not involve verbalization of pitch relational concepts measures more accurately the genuine sensitivity of children to pitch. For instance, [Andrews and Madeira \(1977\)](#) showed that 6- to 8.5-y.o. children, who made many errors in indicating verbally which of two tones (one octave apart) was higher or lower in pitch, could nevertheless perform almost perfectly when the task was achieved by conditioning children to associate the pitch of each tone with the size of a pig (big or small) that the child had to move in its respective barn (big or small). The odd-man-out paradigm may not measure the direction of pitch change per se, but (a) it does not require children to verbalize or even understand the concepts of “high” and “low” pitch, and (b) it still provides information about the children’s ability to discriminate between two pitches. Further, unpublished studies in our lab indicate that in adults, results obtained with a similar paradigm are predictive of performance in a task in which listeners are asked to indicate the direction of pitch change. Therefore the present experiments used the odd-man-out paradigm in both the AMR discrimination (experiment 1) and the F0 discrimination (experiment 2) task. The experimental protocol was designed as a simple video game with picture-pointing responses and feedback, known to enhance accuracy ([Smith and Hodgson, 1970](#)), so that children were kept as engaged as possible with the task.

II. GENERAL EXPERIMENTAL METHOD

A. Listeners

Seventeen children, 9 male and 8 female, between 6.5 and 15.2 y.o., took part in the two experiments. The children (with parental consent) were paid for their participation. They all had pure tone thresholds less than 15 dB HL at frequencies between 0.25 and 8 kHz. None of the children had history of hearing problems. All children were given one or two practice blocks of 10 trials, prior to data collection, to ensure that they understood the task with the easiest condition (a 12-semitones difference in AMR or a 2-semitones difference in F0). The criterion performance for starting data collection was at least 8 trials correct out of 10 during one practice block.

B. Stimuli and conditions

For the AMR discrimination task, broadband Gaussian white noise was used as basis of all stimuli. The reference signals were always sinusoidally modulated at 100 Hz. The target signals were sinusoidally modulated at 0, 2, 4, 6, 8, 10, and 12 semitones above 100 Hz (100, 112.2, 126, 141.4, 158.7, 178.2, 200 Hz), resulting in seven experimental conditions. Consistent cues in the fine structure of the noise might be used by listeners in such an experiment. Rather than use a different random fine structure for each of the three intervals (which might have distracted listeners), the noise carrier was identical for the three intervals within a trial but was refreshed from one trial to the next.

For the F0 discrimination task, broadband sine-phase harmonic complexes were used as the basis of all stimuli, which provided very salient pitches. The reference signals always had a 100-Hz F0. The target signals had F0s at 0, 1/8, 1/6, 1/4, 1/2, 1, and 2 semitones above 100 Hz (100, 100.7, 101, 101.4, 102.9, 105.9, 112.2 Hz), resulting again in seven experimental conditions. All complexes had partials up to 20 kHz, such that the total number of partials decreased as F0 increased.

All stimuli were 500-ms long (except in the practice session in which they were 1-s long) and gated by 10-ms ramps (half a period of a raised cosine). In both tasks, the level of each stimulus, regardless of its AMR or its F0, was first equalized at 65 dB SPL. The loudness of a sound, however, also depends on the spectral distribution of its intensity. For the harmonic complexes, because the reference and target stimuli had slightly different spectral distributions of intensity, they could potentially be perceived at different loudness, thereby providing an additional cue for discrimination. The harmonic complexes were passed through the loudness model described by [Moore et al. \(1997\)](#). Figure 1 shows the excitation patterns (left panels) and the specific loudness patterns (right panels) for the two complexes that had the largest difference in F0. The model predicts that these two stimuli are equally loud (21.9 sones). Note that even a complex with a F0 of 200 Hz at 65 dB SPL is predicted to be perceived almost equally loud (21.7 sones). This model, however, only applies to steady-state signals. For the amplitude-modulated noise stimuli, it remained unclear whether the temporal

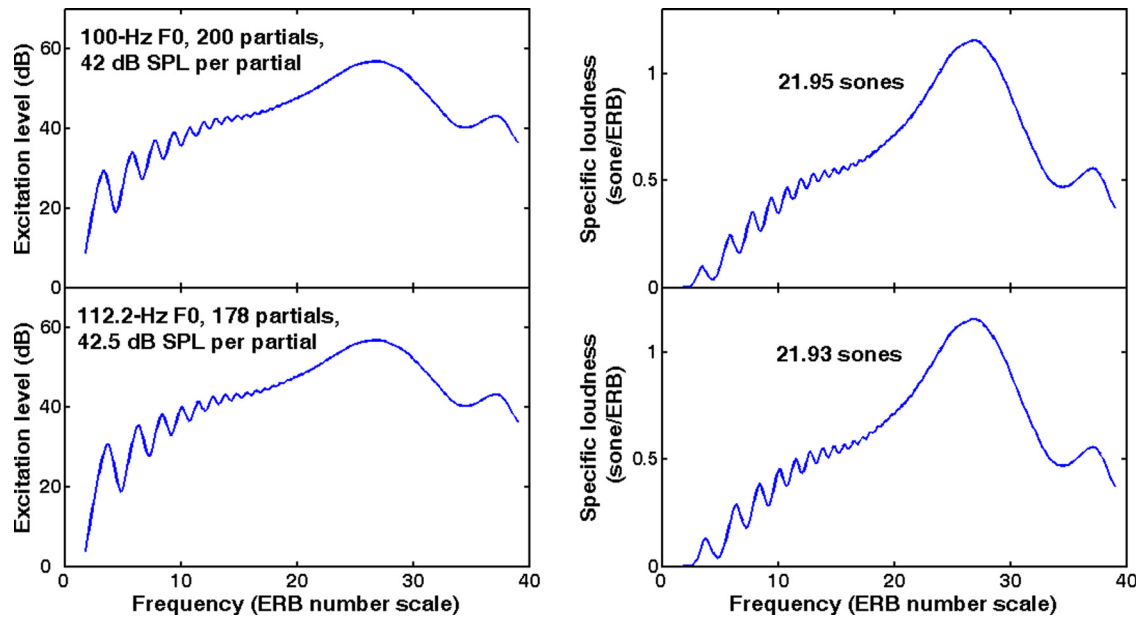


FIG. 1. (Color online) Patterns of excitation (left) and specific loudness (right) for two broadband (up to 20 kHz) harmonic complexes based on a F0 of 100 Hz (top) or two semitones above (bottom), at 65 dB SPL, based on the loudness model developed by Moore *et al.* (1997).

fluctuations caused any difference in loudness depending on their modulation rate. For this reason, a random level perturbation was added to each interval in both experiments, chosen from a uniform distribution of ± 1 dB, in order to discourage listeners from using loudness as alternative cues. The loudness model predicted that this level roving caused the harmonic complexes to vary in loudness between 20.7 and 23.2 sones.

C. Procedure

The method of constant stimuli was used in both tasks in a three-interval, three-alternative-forced-choice (3I-3AFC) procedure. Each of the seven conditions was tested 10 times, resulting in 70 trials per session. In each session, the 70 trials were presented in a random order, such that possible effects of perceptual learning or fatigue did not affect one particular condition. In each trial, the listener heard three intervals: two with the same AMR or same F0 and the target interval with a higher AMR or F0. The target interval was placed with equal probability in the first, second or third interval. The listener was asked to report which interval was different. Listeners participated in two, three, or four sessions (140–280 trials), which were separated by breaks as needed but all performed within a single day.

A child-friendly interface was built to collect data while keeping the youngest listeners reasonably entertained. The interface was displayed on a monitor inside the booth. Before starting one experimental session, the listener was asked to choose among different animated animals. The experiment was described as a game of sound discrimination. In the beginning of the session, three buttons appeared in the screen with the same chosen animal above each button as shown in Fig. 2. The first, second, and third animal became sequentially animated over the duration of the first, second, and third interval. Once the third interval stopped, the three buttons

and three animals reappeared on the screen, and the listener was asked to click on the button for which the sound was different. After the listener’s response, feedback was provided via the smiley face and some winning points (Fig. 2). Correct responses led to a smile and a 1-point gain. Three correct responses in a row led to a 3-points bonus and a super-smiley face. Incorrect response led to a sad face and no point. Three incorrect responses in a row led to a disappointed face and a 1-point loss. Verbal support and congratulations from the experimenters were given to each listener regardless of their performance. At the end of the session, the listener was rewarded for his/her valuable effort with “Silly Bandz” (synthetic bracelets that are presently popular with children of both genders) and a compensation of \$15.

Finally, although no instruction regarding response time (RT) was given to the listeners, RT was measured for each trial from the instant the three buttons reappeared on the screen to the instant the listener clicked on one button. They were recorded primarily to eliminate unreliable responses

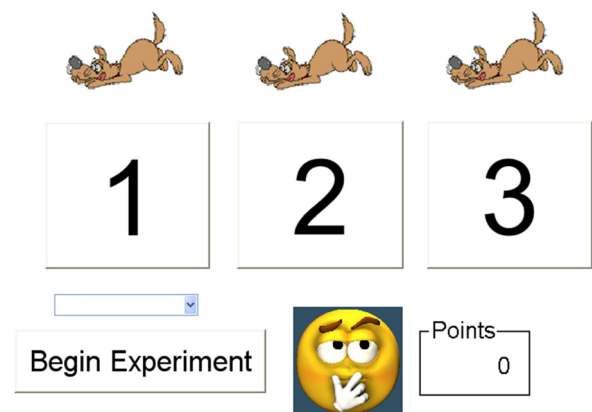


FIG. 2. (Color online) Screen shot of the child-friendly interface used in the 3I-3AFC procedure.

(as detailed further) and ensure that the listeners remained attentive to the tasks. All protocols were approved by the University of Maryland's Institutional Review Board.

D. Equipment

A computer monitor was inside the booth for presentation of the three intervals in a trial and listeners gave their response with the mouse. Depending on the listener's age, one of the experimenters remained in the booth if needed to help with using the mouse or simply to provide moral support. Signals were sampled at 44.1 kHz and 16 bits, digitally mixed, D/A converted by a 24-bit Edirol UA-25 sound card and presented diotically to listeners over Sennheiser HDA 200 headphones in a double-walled IAC sound-attenuating booth within a sound-treated room. None of the listeners had problems wearing the headphones.

III. EXPERIMENT 1. AMR DISCRIMINATION

The first experiment investigated the sensitivity of the 17 school-aged children to a purely temporal pitch-related cue, namely increments in the rate of sinusoidally modulated broadband noise around 100 Hz.

A. Results

To eliminate errors due to excessive inattention, responses provided more than 16 s after presentation of the three intervals were discarded on the basis that children would not be able to remember reliably the intervals they had heard. Adults can accurately reproduce pure tones (by adjusting the knob of a tone generator) following a silent interval of up to 16 s (Ross *et al.*, 2004). Beyond about 15 s, Bachem (1954) found that pitch memory is progressively disrupted. Therefore, when response time exceeded 16 s (perhaps because the child was momentarily inattentive or mistakenly thought he/she already provided a response), that particular response was not considered reliable. Over all trials and listeners, only two trials were discarded on this basis. Data were subsequently averaged over the several sessions that each listener performed.

Psychometric functions were fitted using the *psignifit* toolbox version 2.5.6 for MATLAB (see <http://bootstrap-software.org/psignifit/>) that implements the maximum-likelihood method described by Wichmann and Hill (2001). Contrary to the three other underlying functions offered by this toolbox (logistic, Gumbel, and cumulative Gaussian), the Weibull function is only defined for positive or zero values of x . In the present case, x represents the difference of AMR or difference of F0 and therefore negative values of x are meaningless. So the only valid model underlying the fitting of the present data was the Weibull function, described by the following formula in which the lower bound was fixed at chance level (33.3% in a 3I-3AFC).

$$\Psi(x; \alpha, \beta, \lambda) = \frac{1}{3} + \left(\frac{2}{3} - \lambda\right) \left(1 - \exp^{-(x/\alpha)^\beta}\right), \quad 0 \leq x < \infty$$

where Ψ is the percent correct score, x the difference of AMR or F0 in semitones, α and β the parameters influencing

the shape of the Weibull function and λ the lapse rate that was constrained with a flat Bayesian prior between 0% and 25%. A lapse rate of 0% results in an upper asymptote of 100%; larger lapse rates result in a lowered asymptote. Parameters derived from the psychometric function included discrimination threshold (in semitones, the half-way point between the lower and upper asymptotes) and slope (in % per semitone). Note that as a comparison, the fitting procedure was also run with the three other underlying functions (accepting negative values for x) and the extracted thresholds and slopes were very similar across the four models ($r^2 > 0.95$ for threshold and $r^2 > 0.87$ for slope in experiment 1, $r^2 > 0.90$ for threshold, and $r^2 > 0.70$ for slope in experiment 2). To obtain confidence intervals, 2000 Monte-Carlo simulations were generated from the best-fitting Weibull psychometric function. Threshold, slope, and lapse rate were extracted for each simulation, resulting in a distribution for each parameter, from which the 16% and 84% quantiles were chosen to provide confidence intervals corresponding roughly to plus and minus one standard deviation from the mean (assuming an approximately Gaussian distribution). As an example, Fig. 3 presents the mean data for one listener (6.75 yr of age) with the fitted psychometric function, resulting in a threshold of 3.0 semitones, a slope of 15.5% per semitone and a lapse rate of 0%.

Goodness-of-fit was evaluated using the method advocated by Wichmann and Hill (2001). All fits were within the 95% confidence limits of the Monte-Carlo-generated deviance distributions. All fits, except one (13.65-y.o. listener, whose lack of fit was partly due to a score of 50% at 0 semitone), were within the 95% confidence limits of the distribution of the Monte-Carlo-generated correlation coefficients between the deviance residuals and the percent correct predicted scores. Analysis of the deviance residuals as a

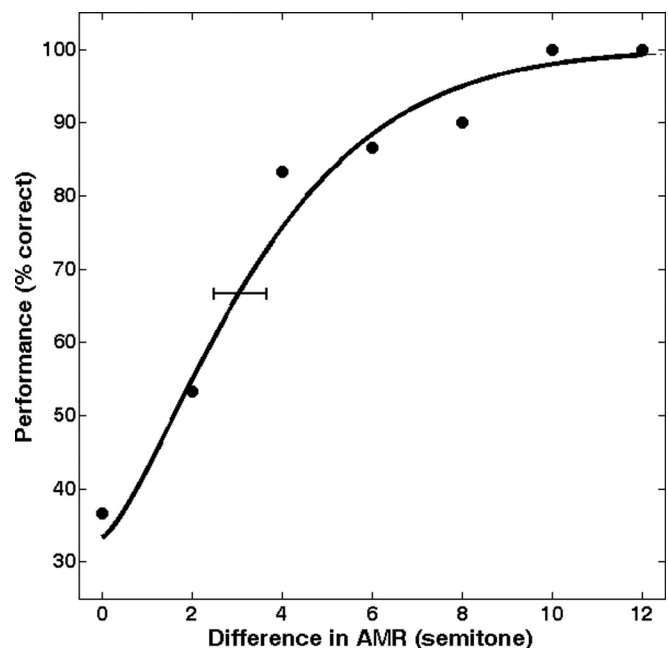


FIG. 3. Percentage correct of the AMR discrimination task, for one listener averaged over three sessions (210 trials). The line represents the psychometric function fitted to the data using the maximum-likelihood technique.

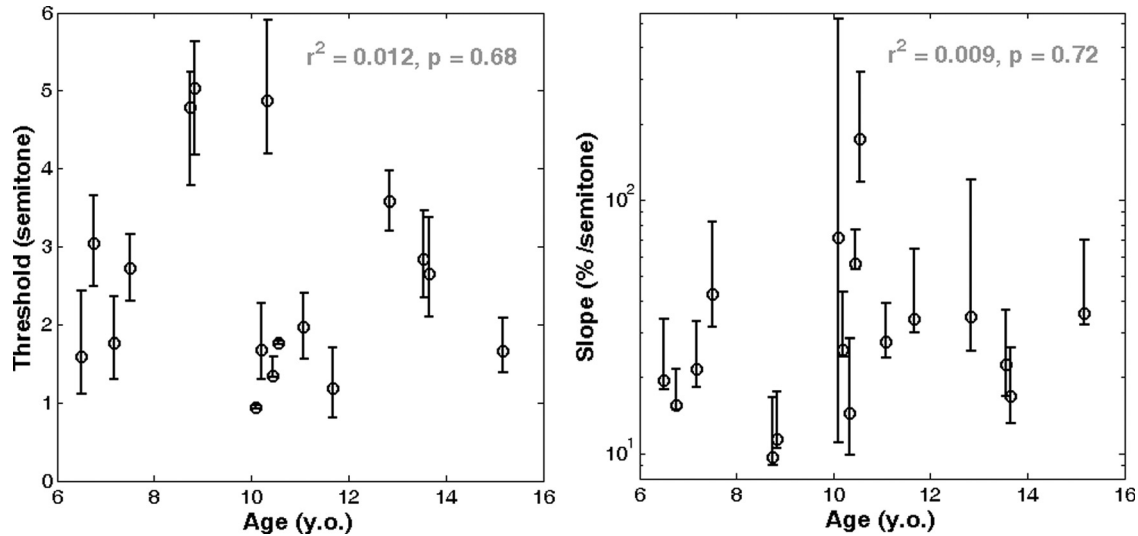


FIG. 4. AMR discrimination threshold (left) and slope (right) of the psychometric function for all listeners in experiment 1.

function of temporal order was not performed because the 70 trials of a given session were presented in a random order. Thus perceptual learning might have occurred but could not have systematically benefitted one particular experimental condition. The lapse rate was between 0% and 5.5% for all listeners, except one (10.33-y.o. listener) who had a lapse rate of 10.3%. Figure 4 shows threshold (left panel) and slope (right panel) for all listeners. There were large individual differences, but threshold did not vary systematically with age, with $r^2 = 0.01$ ($p = 0.68$). In Fig. 4 (right panel), the slope axis is on a logarithmic scale for clarity, but neither slope nor the logarithm of the slope were correlated with age, with $r^2 = 0.01$ ($p = 0.72$) and $r^2 = 0.04$ ($p = 0.43$), respectively.

B. Discussion

The maximum-likelihood technique does not define threshold at a given level of performance but at half way between the lower and upper asymptote. Ideally, i.e., for nine children who did not lapse, threshold corresponded to a

performance of 66.6%, which, in signal detection theory, was equivalent to a d' of 1.134 for a 3I-3AFC procedure. As lapse rate increased, the upper asymptote decreased and threshold corresponded to a lower level of performance. For seven listeners, lapse rate was below 5.5%, so that performance at threshold was between 64% and 66.6% and d' was between 1.043 and 1.134. For the 10.33-y.o. listener, threshold corresponded to a performance of 61.5% and a d' of 0.96. To compare the present data on a common baseline and with other results in the literature, the present thresholds were standardized to a common d' value of 0.77 (70.7% performance in a 2I-2AFC task), by assuming that d' is proportional to threshold. For instance, a threshold of 2 semitones at a d' of 1.134 corresponds to 1.36 semitones at a d' of 0.77. Figure 5 (left panels) shows the set of standardized thresholds, which is directly comparable to AMR discrimination thresholds that would be obtained in a 2-down/1-up adaptive procedure (Levitt, 1971). Again, there was no significant correlation with age ($r^2 < 0.01$). The mean standardized threshold over the 17 listeners was 1.8 semitones.

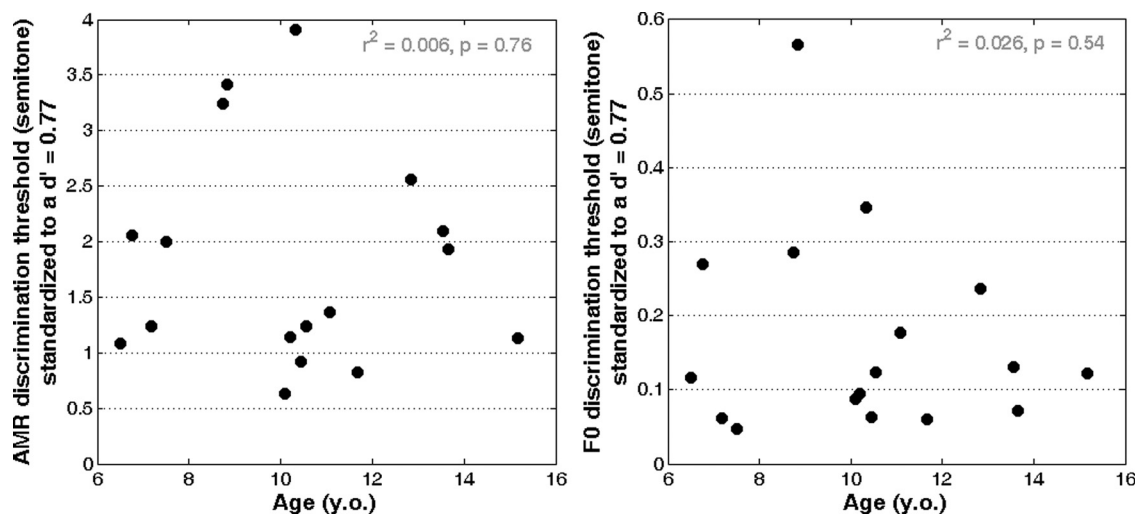


FIG. 5. AMR discrimination threshold (left) and F0 discrimination threshold (right) standardized at a d' value of 0.77 for comparison with thresholds obtained in a 2-down/1-up adaptive procedure.

Rocheron *et al.* (2002) also investigated AMR discrimination in four NH children, between 10 and 15 y.o. At 70.7% performance in a 2I-2AFC task and a AMR reference of 128 Hz, threshold was about 1.6 semitones. Therefore the thresholds obtained from the present data were roughly consistent with the results of Rocheron *et al.* (2002).

Formby (1985) found that threshold (defined as 70.7% correct in a 2I-2AFC task) for 100-Hz AM broadband noise was about 1 semitone in NH adults. Similar results were obtained by Hanna (1992) using noise carriers with a smaller bandwidth. Therefore it appears that children have, on average, slightly poorer sensitivity to AMR than adults. Analyzing individual thresholds in Fig. 5 (left panels), eight listeners had thresholds of 2 semitones (twice the adults' threshold reported by Formby or by Hanna) and above. Thus there were relatively large individual differences in the sensitivity to AMR. However, there was no developmental effect: The children who had adult-like sensitivity were not necessarily the oldest listeners. For instance, the youngest listener (6.5 yr of age) had a standardized threshold of only 1.08 semitones. Therefore the present data suggest that children's sensitivity to temporal pitch cues does not improve systematically beyond 6 yr of age.

Burns and Viemeister (1976) provided convincing evidence that envelope modulations elicit a real sensation of pitch because not only could listeners recognize the difference between two modulation rates as a musical interval, they could also recognize melodies the "notes" of which corresponded to those rates. The salience of this temporal pitch, however, varied with modulation rate, being strongest between 100 and 250 Hz and barely audible below 50 Hz or above 850 Hz, thus defining the existence region of temporal pitch (Burns and Viemeister, 1976). Modulation rates used in experiment 1 varied between 100 and 200 Hz, so they were relatively clear pitch-like percepts for this category of sounds.

IV. EXPERIMENT 2. F0 DISCRIMINATION

The second experiment investigated the sensitivity of the same 17 school-aged children to small differences in the F0 of equal-amplitude sine-phase harmonic complexes.

A. Results

Over all trials and listeners, only three trials were discarded because the responses were provided more than 16 s after presentation of the three intervals. Data were processed in exactly the same way as in experiment 1. Figure 6 presents, as an example, the mean data for one listener (11.08 yr of age) with the fitted psychometric function (Weibull), resulting in a threshold of 0.26 semitones, a slope of 121% per semitone, and a lapse rate of 0%. All fits were within the 95% confidence limits of the Monte-Carlo-generated deviance distributions. All fits, except two (10.10- and 13.55-y.o. listeners, whose data were not well fit due to scores of 20% and 13.3%, respectively, at 0 semitone), were within the 95% confidence limits of the distribution of the correlation coefficients between the deviance residuals and the percent correct predicted scores. Again, perceptual learning might have

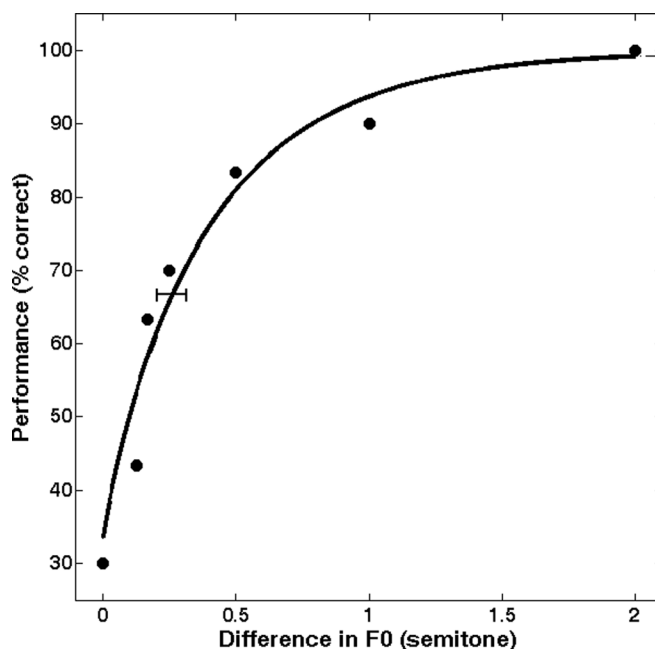


FIG. 6. Percentage correct of the F0 discrimination task, for one listener averaged over three sessions (210 trials). The line represents the psychometric function fitted to the data using the maximum-likelihood technique.

occurred but deviance residuals were not analyzed as a function of temporal order because the seven experimental conditions were shuffled. The lapse rate was between 0% and 4% for all listeners, except two (6.75- and 10.33-y.o. listeners) who had a lapse rate of 18% and 20%. Figure 7 shows threshold (left panel) and slope (right panel) for all listeners.

There were large individual differences but threshold did not vary systematically with age, with $r^2 < 0.02$ ($p = 0.63$). In Fig. 7 (right panel), the slope axis is on a logarithmic scale for clarity, but neither slope nor the logarithm of the slope were correlated with age with $r^2 < 0.01$ ($p = 0.85$) and $r^2 = 0.01$ ($p = 0.69$), respectively.

B. Discussion

For 11 children who did not lapse, threshold corresponded to a performance of 66.6% and a d' of 1.134. For four of the children, lapse rate was below 4%, so that threshold corresponded to a performance between 64.8% and 66.6% and d' was between 1.07 and 1.134. For two listeners (6.75 and 10.33 yr of age), who lapsed substantially, threshold corresponded to a performance of 56.6% and 57.6% and d' was 0.80 and 0.83. The thresholds were again standardized to a common value of d' , equal to 0.77. Figure 5 (right panels) shows the set of standardized thresholds that is directly comparable to F0 discrimination thresholds that would be obtained in a 2-down/1-up adaptive procedure (Levitt, 1971). There was no significant correlation with age ($r^2 < 0.03$) and the mean standardized threshold over the 17 listeners was 0.17 semitones.

Moore and Glasberg (1990) reported that adults can discriminate a difference of about 0.1 semitones, measured at 79.4% performance in a 2I-2AFC task ($d' = 1.16$) and a F0 reference of 100 Hz. Assuming proportionality between threshold and d' , their threshold corresponds to 0.07

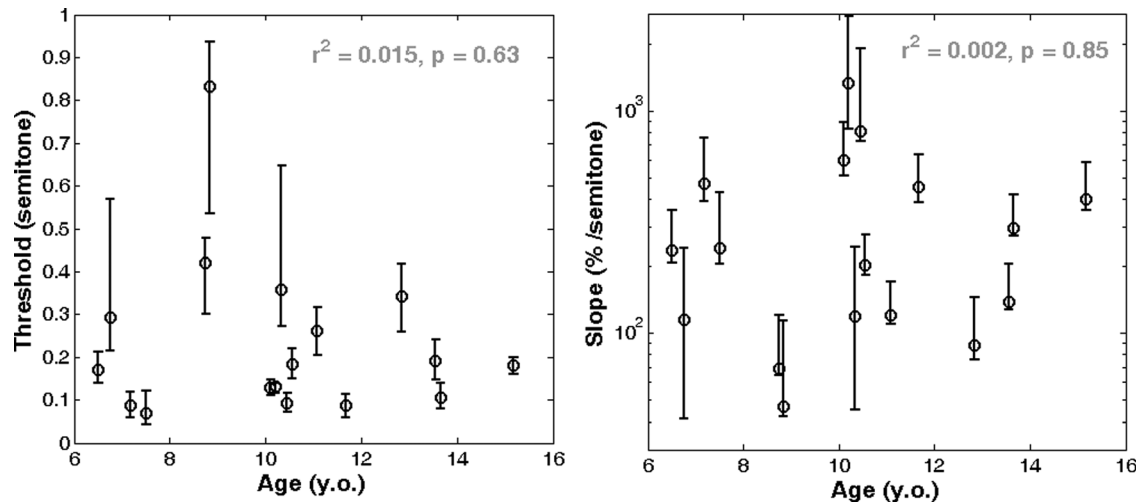


FIG. 7. F0 discrimination threshold (left) and slope (right) of the psychometric function for all listeners in experiment 2.

semitones for a d' of 0.77. Therefore, it appears that children have, on average, slightly poorer sensitivity to F0 than adults. Analyzing individual thresholds in Fig. 5 (right panels), six listeners had thresholds higher than twice the adults' threshold reported by Moore and Glasberg. Thus there were relatively large individual differences in the sensitivity to F0, but no developmental effect: The children who had adult-like sensitivity were not necessarily the oldest listeners. Therefore, the present data suggest that children's sensitivity to F0 does not improve systematically beyond 6 yr of age.

V. GENERAL DISCUSSION

A. Processing inefficiency and individual differences

One finding is recurrent in the pediatric literature, namely that children's basic auditory abilities are limited by "processing inefficiency." In many psychoacoustic tasks, ranging from simple discrimination of intensity and frequency (Maxon and Hochberg, 1982; Jensen and Neff, 1993) to more challenging tasks, such as speech recognition (Elliott, 1979; Nábělek and Robinson, 1982), children perform more poorly than adults. These deficits are sometimes attributed to increased internal noise. In some cases, the internal noise stands for an explicit jitter in peripheral encoding as for intensity discrimination (Schneider *et al.*, 1989; Buss *et al.*, 2006; Buss *et al.*, 2009) and in other cases, it stands for inefficiency of the central auditory system in extracting useful information (Hall and Grose, 1994; Stuart *et al.*, 2006), especially when the task involves informational masking (Wightman *et al.*, 2003; Hall *et al.*, 2005). In the present experiments, it is not quite possible to dissociate the developmental effects of processing efficiency from developmental effects of the particular auditory ability under investigation, because both tasks were tightly related to the processing of pitch-related cues. Nevertheless, several considerations can be made by comparing the psychometric parameters across the two tasks.

First, the lapse rate was not correlated across the two tasks (with one exception: The 10.33-y.o. listener had a large

lapse rate in both tasks). Thus the few attentional deficits seemed specific to the given task. Moreover, analyses showed that within each experiment, lapse rate did not covary with age (data not shown). Therefore there was neither evidence that some children suffered from global attentional deficits nor that young children suffered more from attentional deficits than older children, over the age span of 6–16 yr. The experimental protocol was designed to keep listeners involved in the task: Children were eager to gain as many points as possible and compete with each other when siblings attended the same test session. Furthermore, analysis of RT data showed that in both experiments children took about twice as much time to respond correctly when the cue was barely discriminable as when the cue was the most discriminable. The fact that children were more cautious to respond as the task difficulty increased is evidence that they actively attempted to perform well. Thus it appears that the protocol successfully reinforced a sustained level of attention.

Second, standardized thresholds between the two tasks were significantly correlated ($r^2 = 0.64$, $p < 0.001$), as illustrated in Fig. 8. This correlation might reflect the possibility that both tasks depend to some extent on a common processing efficiency or on a common underlying pitch-processing mechanism. Whether or not individual differences were caused by processing inefficiency, this correlation shows that listeners who were good at discriminating modulation rates were also good at discriminating close F0s.

Third, no developmental effect on threshold or on slope was observed in either of the two experiments. Some children displayed a considerably poorer sensitivity to pitch-related cues than adults, but age was not a relevant factor. In Fig. 8, eight listeners had thresholds scattered around the thresholds obtained with adults. Nine other listeners had poorer thresholds in one task or the other or both. These nine listeners were either inefficient in processing pitch-related cues or their sensitivity to pitch was still developing. Because they did not belong to a common age group, there was no evidence that the processing of pitch-related cues improves over the age span of 6–16 yr. Finally, individual

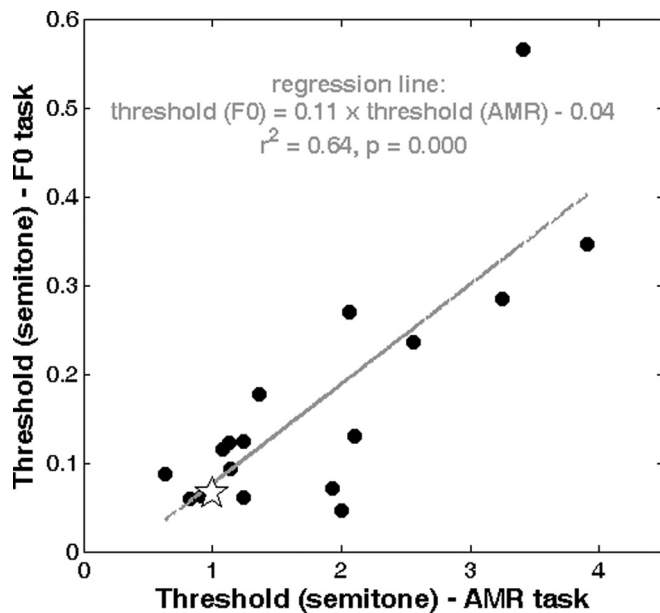


FIG. 8. Significant correlation between the standardized ($d' = 0.77$) thresholds extracted in the two tasks. Thresholds were about nine times finer for F0 discrimination than for AM rate discrimination. The star corresponds to the adults' performance taken from results in the literature (text).

differences in the sensitivity to pitch-related cues might also be partly accounted for by different exposures to musical education, but such data had not been collected. Among the nine males and eight females, no gender difference was apparent.

B. Sensitivity to temporal pitch poorer than sensitivity to spectro-temporal pitch

The results of the present experiments confirmed that the sensitivity of school-aged children to the F0 of sine-phase harmonic complexes is much finer than that of AMR by a factor of about 9 (slope of the regression line in Fig. 8). This difference in sensitivity is expected for at least three reasons. First, there are resolved partials in the excitation pattern of a harmonic complex that are known to contribute largely to pitch perception. In contrast, the long-term excitation pattern of an amplitude-modulated broadband noise does not vary with changes in modulation rate. Second, the fine structure a harmonic complex is periodic within each auditory filter, whereas that of an amplitude-modulated noise is not. Third, the envelope modulations of a harmonic complex can be very peaky at the output of basal auditory filters when all partials interact in phase. In contrast, the envelope modulations of amplitude-modulated noise are sinusoidal within each filter. Patterson *et al.* (1978) showed that the shape of the AM could influence the salience of the temporal pitch (square-wave versus sinusoidal).

NH children are not very sensitive to periodicity cues present in the sinusoidally modulated temporal envelope. In contrast, it is possible that early implanted children process similar cues more efficiently, i.e., brain plasticity might overcome the current limitations of CI devices to provide pitch-related cues. Further work with this population is needed to understand whether or not these implanted

children exhibit a better sensitivity to purely temporal pitch than the NH children of the present study and if so, how comparable they are to the sensitivity of NH children to F0.

VI. SUMMARY

The present experiments estimated psychometric functions for discrimination of AMR and F0 in school-aged children. Stimuli were respectively broadband noise bursts and sine-phase harmonic complexes. A 3I-3AFC constant-stimulus procedure with a child-friendly interface was used to estimate percent correct performance. Many children demonstrated adult-like sensitivity in both tasks. The sensitivity of some children was poorer than that of adults, but those were not necessarily the youngest listeners. As has been previously observed in adults, children had a finer sensitivity to F0 than to AMR.

ACKNOWLEDGMENTS

This work was supported by NIH Grants No. R01 DC004786, No. R01 DC004786-08S1, and No. R21 DC011905 to M.C. We wish to thank Fred Wightman and an anonymous reviewer for their thoughtful comments on this manuscript. We are very grateful to all listeners for their support and their interest in our experiments.

- Allen, P., Wightman, F., Kistler, D., and Dolan, T. (1989). "Frequency resolution in children," *J. Speech Hear. Res.* **32**, 317–322.
- Andrews, F. M., and Diehl, N. C. (1970). "Development of a technique for identifying elementary school children's musical concepts," *J. Res. Music Ed.* **18**, 214–222.
- Andrews, M. L., and Madeira, S. S. (1977). "The assessment of pitch discrimination ability in young children," *J. Speech Hear. Disord.* **42**, 279–286.
- ANSI. (1994). S1.1, *American National Standard Acoustical Terminology* (Acoustical Society of America, New York), 111–1994, 34.
- Bachem, A. (1950). "Tone height and tone chroma as two different pitch qualities," *Acta Psychol.* **7**, 80–88.
- Bachem, A. (1954). "Time factors in relative and absolute pitch determination," *J. Acoust. Soc. Am.* **26**, 751–753.
- Burns, E. M., and Viemeister, N. F. (1976). "Nonspectral pitch," *J. Acoust. Soc. Am.* **60**, 863–868.
- Buss, E., Hall, J. W., III, and Grose, J. H. (2006). "Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli," *J. Acoust. Soc. Am.* **120**, 2777–2788.
- Buss, E., Hall, J. W., III, Grose, J. H. (2009). "Psychometric functions for pure tone intensity discrimination: Slope differences in school-aged children and adults," *J. Acoust. Soc. Am.* **125**, 1050–1058.
- Carlyon, R. P., and Shackleton, T. M. (1994). "Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms?" *J. Acoust. Soc. Am.* **95**, 3541–3554.
- Clarkson, M. G., and Clifton, R. K. (1985). "Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental," *J. Acoust. Soc. Am.* **77**, 1521–1528.
- Clarkson, M. G., and Rogers, E. C. (1995). "Infants require low-frequency energy to hear the pitch of the missing fundamental," *J. Acoust. Soc. Am.* **98**, 148–154.
- Cooper, N. (1994). "An exploratory study in the measurement of children's pitch discrimination ability," *Psych. Music.* **22**, 56–62.
- Cooper, R. P., and Aslin, R. N. (1994). "Developmental differences in infant attention to the spectral properties of infant-directed speech," *Child Dev.* **65**, 1663–1677.
- Cooper, R. P., Abraham, J., Berman, S., and Statska, M. (1997). "The development of infants' preference for motherese," *Infant Behav. Dev.* **20**, 477–488.
- Costa-Giomi, E., and Descombes, V. (1996). "Pitch labels with single and multiple meanings: A study with French-speaking children," *J. Res. Music Ed.* **44**, 204–214.

- Crystal, D. (1979). "Prosodic development," in *Language Acquisition*, edited by P. J. Fletcher and M. A. Garman (Cambridge University Press, London), pp. 33–48.
- Demany, L., and Armand, F. (1984). "The perceptual reality of tone chroma in early infancy," *J. Acoust. Soc. Am.* **76**, 57–66.
- Elliott, D. N., and Trahiotis, C. (1972). "Cortical lesions and auditory discrimination," *Psych. Bull.* **77**, 198–222.
- Elliott, L. L. (1979). "Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability," *J. Acoust. Soc. Am.* **66**, 651–653.
- Fernald, A. (1985). "Four-month-old infants prefer to listen to motherese," *Infant Behav. Dev.* **8**, 181–195.
- Fernald, A., and Kuhl, P. (1987). "Acoustic determinants of infant preference for motherese," *Infant Behav. Dev.* **10**, 279–293.
- Formby, C. (1985). "Differential sensitivity to tonal frequency and to the rate of amplitude modulation of broadband noise by normally hearing listeners," *J. Acoust. Soc. Am.* **78**, 70–77.
- Gilley, P. M., Sharma, A., and Dorman, M. F. (2008). "Cortical reorganization in children with cochlear implants," *Brain Res.* **1239**, 56–65.
- Gockel, H., Carlyon, R. P., and Plack, C. J. (2005). "Dominance region for pitch: effects of duration and dichotic presentation," *J. Acoust. Soc. Am.* **117**, 1326–1336.
- Grossmann, T., Oberecker, R., Koch, S. P., and Friederici, A. D. (2010). "The developmental origins of voice processing in the human brain," *Neuron* **65**, 852–858.
- Hair, H. I. (1981). "Verbal identification of music concepts," *J. Res. Music Ed.* **29**, 11–21.
- Hall, J. W., III, Buss, E., and Grose, J. H. (2005). "Informational masking release in children and adults," *J. Acoust. Soc. Am.* **118**, 1605–1613.
- Hall, J. W., III, and Grose, J. H. (1991). "Notched-noise measures of frequency selectivity in adults and children using fixed-masker-level and fixed-signal-level presentation," *J. Speech Hear. Res.* **34**, 651–660.
- Hall, J. W., III, and Grose, J. H. (1994). "Development of temporal resolution in children as measured by the temporal modulation transfer function," *J. Acoust. Soc. Am.* **96**, 150–154.
- Hanna, T. E. (1992). "Discrimination and identification of modulation rate using a noise carrier," *J. Acoust. Soc. Am.* **91**, 2122–2128.
- Heffner, H., and Whitfield, I. C. (1976). "Perception of the missing fundamental by cats," *J. Acoust. Soc. Am.* **59**, 915–919.
- Irwin, R. J., Stillman, J. A., and Shade, A. (1986). "The width of the auditory filter in children," *J. Exp. Child Psych.* **41**, 429–442.
- Ives, D. T., and Patterson, R. D. (2008). "Pitch strength decreases as F0 and harmonic resolution increase in complex tones composed exclusively of high harmonics," *J. Acoust. Soc. Am.* **123**, 2670–2679.
- Jensen, J. K., and Neff, D. L. (1993). "Development of basic auditory discrimination in preschool children," *Psych. Sci.* **4**, 104–107.
- Johnsrude, I. S., Penhune, V. B., and Zatorre, R. J. (2000). "Functional specificity in the right human auditory cortex for perceiving pitch direction," *Brain* **123**, 155–163.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Lorenzi, C., Dumont, A., and Füllgrabe, C. (2000). "Use of temporal envelope cues by children with developmental dyslexia," *J. Speech Hear. Res.* **43**, 1367–1379.
- Maxon, A. B., and Hochberg, I. (1982). "Development of psychoacoustic behavior: sensitivity and discrimination," *Ear Hear.* **3**, 301–308.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). "A precursor of language acquisition in young infants," *Cognition* **29**, 143–178.
- Moore, B. C. J., and Glasberg, B. R. (1990). "Frequency discrimination of complex tones with overlapping and non-overlapping harmonics," *J. Acoust. Soc. Am.* **87**, 2163–2177.
- Moore, B. C. J., Glasberg, B. R., and Baer, T. (1997). "A model for the prediction of thresholds, loudness, and partial loudness," *J. Audio Eng. Soc.* **45**, 224–240.
- Moore, J. K., and Linthicum, F. H., Jr. (2007). "The human auditory system: A timeline of development," *Int. J. Audiol.* **46**, 460–478.
- Nábělek, A. K., and Robinson, P. K. (1982). "Monaural and binaural speech perception in reverberation for listeners of various ages," *J. Acoust. Soc. Am.* **71**, 1242–1248.
- Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., and Fink, N. E. (2010). "Spoken language development in children following cochlear implantation," *J. Am. Med. Assoc.* **303**, 1498–1506.
- Olsho, L. W. (1985). "Infant auditory perception: Tonal masking," *Infant Behav. Dev.* **8**, 371–384.
- Patterson, H. D., Johnson-Davies, D., and Milroy, H. (1978). "Amplitude-modulated noise: The detection of modulation versus the detection of modulation rate," *J. Acoust. Soc. Am.* **65**, 1904–1911.
- Rocheron, I., Lorenzi, C., Füllgrabe, C., and Dumont, A. (2002). "Temporal envelope perception in dyslexic children," *Neuroreport* **13**, 1683–1687.
- Ross, D. A., Olson, I. R., Marks, L. E., and Gore, J. C. (2004). "A nonmusical paradigm for identifying absolute pitch possessors," *J. Acoust. Soc. Am.* **116**, 1793–1799.
- Schneider, B. A., Trehub, S. E., Morriongiollo, B. A., and Thorpe, L. A. (1989). "Developmental changes in masked thresholds," *J. Acoust. Soc. Am.* **86**, 1733–1742.
- Sergeant, D., and Boyle, J. D. (1980). "Contextual influences on pitch judgement," *Psych. Music* **8**, 3–15.
- Shackleton, T. M., and Carlyon, R. P. (1994). "The role of resolved and unresolved harmonics in pitch perception and frequency modulation," *J. Acoust. Soc. Am.* **95**, 3529–3540.
- Smith, K. E., and Hodgson, W. R. (1970). "The effects of systematic reinforcement on the speech discrimination responses of normal and hearing impaired children," *J. Auditory Res.* **10**, 110–117.
- Stalinski, S. M., Schellenberg, E. G., and Trehub, S. E. (2008). "Developmental changes in the perception of pitch contour: distinguishing up from down," *J. Acoust. Soc. Am.* **124**, 1759–1763.
- Stuart, A., Givens, G. D., Walker, L. J., and Elangovan, S. (2006). "Auditory temporal resolution in normal-hearing preschool children revealed by word recognition in continuous and interrupted noise," *J. Acoust. Soc. Am.* **119**, 1946–1949.
- Thompson, N. C., Cranford, J. L., and Hoyer, E. (1999). "Brief-tone frequency discrimination by children," *J. Speech Lang. Hear. Res.* **42**, 1061–1068.
- Trainor, L. J., Austin, C. M., and Desjardins, R. N. (2000). "Is infant-directed speech prosody a result of the vocal expression of emotion?" *Psych. Sci.* **11**, 188–195.
- Van Zee, N. (1976). "Responses of kindergarten children to musical stimuli and terminology," *J. Res. Music Ed.* **24**, 14–21.
- Whitfield, I. C. (1980). "Auditory cortex and the pitch of complex tones," *J. Acoust. Soc. Am.* **67**, 644–647.
- Wichmann, F. A., and Hill, N. J. (2001). "The psychometric function. I. Fitting, sampling and goodness-of-fit," *Percept. Psychophys.* **63**, 1293–1313.
- Wightman, F. L., Callahan, M. R., Lutfi, R. A., Kistler, D. J., and Oh, E. (2003). "Children's detection of pure-tone signals: informational masking with contralateral maskers," *J. Acoust. Soc. Am.* **113**, 3297–3305.